UNIVERSITÀ
DEGLI STUDI
FIRENZE

Università degli Studi di Firenze
Dipartimento di Ingegneria dell'Informazione (DINFO)
Corso di Dottorato in Ingegneria dell'Informazione
Curriculum: Automatica, Ottimizzazione e Sistemi Complessi

Dynamics on complex networks: modeling the form to shape the substance

Candidate
Giulia Cencetti

Supervisors
Prof. Duccio Fanelli
Prof. Franco Bagnoli
Prof. Giorgio Battistelli
Prof. Luigi Chisci

PhD Coordinator
Prof. Luigi Chisci

ciclo XXXI, 2015-2018

Università degli Studi di Firenze, Dipartimento di Ingegneria
dell'Informazione (DINFO).

Thesis submitted in partial fulfillment of the requirements for the degree
of Doctor of Philosophy in Information Engineering.

Contents

Contents **iii**

Introduction **1**

1 Dynamics on complex networks: from mobility to equilibrium stability **7**
1.1 Complex networks 7
1.2 Random walk on networks 10
1.3 Reaction-diffusion systems 13
1.4 Equilibria stability in reaction-diffusion systems 15

2 Pattern invariant networks **19**
2.1 Dispersion relation and pattern formation 20
2.2 Topology modification 22
2.2.1 Eigenmode randomization 22
2.3 Local rewiring 24
2.4 Results and discussion 25
2.4.1 The Ginzburg-Landau model 25
2.4.2 Eigenmode randomization 25
2.4.3 Local rewiring 31

3 Spectral control for reaction-diffusion systems on networks **35**
3.1 Complex Ginzburg-Landau equation: linear stability analysis 37
3.2 Global topological control 43
3.3 Controlling the instability on balanced directed networks . . 47

4 Spectral control for ecological systems **55**
4.1 The model 56

4.2 Topological control scheme 58
4.3 Stabilizing a predator-prey dynamics 65
4.4 Generalization: a larger ecosystem 70

5 Control by adding one node to the network 75
5.1 Adding a species to enforce stable equilibria in a multidimensional system 77
5.2 Testing the control method: from synthetic gene network to real microbiota dataset 81
5.2.1 Stabilizing an unstable fixed point by means of an external controller (Case A). 85
5.2.2 Acting with one species of the pool to damp the concentration of the pathogens (Case B). 86
5.2.3 Driving to extinction one species, the other being the target of the control (Case C). 89
5.2.4 Controlling the Microbiota network: exploiting a transient control to drive the system towards an existing stable fixed point (Case D) 90
5.3 Generalized control 92
5.3.1 Strategy 1: single $\boldsymbol{\beta}$ approach 94
5.3.2 Strategy 2: optimization 95
5.3.3 Strategy 3: least squares 97
5.3.4 Strategy 4: LASSO regularization 98

6 Random walk and competition between traps 103
6.1 Competition between traps on asymmetric regular lattices . 106
6.2 Traps on a directed disordered network and the optimization problem 112

7 Reactive random walk 121
7.1 Model 122
7.2 Analytical derivation of the stationary state 131
7.3 Measures of functional ranking 134
7.4 Detecting network symmetries 137
7.5 Measuring degree correlations 139

Conclusions 147

Appendices 152

A Spectral stabilization: supplementary details **153**
A.1 Eigenmode randomization for directed networks 153
A.2 Local rewiring: acceptance threshold 158

B Mathematical details on ecosystem stability **159**
B.1 Couplings in the original system 159
B.2 Extending the analysis to account for a generalized nonlinear reaction term 161
B.3 The controlled matrix $\boldsymbol{D'}$ is real and zero-row-sum as $\boldsymbol{D}$ is 162
B.4 On the conditions of controllability 164
B.5 Controlling without modifying the fixed point 167

C Control adding one node: inside the model and the controllability **169**
C.1 Genetic network model 169
C.2 On the controllability condition 170

D Traps competition: supplementary details **175**
D.1 Explicit solution of the 1D Fokker-Planck with two absorbing boundaries 175
D.2 The case of a generic network: details of the calculation. . . 176

E Publications **179**

Bibliography **181**

Complex systems are, by definition, systems composed of many, possibly simple, elements interacting with each other. Such interactions can give rise to unexpected and remarkable collective behaviors. A vast plethora of physical phenomena occurring in nature can indeed be ascribed to collective dynamics which spontaneously emerge at the macroscopic level in systems made up of microscopically interacting constituents [48, 93]. This is a widespread observation in nature which, fertilized in a cross-disciplinary perspective, can ideally embrace distinct realms of investigation. Convection instabilities in fluid dynamics, weak turbulences and defects are among the examples that testify on the inherent ability of physical systems to yield coherent dynamical behaviors [58]. Insect swarms and fish schools exemplify the degree of spontaneous coordination that can be reached in ecological applications [140], while rhythm production and the brain functions refer to archetypical illustrations drawn from biology and life science in general [13, 17, 75, 142, 182, 202].

In many cases of interest, it is useful to schematize a complex system at the microscopic level as a network, where the nodes represent the components and the links their interactions [117, 143]. The architecture of numerous systems can be well described in terms of complex networks. Network theory is therefore able to provide new insights into many real world problems, with applications in biology, social sciences, transport infrastructures, communications, financial markets, and more.
In the simplest implementation of graph theory the network nodes configure as the theoretical representation of physical space regions, e.g. cities, urban areas, foraging spots, or even virtual spaces like web sites. In this configuration, the connections between nodes therefore represent the possible ways an agent (a citizen, an animal, a web surfer...) can use to move when exploring the network. However, graph theory can also provide a framework of coupling for general systems composed by individual entities which, either individually or grouped in families, can be schematized by mutually interacting vertices. Irrespectively of the specific realm of investigation, the topological structure of the network of interactions is therefore exemplified by directed or indirected edges among nodes [27, 28]. These fundamental ingredients, flexibly combined and properly integrated, are at the roots of any plausible mathematical model targeted to community interactions [37], from ecology [47] to neuroscience [16, 106], passing through genetic, human health [128], and a full load of man-made techno-

logical applications [169].
During the last years, network theory has proven to be a convenient, often crucial instrument not only to model the structure of many complex systems, but also to properly describe the dynamical processes they are involved into. Complex networks will configure, in our representation scheme, as the discrete irregular supports for specific dynamical processes built to describe, in a mathematical framework, the behavior of real world phenomena [21]. For each node we can indeed define a multidimensional variable whose evolution in time, modeled by resorting to a collection of ordinary differential equation, will represent one of the main characters of our analysis.
The dynamics can either simply stem from interactions, i.e. how adjacent nodes affect each others, like in random walk, spreading, diffusion or other linear processes, or it can additionally involve a local reaction. The single nodes of the network indeed, instead of being merely crossing points, could represent nonlinear dynamical systems, each one evolving according to specific differential equations. For instance, when the entities that compose the whole set are subject to specific self-reactions and, at the same time, diffuse across the embedding spatial medium, the system goes under the name of reaction-diffusion [10, 16, 139, 142, 155, 156, 192]. Examples are invasion models in ecology [92], epidemic spreading [138], and also the celebrated Turing patterns that arise, for instance, from the dynamical interplay between reaction and diffusion in a chemical system [192].
If we consider for a moment the dynamics that each node would exhibit if it was isolated, we observe that the long-term behavior of each of these generic dynamical systems is usually given by stable equilibria like fixed points and limit cycles or chaotic attractors. However, once the isolated systems are coupled together, so becoming the nodes of a network, nothing much can be a priori said about their collective behavior [183]. We know, for instance, that, if the single node dynamics admits a fixed point as an equilibrium, the insertion of couplings between them will lead to the emergence of static patterns, however these can be characterized by a wide number of locally stable equilibria [179]. At the same time, curiously, a network of identical chaotic systems can synchronize to the same solution [163]. Instead, if a set of uncoupled dynamical systems evolve towards limit cycles, once connected they can either synchronize or generate patterns that depend on the topology of the underlying network [45].

This rises one of the main points of our analysis: why are we interested in characterizing network structure? Because architecture always affects the dynamical behavior, the form always affects the substance. For instance, the topology of species interaction networks influences the ecosystems stability, as well as the structure of social networks shapes the spread of epidemics and information, and the network of power grids affects the robustness of power transmission [21, 183]. In general, the observed dynamical behavior is intimately connected to the underlying topology. Ingenious techniques have been proposed to reconstruct the topology of a given network from direct inspection of its emerging dynamics, combining fundamental [14,41,52] and applied expertise [42,126]. However, a general theory to fully resolve the subtle interplay between network structure and ensuing dynamics is still lacking and is considered today as one of the great challenges of network science. Our way to contribute to this open question involves thoroughly studying the repercussion of topology on the emerging dynamics with the twofold goal of predicting the system performance and, mainly, of influencing such functional behavior and steering it at will. Instead of altering the local dynamics on each network node, we focus here on devising possible innovative ways to customize the effective network structure in order to obtain various desired effects [149].

In the next chapters, after Chapter 1 which provides a brief introduction of the mathematical tools needed in the subsequent chapters, we will directly go inside the topic of topology-reflecting dynamics. In Chapter 2 we propose two different techniques to modify the network structure of a reaction-diffusion model, while preserving its dynamical behavior. In the region of parameters where the homogeneous solution gets spontaneously destabilized, perturbations grow along the unstable directions made available across the network of connections, yielding irregular spatio-temporal patterns. We exploit the spectral properties of the Laplacian operator associated to the graph in order to modify its topology, while preserving the unstable manifold of the underlying equilibrium. The resulting network is isodynamic to the former, meaning that it reproduces the dynamical response to a perturbation, as displayed by the original system. This work opens up a new perspective on the multiple possibilities for identifying the family of discrete supports that generate equivalent dynamical responses on a generic reaction-diffusion system.

The same fundamental idea is pursued in Chapter 3 where the spectral control is used to reach a different goal: instead of reproducing irregular network patterns, we aim to dampen the perturbations which bear instability, avoiding pattern formation. We thus act on the topology of the inter-node contacts (without altering the dynamical parameters of the model) in order to achieve stabilization of a synchronized dynamics where all the nodes reach consensus and oscillate together.
As anticipated before, dynamics on networks is central also when the system being examined is not spatially extended. In this case, the interactions, as encoded in the structure of the assigned network, might follow non-diffusive rules. In ecosystems, for instance, inter-species interactions are assumed to be mediated by pairwise, hence quadratic, exchange, to some level of approximation. Each population (species) is characterized by a self-reaction dynamics, typically described via a suitable nonlinear function of the concentration amounts, and different populations can be abstractly assigned to given nodes of a virtual graph [187]. The network then represents the interactions between different species and the sign of the weighted entries of the associated adjacency matrix define the nature of coupling (competitive, cooperative, predator-prey, etc...). The concept of stability is in such systems of paramount importance as it relates to resilience, the ability of the system to counteract external perturbations that would tend to get away from the existing equilibrium. It is therefore crucial to devise possible strategies aimed at interfering with the system of interest so as to enforce the desired stability [56, 70, 81, 188]. This topic is analyzed in detail in chapter 4.
By playing with ecosystem stability analysis, we can address different applications, for example, hostile pathogens could be forced to go extinct: the stability of the attained equilibrium would efficaciously shield from subsequent harmful invasion and outbreaks. Alternatively, it could prove vital to robustly enhance the expression of species identified as beneficial for the system at hand. Building on these premises, in Chapter 5 we develop and test a general control strategy targeted to systems consisting of a large number of components that interact through a complex network. By inserting one additional species, the controller, which configures as a further node of the collection, we will be able to locally manipulate the asymptotic dynamics of the system, in terms of existence and stability of the allowed fixed points.

In Chapter 6, a different control method is devised, where the process under exam is a random walk and we will analyze the way a set of microscopical agents asymptotically distribute on the different nodes when traveling a complex network. Random walks are one of the most fundamental types of stochastic processes. They can be used to model numerous phenomena and to extract information about the network structure. In this chapter we do not modify the network structure at the level of nodes and edges, but we deal with the role that specific nodes play in the random walk process. In particular we insert absorbing sinks for agents in the graph and study their performances so as to identify strategic positioning and consequently customize the flux of agents on graphs.
While random walks are the basic ingredient to describe mobility, they do not take into account the possibile interactions between agents present in the same node of a network. The last chapter before conclusions, Chapter 7 is devoted to study a process, called "reactive random walk" where, inspired by reaction-diffusion, we take up the challenge to associate reaction terms to the random walk process on network. Generalized biased random walkers not only navigate the system, but also interact when they meet at the nodes of the network. Contrarily to conventional random walk, for these walkers the probability of relocation between adjacent nodes is also sensitive to local reactions, which ultimately confers to each node a self-identity. For such a reason, the occupation probability of a given node depends not only on the connectivity pattern but also on the ability of the node itself to attract walkers. This last property can be tuned at will by properly shaping the reaction term, and this enables in turn to highlight different characteristics of the network structure. We will in particular focus on: *(i)* the definition of a novel functional centrality measure, *(ii)* the issue of revealing hidden symmetries in a graph and *(iii)* the problem of characterizing node degree-degree correlations in complex networks.
Finally, in the last chapter we will sum up and draw our conclusions, also discussing possible further extensions of the proposed models and techniques. Relevant technical material is provided in the Appendices.

Chapter 1

Dynamics on complex networks: from mobility to equilibrium stability

1.1 Complex networks

The mathematical branch devoted to the study of complex networks is called graph theory [117, 143], where with the word *graph* we refer to the mathematical structure implemented to represent pairwise relations between a set of items. The networks that model real systems, be they biological, social, environmental or other, are typically represented as a set of nodes, or vertices, connected by links, or edges. This general framework includes a huge amount of samples, thus giving an (initial) idea of the potential versatility of networks. Furthermore, the edges can be characterized by different weights quantifying the strength of the connections, in this case the network is called *weighted*, and *signed* if the weights can also assume negative values. Moreover a network is said *directed* if each edge is associated to a specific direction, implying a non symmetric relation between connected nodes. Each node is identified by an index $i = 1, ..., N$ and the mathematical tool which allows us to describe a network is the *adjacency matrix*, $\boldsymbol{A} \in \mathbb{R}^{N \times N}$, whose generic entry A_{ij} is equal to the weight of the link from node j to node i and is 0 if there is no connection. For undirected networks, $\boldsymbol{A}$ is a symmetric matrix and is said *binary* if its entries are only zeros and ones. Clearly, there is a one-to-one correspon-

dence between networks and adjacency matrices. We define *connectivity* or *degree* the N-dimensional vector $\boldsymbol{k}$ whose elements account for the total number of links involving each node of the network. For directed graphs two kinds of connectivities exist: the in-degree $k_i^{in} = \sum_j A_{ij}$ and the out-degree $k_i^{out} = \sum_j A_{ji}$, which coincide for undirected networks.

The concept of complex networks is very general and each graph is different from the others, but it is useful to group them into families that share some common characteristics. Let us shortly introduce some of the different classes of networks [182]. The simplest one is represented by regular lattices, characterized by a highly regular geometrical structure periodically repeated, like the arrangement of atoms, ions or molecules in a crystalline material.
A more interesting class is that of *random graphs*, which for example can be generated following the Erdös-Rènyi (ER) model [60]. These graphs essentially are anything but regular and are built by adding random edges to a set of N nodes initially disconnected. It is necessary to add a large enough number of links (at least of the order of N) in order to end up with a completely connected graph, i.e. where all nodes are reachable in a finite number of steps. These kinds of networks have been for instance used to represent gene networks [108], ecosystems [133] and often man-made networks like the electricity network, which exhibit sufficient randomness because their structure has often developed over time subject to random constraints [107].
A network structure that falls in the middle between regular lattices and random graphs is represented by *small-world networks*. The building recipe proposed by Watts and Strogatz [200] consists of a random rewiring (with a probability $0 < p < 1$) of the edges of a regular lattice. This process endows the network with long-range connections, actually decreasing the distance between vertices and "making the world smaller", phenomenon also referred to as six degrees of separation [83, 190]. Many empirical examples of small-world networks have been found in different fields, like biology, social sciences and business [1, 99, 111, 144, 198]. In general, as a result of the combination of local clustering and long-distance contacts, this family of graphs is very rich and interesting also from the theoretical point of view, as we will see in the following.
A last network structure that is worth citing is the *scale-free*, correspond-

ing to a topology characterized by an exponential degree distribution $p_k \propto k^{-\gamma}$. This indicates the presence of many nodes with small degree and a few highly connected nodes, or hubs. Examples of these topologies are surprisingly common in nature, from the Internet backbone [62], to metabolic reaction networks [99], passing through social and transportation connections [19], where the exponent γ always exhibits a value between 2 and 3. Many scientists have wondered about the functional advantages of scale-free architectures and have invented building models, the most famous one being the preferential attachment method by Albert and Barabasi [20], based on linking nodes according to a degree-dependent probability, following the concept that "the rich (of connections) gets richer".

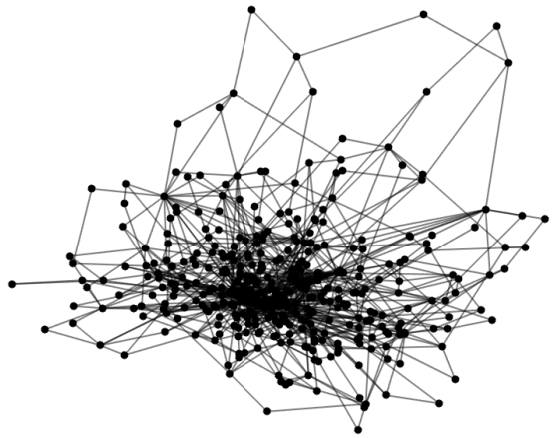


Figure 1.1: Example of scale-free network.

1.2 Random walk on networks

Let us suppose now that a microscopic agent, a walker, travels a network, moving from node to node by using the existing connections. Let us suppose also that the only information he/she has access to is local and is represented by the nodes which are adjacent to the one he/she is currently occupying. Then, if more than one outgoing edge is at his/her disposal, the choice will be random (or proportional to the couplings strength for weighted networks), irrespective of the global network structure, since the walker does not have a global perspective of the network and it is hence impossible for him/her to make far-sighted decisions. This mobility process may at first glance appear trivial and not interesting but is at the base of the random walk concept, and turns out to be a very common process as well as exceptionally useful for the purpose of unveiling the network topology.

Probing a random walk process on a complex network defines a topic of paramount importance and inter-disciplinary breath [3, 21, 28]. Random walks probably are the simplest among the many dynamical processes which have been studied on networks. Since the pioneering works of Pearson [161], who also coined the term, random walks have been extensively investigated in different fields ranging from probability theory to statistical physics and computer science, and have found a number of practical applications, on both regular lattices and (especially in the last couple of decades) on networks with a variety of structures. Typical examples, inspired by applications, are represented by financial markets [18], epidemic spreading [7, 158], social contagions [194], animal movement [43], decision-making in the brain [105, 199], and also swarming behaviors mediated by the diffusive sharing of resources, in robotics [170] as well as in insects [78, 109].

A random walk on a network consists of an agent that performs local hops from one node to one of its neighbours, producing in this way random sequences of adjacent nodes [4, 151, 204]. They are generally introduced as a discrete time process governed by the system of equations

$$x_i(n+1) = \sum_j \Pi_{ij} x_j(n) \tag{1.1}$$

where $x_i(n)$ denotes the probability that node i is visited at time step n and $\mathbf{\Pi}$ is the transition matrix between nodes. The stationary distribution

$\boldsymbol{x}^* = \lim_{n\to\infty} \boldsymbol{x}(n)$ satisfies the equation $\boldsymbol{x^*} = \Pi\boldsymbol{x^*}$ and, for undirected networks $\Pi_{ij}x_j^* = \Pi_{ji}x_i^*$, meaning that the flow of probability in each direction must equal each other at equilibrium, phenomenon which takes the name of *detailed balance* [173]. The standard and simpler definition of the transition matrix is $\Pi_{ij} = A_{ij}/k_j$, implying that from node j one can only move to its adjacent nodes and the transition probability is normalized with the degree, i.e. if node j is highly connected the probability of falling on one of its adjacent nodes is smaller than in the case where j has few neighbors. In this case the stationary distribution is proportional to the degree of nodes: $x_i^* = k_i/2K$, where K is the total number of links.
Switching from discrete to continuous time when the spatial support is discrete, as in the case of a network, is not trivial. The main point is to set the time scale, which is no longer simply defined by discrete steps. Two different types of continuous-time processes can be defined: the node-centric random walk, to which in the following we will simply refer as random walk, and the edge-centric random walk, also know as *diffusion* on network [132]. More specifically, in the node-centric version we consider that a walker sitting on a node waits until the next move for a time τ, where τ is a random variable. If we assume that there are independent, identical Poisson processes at each node of the graph such that the walkers jump at a constant rate, the corresponding continuous-time process is governed by:

$$\dot{x}_i = \sum_j (\Pi_{ij} - \delta_{ij})x_j \equiv \sum_j L_{ij}^{\mathrm{RW}} x_j. \tag{1.2}$$

where δ_{ij} is a Kronecker delta and $L_{ij}^{\mathrm{RW}} = \Pi_{ij} - \delta_{ij}$ is the (ij) element of the *random walk Laplacian*. The concept of Laplacian operator is general and is always characterized by $N-1$ eigenvalues with negative real part and one null eigenvalue, associated to a basis of N eigenvectors, which encode the network structure. For undirected networks the eigenvalues are real and the eigenvectors constitute a orthonormal basis. If we examine the processes governed by this Laplacian we discover that the trajectories are statistically the same as those of the discrete time random walk. In particular, the stationary state is obtained by setting $\dot{x}_i$ equal to zero, which gives $\sum_j L_{ij}^{\mathrm{RW}} x_j^* = 0$, thus yielding the same stationary point of the discrete time version. This also corresponds to the right eigenvector of matrix $\boldsymbol{L}^{\mathrm{RW}}$ associated to the maximum eigenvalue, $\Lambda = 0$. Node-centric random walks are considered to be at the heart of several real-

world dynamical systems, like diseases spreading [44, 97, 197], foraging of animals [172], innovation growth [96] and more. They have also found applications in the context of metapopulation models [86,122,148,157,193], where the nodes of the network represent discrete patches occupied by members of a local population, and the random walk process describes the migration from patch to patch.
In diffusion processes, or fluid models [4, 91, 171], a step occurs when the walker decides to move to another node by using one of the outbound edges of its vertex, or in other words, when an edge is activated. Clearly, the more connected is the starting node the larger the set of options that can be alternatively selected to jump away. The walker therefore leaves a node with large degree more quickly than a node with small degree, and the transition rate for a walker starting from node i is equal to k_i. The occupation probability evolves in this case according to:

$$\dot{x}_i = \sum_j (A_{ij} - k_j \delta_{ij}) x_j \equiv \sum_j L^D_{ij} x_j \tag{1.3}$$

which defines another Laplacian operator with $L^D_{ij} = A_{ij} - k_j \delta_{ij}$, associated to diffusion. More in general, the diffusion process refers to the flow of a (material or immaterial) substance, on a continuous or discrete support, from regions of high concentration to regions of low concentration. This process inevitably yields a space-homogeneous redistribution of the density, which is forcefully subject to detailed balance constraints. When diffusion occurs on a network, the system evolves towards an asymptotic state where all nodes are equally activated, often termed as synchronization [12, 167] or consensus [53]. The homogeneous stationary distribution associated to a purely diffusive process indeed corresponds to the normalized right eigenvector of L^D associated to 0, represented by an N-dimensional vector with all entries equal to $\frac{1}{N}\sum_i x_i(0)$. All remaining eigenvalues are real and negative, associated to N orthonormal eigenvectors.

Let us observe that the Laplacian matrix is named after the similarity of mobility equations (1.2) and (1.3) with the well-known heat equation $\dot{u}(\boldsymbol{x}, t) = \nabla^2 u(\boldsymbol{x}, t)$ which describes the distribution of heat in a given region over time. When the spatial support is represented by a complex network, the Laplace operator ∇^2 is substituted by the Laplacian matrix. The latter is bound to take into account a non-trivial underlying space, thus becoming a powerful instrument which completely defines the con-

sidered graph.

All the above considerations can be easily extended to directed networks, that represent the necessary support for many important applications. Human mobility flows on veritable networks with asymmetric edges between nodes, as often roads can be trodden in one direction only. Information spreads on Internet, the cyberspace being *de facto* schematized as a network with asymmetric routing of the links. Directed networks are characterized by asymmetric Laplacian matrices and consequently complex eigenvalues. Observe also that the Laplacian operator for diffusion on directed networks is defined as $L^D_{ij} = A_{ij} - k^{in}_i \delta_{ij}$, with uniform eigenvector corresponding to the null eigenvalue.

1.3 Reaction-diffusion systems

While diffusion and random walks are the basic ingredients to describe mobility, they do not take into account interactions between agents at the node of a network. These are typically described by a local dynamics, which can be different for each node. Local dynamics have been routinely coupled with diffusive processes to describe the self-consistent evolution of mutually coupled species, when subject to the combined influence of diffusion and reaction terms [10, 16, 139]. These processes go under the name of *reaction-diffusion* and are usually associated to time-evolution equations of the form

$$\dot{x}_i = f_i(x_i) + \sum_j L^D_{ij} x_j$$

where f_i is a generic non-linear function accounting for the local reactions. If the function f_i is identical for each node of the network, the homogeneous state is still a solution of the system.
Examples of real-world situations that can be described by reaction-diffusion processes on networks are represented by the spreading of infectious diseases [7, 95], animal migration among different habitats [8, 136], metapopulation models in general [44], or traffic behavior in technological systems such as the Internet [160]. In the continuum limit, where instead of a network the spatial support is represented by a continuous space region, the interplay between local dynamics and diffusing coupling can give rise

to intriguing phenomena like the formation of travelling waves as well as other self-organized patterns like stripes, generally known as "Turing patterns". These made their first appearance in 1952 after the seminal paper by Alan Turing [192] and have subsequently been theoretically analyzed and experimentally confirmed for various chemical, biological, and ecological systems. They are also supposed to be connected to morphogenesis in biology [87] and to animal coats and skin pigmentation [134, 139]. The important idea proposed by Alan Turing is that a state characterized by local stability to small perturbations can become unstable when a spatial coupling, mediated by diffusion, is introduced. In other words, if, in absence of diffusion, the system under exam tends to a linearly stable uniform steady state, then, under certain conditions, spatially inhomogeneous patterns can evolve by diffusion-driven instability, even though diffusion is usually considered a stabilizing and homogenizing process. This phenomenon is called Turing instability. Even on network-organized systems, collective behaviors consisting of spontaneously emerging spatio-temporal patterns are observed. It has been firstly pointed out by Othmer and Scriven in 1971 [155], trying to describe multi-cellular morphogenesis. They indeed conjectured that the differentiation between cells could be induced by the diffusion on the network of inter-cellular connections where the morphogens are transported. A theoretical study focusing on nonlinear patterns emerging in large random networks has more recently been performed by Nakao and Mikhailov [142], where a network-organized activator-inhibitor system is considered. Here the combined effect of local reaction and network diffusion leads to spontaneous differentiation of the network nodes into activator-rich and activator-low groups. A generalization of this work has been proposed by Asllani, Di Patti and Fanelli [17] where also finite size fluctuations have been introduced into the model. Stochasticity has proven to have the effect of allowing the appearance of self-organized patterns also outside the region of parameters associated to the deterministic Turing instability.

A typical example of reaction-diffusion system, that we will use in chapters 2 and 3 is represented by a set of coupled oscillators which can be described by the Complex Ginzburg-Landau equation (CGLE) [11, 11, 71, 195], a prototypical model for nonlinear physics, whose applications range from superconductivity, superfluidity and Bose-Einstein condensa-

tion to liquid crystals and strings in field theory. Let us assume that the function f_i, which specifies the oscillator dynamics, also depends on a bifurcation parameter μ, such that the system undergoes a supercritical Hopf bifurcation at $\mu = 0$: there exists a linearly stable fixed point when $\mu < 0$ and a stable limit cycle when $\mu > 0$. Near the supercritical Hopf bifurcation, we can derive the CGLE from this general model, following the method of multiple timescales used by Kuramoto [115]. In particular, we can extract slow spatio-temporal dynamics of the complex oscillations amplitude near the bifurcation point. The only difference from Kuramoto's derivation [115] is that the oscillators are discrete and the diffusion term has a different form [141]. We end up with the following approximate equation governing the dynamics of each node j:

$$\dot{x}_j = x_j - (1 + ic_2)|x_j|^2 x_j + (1 + ic_1)K \sum_{k=1}^{N} L^D_{jk} x_k \tag{1.4}$$

where i denotes the imaginary unit; c_1 and c_2 are real parameters, which can be externally assigned; K is a suitable parameter setting the coupling strength.

1.4 Equilibria stability in reaction-diffusion systems

It is important at this stage to introduce the concept of stability together with the standard tools which are commonly used to analyze the stability conditions of a multi-species reaction-diffusion system on a complex network.

Let us suppose that we deal with a generic system composed of N identical entities linked through a complex network. We consider the general case where at time t the activity of node j is described by an m-dimensional variable $\boldsymbol{x}_j(t) \in \mathbb{R}^m$ (vector notation). Starting from a specific initial state $\boldsymbol{x}_j(0)$, the dynamics of $\boldsymbol{x}_j$ evolves according to

$$\dot{\boldsymbol{x}}_j = \boldsymbol{\mathcal{F}}(\boldsymbol{x}_j) + K \sum_{k=1}^{N} L^D_{jk} \boldsymbol{x}_k \qquad j = 1, \ldots, N \tag{1.5}$$

where the function $\boldsymbol{\mathcal{F}} : \mathbb{R}^m \to \mathbb{R}^m$ is identical for each node of the network and K is a constant representing the diffusion coefficient. From the above

considerations we know that such system always admits as an equilibrium a homogeneous state $\boldsymbol{x}^*$, which can be either stationary (a fixed point) or time-dependent (a limit cycle). The stability of this equilibrium can be assessed by perturbing the homogeneous state with a non-uniform small disturbance $\boldsymbol{\delta x}_j$, thus analyzing the system in the perturbed state $\boldsymbol{x}_j = \boldsymbol{x}^* + \boldsymbol{\delta x}_j$. Consequently, by Taylor expanding equations (1.5) to the linear order in the perturbation, we end up with:

$$\dot{\boldsymbol{\delta x}}_j = \boldsymbol{\partial_x \mathcal{F}}(\boldsymbol{x}^*)\boldsymbol{\delta x}_j + K\sum_{k=1}^{N} L^D_{jk}\boldsymbol{\delta x}_k \equiv \sum_{k=1}^{N} J_{jk}\boldsymbol{\delta x}_k \tag{1.6}$$

where we used the diffusion Laplacian property $\sum_k L^D_{jk} = 0$, $\boldsymbol{\partial_x \mathcal{F}}(\boldsymbol{x}^*)$ is an m-dimensional matrix of derivatives, and the $Nm \times Nm$ matrix $\boldsymbol{J}$ is denoted as *Jacobian*. By exploiting the Laplacian eigenvector basis it is possible to convert $\boldsymbol{J}$ in a block-diagonal matrix, thus decoupling the equations and significantly reducing the mathematical complexity of the system. The expansion of the state vector $\boldsymbol{\delta x}_j$ on such basis represents an analogy with the Fourier series expansion, that one would obtain if the network were a regular lattice. We then diagonalize the Laplacian, $\sum_k L^D_{jk}\phi^{(\alpha)}_k = \Lambda^{(\alpha)}\phi^{(\alpha)}_j$, and express the perturbation on the new basis:

$$\boldsymbol{\delta x}_j(t) = \sum_{\alpha=1}^{N} \boldsymbol{c}_\alpha(t)\phi^{(\alpha)}_j, \tag{1.7}$$

where $\boldsymbol{c}_\alpha \in \mathbb{R}^m$. Consequently, eqs. (1.6) can be decoupled as:

$$\sum_{\alpha=1}^{N} \dot{\boldsymbol{c}}_\alpha(t)\phi^{(\alpha)}_j = \boldsymbol{\partial_x \mathcal{F}}(\boldsymbol{x}^*)\sum_{\alpha=1}^{N} \boldsymbol{c}_\alpha(t)\phi^{(\alpha)}_j + K\sum_{\alpha=1}^{N} \Lambda^{(\alpha)}\boldsymbol{c}_\alpha(t)\phi^{(\alpha)}_j \tag{1.8}$$

and using the linear independence of the eigenvectors, it is reduced to a set of N m-dimensional systems indexed by α:

$$\dot{\boldsymbol{c}}_\alpha(t) = \boldsymbol{J}_\alpha \boldsymbol{c}_\alpha(t) \tag{1.9}$$

where the $m \times m$ matrix $\boldsymbol{J}_\alpha \equiv \boldsymbol{\partial_x \mathcal{F}}(\boldsymbol{x}^*) + K\Lambda^{(\alpha)}$ is the α-th diagonal block of $\boldsymbol{J}$. A distinction becomes necessary at this stage: if $\boldsymbol{x}^*$ is a fixed point the Jacobian is constant in time and a simple solution of (1.9) takes the form

$$\boldsymbol{c}_\alpha(t) = \boldsymbol{c}_\alpha(0)e^{\lambda^{(\alpha)}t} \tag{1.10}$$

and it is straightforward that in order to admit a non-trivial solution of the system the determinant of $\boldsymbol{J}_\alpha - \lambda^{(\alpha)}\boldsymbol{I}_m$ must be zero. From this condition we obtain $\boldsymbol{\lambda}^{(\alpha)}$ which configures as the set of eigenvalues of the block $\boldsymbol{J}_\alpha$. The maximum eigenvalue, corresponding to the largest Lyapunov exponent of eq. (1.9), is also denoted as Master Stability Function [94,162]. If instead the Jacobian is periodic in time due to the limit cycle solution ($\boldsymbol{J}_\alpha(t+T) = \boldsymbol{J}_\alpha(t)$), the Floquet theory [40, 82, 130] comes into play. This latter allows us to again express the temporal dependence of the perturbations $\boldsymbol{\delta x}$ as an exponential function: $e^{\mu_k^{(\alpha)}t}$, but here $\mu_k^{(\alpha)}$ are the *Floquet exponents*, defined by $\mu_k^{(\alpha)} = \log(\rho_k^{(\alpha)})/T$ with $\rho_k^{(\alpha)}$ eigenvalues of the constant matrix $\boldsymbol{B}$ such that $\det(\boldsymbol{B}) = \exp\big[\int_0^T \mathrm{Tr}(\boldsymbol{J}_\alpha(t))dt\big]$.
In both, the static and the periodic case, the stability of the system is assessed by the component of respectively $\boldsymbol{\lambda}^{(\alpha)}$ or $\boldsymbol{\mu}^{(\alpha)}$ with maximum real part. Its relation with the Laplacian eigenvalues is called the *dispersion relation* and represents the network extension of the dispersion relation defined on a regular (continuous or discrete) medium. The analogy relies on the identification of the eigenvalues and eigenvectors of the network Laplacian with the Fourier wavelengths and modes used when the support is periodic.

Finally, the exponential time evolution of the linear perturbations implies that the system can only be stable if all the exponents $\lambda^{(\alpha)}$ or $\mu^{(\alpha)}$ are negative, because if even just one of them has a positive real part, then the perturbation grows through the direction associated to the corresponding eigenvector, resulting in an irregular behavior.

Let us observe that the above stability analysis can be extended to the case where the coupling is not necessarily diffusive but still dependent on the concentration difference between adjacent nodes by means of a function $\boldsymbol{\mathcal{G}}$, like:

$$\dot{\boldsymbol{x}}_j = \boldsymbol{\mathcal{F}}(\boldsymbol{x}_j) + K\sum_{k=1}^{N} A_{jk}\boldsymbol{\mathcal{G}}(\boldsymbol{x}_k - \boldsymbol{x}_j). \tag{1.11}$$

Indeed, during the linearization process the Laplacian matrix emerges anyway and we again obtain a system of the form $\boldsymbol{\delta}\dot{\boldsymbol{x}}_j = \sum_{k=1}^{N} J_{jk}\boldsymbol{\delta x}_k$, with the difference that the Jacobian matrix results defined by:

$$\boldsymbol{J} = \boldsymbol{\partial}_{\boldsymbol{x}}\boldsymbol{\mathcal{F}}(\boldsymbol{x}^*)\boldsymbol{I}_N + K\boldsymbol{\partial}_{\boldsymbol{x}}\boldsymbol{\mathcal{G}}(\boldsymbol{0})\boldsymbol{L}^{\boldsymbol{D}}, \tag{1.12}$$

see eq. (1.6) for a comparison.

Let us say a brief word on the possibility that some of the Jacobian eigenvalues have algebraic multiplicity m_a larger than one and geometric multiplicity $m_g \neq m_a$. In this case the Jacobian eigenvectors do not constitute a basis and the solution takes a different form. Let us consider for instance the case where we only have $N-1$ independent eigenvectors $\boldsymbol{\phi}^{(1)}, ..., \boldsymbol{\phi}^{(N-1)}$ where $\Lambda^{(1)}$ is characterized by $m_a = 2$ and $m_g = 1$. The basis that we choose for the expansion should be $\boldsymbol{\xi}^{(0)}, \boldsymbol{\xi}^{(1)} \equiv \boldsymbol{\phi}^{(1)}, ..., \boldsymbol{\xi}^{(N-1)} \equiv \boldsymbol{\phi}^{(N-1)}$ where $\boldsymbol{\xi}^{(0)}$ is a generic vector which is linearly independent on the others. The analogous of (1.7) will take the form:

$$\boldsymbol{\delta x}_j(t) = [\boldsymbol{c}_0 \xi_j^{(0)} + (\boldsymbol{c}_1 + \boldsymbol{a}t)\xi_j^{(1)}]e^{\lambda^{(1)}t} + \sum_{\alpha=2}^{N-1} \boldsymbol{c}_\alpha e^{\lambda^{(\alpha)}t} \xi_j^{(\alpha)} \tag{1.13}$$

with $\boldsymbol{c}_0, \boldsymbol{c}_1, ..., \boldsymbol{c}_{N-1}, \boldsymbol{a}$ arbitrary constant vectors $\in \mathbb{R}^m$ to be fixed by the initial conditions.

In general, the Jacobian eigenvalues degeneracy has the effect of modifying the exponential time behaviour by introducing a multiplicative polynomial in t. The asymptotic state still only depends on the eigenvalues and the system is stable if all their real parts are negative. However, also in the stable scenario, during the transient before equilibrium the state is governed by the polynomial terms. These latter may generate short time amplifications of the perturbations, with the consequent risk of moving away from the basin of attraction of the equilibrium under exam [137].

In the next chapters we will deal with different kinds of dynamical systems on networks, not only reaction-diffusion, and we will develop techniques to affect and modify the system stability in a desired way. In all cases the eigenvectors decomposition and the subsequent dispersion relation will represent the main characters of our analysis, mostly based on the Laplacian spectral properties.

Chapter 2

Pattern invariant networks

Patterns are ubiquitous in nature and arise in different contexts, ranging from chemistry to physics, passing through biology and life sciences [192]. The mathematics that underlies pattern formation focuses on the dynamical interplay between reaction and diffusion processes. In particular, when the reaction-diffusion system is hosted on a graph, the Turing pattern counterpart is represented by a non-homogeneous filling of the different nodes of the network. The activation motif can be static or evolving in time. In the following we will use the word "pattern" referring to the (regular or irregular) activation scheme of the different nodes on a finite window of time.

This chapter is devoted to explore the subtle interplay between network structure and ensuing dynamics. The focus is put on reaction-diffusion systems with identical single node dynamics, as introduced in the previous chapter, where the coupling between the units always admits a homogeneous state where all the nodes are synchronized. If such equilibrium proves unstable, any arbitrary perturbation grows through the unstable directions giving rise to irregular spatio-temporal patterns [16, 155, 168, 192]. While in the homogeneous state no information is flowing through the network, the result of instability may turn out to be of crucial interest from the point of view of network characterization. This irregular behaviour indeed represents the *functional* response to a given input which can be traced back to the *structural* characteristics of the underlying network [32]. Whereas, in general, a random change on the network topology

would lead to a modification of its dynamical response, it is also presumable that a specific outcome is not unique from a particular network, so that similar patterns might arise from different *isodynamic* topologies. The identification of different compatible structures that give rise to the same dynamical behavior represents an important leap forward in the study of complex networks dynamics. For instance, it paves new roads to devise network reconstruction protocols, in cases where more than one network can correspond to the same dynamical output. Upon analyzing the common topological properties that contain the dynamical information of the system, one might overlook the inaccessible details of the topology.

Here, we propose two different methods to generate, given a specific network, a second network isodynamic to the first one, so that they share the same dynamical response. For this purpose, we exploit the idea that just part of the degrees of freedom stored in the topology proves significant for the emergence of a specific dynamical behavior [9, 79, 85]. The first method is exact, since it directly acts on the subspace generated by the Laplacian eigenvectors characterizing the stability of the equilibrium state. On the other hand, the second technique relies on a Monte-Carlo algorithm that allows for a neater control in terms of network topology. In brief, the goal of this analysis is to prove that it is possible to provide two (or more) structurally different complex networks so that furnishing the same input can lead to the same output.

2.1 Dispersion relation and pattern formation

We begin by analyzing the mechanism which leads to pattern formation and how the switching on of irregularities is related to the network topology.

Let us consider a generic system composed of N identical entities linked through a complex network. At time t the activity of node j is described by an m-dimensional variable $\boldsymbol{x}_j(t) \in \mathbb{R}^m$. Starting from a specific initial state $\boldsymbol{x}_j(0)$, the dynamics of $\boldsymbol{x}_j$ evolves according to

$$\dot{\boldsymbol{x}}_j = \boldsymbol{\mathcal{F}}(\boldsymbol{x}_j) + K \sum_{k=1}^{N} A_{jk} \boldsymbol{\mathcal{G}}(\boldsymbol{x}_k - \boldsymbol{x}_j) \qquad j = 1, \ldots, N \tag{2.1}$$

where $\boldsymbol{A}$ denotes the adjacency matrix of the network, and $\boldsymbol{\mathcal{F}} : \mathbb{R}^m \to \mathbb{R}^m$ and $\boldsymbol{\mathcal{G}} : \mathbb{R}^m \to \mathbb{R}^m$ are generic continuous functions. The first term on the right hand side of the equation describes the self-dynamics of each individual node j, and is referred to as *reaction term*. The second term instead, defines the interaction of node j with the other nodes of the network: the existence of a coupling is set by the adjacency matrix $\boldsymbol{A}$, while the interaction shape is established by function $\boldsymbol{\mathcal{G}}$. Our analysis requires two assumptions: (i) there exists an equilibrium $\boldsymbol{x}_j = \boldsymbol{x}^* \ \forall j$ for the uncoupled problem $\dot{\boldsymbol{x}_j} = \boldsymbol{\mathcal{F}}(\boldsymbol{x}_j)$, and (ii) function $\boldsymbol{\mathcal{G}}$ annihilates in zero, i.e. $\boldsymbol{\mathcal{G}}(\boldsymbol{0}) = \boldsymbol{0}$. The uncoupled equilibrium $\boldsymbol{x}^*$ can be either a stationary fixed point or a limit cycle. Whatever is the case, the second condition ensures that $\boldsymbol{x}^*$ becomes a *homogeneous solution* of the system (2.1). A wide class of models corresponding to this setup are *reaction-diffusion* systems, where $\boldsymbol{\mathcal{G}}$ is a linear function. This constraints can be also generalized in order to include networks of homogeneous Kuramoto-Daido systems with arbitrary coupling functions [49–51].

At the homogeneous equilibrium $\boldsymbol{x}^*$ no information is flowing through the network. However, if such state turns out to be unstable, a generic small perturbation $\boldsymbol{\delta x}$ of $\boldsymbol{x}^*$ can give rise to an irregular spatio-temporal pattern. Linearization of (2.1) around $\boldsymbol{x}^*$ provides the time evolution of the perturbation $\boldsymbol{\delta x}_j \in \mathbb{R}^{Nm}$ as a system of Nm linear ordinary differential equations,

$$\dot{\boldsymbol{\delta x}} = \boldsymbol{J}(\boldsymbol{x}^*)\boldsymbol{\delta x}, \tag{2.2}$$

where $\boldsymbol{J} \in \mathbb{R}^{Nm \times Nm}$ is the Jacobian matrix of the system. Therefore, the stability analysis of $\boldsymbol{x}^*$ amounts to studying the high-dimensional operator $\boldsymbol{J}(\boldsymbol{x}^*)$. Nevertheless, as explained in chapter 1, it is possible to link the diagonalization of the Jacobian to that of a simpler operator. In the linearized regime, the flow of the quantity $\boldsymbol{\delta x}_j$ to another node i of the underlying network is described by $L^D_{ij} = A_{ij} - k_i^{(in)}\delta_{ij}$, where $k_i^{(in)}$ is the in-degree of node i and δ_{ij} the Kronecker delta. The resulting $N \times N$ matrix $\boldsymbol{L^D}$ is known as *Laplacian* of the network.

By expressing $\boldsymbol{\delta x}_j$ in the basis of the Laplacian eigenvectors $\{\boldsymbol{\phi}^{(\alpha)}\}_{\alpha=1}^N$, and making use of the corresponding eigenvalues $\{\Lambda^{(\alpha)}\}_{\alpha=1}^N$ it is possible to decouple the Nm equations from system (2.2), thus reducing the problem to N uncoupled m-dimensional systems indexed by α (see chapter 1). Each of the reduced systems is described by the reduced Jacobian $\boldsymbol{J}_\alpha \equiv \boldsymbol{\partial_w \mathcal{F}}(\boldsymbol{x}^*) + \boldsymbol{\partial_w \mathcal{G}}(\boldsymbol{0}) K \Lambda^{(\alpha)}$. If $\boldsymbol{J}_\alpha$ is time independent, as when $\boldsymbol{x}^*$ is a fixed

point, the stability analysis is simply assessed by its m eigenvalues $\boldsymbol{\lambda}^{(\alpha)} = (\lambda_k^{(\alpha)})$. Therefore, the stability of $\boldsymbol{x}^*$ in the full Nm dimensional problem is ultimately controlled by the dispersion relation $\lambda^{(\alpha)} \equiv \max_k(\lambda_k^{(\alpha)})$. If instead the Jacobian has a periodic dependence on time, one needs to obtain the m *Floquet exponents* $\boldsymbol{\mu}^{(\alpha)} = (\mu_k^{(\alpha)})$ of the system [40, 82, 130], which represent the analogue of $\boldsymbol{\lambda}^{(\alpha)}$ for a time-dependent Jacobian (see 1). In both instances, the dispersion relation assesses the stability of $\boldsymbol{x}^*$ and its dependence on the Laplacian eigenvalues, as explained in chapter 1.

If the dispersion relation of a subportion N_c of the total N eigenvalues $\Lambda^{(\alpha)}$ is positive, then any perturbation would grow through the unstable modes. The resulting irregular spatio-temporal pattern thus represents the unpredictable *response* of the system to a specific *input signal*. It is however reasonable to suppose that most of the relevant information is stored in the unstable manifold of the homogeneous solution $\boldsymbol{x}^*$, as it will be proven later on.

2.2 Topology modification

In this section we propose two methods to modify the discrete support of the system while preserving the relevant directions for the emergence of the pattern.

2.2.1 Eigenmode randomization

The first method for network modification consists of preserving a subportion n of the N total eigenmodes of the original network Laplacian whereas all the others are modified. Let us suppose, without loss of generality, that the subset of modes to be left invariant are the first n. We define the diagonal matrix $\tilde{\boldsymbol{\Lambda}}$, such that

$$\tilde{\Lambda}^{(\alpha)} = \begin{cases} \Lambda^{(\alpha)} & \text{if} \quad \alpha \leq n \\ \Lambda^{(\alpha)} + \delta\Lambda^{(\alpha)} & \text{if} \quad \alpha > n. \end{cases}$$

The corresponding eigenvectors are modified by performing a change of basis

$$\tilde{\boldsymbol{\Phi}} = \boldsymbol{\Phi} \left(\begin{array}{c|c} I_n & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{R}_{N-n} \end{array} \right)$$

where $\mathbf{\Phi}$ is a matrix whose columns are the eigenvectors of $\boldsymbol{L}^D$, I_n is the identity matrix of dimension n and $\mathbf{R}_{N-n} \in SO(N-n)$ is a random rotation matrix. In order to practically obtain such rotation we perform QR decomposition of a random matrix. Taking care of preserving the uniform mode, which corresponds to an identical perturbation acting independently on each node, the transformation

$$\tilde{\mathbf{L}}^{\mathbf{D}} := \tilde{\mathbf{\Phi}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{\Phi}}^{-1}$$

defines a new Laplacian matrix, indeed:
(i): **The elements of $\tilde{\mathbf{L}}^{\mathbf{D}}$ are real**. In the symmetric case, the eigenvalues and their corresponding corrections are real, as well as the eigenvectors and the rotation matrix $\mathbf{R}$. Hence, the elements of $\tilde{\mathbf{\Lambda}}$, $\tilde{\mathbf{\Phi}}$, $\tilde{\mathbf{\Phi}}^{-1}$ are real and, consequently, $\tilde{L}^D_{ij} \in \mathbb{R}\ \forall i,j$. The directed case is more complicated to treat and is discussed in Appendix A.
(ii): **Each column of $\tilde{\mathbf{L}}^{\mathbf{D}}$ sums up to zero.** Observe that

$$\begin{aligned}\Lambda^{(\alpha)}\phi_i^{(\alpha)} &= \sum_j L^D_{ij}\phi_j^{(\alpha)} = \sum_j (A_{ij} - \delta_{ij}k_j)\phi_j^{(\alpha)} = \\ &= \sum_j A_{ij}\phi_j^{(\alpha)} - k_i\phi_i^{(\alpha)} = \\ &= \sum_j A_{ij}\phi_j^{(\alpha)} - \sum_l A_{li}\phi_i^{(\alpha)}\end{aligned} \tag{2.3}$$

Summing over i one obtains:

$$\Lambda^{(\alpha)}\sum_i \phi_i^{(\alpha)} = \sum_{i,j} A_{ij}\phi_j^{(\alpha)} - \sum_{l,i} A_{li}\phi_i^{(\alpha)} = 0 \ \ \forall\alpha, \tag{2.4}$$

thus the sum of elements of each Laplacian's eigenvector corresponding to an eigenvalue different from zero is identically equal to zero. This is also true for the modified eigenvectors, since $\Lambda^{(\alpha)}\sum_i \Phi_{i\alpha} = 0$ implies $\tilde{\Lambda}^{(\alpha)}\sum_i \tilde{\Phi}_{i\alpha} = \tilde{\Lambda}^{(\alpha)}\sum_l R_{l\alpha}\sum_i \Phi_{il} = 0$, assuming that $\Lambda^{(\alpha)} \neq 0$ implies $\tilde{\Lambda}^{(\alpha)} \neq 0$. This observation can be used to conclude the proof considering that the (i,l) entry of the controlled Laplacian can be written as

$$\begin{aligned}\sum_i \tilde{L}^D_{il} &= \sum_{i,j} \tilde{\Phi}_{ij}\tilde{\Lambda}_{jj}(\tilde{\Phi}^{-1})_{jl} = \\ &= \sum_j (\tilde{\Phi}^{-1})_{jl}\tilde{\Lambda}_{jj}\sum_i \tilde{\Phi}_{ij}\end{aligned} \tag{2.5}$$

and that the zero eigenvalue is associated to a null correction.
(iii): If $\mathbf{L}^D$ is symmetric, then also $\tilde{\mathbf{L}}^D$ is symmetric. It is enough to prove that the corrections $\tilde{\mathbf{L}}^{\mathbf{D}}$ are bound to be symmetric. Indeed, when $\mathbf{L}^D$ is symmetric, matrix $\mathbf{\Phi}$ is orthogonal ($\mathbf{\Phi}^{-1} = \mathbf{\Phi}^T$). Hence:

$$\begin{aligned} \tilde{L}^D_{il} &= \sum_j \tilde{\Phi}_{ij}\tilde{\Lambda}_{jj}(\tilde{\Phi}^{-1})_{jl} = \\ &= \sum_j \tilde{\Phi}_{ij}\tilde{\Lambda}_{jj}\tilde{\Phi}_{lj} = \\ &= \sum_j \tilde{\Phi}_{lj}\tilde{\Lambda}_{jj}(\tilde{\Phi}^{-1})_{ji} = \tilde{L}^D_{li} \end{aligned} \tag{2.6}$$

which concludes the proof.

2.3 Local rewiring

The previous method does not provide any control whatsoever on the topological modifications introduced in the new Laplacian. For this reason, we propose an alternative route that acts at the level of single nodes. For many network structures, the Laplacian eigenvectors are well localized on the network, *i.e.,* their coordinates in the original vector space mostly involve a small subset of nodes, different for each eigenvector [89]. Therefore, modifying some eigenvectors means acting mainly on the connections of a specific subnetwork, and vice versa. With this method we aim to identify and modify the links among nodes which are poorly involved on the n-dimensional manifold related to pattern formation.

We rely on a Monte-Carlo algorithm that proceeds as follows. Given the original network, we choose a random non-diagonal entry of the adjacency matrix, A_{ij}. If the entry indicates the existence of a link between node i and node j, such link is removed, otherwise it is created. We then compare the n eigenvalues and eigenvectors of the modified network with those of the original one, and only if they prove similar according to a chosen threshold τ the change is accepted (see Appendix for further details). The process is repeated selecting new random entries over the modified adjacency matrix until a desired number of links have been changed or the method fails to detect new entries that lead to small error.

2.4 Results and discussion

2.4.1 The Ginzburg-Landau model

In order to test the different methods we use the CGLE, introduced in the previous chapter, as a reference model for the dynamics of each node. The system is hence composed by a set of oscillators, described in terms of their complex amplitude, occupying the nodes of an undirected network with a diffusive coupling (for a case with a directed network see Appendix). The equation governing the dynamics of each node j is

$$\dot{x}_j = x_j - (1+ic_2)|x_j|^2 x_j + (1+ic_1)K\sum_{k=1}^{N} L^D_{jk} x_k \tag{2.7}$$

with $c_1, c_2, K \in \mathbb{R}$. We focus our analysis on the limit-cycle globally synchronous state, $x_j(t) = x^*(t) := e^{-ic_2 t}$.

Even though x^* is time-dependent, the Jacobian $\boldsymbol{J}_\alpha$ is constant in time, and following the procedure indicated in the Appendix, we obtain the dispersion relation characterizing the linear stability analysis of the limit cycle solution,

$$\lambda(-K\Lambda^{(\alpha)}) = K\Lambda^{(\alpha)} - 1 + \sqrt{-c_1^2K^2(\Lambda^{(\alpha)})^2 + 2c_1c_2K\Lambda^{(\alpha)} + 1}. \tag{2.8}$$

Depending on the system parameters, function λ might have a positive part, corresponding to the smaller Laplacian eigenvalues in absolute value. Therefore, a particular network with N nodes might have $N_c < N$ Laplacian eigenvalues associated to unstable modes. From now on, we assume that the Laplacian eigenvalues are sorted in descending order, $\Lambda^{(0)} = 0 > \Lambda^{(1)} > \cdots > \Lambda^{(N)}$, so that the N_c unstable modes correspond to the Laplacian eigenvalues with index ranging from 1 to N_c.

2.4.2 Eigenmode randomization

Erdös-Rènyi networks

We first test the eigenmode randomization method using as original topology an Erdös-Rènyi (ER) network with $N = 100$ nodes and average node degree $\langle k \rangle = 3.5$. For system parameters $K = 1$, $c_1 = 1$, and $c_2 = -3$, the dispersion relation associated to this network has $N_c = 44$ modes corresponding to unstable directions (see red dots in Fig. 2.1). We integrate

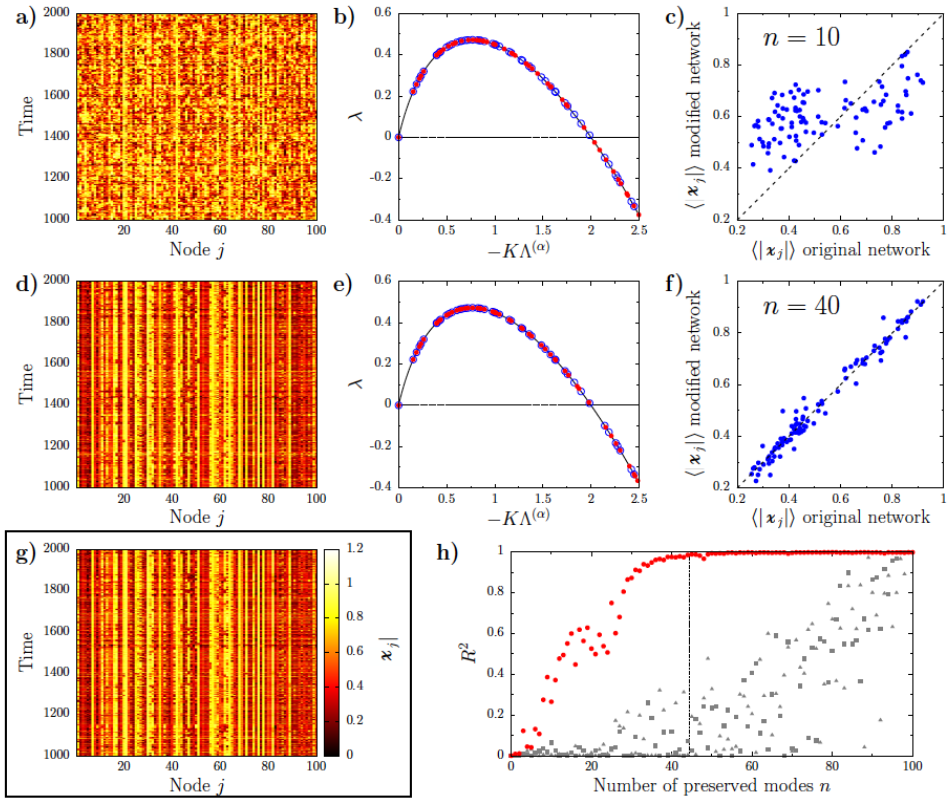


Figure 2.1: Results corresponding to an Erdös-Rènyi undirected network with $N = 100$ nodes and average degree $\langle k \rangle = 3.5$. (a-f) Outcome of the implementation of the eigenmode randomization preserving 10 (a-c) and 40 (d-f) modes. (a,d): Time evolution of the modulus for each node of the modified network. (b,e): Section of the dispersion relation λ showing the unstable eigenvalues for the original (red dots) and modified network (blue circles). A full plot of λ for the original network is depicted in the Appendix. (c,f): Relation between the time-averaged modulus of each node on the modified network and that of the original network. (g): Time evolution of the modulus of each node of the original network. (h): Squared correlation coefficient between the time-average modulus of each node on the modified network and that of the original network for different number of preserved modes. Purple circles correspond to networks where the first n Laplacian eigenvalues are preserved. Grey squares and grey triangles are the outcome of two different realizations where the n preserved modes of the modified networks are selected at random.

the system using as initial condition a randomly generated small perturbation of the homogeneous limit-cycle solution. After a short transient, the system reaches a dynamical regime characterized by an irregular spatio-temporal pattern depicted in figure 2.1(g). Notice that throughout the chapter we focus only on each oscillator modulus, since the frequency of rotation does not seem to contain relevant spatial structure.

We aim to generate a new topology that reproduces such dynamical behavior by leaving invariant a subset of the original network modes. Taking into account that the Laplacian eigenvalues are sorted in descending order, we preserve the first n modes and modify the rest using the eigenmode randomization method. As an illustrative example, in Fig. 2.1(a) we show the pattern resulting from a network where only the first $n = 10$ modes of the original topology are preserved and all the others have been randomized. Manifestly, such outcome has little similarity with that of the original network (*cf.* fig. 2.1(g)). In Fig. 2.1(b) we plot a zoom on the unstable part of the dispersion relation for the modified network (blue circles), and that of the original topology (red dots). Overall is plausible to explain the disagreement between the dynamical behavior of the two networks from the large difference between the corresponding tangent space of the synchronized solution. Repeating the procedure with $n = 40$ preserved modes instead, leads to the pattern from Fig. 2.1(d). In this case, the behavior of the modified topology resembles much better that of the original network, and observation which is to be traced to the similarity between the dispersion relations of the respective unstable manifolds (see Fig. 2.1(e)).

In order to make the comparison between different patterns more transparent, we compute the time-average modulus of each node, $\langle |x_j| \rangle$ after discarding a transient of 1000 time units. In Figs. 2.1(c) and (f) we report the resulting mean node activity of the modified networks versus that of the original one for $n = 10$ and $n = 40$ respectively. An outcome from two identical patterns would lie exactly on the diagonal, while two completely independent processes would provide a random collection of points. One can then quantify the similarity between the modified and original patterns in terms of squared Pearson correlation coefficient R^2, which is 0.38 for $n = 10$ and 0.97 for $n = 40$. In Fig. 2.1(h) we show the outcome of repeating this analysis systematically for increasing number of preserved modes n. When the modes to be preserved are selected to be the first n,

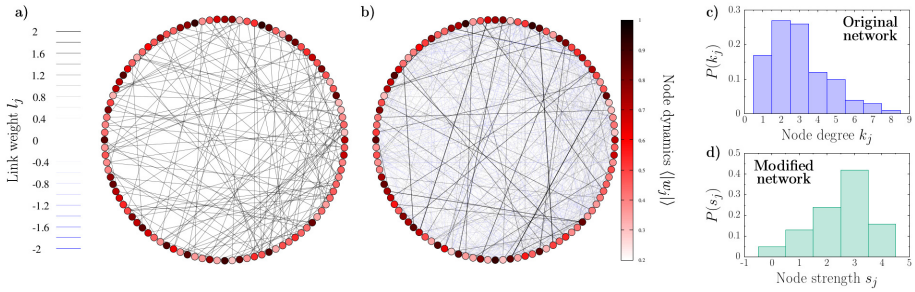


Figure 2.2: (a) Original ER topology. (b) Network resulting from preserving all the unstable modes. The color of the nodes correspond to the time average modulus of x_j, each link thickness is associated to its weight and blue links indicate negative contributions. (c) Node degree distribution of the original network. (d) Node strength of the modified network.

the correlation between the patterns of the modified and original networks increases quickly with n, reaching a very good agreement when approximately all the 44 unstable modes are preserved (see red circles). On the other hand, if the n invariant modes are randomly selected among all the N, even when a large number of directions are maintained, there is no guarantee that the modified network will respond similarly to the original one (see gray triangles and squares).

The eigenmode randomization technique does not provide, *a priori* any information about the structure of the resulting network. In Fig. 2.2(a) and (b) we show the topologies of the original network and a modified version where all the 44 unstable modes have been preserved. The first obvious difference one can see is the existence of weighted links, including negative ones (see blue edges), which are absent in the original network. The adjacency matrix loses its sparsity and the network becomes highly connected, although most of the new links are very weak. In order to compare the two networks we focus on the *strength* s_j of each node, defined as $s_j = \sum_{m=1}^{N} A_{jm}$, which extends the concept of node degree to weighted topologies. In Fig. 2.2(c) and (d) we plot the degree distribution of the original network, and the strength distribution of the modified topology, respectively. Although the distribution of the new network looks different from the original one, it does not present any particular structure, as

expected for a random topology. In fact, the average strength of the new network coincides with the average degree of the initial support.

Scale-Free networks

In order to further investigate the relation between the Laplacian eigenmodes and network topology, we next move to a case where the network has a specific structure. For this purpose we consider Scale-Free networks (SF), characterized by a typical power law degree distribution. In Fig. 2.3(a) we show an example of such degree distribution for a network composed of 150 nodes. Fixing the parameters of the CGLE as $K = 1$, $c_1 = 1.2$, and $c_2 = -10$, the dispersion relation of this network shows $N_c = 76$ unstable modes. We apply the eigenmode randomization technique to generate a new support taking care to preserve all the unstable directions, but none of the stable ones ($n = N_c$). The patterns resulting from inserting the same perturbation to the synchronized solution on both networks are highly alike, as can be seen from the correlation figure 2.3(d). Differently from the ER case, here the modulus of each node gets stationary after some transient, so such time-average activity is free from statistical fluctuations.

The topology of the new network is, again, highly connected and involves negative links. Morevoer, the node strength distribution (see Fig. 2.3(b)) does not preserve the scale-free structure of the original topology. To identify the nature of this strong changes, in Fig. 2.3(c) we plot a one to one comparison between the degree of each node of the original support and the corresponding strength on the new topology. This analysis reveals that nodes with smaller degree on the original setup have a comparable strength after the modification, whereas the hubs become much weaker.

The localization properties of the Laplacian eigenvectors allow us to understand this situation. Inspired by the analysis performed by Hata and Nakao in [89], we depict in Fig. 2.3(f) the absolute values of the vector components $|\phi_i^{(\alpha)}|$, where nodes are sorted according to their degree in descending order ($k_1 \geq k_2 \geq ... \geq k_N$), whereas the eigenmode index α follows the usual descending order of the eigenvalues. The almost diagonal behavior indicates that the eigenvectors associated to the eigenvalues with smaller absolute value mostly involve the less connected nodes of the network.

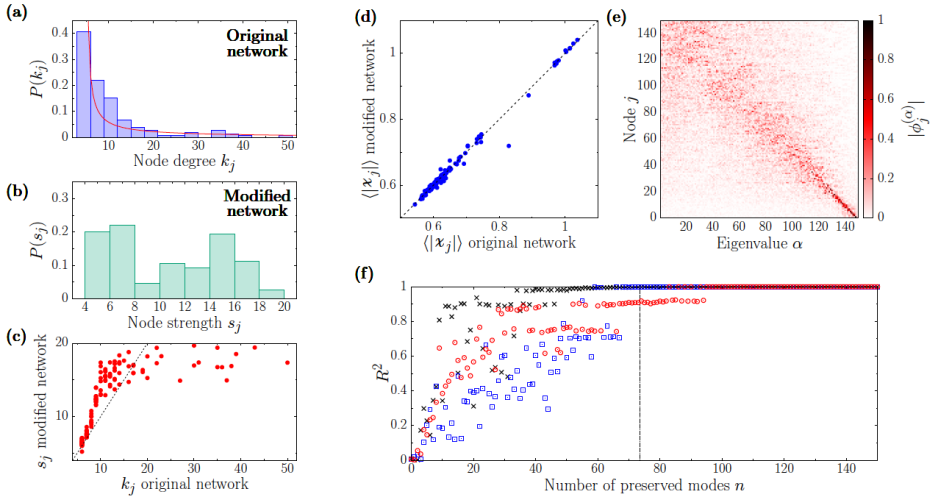


Figure 2.3: Outcome of the eigenmode randomization technique using as original network a scale-free topology with $N = 150$ nodes. (a) Degree distribution of the original network (blue boxes). The continuous red curve corresponds to a numerical fit of the data to an exponential shape $P(k) = ck^{-\beta}$, with $c = 0.23$ and $\beta = 0.85$. (b) Strength distribution on a modified version of the network where the 76 unstables modes have been preserved. (c) Node strength s_j of the modified network with respect to the node degree of the same node k_j in the original network. The black dashed line shows the case $s_j = k_j$ for eye guide. (d) Correlation between the dynamical patterns of the original and modified networks, with resulting squared correlation coefficient $R^2 = 0.99$. (e) Eigenvectors of the original Scale-Free network Laplacian. The eigenvalues are sorted in descending order according to the corresponding Laplacian eigenvalues, and the nodes have been sorted in descending order with respect to their degree. The color represents the absolute value of each eigenvector component. (f) Squared correlation coefficient as the number of preserved modes n increases. Red circles, blue squares and black crosses indicate the results corresponding to three different initial conditions. The vertical dashed line indicates the threshold between unstable and stable modes.

This implies that all the networks generated with the eigenmode randomization procedure will mostly differ from the original one for what concerns the nodes characterized by a large degree k_j. For this reason the

characteristic shape of the degree distribution is not preserved in the second network. Nevertheless, in dynamical systems with a different shape of the dispersion relation it might be possible to maintain the tail of the distribution if the stable modes are to be found among the first Laplacian eigenmodes (see a specific example in the Appendix).

Finally, we repeat the procedure of systematically increasing the number of invariant modes n from 1 to 150 and tracking the resulting correlation coefficient, Fig. 2.3(f). For each new network we analyze the dynamical patterns resulting from inserting three different randomly generated perturbations. As for the ER case, the correlation between the dynamics of the original and modified topologies increases quickly with n, being nearly optimal when all the unstable modes are preserved (see vertical dashed line). Nevertheless, the choice of the initial condition is clearly relevant: whereas in one case preserving around 40 modes already provides a good agreement (see black crosses), in other situations one might need to preserve also a considerable number of stable directions to reproduce the original pattern (see red circles).

2.4.3 Local rewiring

The second method we present offers the advantage of a more flexible control on the resulting network topology so that, for instance, one can keep the network binary. On the other hand, one needs to allow for some variability also on the unstable manifold due to the non perfect localization of the Laplacian eigenmodes on a specific subnetwork.

We first apply the local rewiring technique on a ER network with $N = 400$ nodes and average degree $\langle k \rangle = 20$. At each step the tolerance of difference between the unstable manifolds of the original and modified networks is set to $\tau = 10^{-1}$. For system parameters $K = 0.15$, $c_1 = 1$, and $c_2 = -3$, the system displays 64 unstable modes. The algorithm stops after 100 positions on the adjacency matrix have been changed. For comparison purposes, we also created three different networks where 100 entries on the adjacency matrix have been changed totally at random. Blue circles in Fig. 2.4(a) show the correlation between the patterns of the original and the resulting network, whereas the gray symbols correspond to that of the random rewired graphs. The dynamical behavior of the network obtained using the local rewiring method provides a better agreement with the original pattern. We obtain similar results when

using a SF network generated with the Barabási-Albert algorithm (see Fig. 2.4(b)). Although these results change upon inserting different initial perturbations, the rewiring technique outperforms the randomly modified networks in most of the cases (see Appendix for additional results).

In terms of topology changes, in Figs. 2.4(c) and (d) we plot the total number of links that have been modified for each node of the network. It is clear that most of the changes correspond to nodes with larger degree which, as shown in the previous section, mostly involve the Laplacian eigenvectors associated to stable directions in the case of the CGLE.

In conclusion, in the present chapter we have devised two specific strategies of network generation so as to mimic the pattern obtained by a former sample network when both are subject to the same dynamical system. The first method is analytical, and provides isodynamic networks at the cost of changing important topological features of the original graph. The second technique instead makes use of a Monte-Carlo algorithm, allowing for a larger control in terms of network topology, but providing less accurate results. Both methods rely on a preliminary identification of a manifold generated by the Laplacian eigenvectors associated to homogeneous solution instability, which has to be preserved during the modifications. The localization properties of the Laplacian eigenvectors in some cases make it possible to identify a subnetwork or a set of nodes which are (recognized to be) practically irrelevant for pattern formation. The consequence is that a specific activation pattern can be associated not only to a single network but to a class of networks having in common the relevant substructures and differentiating for the rest. Hence, the above analysis also opens the perspective to multiple possibilities in network reconstruction.

The proposed network generation method, based on the discovery of pattern-invariant networks, sheds light on the strong dependence between network topology and dynamical patterns. This argument will be further investigated in the next chapter where we will focus on how the network structure can be adjusted in order to induce the spontaneous patterns to dampen.

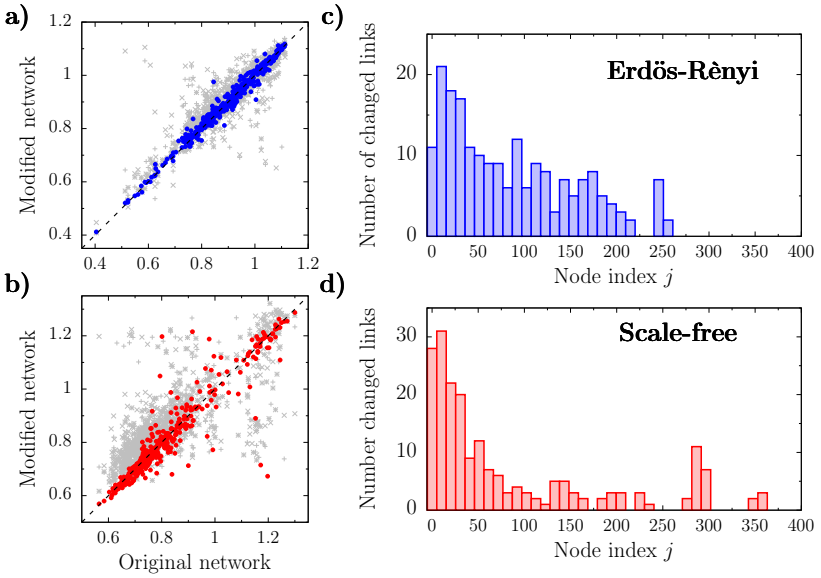


Figure 2.4: Results of the local rewiring method applied to a ER with average degree $\langle k \rangle$=20 and 64 unstable modes (a,c), and to a SF network with 63 unstable modes (b,d). Both topologies consist of $N = 400$ nodes. At each step the tolerance error of the algorithm is $\tau = 0.1$. In total, the new networks present 100 modified links with respect to the original topologies. (a,b) Correlation between patterns of the original and modified networks. Colored circles correspond to the results obtained using the local rewiring algorithm. Gray pluses, crosses, and stars correspond to three different networks generated by random rewiring as many links as the network outcoming from the local rewiring procedure. The squared coefficient correlations for the modified ER network is 0.98 (blue circles) whereas the randomly rewired networks provide 0.67, 0.62, and 0.94. The modified SF network has R^2 =0.86 (red circles), whereas the random networks provide 0.49, 0.50, and 0.67. (c,d) Number of links changed for each node of the network. The nodes are sorted in descending degree order.

Chapter 3

Spectral control for reaction-diffusion systems on networks

In the previous chapter we analysed the spontaneous emergence of patterns in reaction-diffusion systems and how similar patterns can be observed with different network topologies.

In many cases of interest it is however important to oppose the natural tendency to the formation of patterns, by preserving (or recovering) the synchronized state [135, 167]. Synchronization plays indeed a pivotal role in many branches of science: the efficient coordination of a multitude of events is often decisive to have a system operated as a unison orchestra. In an alternating current electric power grid, one needs to match the speed and frequency of any given generator to the other sources of the shared network [56, 57, 102]. In neuroscience, patterns of synchronous firings are promoted by dedicated neuronal feedbacks. Circadian rhythms are another example that certifies the ubiquitous tendency towards entrainable oscillations as displayed by a vast plethora of biological processes [69, 75]. In computer science, synchronization is customarily referred to as consensus [152], a form of final agreement, stationary or time dependent, which is reached by a crowd of interacting agents.

Given these premises it is in general important to devise suitable control strategies to stabilize, and possibly preserve, the synchronous regime. The control is classically applied to the reactive component of the dy-

namics, and ultimately shapes the local interaction between constitutive elements [90]. Global, mean field term can be also accommodated for so as to induce the sought behavior. When the dynamics flow on a network, topology matters and does play a prominent role in eliciting the instability [16, 46]. This observation motivates the search of alternative control protocols, which leave the reaction part unchanged, while acting on the underlying web of inter-nodes connections [15, 178].

We here deal with systems like 1.11, however to illustrate the proposed method we shall again operate in the framework of the Complex Ginzburg-Landau Equation (CGLE). The CGLE admits a uniform fully synchronized solution, the spatially extended replica of the periodic orbit displayed by the system in its a-spatial version, provided the Laplacian is balanced (equal incoming and outgoing connectivity) [141]. Hereafter, we shall assume that the nodes of the network where oscillators lie are initially paired (and the reaction parameters set) so as to make the system unstable to externally injected, non homogeneous, perturbations. The network of connections is then globally reshaped (keeping the reaction parameters unchanged) to regain the stability of the synchronized, time dependent, solution. We will then move forward to considering a system of coupled (real) Ginzburg-Landau equations [121], which admits a stationary stable fixed point. Turing-like instabilities will be controlled, hence formally prevented, with a supervised intervention targeted to the net of interlaced couplings which is again based on the identification of pattern-relevant (and irrelevant) Laplacian eigenmodes.

This chapter is organized as follows. In the next section we will introduce the CGLE and carry out a linear stability analysis to delineate the conditions that make the spatially extended homogenous limit cycle solution stable. We will in particular elaborate on the remarkable differences that arise when the system involves a finite and discrete collection of interlinked oscillators, as opposed to the reference case where the population of elementary constituents is made infinite and continuous. In Section 3.2 we will provide the mathematical basis for the proposed control method. The approach will be successfully tested by operating with the CGLE and assuming a symmetric network. In Section 3.3, we will consider a directed, although balanced, network of couplings and extend to this setting the analysis.

3.1 Complex Ginzburg-Landau equation: linear stability analysis

Consider an ensemble made of N nonlinear oscillators and label with x_i their associated complex amplitude, where $i = 1, ..., N$. Each individual oscillator obeys to a Stuart Landau model, which combines linear and nonlinear (cubic) contributions. In addition, we assume the oscillators to be mutually coupled via a diffusive-like interaction which is mathematically exemplified via the discrete Laplacian operator. In this chapter we will deal with symmetric or balanced and directed networks, hence $k_i^{out} = k_i^{in} \equiv k_i$. The spatially extended system can therefore be cast in the form of CGLE:

$$\frac{d}{dt}x_j = x_j - (1 + ic_2)|x_j|^2 x_j + (1 + ic_1)K \sum_k L_{jk}^D x_k \quad j = 1, ..., N \quad (3.1)$$

where $c_1, c_2, K \in \mathbb{R}$ and $x_j \in \mathbb{C}$. We shall begin by considering a symmetric adjacency matrix and postpone to a later stage the case of a directed, though balanced, network of couplings.

Hereafter, we will study the equilibria of the system and their stability. It is however important to make a brief digression on the continuum limit version of the CGLE, which will be useful to separate local and global contributions so as to understand how the topology is at stake to stabilize or destabilize the considered equilibria.

Let us start by considering a regular lattice, embedded on a Euclidean space of arbitrary dimension. By performing the continuum limit, i.e assuming the linear distance between neighbor nodes to asymptotically vanish, one can formally replace the discrete variable $x_j(t)$ $(j = 1, 2, \cdots, N)$ with its continuous counterpart $x(\boldsymbol{r}, t)$. Here, $x(\boldsymbol{r}, t) \in \mathbb{C}$ and $\boldsymbol{r}$ identifies the space location. Under these conditions, the discrete operator $\boldsymbol{L^D}$ transforms into ∇^2, the standard Laplacian on a continuous support. For this reason, and with a slight abuse of language, we shall often employ the adjective spatial to tipify the nature of the coupling, even when the network of oscillators is not necessarily bound to a physical space.

As a preliminary remark we note that $x^{LC}(t) = e^{-\mathrm{ic_2}t}$ is a homogenous solution of the CGLE, both in its discrete or continuous, spatially extended, versions. This latter can be referred to as the limit cycle (LC) solution, since it results from a uniform, fully synchronized, replica of the periodic orbit displayed by the system in its a-spatial ($K = 0$) version.

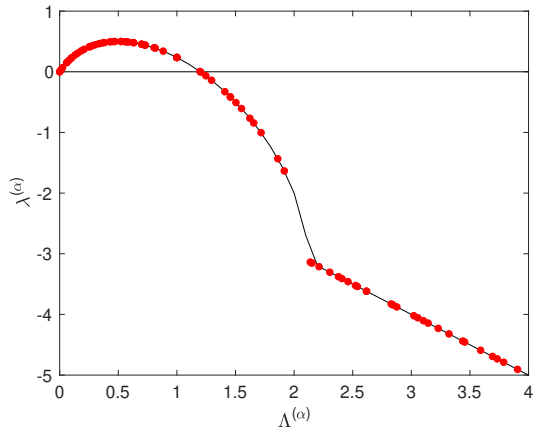


Figure 3.1: Continuous (solid line) and discrete (red dots) dispersion relation. The dots correspond to a CGLE with $c_1 = -1.8$, $c_2 = 1.6$, $K = 1$ for a network composed of 100 nodes, generated from the Watts-Strogatz method with rewiring probability 0.8.

In the remaining part of this chapter, we shall determine the stability of the LC solution. We will deal at first with the continuous version of the model and re-derive for completeness the conditions for the onset of the so called Benjamin-Feir instability [25, 184]. The peculiarities that stem from assuming a discrete and heterogeneous web of symmetric couplings will be also reviewed.

To assess the stability of the LC solution we introduce a non homogeneous perturbation, both in phase and amplitude:

$$x(\boldsymbol{r},t) = x^{LC}(t)[1+\rho(\boldsymbol{r},t)]e^{\mathrm{i}\theta(\boldsymbol{r},t)}. \tag{3.2}$$

Linearizing around the LC ($\rho(\boldsymbol{r},t)=0$, $\theta(\boldsymbol{r},t)=0$) one readily obtains:

$$\frac{d}{dt}\begin{bmatrix}\rho\\ \theta\end{bmatrix} = \begin{bmatrix}-2 & 0\\ -2c_2 & 0\end{bmatrix}\begin{bmatrix}\rho\\ \theta\end{bmatrix} + K\begin{bmatrix}1 & -c_1\\ c_1 & 1\end{bmatrix}\nabla^2\begin{bmatrix}\rho\\ \theta\end{bmatrix}. \tag{3.3}$$

To solve the above linear problem we perform a space-time Fourier transform:

$$\begin{aligned}\rho(\boldsymbol{r},t) &= \int\int d\omega d\boldsymbol{k} e^{\mathrm{i}\omega t}e^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\rho_{\boldsymbol{k}}\\ \theta(\boldsymbol{r},t) &= \int\int d\omega d\boldsymbol{k} e^{\mathrm{i}\omega t}e^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\theta_{\boldsymbol{k}}.\end{aligned} \tag{3.4}$$

A straightforward calculation returns the following condition that should be matched as a necessary consistent requirement for the linear problem to admit a non-trivial solution:

$$\det \begin{bmatrix} \lambda + 2 + Kk^2 & -Kc_1k^2 \\ 2c_2 + Kc_1k^2 & \lambda + Kk^2 \end{bmatrix} = 0 \tag{3.5}$$

with $\lambda = \mathfrak{i}\omega$ and $k = |\boldsymbol{k}|$. The quantity λ hence assesses the linear growth rate associated to the k-th mode. Without losing generality we will hereafter set the coupling constant to unit ($K = 1$) and proceed with the calculation to determine the root of the characteristic polynomial with largest real part:

$$\lambda(k^2) = -k^2 - 1 + \sqrt{-c_1^2k^4 - 2c_1c_2k^2 + 1}. \tag{3.6}$$

The perturbation that shakes the homogenous and time dependent solution $x^{LC}(t)$ gets exponentially magnified in the linear regime of the evolution provided that the real part of λ (the dispersion relation λ_{Re}) is positive. Notice that $\lambda(0) = \lambda_{Re}(0) = 0$, as expected, based on an obvious argument of internal coherence. Expanding (3.6) for small k returns $\lambda_{Re} \simeq -(1 + c_1c_2)k^2$. The stability of the synchronized LC solution is therefore lost when $1 + c_1c_2 < 0$, the standard condition for the onset of the Benjamin-Feir instability.

We now turn to considering the case of a heterogenous, although symmetric, network of connections among oscillators. To investigate the conditions to be met for a symmetry breaking instability of the homogeneous LC solution, we proceed in analogy with the above and set $x_j(t) = x^{LC}(t)[1+\rho_j(t)]e^{i\theta_j(t)}$, with a clear meaning of the symbols. Plugging this latter expression in the CGLE (3.1) and expanding to the first order in the perturbation amount, one obtains the obvious generalization of system (3.3):

$$\frac{d}{dt}\begin{bmatrix} \rho_j \\ \theta_j \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ -2c_2 & 0 \end{bmatrix}\begin{bmatrix} \rho_j \\ \theta_j \end{bmatrix} + \begin{bmatrix} 1 & -c_1 \\ c_1 & 1 \end{bmatrix} \sum_k L^D_{jk} \begin{bmatrix} \rho_k \\ \theta_k \end{bmatrix}. \tag{3.7}$$

where we recall that $K = 1$. For regular lattices, the Fourier transform is usually invoked to solve the system of equations homologous to (3.7). This amounts to expanding the spatial perturbations on a set of planar

waves, the eigenfunctions of the continuous Laplacian operator. When the system is instead defined on a network, an analogous procedure can be employed. To this end, we define the eigenvalues and eigenvectors of the discrete Laplacian operator:

$$\sum_j L^D_{ij}\phi^{(\alpha)}_j = \Lambda^{(\alpha)}\phi^{(\alpha)}_i \quad \alpha = 1, ..., N \tag{3.8}$$

When the network is undirected, the Laplacian operator is symmetric. Therefore, the eigenvalues $\Lambda^{(\alpha)}$ are real and the eigenvectors $\boldsymbol{\phi}^{(\alpha)}$ form an orthonormal basis. This condition needs to be relaxed when dealing with the more general setting of a directed graph, as we shall discuss in the second part of the paper[1]. The symmetric Laplacian matrix $\boldsymbol{L^D}$ has a single zero eigenvalue $\Lambda^{(\alpha=1)}$ corresponding to the uniform eigenvector and all other eigenvalues are negative. The indices α are sorted so as to satisfy $0 = \Lambda^{(1)} > \Lambda^{(2)} \geq \cdots \geq \Lambda^{(N)}$.

The inhomogeneous perturbations ρ_j and θ_j can be expanded as:

$$\begin{bmatrix} \rho_j \\ \theta_j \end{bmatrix} = \sum_{\alpha=1}^{N} \begin{bmatrix} \rho^{(\alpha)} \\ \theta^{(\alpha)} \end{bmatrix} e^{\lambda t}\phi^{(\alpha)}_j. \tag{3.9}$$

By inserting (3.9) in (3.7) and making use of relation (3.8) one eventually gets a condition formally equivalent to expression (3.5). As an important difference, the eigenvalues of the continuous Laplacian, $-k^2$, are replaced by the discrete (real and negative) quantities $\Lambda^{(\alpha)}$, the eigenvalues of the discrete Laplacian. Insisting on the analogy, it is of immediate evidence that the instability rises for a CGLE defined on a symmetric network when $1 + c_1c_2 < 0$, a dynamical condition identical to that obtained when operating under the continuous, by definition regular, viewpoint. The quantity $\Lambda^{(\alpha)}$ constitutes the analogue of the wavelength for a spatial pattern in a system defined on a continuous regular lattice. It is this latter quantity which determines the spatial characteristic of the emerging patterns, when the system is defined on a heterogeneous complex support.

In figure 3.1 the dispersion relation λ_{Re} is plotted versus $-\Lambda^{(\alpha)}_{Re}$, for a specific choice of c_1 and c_2, so that $1 + c_1c_2 < 0$. The solid line refers to the continuum setting ($-\Lambda^{(\alpha)}_{Re} \to k^2$,), while circles are obtained when

[1] A diagonalizable and connected Laplacian matrix is instead a minimal requirement to be satisfied by our analytical treatment both in the symmetric and in the directed case.

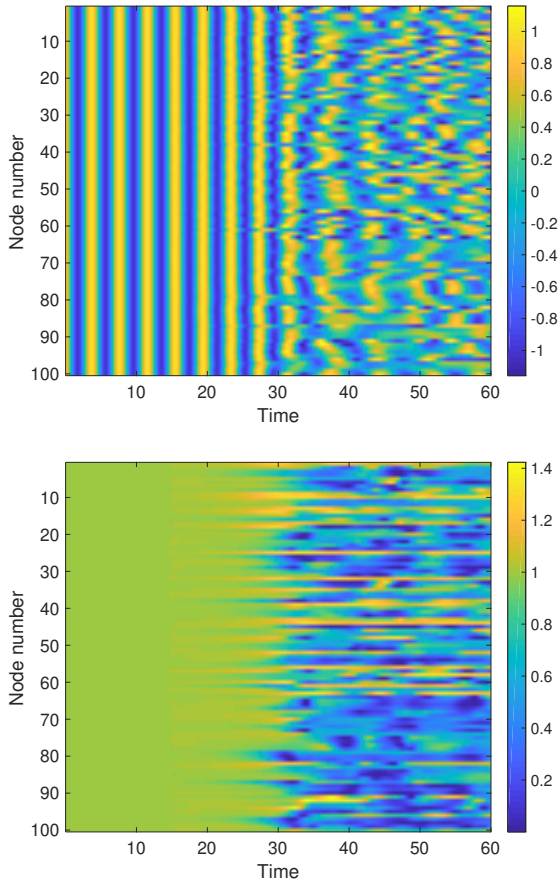


Figure 3.2: Evolution of $\boldsymbol{x}_{Re}$ (upper panel) and $|\boldsymbol{x}|^2$ (lower panel) versus time, assuming a uniform LC initial condition. At time $\tau_1 = 15$, a non homogeneous perturbation is inserted and the synchronized state is consequently disrupted. Here, $c_1 = -1.8$, $c_2 = 1.6$. The nonlinear oscillators are mutually linked via the Watts-Strogatz network, used in depicting the discrete dispersion relation of figure 3.1.

operating with the CGLE, hosted on a Watts-Strogatz network [200]. As anticipated, the discrete collection of points which defines the dispersion relation when a symmetric, finite and heterogenous network of coupling is accommodated for, follows the same profile which applies to the limiting continuum setting.

Let us observe that the continuous curve representing the continuum limit only depends on the reaction function f. The byproduct of the Laplacian (and therefore of the network topology) relies on how the dots of the discrete dispersion relation distribute along the curve.

In figure 3.2 the time evolution of the system is displayed for a choice of the parameters which corresponds to the unstable dispersion relation of figure 3.1. After a given time the synchronized LC solution is perturbed by insertion of an external source of non homogenous disturbance. This latter grows, as predicted by the linear stability analysis, and yields the irregular patterns displayed for both $\boldsymbol{x}_{Re}$ and $|\boldsymbol{x}|^2$.

Back to Fig. 3.1, it is however important to realize that the instability actually takes place only when at least one eigenvalue $-\Lambda_{Re}^{(\alpha)}$ exists in the range where λ_{Re} is positive. If the ensemble of discrete modes, which ultimately reflects the topology of the imposed couplings, populates the portion of the dispersion relation with $\lambda_{Re} < 0$, no instability can develop, even if $1 + c_1 c_2 < 0$. Stated differently, the spectral gap, i.e. the difference between the moduli of the two largest eigenvalues of the Laplacian operator, $|\Lambda^{(2)}| - |\Lambda^{(1)}| \equiv |\Lambda^{(2)}|$, should be larger than $-2(c_1 c_2 + 1)/(1 + c_1^2)$, the non trivial root of (3.6), for the instability to take place.

This observation has been exploited by Nakao in [141] to propose a novel control strategy aimed at suppressing the Benjamin-Feir instability and thus preserving the initial synchronized regime for a CGLE defined on a symmetric network support. Imagine to start with an unstable condition, which in turn implies to operate with a suitable choice for both the reaction parameters and the network specificity. The key idea of [141] is to randomly rewire the network so as to make the second eigenvalue progressively more negative. Random moves are accepted or rejected following a Metropolis scheme. The numerical procedure converges to a (globally) modified network which has no eigenvalue in the range where $\lambda_{Re} > 0$. The control is topological since it only affects the couplings that link the oscillators, without acting on the dynamical parameters c_1 and c_2. In practice, the discrete network-like system can be made stable for a choice of

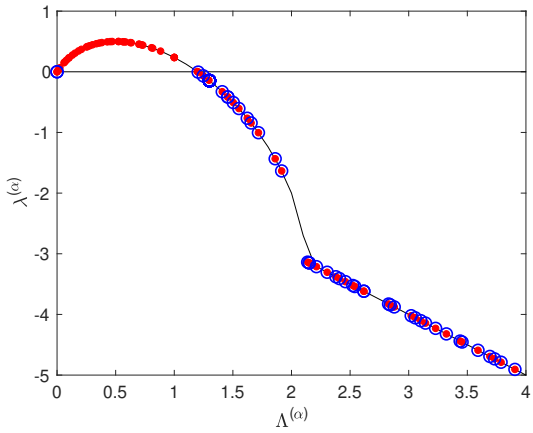


Figure 3.3: Dispersion relation: red dots correspond to the original system of Fig. 3.1 while blue circles show the dispersion relation obtained for the controlled adjacency matrix.

the parameter that would drive a Benjamin-Feir instability in the continuum limit. Building on these intriguing observations, we will here devise an analytical approach that enables to implement a similar control protocol, without resorting to an iterative, numerically supervised, rewiring. Importantly, the method that we shall introduce here can be successfully extended to the general case where a directed network of connections is assumed to hold. The next section is devoted to discussing the proposed method.

3.2 Global topological control

We here aim at developing an appropriate control strategy which acts on the global network of connections, leaving unchanged the dynamical parameters of the model. The method that we shall hereafter discuss takes inspiration from the seminal work of Nakao [141]. There it was shown that a numerically supervised rewiring of the inter-oscillators couplings can stabilize the CGLE, thus preserving the consensus state. Building on similar grounds, we will provide hereafter an analytical procedure to achieve the sought stabilization. The proposed method allows to imme-

diately generate the controlled matrix of contacts, without involving any iterative scheme, thus being free from concerns on the numerical convergence. Starting from a condition of instability, as displayed in Fig. 3.1, we wish to modify the spectrum of the Laplacian operator so as to force the finite and discrete collection of modes to populate the negative branch of the dispersion relation λ_{Re}.

As emphasized in the previous section, when the network is undirected the discrete dispersion relation superposes to the continuum one (see the solid line in Fig. 3.1). The instability localizes on a finite set of modes, those falling on the positive bump of the curve $\lambda_{Re}(k^2)$. Is it possible to alter the network topology so as to make the (negatively defined and real) eigenvalues larger in absolute value than $-2(c_1c_2+1)/(1+c_1^2)$, the point where the parabola $\lambda_{Re}(k^2)$ crosses the horizontal axis, so turning negative? In a figurative sense, we want to slide the discrete points of Fig. 3.1 onto the curve, as beads on a cord, causing them to reach its negative branch. To answer this question we rely on the network generation method exposed in the previous chapter readjusted for this problem where, for simplicity, the Laplacian eigenvectors are preserved unchanged.

Let us start by defining the $N \times N$ matrix $\boldsymbol{\Phi}$ whose columns are the eigenvectors $\boldsymbol{\phi}^{(1)}, \cdots, \boldsymbol{\phi}^{(N)}$ of the Laplacian operator $\boldsymbol{L^D}$. Hence $\boldsymbol{\Lambda} = \Phi^{-1}\boldsymbol{L^D}\Phi$, where $\boldsymbol{\Lambda}$ is the diagonal matrix formed by the eigenvalues $\Lambda^{(1)}, ..., \Lambda^{(N)}$. We then calculate the minimal corrections $\delta\Lambda^{(\alpha)}$ ($\alpha = 1, \cdots, N$) that need to be imposed to shift the eigenvalues $\Lambda^{(\alpha)}$ on the stable side of the dispersion relation. The computed corrections are then organized in a diagonal matrix $\boldsymbol{\delta\Lambda}$, so that $\delta\Lambda_{\alpha\alpha} = \delta\Lambda^{(\alpha)}$, for $\alpha = 1, \cdots, N$. As a matter of fact, and to keep the formulation general, $\delta\Lambda^{(\alpha)} \in \mathbb{C}$. For the case of a symmetric network that we are bound to explore within this section, the quantities $\delta\Lambda^{(\alpha)}$ are however real and negative. When the original $\Lambda^{(\alpha)}$ falls in the region of stability, the corresponding correction $\delta\Lambda^{(\alpha)}$ is set to zero.

The next step of the procedure is to perform the following transformation

$$\boldsymbol{\delta L^D} = \Phi\boldsymbol{\delta\Lambda}\Phi^{-1} \tag{3.10}$$

and define the controlled matrix $\tilde{\boldsymbol{L}}^{\boldsymbol{D}} = \boldsymbol{L^D} + \delta\boldsymbol{L^D}$. By construction the eigenvalues of $\tilde{\boldsymbol{L}}^{\boldsymbol{D}}$ (with the only exception of the zero eigenvalue, $\alpha = 1$) are smaller than $2(c_1c_2+1)/(1+c_1^2)$.

As it is proven in the previous chapter the matrix $\tilde{\boldsymbol{L}}^{\boldsymbol{D}}$ can be inter-

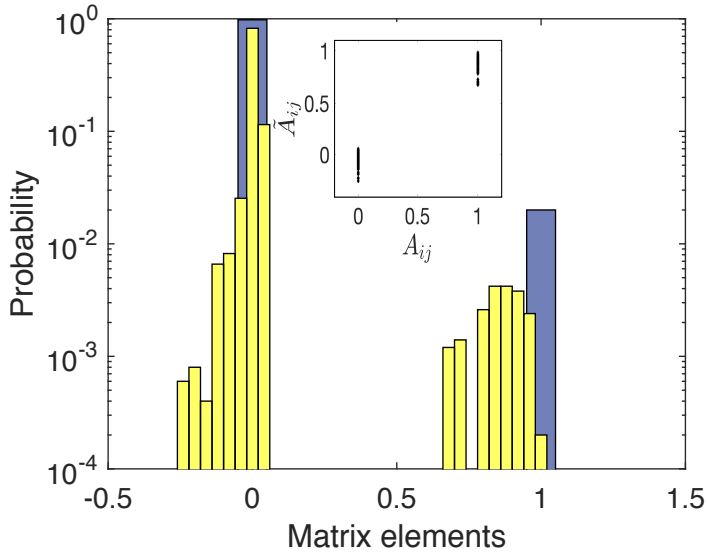


Figure 3.4: Main panel: distribution of the elements of the adjacency matrix before (large dark bins) and after (light small bins) the control. Initially the distribution displays two peaks localized in 0 and 1, reflecting the choice of a binary matrix of contacts. The controlled adjacency matrix is still bimodal, but the peaks are now smoothed out. Importantly, negative connections, pointing to inhibitory loops, should be accommodated for in the rewired weighed network. Inset: the elements of the controlled adjacency matrix $(\tilde{A})_{ij}$ are plotted vs. the original adjacency matrix A_{ij} and a clear correlation is displayed. The control manifests as a rather local modification of the weights, strong (resp. weak) couplings being preserved under the imposed rewiring.

preted as a Laplacian operator, namely (i) the entries of the matrix are real and (ii) every row of the matrix sums up to zero, $\sum_i (\tilde{L}^D)_{ij} = 0 \; \forall j$). As an important complement, in the Appendix it is also shown that symmetry and balancedness are perpetuated from $\boldsymbol{L^D}$ to $\tilde{\boldsymbol{L}}^{\boldsymbol{D}}$.

Hence, the obvious conclusion is that we have generated a modified adjacency matrix $\tilde{\boldsymbol{A}}$, hidden inside $\tilde{\boldsymbol{L}}^{\boldsymbol{D}}$, which should engender a negative dispersion relation (when employed in the CGLE, at fixed c_1 and c_2), thus preserving the stability of the synchronized configuration.

In the remaining part of this section we will test the proposed control scheme assuming a symmetric matrix of inter-nodes couplings. In the next Section we will turn to discussing the more general case of a directed, although balanced, adjacency matrix. To demonstrate the adequacy of the technique, we will assume the setting depicted in Fig. 3.1: the parameters (c_1, c_2) and the underlying network of contact are chosen so as to make the system unstable to external non homogeneous perturbations. By rewiring the network following the strategy outlined above we obtain the dispersion relation represented by blue circles in Fig. 3.3. The circles stand for the discrete dispersion relation and populate the negative portion of the continuous curve: the instability has been hence removed, by solely acting on the topology of the graph. This latter was initially assumed of the binary type: the entries of the adjacency matrix are therefore a collection of zeros and ones. The elements of the controlled matrix are still characterized by a bimodal distribution, as displayed in Fig. 3.4. Each element of the controlled adjacency matrix $(\tilde{A})_{ij}$ takes a value close to the initial entry A_{ij}. In practice, the control returns a local adjustment of the weights, strong (resp. weak) couplings being preserved under the imposed rewiring. Interestingly, negative coupling constants appear as a result of the continuous smoothing of the peak initially localized in zero. Inhibitory interactions should be hence at play for an effective stabilization of the dynamics. In general the new network appears to be much more connected, as it is also shown by its spectral gap, larger in absolute value than the one corresponding to the original network.

To provide a numerical validation of our conclusion, we evolved for a transient the CGLE assuming the original, unstable and binary, adjacency matrix. When the imposed perturbation has grown to become significant, we instantaneously switched to the controlled Laplacian. As shown in Fig. 3.5, the perturbation fades progressively away and the synchronous

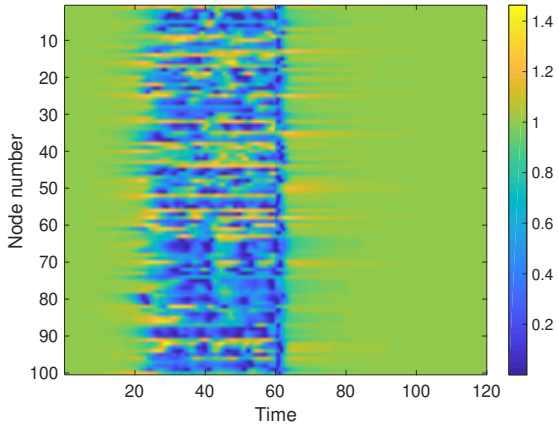


Figure 3.5: $|\boldsymbol{x}|^2$ vs. time. The system assumes initially a binary matrix of connections and it is unstable to external non homogenous perturbation. At time $\tau_1 = 10$ the LC is perturbed, and the injected disturbances grow, yielding the expected loss of synchronization. At time $\tau_2 = 60$ the adjacency matrix is instantaneously controlled, according to the exposed scheme. The perturbation is then re-absorbed and the consensus state recovered.

dynamics is eventually restored. The proposed control scheme was originally devised to contrast the onset of instability and, as such, targeted to the linear regime of the evolution. As shown in Fig. 3.5, the method proves however effective in stabilizing the system also at relatively large time, when nonlinearities are at play.

3.3 Controlling the instability on balanced directed networks

Let us now turn to considering the case of a CGLE defined on a directed, heterogeneous although balanced (for each node the sum of incoming weights coincides with the sum of outgoing weights) network. Before discussing the application of the control technique introduced in the previous section, we will review the conditions that determine the emergence of instability.

When reaction-diffusion systems are placed on directed, hence asym-

metric graphs, patterns can develop, even if they are formally impeded on a symmetric, continuum or discrete, spatial support. Directionality matters and proves indeed fundamental in shaping the emerging patterns. The conditions for the asymmetry driven instability, reminiscent of a Turing like mechanism, for a multi-species reaction diffusion model evolving on a directed graph have been discussed in [16]. In this latter case, the perturbation acts on a homogeneous fixed point, a time independent equilibrium for the reaction dynamics. In [54] the analysis has been extended to the setting where the unperturbed homogeneous solution is a LC and thus depends explicitly on time. In the following, for the sake of consistency, we will go through the analysis of [54] to eventually obtain the conditions that instigate the topological instability of a time-dependent solution of the LC type.

By perturbing $\boldsymbol{x}^{LC}(t)$ as discussed in the first Section, one eventually ends up with the self-consistent condition:

$$\det \begin{bmatrix} -2+\Lambda^{(\alpha)}-\lambda & -c_1\Lambda^{(\alpha)} \\ -2c_2+c_1\Lambda^{(\alpha)} & \Lambda^{(\alpha)}-\lambda \end{bmatrix} = 0 \tag{3.11}$$

which is equivalent to $\det\left(\mathbf{J}_\alpha - \lambda \mathbb{I}_2\right) = 0$ with:

$$\mathbf{J}_\alpha = \begin{pmatrix} -2+\Lambda^{(\alpha)} & -c_1\Lambda^{(\alpha)} \\ -2c_2+c_1\Lambda^{(\alpha)} & \Lambda^{(\alpha)} \end{pmatrix} \quad .$$

Recall that for an asymmetric network, the Laplacian eigenvalues $\Lambda^{(\alpha)}$ are complex. Furthermore, $\Lambda^{(\alpha)}_{Re} < 0$, since the spectrum of the Laplacian matrix falls in the left half of the complex plane, according to the Gerschgorin theorem [23]. Simple calculations yield:

$$\begin{aligned} (\mathrm{tr}\mathbf{J}_\alpha)_{Re} &= -2 + 2\Lambda^{(\alpha)}_{Re} \\ (\mathrm{tr}\mathbf{J}_\alpha)_{Im} &= 2\Lambda^{(\alpha)}_{Im} \\ (\det\mathbf{J}_\alpha)_{Re} &= -2\Lambda^{(\alpha)}_{Re} + (\Lambda^{(\alpha)}_{Re})^2 - (\Lambda^{(\alpha)}_{Im})^2 \\ &\quad - 2c_1c_2\Lambda^{(\alpha)}_{Re} + c_1^2\left[(\Lambda^{(\alpha)}_{Re})^2 - (\Lambda^{(\alpha)}_{Im})^2\right] \\ (\det\mathbf{J}_\alpha)_{Im} &= -2\Lambda^{(\alpha)}_{Im} + 2(1+c_1^2)\Lambda^{(\alpha)}_{Re}\Lambda^{(\alpha)}_{Im} \\ &\quad - 2c_1c_2\Lambda^{(\alpha)}_{Im}\Lambda^{(\alpha)}_{Im} \end{aligned} \tag{3.12}$$

with a clear meaning of the chosen notation.

From (3.11), one gets:

$$\lambda = \frac{1}{2}\left[(\mathrm{tr}\mathbf{J}_\alpha)_{Re} + \gamma\right] + \frac{1}{2}\left[(\mathrm{tr}\mathbf{J}_\alpha)_{Im} + \delta\right]\mathrm{i} \tag{3.13}$$

where:

$$\gamma = \sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}} \tag{3.14}$$

$$\delta = sgn(b)\sqrt{\frac{-a + \sqrt{a^2 + b^2}}{2}} \tag{3.15}$$

and:

$$\begin{aligned} a &= [(\mathrm{tr}\mathbf{J}_\alpha)_{Re}]^2 - [(\mathrm{tr}\mathbf{J}_\alpha)_{Im}]^2 - 4(\mathrm{det}\mathbf{J}_\alpha)_{Re} \\ b &= 2(\mathrm{tr}\mathbf{J}_\alpha)_{Re}(\mathrm{tr}\mathbf{J}_\alpha)_{Im} - 4(\mathrm{det}\mathbf{J}_\alpha)_{Im}. \end{aligned} \tag{3.16}$$

As discussed in [16,54], diffusion driven instabilities arise also when $\mathrm{tr}(\mathbf{J}_\alpha)_{Re} < 0$, as opposed to what happens when the system evolves on a symmetric spatial support. In fact, $\lambda_{Re} > 0$ if:

$$|(\mathrm{tr}\mathbf{J}_\alpha)_{Re}| \leqslant \sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}} \tag{3.17}$$

a condition that can be met for $\mathrm{tr}(\mathbf{J}_\alpha)_{Re} < 0$, if the network of interactions is made directed and, consequently, an imaginary component of the Laplacian spectrum is accommodated for. A straightforward, though lengthy, calculation allows one to derive the following compact condition for the topology instability to occur:

$$S_2(\Lambda_{Re}^{(\alpha)}) \leq S_1(\Lambda_{Re}^{(\alpha)}) \left[\Lambda_{Im}^{(\alpha)}\right]^2 \tag{3.18}$$

where

$$\begin{aligned} S_2(\Lambda_{Re}^{\alpha}) &= C_{2,4}(\Lambda_{Re}^{(\alpha)})^4 - C_{2,3}(\Lambda_{Re}^{(\alpha)})^3 + C_{2,2}(\Lambda_{Re}^{(\alpha)})^2 \\ &\quad - C_{2,1}\Lambda_{Re}^{(\alpha)} \\ S_1(\Lambda_{Re}^{\alpha}) &= C_{1,2}(\Lambda_{Re}^{\alpha})^2 - C_{1,1}\Lambda_{Re}^{(\alpha)} + C_{1,0} \end{aligned} \tag{3.19}$$

with

$$\begin{aligned} C_{2,4} &= 1 + c_1^2 \\ C_{2,3} &= 4 + 2c_1c_2 + 2c_1^2 \\ C_{2,2} &= 5 + 4c_1c_2 + c_1^2 \\ C_{2,1} &= 2 + 2c_1c_2 \\ C_{1,2} &= c_1^4 + c_1^2 \\ C_{1,1} &= 2c_1^3c_2 + 2c_1^2 \\ C_{1,0} &= c_1^2(1 + c_2^2) \qquad . \end{aligned} \tag{3.20}$$

Notice that (3.18) reduces to $S_2(\Lambda_{Re}^{(\alpha)}) \leq 0$ when dropping the imaginary components of $\Lambda^{(\alpha)}$ or, equivalently, when assuming a symmetric network of couplings. Expanding the solution for small $\Lambda_{Re}^{(\alpha)}$, assumed as a continuum variable, one readily gets $1 + c_1c_2 < 0$, i.e. the standard condition for the Benjamin-Feir instability on a symmetric support.

To gain insight into the above analysis, we generate a directed and balanced network, via a suitable modification of the Newman-Watts (NW) algorithm [147]. We begin from a substrate L-regular ring made of N nodes and add, on average, NLp long-range directed links. Here, $p \in [0, 1]$ is a probability that quantifies the amount of introduced long-range links. To keep the network balanced, the insertion of a long-range link stemming from node i is followed by a fixed number (3 is our arbitrary choice) of additional links to form a loop that closes on i [16].

In Fig. 3.6(a) the dispersion relation λ_{Re} is plotted as a function of $-\Lambda_{Re}^{(\alpha)}$. The black solid line refers to the limiting case of a symmetric (and continuum) support: the reaction parameters (c_1, c_2) are chosen so as to prevent the instability to develop since $\lambda_{Re} < 0$. The online circles refer instead to the directed case: the points abandon the solid curve and lift above zero, signaling a topology driven instability of the uniform LC solution.

In Fig. 3.6(b) the same situation is illustrated in the reference plane $(\Lambda_{Re}^{(\alpha)}, \Lambda_{Im}^{(\alpha)})$. Once the reaction parameters c_1 and c_2 have been assigned, one can calculate the coefficients $C_{1,q}(q = 0, 1, 2)$ and $C_{2,q}(q = 0, ..., 4)$ via (3.20). The inequality (3.18) allows us to draw the domain of instability, depicted as a shaded region in Fig. 3.6(b). Each eigenvalue (blue circles) of the discrete Laplacian corresponds to a localized point in the plane $(\Lambda_{Re}^{(\alpha)}, \Lambda_{Im}^{(\alpha)})$. The instability develops when at least one non-null

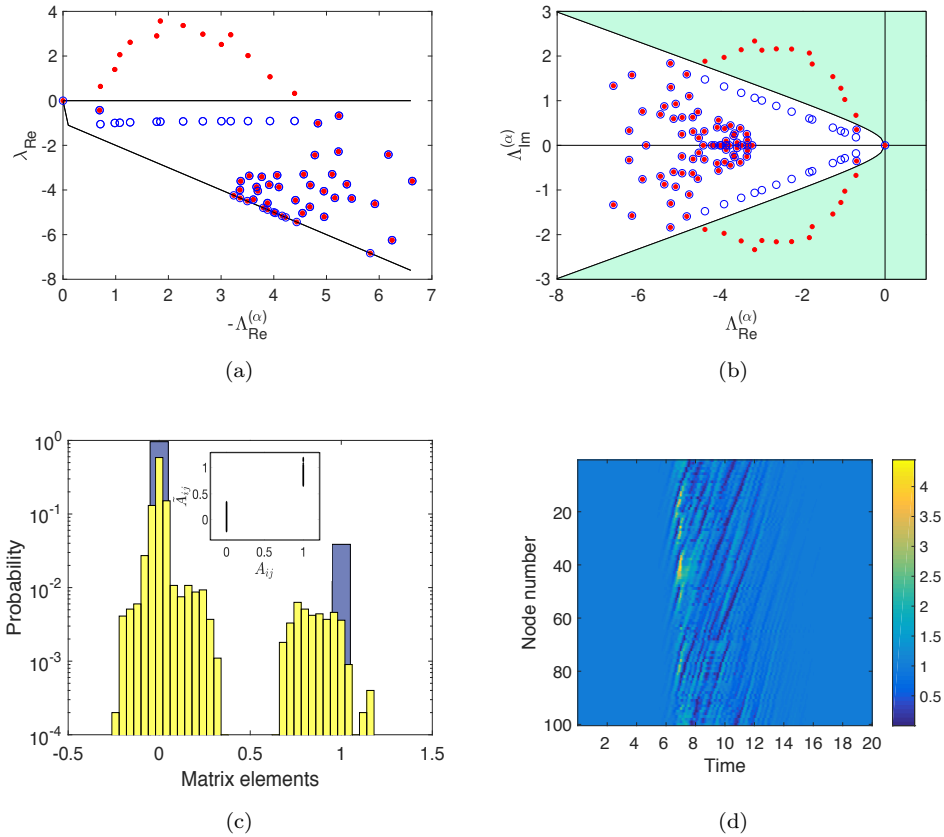


Figure 3.6: (a): The dispersion relation λ_{Re} as a function of $-\Lambda_{Re}^{(\alpha)}$. The solid line stands for the continuum dispersion relation. The red points are obtained for the CGLE defined on a NW directed balanced network with $p = 0.27$. The blue circles represent the dispersion relation obtained for the controlled matrix. Here, $c_1 = 3$ and $c_2 = 2.4224$. (b): Eigenvalues positioned in the complex plane. The shaded area represents the instability region obtained from (3.18). The eigenvalues in this region correspond to the unstable modes, characterized by $\lambda_{Re} > 0$, in panel (a). (c): Analogous of Fig. 3.4. Also for the directed case, a clear correlation between the two is observed. The control induces a rather local modification of the couplings, strong (resp. weak) couplings being preserved under the imposed rewiring. (d): $|\boldsymbol{x}|$ vs. time. The system corresponding to the above matrix of connections is unstable to external non homogenous perturbation. At time $\tau_1 = 5$ the LC is perturbed, and the injected disturbances develop, yielding a loss of synchronization, as predicted by the linear stability analysis. At time $\tau_2 = 7$ the adjacency matrix is instantaneously controlled, as follows the devised scheme. The perturbation is consequently re-absorbed and the synchronized configuration recovered.

eigenvalue enters the shaded region. For an undirected graph, the points are distributed on the (horizontal) axis, thus outside the region deputed to the instability. When the graph turns asymmetric the imaginary component of $\Lambda^{(\alpha)}$ promotes an instability, which bears a direct imprint of the network topology. As usual, the instability will eventually unfold complex pattern, in the nonlinear regime of the evolution.

Starting from this setting, and to restore the synchronization, one can rewire the network connections according to the control procedure outlined in the preceding Section [2]. In this case, one needs to operate in the complex plane $(\Lambda^{(\alpha)}_{Re}, \Lambda^{(\alpha)}_{Im})$, and act simultaneously on the imaginary component of $\Lambda^{(\alpha)}$, to force the eigenvalues outside the region of instability [3] . In other words the elements of the diagonal matrix $\tilde{\mathbf{\Lambda}}$ which encodes for the imposed shifts are, in general, complex. For the case at hand, the spectrum of the controlled Laplacian operator is displayed in Figs. 3.6(a) and 3.6(b) with red stars. The dispersion relation λ_{Re} is now consistently negative, reflecting the fact that stabilization has been enforced into the model. Similarly, stars populate the domain of stability in Fig. 3.6(b) without invading the shaded portion of the plane.

As for the preceding case, the initial adjacency matrix is assumed binary. The elements of the controlled matrix still display a bimodal distribution (see Fig. 3.6(c)): inhibitory coupling are at play as for the case of a symmetric support.

To conclude this section we provide a numerical validation of the implemented method. In Fig. 3.6(d) we initially evolve the perturbation assuming the unstable and directed adjacency matrix. Then, when the perturbation has evolved in a nonlinear *quasi*-wave, the Laplacian is instantaneously mutated into its controlled counterpart. The perturbation damps and the system regains the initial homogenous consensus state. We again remark that the control is also effective when acted far from the linear regime of the evolution, when nonlinearities are presumably playing a

[2]Remember that the mathematical proofs provided in the previous section hold in general, assuming a directed (balanced and diagonalizable) Laplacian

[3]The control can be implemented in different ways. The eigenvalues can be moved for instance horizontally, by acting on their real component, vertically by modifying their imaginary part, or diagonally, by resorting to a linear combination of the two aforementioned strategies. Real eigenvalues lay on the horizontal axis: when falling in the region of instability, they are transferred into the stable domain by sliding them horizontally, namely by imposing a real correction, the imaginary part proving, in this respect, useless.

role.

As a final mandatory remark, we emphasize that the developed control strategy holds in general, beyond the application to the CGLE here considered for purely pedagogical reasons. Indeed, a formally identical scheme can be applied to stabilizing homogenous time-independent fixed points, so preventing the classical Turing-like route to patterns to eventually take place. This extension is discussed for completeness in [39]by employing an *ad hoc* multispecies framework which takes inspiration from the Ginzburg-Landau reference model.

The paradigmatic approach to pattern formation deals with a set of reaction-diffusion equations: an initial homogenous equilibrium, constant or time-dependent, can turn unstable via a symmetry breaking instability, instigated by the external injection of a non homogenous disturbance. A non trivial interplay between reaction and diffusion terms, first imagined by Alan Turing in his seminal paper on morphogenesis, is ultimately responsible for the growth of the imposed perturbation. This event takes place for specific choices of the parameter setting and preludes the outbreak of the fully developed patterns. When the reaction-diffusion system is hosted on a network support, the inherent discreteness and the enforced degree of imposed asymmetry matter in determining the conditions that make the route to patterns possible. The vital role which is played by the topology of the underlying networks of contacts can be efficiently exploited to control the instability and so contrast the drive to pattern formation. In this chapter we have elaborated along these lines by devising a suitable control strategy that enforces stabilization, via a supervised redefinition of the inter-node couplings. The idea is to modify the spectrum of the Laplacian by altering the matrix of connections so as to confine the active modes outside the region of instability. The method builds on the work of Nakao [141] who numerically showed that an effective stabilization can be achieved by link-rewiring. As in [141], the Complex Ginzburg-Landau equation has been here assumed as a reference model, to provide a probing test for the newly proposed approach. In this case the control stabilizes the synchronous limit cycle uniform solution. A multispecies system that couples together two real Ginzburg-Landau equations has been also considered, to assess the performance of the method in presence of a homogeneous stationary stable fixed point. When the adjacency matrix

is symmetric, the discrete points that constitute the unstable portion of the dispersion relation are moved along the continuum parabola which embodies the characteristic of stability in the idealized continuum limit. Conversely, when the connections are asymmetric, though balanced, different strategies can be implemented to achieve the sought stabilization. One can in general act on the imaginary and real components of the spectrum of the Laplacian operator, integrating such independent moves as desired. Numerical checks confirmed the effectiveness of the proposed scheme. The specific choice for the inter-nodes diffusing couplings defined by the Laplacian operator however requires the existence of a *homogenous*, fixed or time-dependent solution for the system to be controlled.

The above analysis configures as an innovative method for controlling reaction-diffusion systems which allows to reach consensus by solely acting on the network topology. The idea is again based on a supervised modification of the Laplacian spectrum and consequently of the discrete dispersion relation which determines the stability of the synchronized equilibrium. Let us observe that the synchronized oscillating solution can prove unstable in the continuum limit but stable in the discrete space due to the network topology. By contrast it has formerly been proven [16] that the network structure (if directed) can be responsible for desynchronization when in the continuum space the limit cycle results stable.

The mathematical technique which has been developed and adopted in the last two chapters paves the way to a series of possible applications since it allows to develop several procedures aiming to network modification based on spectral properties. One example is given in the next chapter where, moving from the context of reaction-diffusion systems, we relax the linear diffusive constraint in the model to admit a more generic coupling. This will represent an increase in the system explorability by broadening the spectrum of possible equilibrium solutions, also including *inhomogenous* states. The generalized stabilization method will then become suitable for studying different systems, from natural, to man-made applications. In the next chapter we will envisage and investigate the important topic of ecosystem stability.

Chapter 4

Spectral control for ecological systems

The stability of ecosystems, or their resilience, is essential in many respects. In the past decades many efforts have been made to understand the principles that rule the stability of a complex ecosystem. These concepts have been originally addressed by Robert May [133] in a seminal paper that paved the way to a completely new field of investigation, still very productive and fertile. These studies resulted in multiple attempts of providing methods to make an ecological community structurally stable [70, 80, 101, 185–188, 196]. May [133] analysed a system described by N variables (N interacting species) obeying a set of differential equations. The stability analysis is performed by linearising the equations in the neighbourhood of an equilibrium point, whose stability depends on the spectrum of the interaction matrix. May's analysis focuses on this latter, eventually bearing to the challenging statement that, in short, the more complex the more unstable is the system. Recent work by Allesina et al. [6] provides an implementation of May's ecosystem accounting for well defined (non random) interactions. These are competitive, mutualistic and predator-prey and, according to Allesina et al., they play different roles in the stability of the ecosystem. Remarkably, the presence of predator-prey relations has a stabilizing effect. In Coyte et al. [47] stability of a microbioma ecosystem is obtained by allowing for sufficiently weak couplings.

Most of the mathematical models used to study ecosystems employ a

set of equations which can be thought as a generalization of the model introduced in Chapters 2 and 3. In particular, we here exploit the concept of network, not representing a physical allocation and displacement, but instead to describe the set of interactions within the ecosystem. The edges will then quantify the intensity of the (positive or negative) relations between different species, each one represented by a different node. In parallel, the reaction term of the equations will mimic the intrinsic time evolution of each single population, independently on the others.
Starting from these premises, in this chapter we provide an alternative approach to ecological stability by developing a self-consistent mathematical strategy which implements and generalises the spectral control algorithm introduced in chapter 2 and 3. The method builds on the previously developed technique (also illustrated in [39]) and extends its domain of applicability, beyond diffusion mediated (linear) processes to the interesting setting where pairwise, hence non-linear, non-local interactions are considered.

In doing so, we will contribute to identifying the key topological features that should be possessed by a stable (resilient) ecological network. To anticipate our findings we will show that predator-prey interactions exert a beneficial role in terms of stability in qualitative agreement with the results reported in [6]. Further weak interactions tend to favour the overall stabilization, as observed in [47].
The chapter is organized as follows. In Section 4.1 we will introduce the model and define the reference mathematical setting. In Section 4.2 the control scheme is discussed in detail and we will also show how the spectral modifications trace back to actual changes in the matrix of interactions. In Section 4.3 the method is applied to a simple bidimensional predator-prey model, while the extension to arbitrarily large systems is discussed in Section 4.4 before summing up and drawing our conclusions. Relevant technical material is provided in the Appendices.

4.1 The model

We shall hereafter consider the coupled evolution of N species and denote by x_i, $i = 1, ..., N$, their associated concentrations. Each of the N nodes of the network therefore here represents a different species. We will operate under the deterministic viewpoint and deliberately omit any source of

stochastic disturbance, be it endogenous (demographic noise) or exogenous (external perturbation). The evolution of the ecosystem is hence described by the following set of first-order differential equations:

$$\dot{x}_i = x_i(r_i - s_i x_i) + x_i \sum_{j \neq i} A_{ij} x_j \quad i, j = 1, ..., N. \tag{4.1}$$

The self-reaction term is assumed logistic, for pedagogical reasons. We will subsequently relax this working hypothesis and generalize the analysis so as to account for an extended family of nonlinear reaction terms. In the above equations, r_i stands for the intrinsic growth rate of species i, while s_i is inversely proportional to the assigned carrying capacity. Matrix $\boldsymbol{A}$, in general asymmetric, defines the relations among species, delineating the interaction network. More specifically, the scalar entry A_{ij} encodes the effect exerted by species j on species i. The magnitude of A_{ij} weights the strength of the interactions. The sign of A_{ij}, respectively A_{ji}, defines the specific nature of the interaction between species i and j. Adopting a wording which is inspired to ecological applications: exploitation $(+,-)$, competition $(-,-)$, cooperation $(+,+)$, commensalism $(+,0)$, amensalism $(-,0)$ or null interaction $(0,0)$. The coupling among species is shaped by a quadratic term, which scales like the product of relative concentrations. This implies assuming the interaction to be mediated by pair exchanges, as it is customarily the case in ecology. On a more fundamental level, the coupling term here introduced will enable us to generalize the analysis reported in [39] beyond standard diffusion.

Let us denote by $\boldsymbol{x^*}$ the fixed point of the dynamics and assume for the sake of simplicity that all entries x_i^* are different from zero. In formulae we have:

$$r_i - s_i x_i^* + \sum_{j \neq i} A_{ij} x_j^* = 0 \quad i = 1, ..., N \tag{4.2}$$

which essentially defines the sole non trivial equilibrium eventually attained by our system. As mentioned before, we are concerned with the resilience of the system, namely its inherent capability to recover from perturbations. Stated it differently, we shall elaborate on the conditions which make the fixed point stable, according to a linear stability analysis. To this end, it is convenient to introduce the rescaled variable $y_i \equiv x_i / x_i^*$. In the new variables the fixed point is homogeneous and reads $y_i^* = 1$, $\forall i$.

The dynamics of system (4.1) can be cast in the form:

$$\dot{y}_i = y_i(r_i - \tilde{s}_i y_i + \sum_{j \neq i} B_{ij} y_j) \tag{4.3}$$

where $\tilde{s}_i = s_i x_i^*$, $B_{ij} = A_{ij} x_j^*$ and the equation (4.2) assumes the form $r_i - \tilde{s}_i + \sum_{j \neq i} B_{ij} = 0$.

To assess the stability of the fixed point, we set $y_i = 1 + v_i$ and Taylor expand at the first order in the perturbation amount. This yields the following linear system for the evolution of the imposed perturbation:

$$\dot{v}_i = -\tilde{s}_i v_i + \sum_{j \neq i} B_{ij} v_j \equiv \sum_j C_{ij} v_j. \tag{4.4}$$

The stability of the fixed point is ultimately controlled by the eigenvalues of the matrix $\boldsymbol{C}$ obtained by adding the elements $-\tilde{s}_i$ on the diagonal of matrix $\boldsymbol{B}$. The system is unstable when at least one eigenvalue of $\boldsymbol{C}$ has a positive real part. In the following, we will discuss a control procedure to stabilise a fixed point, that is initially engineered to be unstable. The method builds on the technique discussed in [39] and aims at reshaping the coupling among interacting species. As a side observation, which will become crucial in the forthcoming analysis, we notice that the fixed point condition ($y_i^* = 1 \ \forall i$) translates into a constraint for $\boldsymbol{C}$: summing the elements of $\boldsymbol{C}$ relative to row i one should recover $-r_i$, i.e. $\sum_j C_{ij} = -r_i$.

4.2 Topological control scheme

The proposed control strategy aims at modifying an initially unstable system of the type described above to yield an equivalent analogue which preserves the form (4.1) while admitting a stable non trivial equilibrium. As it shall be argued hereafter, we can either enforce the stability of the original, assumed unstable, fixed point, or, alternatively, steer the system towards a different equilibrium. In the former case we shall also alter the original carrying capacity (a parameter which indirectly encodes for the interaction with the surrounding environment), while in the latter the envisaged protocol will solely impact the network of interspecies couplings, leaving unchanged individual reaction parameters. In both cases, the intrinsic growth rates r_i, are kept unvaried. Stable ecological networks display remarkable topological characteristics, as discussed in a series of

papers devoted to this topic [6,47]. Operating along this line, we will isolate and discuss a selected gallery of features that appear to be recurrently shared by the ecological networks stabilized as outlined in the following.

Since, by definition, the structure of (4.1) is invariant under the foreseen procedure, inspecting the linear stability of the ensuing equilibrium implies dealing with a system of the type $\dot{v}_i = \sum_j C'_{ij} v_j$, where $\boldsymbol{C'}$ is obtained from $\boldsymbol{C}$, defined as in equation (4.4), via the devised control algorithm. Requiring the sought stability is, in turn, equivalent to constrain the spectrum of $\boldsymbol{C'}$ in the left portion of the complex plane, namely to set the real part of the associated eigenvalues to negative values. Our goal, pursued hereafter, is to elaborate on a rigorous mathematical procedure, which is both anchored to first principles and potentially minimally invasive, to derive $\boldsymbol{C'}$ from $\boldsymbol{C}$. Importantly, the obtained matrix $\boldsymbol{C'}$ should match the condition $\sum_j C'_{ij} = -r_i$, for the homogeneous fixed point to exist in terms of the rescaled variables y_i (recall that, by hypothesis, r_i is frozen to its original value). The effect of the control will be then gauged by tracing the modifications back to the underlying nonlinear framework, i.e. by evaluating the impact produced on the relevant dynamical parameters.

As a first step in the analysis, we write the linear equation (4.4) in the equivalent form:

$$\dot{v}_i = -r_i v_i + \sum_j D_{ij} v_j. \tag{4.5}$$

where the definition of $\boldsymbol{D}$ follows trivially. Recalling that $\sum_j C_{ij} = -r_i$ by virtue of the aforementioned fixed point condition, it is immediate to conclude that $\boldsymbol{D}$ is a zero-row-sum matrix, namely $\sum_j D_{ij} = 0$. The next step is to diagonalize matrix $\boldsymbol{D}$. Formally, we set $\boldsymbol{\Phi}^{-1}\boldsymbol{D}\boldsymbol{\Phi} = \boldsymbol{\Lambda}$, where $\boldsymbol{\Phi}$ is the matrix whose columns are the eigenvectors of $\boldsymbol{D}$ and $\boldsymbol{\Lambda}$ the diagonal matrix formed by the corresponding eigenvalues. Diagonalizability of matrix $\boldsymbol{D}$ is hence a necessary requirement for the method to hold. The idea is now to calculate the (minimal) shifts $\delta\Lambda^{(\alpha)}$, $\alpha = 1, ..., N$, to be applied to the eigenvalues of $\Lambda^{(\alpha)}$ of matrix $\boldsymbol{D}$, for the homogeneous fixed point $y_i^* = 1\ \forall i$ to prove linearly stable. Recall that the stability of this latter fixed point is eventually dictated by the spectrum of $\boldsymbol{C}$ (or, more precisely, by its controlled version $\boldsymbol{C'}$), from which the zero-row-sum counterpart $\boldsymbol{D}$ originates. The needed corrections $\delta\Lambda^{(\alpha)}$ are organized in a

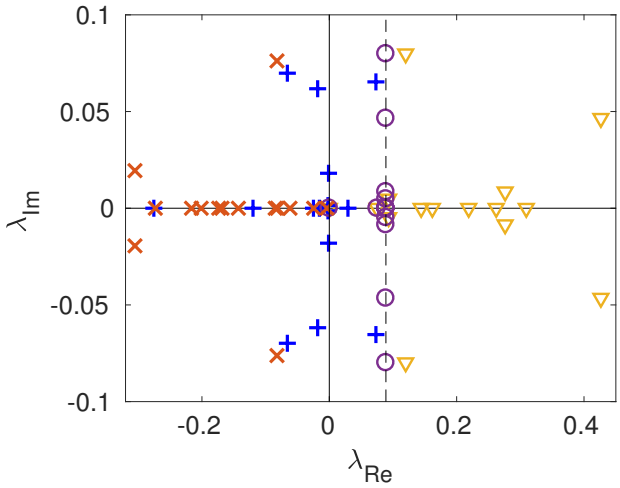


Figure 4.1: The eigenvalues of $\boldsymbol{C}$, (blue) plus symbols, signals an initial instability. This is also seen at the level of the spectrum of $\boldsymbol{D}$, (yellow) triangles, which partially extends beyond the critical vertical (dashed) line, located at r_{min}. Circles (purple) and crosses (red) refer respectively to the eigenvalues of $\boldsymbol{D'}$ and $\boldsymbol{C'}$, pointing to the recovered stability.

$N \times N$ diagonal matrix $\boldsymbol{\delta\Lambda}$ (with $\delta\Lambda_{kk} = \delta\Lambda^{(k)}$) to be added to matrix $\boldsymbol{\Lambda}$. The key point is how to choose the entries of $\boldsymbol{\delta\Lambda}$ for the control to return an effective, moderately intrusive (in terms of the modifications made to the spectrum), stabilization.

To answer this question, we proceed as if the original eigenvalues $\boldsymbol{\Lambda}$ were perturbed by the finite amount $\boldsymbol{\delta\Lambda}$ and recover a matrix $\boldsymbol{D'}$, which displays the modified spectrum, via the inverse transformation $\boldsymbol{D'} \equiv \boldsymbol{\Phi}(\boldsymbol{\Lambda}+\boldsymbol{\delta\Lambda})\boldsymbol{\Phi}^{-1}$. By construction $\boldsymbol{D'}$ commutes with $\boldsymbol{D}$, as the two matrices share the same set of eigenvectors. Notably, the corrections $\boldsymbol{\delta\Lambda}$ can be chosen such that matrix $\boldsymbol{D'}$ is also zero-row-sum, as $\boldsymbol{D}$ is. This is rigorously proven in Appendix B.3, building on the derivation reported in [39]. Moreover, $\boldsymbol{D'}$ is real: this property is also inherited from $\boldsymbol{D}$, as shown again in Appendix B.3. We are thus brought back to the linear problem:

$$\dot{v}_i = -r_i v_i + \sum_j D'_{ij} v_j \tag{4.6}$$

or, equivalently to $\dot{v}_i = \sum_j C'_{ij} v_j$, where:

$$C'_{ij} = \begin{cases} D'_{ij} & \text{if } i \neq j \\ D'_{ii} - r_i & \text{if } i = j. \end{cases} \tag{4.7}$$

By construction $\sum_j C'_{ij} = -r_i$, since $\sum_j D'_{ij} = \sum_j D_{ij} = 0$. This is a crucial observation which makes it possible to interpret $\boldsymbol{C'}$ as the Jacobian matrix associated to a rescaled nonlinear problem of the type (4.3) where parameters $\boldsymbol{r}$ remain unchanged. The zero-row-sum-property of matrix $\boldsymbol{D'}$ is not a necessary condition to obtain stability but represents a useful requirement to help interpreting the results in terms of the original variables.

We now return to discussing the selection of the elements of the shift matrix $\boldsymbol{\delta\Lambda}$. These latter are to be chosen so as to constrain the spectrum of $\boldsymbol{C'}$ to the left hand side of the imaginary plane, thus ensuing the desired stability. From (4.7), it is clear that the eigenvalues of $\boldsymbol{D'}$ are positioned, in the complex plane, on the right of those stemming from matrix $\boldsymbol{C'}$. The relative separation between the two respective spectra can be somehow quantified through $\boldsymbol{r}$. To make this observation rigorous, we recall the celebrated Gershgorin theorem [23]: the eigenvalues of a given matrix are included in disks defined by the elements of the matrix itself. More specifically, the i-th Gershgorin disk of matrix $\boldsymbol{D'}$ corresponds to the relative disk of matrix $\boldsymbol{C'}$, translated to the right by the scalar quantity r_i. Unfortunately, it is not trivial to relate the index i (running on the nodes) to the eigenvalues (sorted with the index α). To enforce stability, and assuming the worst case scenario, we shall assign the (real) shifts $\delta\Lambda^{(\alpha)}$ so that all eigenvalues of $\boldsymbol{D'}$ have their real part smaller than r_{min}, the minimum of all the entries of vector $\boldsymbol{r}$. In practical terms, the imposed corrections $\delta\Lambda^{(\alpha)}$ are chosen as:

$$\delta\Lambda^{(\alpha)} = \begin{cases} R - \Lambda^{(\alpha)}_{Re} & \text{if } \Lambda^{(\alpha)}_{Re} > r_{min} \\ 0 & \text{otherwise} \end{cases} \tag{4.8}$$

where the scalar quantity R has been introduced such that $R < r_{min}$. To help visualizing the whole procedure we report an illustrative example in Fig. 4.1, without insisting on the specific selection of the involved parameters. The eigenvalues of the original matrix $\boldsymbol{C}$ are displayed in the complex plane with (blue) plus symbols: the homogeneous fixed point of the rescaled equations (4.3) is therefore unstable, as the spectrum protrudes

in the right half-plane. The eigenvalues of $\boldsymbol{D}$ are shown with (yellow) triangles and extends on the right of the vertical dashed line, which is traced at r_{min}. The (purple) circles stands for the eigenvalues of the controlled matrix $\boldsymbol{D}'$: as anticipated, they are confined on the left of the vertical dashed line, the closer to the line the less invasive the control imposed on the population of unstable modes. Finally, and as predicted, the spectrum of the controlled Jacobian $\boldsymbol{C}'$ is contained in the negative half-plane, thus implying asymptotic stability.

As it should be clear from the above, the control protocol assumes dealing with constant r_i parameters. Given this constraint, two viable strategies are envisaged to re-parametrize the original system, in light of the outcome of the control scheme. We have in fact

$$D'_{ij} = \begin{cases} r_i - s'_i x_i^{*\prime} & \text{if } i = j \\ B'_{ij} = A'_{ij} x_j^{*\prime} & \text{if } i \neq j \end{cases} \tag{4.9}$$

which allows in principle to define the new coupling strengths, as encoded in $\boldsymbol{A}'$, the novel fixed point $\boldsymbol{x}^{*\prime}$ and the modified inverse carrying capacities $\boldsymbol{s}'$. A first strategy to finalize the transformation suggests leaving the parameters $\boldsymbol{s}$ unchanged, namely $\boldsymbol{s}' = \boldsymbol{s}$. In practical terms, we assume that the reaction parameters, which characterize the dynamics of each species when evolved on an isolated patch, remain unchanged. The ecosystem can achieve stabilization, by just reshaping the underlying networks of interlaced dependencies. From equation (4.9), we have therefore:

$$\begin{aligned} x_i^{*\prime} &= \tfrac{r_i - D'_{ii}}{s_i} \\ A'_{ij} &= \tfrac{D'_{ij}}{x_j^{*\prime}} = \tfrac{D'_{ij} s_j}{r_j - D'_{jj}}. \end{aligned} \tag{4.10}$$

Setting paired interactions as specified by matrix $\boldsymbol{A}'$ guarantees the stability of the associated, and consistently modified, fixed point $\boldsymbol{x}^{*\prime}$. The quantities $x_i^{*\prime}$ should be positive defined (at least when ecological applications are concerned), which in turn translates into the additional requirement

$$r_i - D'_{ii} > 0 \ \forall i. \tag{4.11}$$

The second strategy consists of modifying the parameters $\boldsymbol{s}$, together with the matrix $\boldsymbol{A}$, leaving unchanged the fixed point $\boldsymbol{x}^*$. The carrying capacity $\boldsymbol{s}$ is prone to environmental influences, and, as such, it can be imagined to be tunable with some degree of realism. This is opposed

to the growth parameter $\boldsymbol{r}$, constrained, among other factors, by species genetics, and thus assumed constant throughout the procedure. From equation (4.9):

$$\begin{aligned} s'_i &= \frac{r_i - D'_{ii}}{x^*_i} \\ A'_{ij} &= \frac{D'_{ij}}{x^*_j}. \end{aligned} \tag{4.12}$$

The additional condition $s'_i > 0$ should be imposed, which again amounts to requiring equation (4.11) to hold. In other words, condition (4.11) is a general constraint that the control scheme is bound to verify, for the specific ecologically inspired application, here discussed. We shall refer to condition (4.11) as to the *applicability constraint* and elaborate on its implications hereafter. The idea is to find a suitable value for R in order to match the condition (4.11), and, hence, to make the control scheme applicable. To begin, let us assume that we are allowed to modify all eigenvalues of the considered spectrum, and not just the limited sub-set that triggers the system unstable. Then, it is enough to impose $\delta\Lambda^{(i)} \leq \tilde{s}_{min}$ $\forall i$, which is in principle always possible, in order to automatically verify (4.11). The system is therefore always controllable when all eigenvalues are to be modified.

Consider now the more interesting case where a subset $\mathcal{M}$ of elements of the whole spectrum is the target of the control. One has to face the following restrictions:

- if an index i exists such that $\sum_{j\in\mathcal{M}} \Phi_{ij}\Phi^{-1}_{ji} < 0$ and $\tilde{s}_i + \sum_{j\in\mathcal{M}} \Phi_{ij}\Lambda^{(j)}_{Re}\Phi^{-1}_{ji} - r_{min}\sum_{j\in\mathcal{M}} \Phi_{ij}\Phi^{-1}_{ji} < 0$, then constraint (4.11) is never matched and the system is not controllable with the above discussed technique.

- if an index i exists such that $\sum_{j\in\mathcal{M}} \Phi_{ij}\Phi^{-1}_{ji} > 0$ and $\tilde{s}_i + \sum_{j\in\mathcal{M}} \Phi_{ij}\Lambda^{(j)}_{Re}\Phi^{-1}_{ji} - r_{min}\sum_{j\in\mathcal{M}} \Phi_{ij}\Phi^{-1}_{ji} < 0$ (suppose node indices are sorted so that such an index i is in the subset of nodes $1, ..., \tilde{n}$ with $\tilde{n} < N$), than the applicability condition (4.11) is verified only if the following statement holds true:

$$\max_{i\in[1,\tilde{n}]}(P_i) \leq \min_{i\in[n+1,N]}(P_i) \tag{4.13}$$

where $P_i \equiv r_{min} - \frac{\tilde{s}_i + \sum_{j\in\mathcal{M}} \Phi_{ij}\Lambda^{(j)}_{Re}\Phi^{-1}_{ji}}{\sum_{j\in\mathcal{M}} \Phi_{ij}\Phi^{-1}_{ji}}$ and $[n+1, N]$ numbers the set of indices i for which $\sum_{j\in\mathcal{M}} \Phi_{ij}\Phi^{-1}_{ji} < 0$.

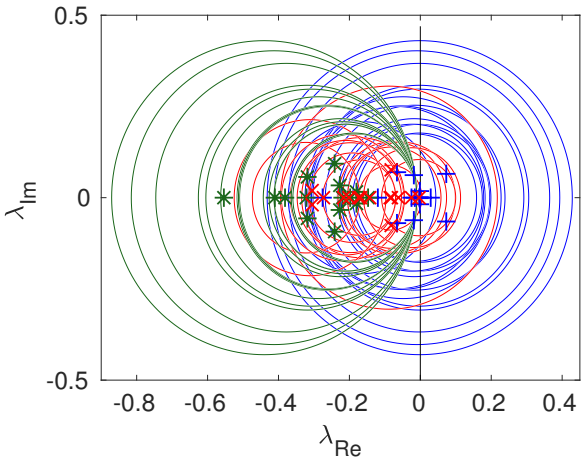


Figure 4.2: Eigenvalues in the complex plane for a particular choice of matrix $\boldsymbol{C}$ (blue plus symbols): the system is unstable as the symbols invade the region with positive λ_{Re}. The (blue) circles represent the associated Gershgorin disks. Red symbols stand for the eigenvalues of matrix $\boldsymbol{C'}$, which fall in the negative part of the complex plane, while their corresponding circles protrude in the positive half-plane. Green symbols refer to the alternative control scheme (4.14): now the Gershgorin circles are contained in the negative half-plane. The spectrum obtained by following this latter route to stability is well inside the region of stability, at variance with that generated by the former approach, which sits at the border of stability. In this respect, the method that we have depicted, and which necessitates reshaping the underlying network of contacts, can be thought as minimally invasive.

The above conditions (derived in Appendix B.4) are to be carefully checked before attempting to control the system under exam via the procedure that we have here illustrated and which is ultimately aimed at recalibrating the weights of the underlying couplings.

Before concluding this Section, we briefly mention an alternative control strategy which builds on the already mentioned Gershgorin theorem. In the above analysis we have reshaped the networks of contacts and altered either the fixed point $\boldsymbol{x^*}$ or the carrying capacities s_i, while preserving the values of the growth factors r_i. The alternative route to stabilization that we shall hereafter discuss follows a dual path: the only

parameters to be tuned are the growth rates r_i. Recalling that $\boldsymbol{C'}$ is real (from definition of (4.7), $\boldsymbol{D'}$ being real), it is clear that each eigenvalue has its real part smaller than $C'_{ii} + \sum_{j \neq i} |C'_{ij}|$, which is the rigthmost point in the complex plane of the i-th Gershgorin disk for matrix $\boldsymbol{C'}$. To enforce stability, we could then require that all the Gershgorin circles are included in the left half-plane:

$$C'_{ii} + \sum_{j \neq i} |C'_{ij}| \leq 0 \Rightarrow D'_{ii} - r_i + \sum_{j \neq i} |D'_{ij|} \leq 0 \Rightarrow D'_{ii} + \sum_{j \neq i} |D'_{ij|} \leq r_i \quad (4.14)$$

which implies that the Gershgorin disks computed for matrix $\boldsymbol{D'}$ should be contained in the semi-plane constrained, from the right, by the vertical line located at r_i for any selected i, or, better, at the left of r_{min}, for all i. The shift of the Gershgorin disks can be performed by modifying the diagonal of matrix $\boldsymbol{C}$, which is equivalent to changing the vector of parameters $\boldsymbol{r}$. This alternative control strategy proves however more invasive in terms of the perturbation that is produced on the original spectrum. This is clearly verified in Fig. 4.2, where the original (unstable by construction) spectrum is compared to those obtained applying the two control strategies outlined above.

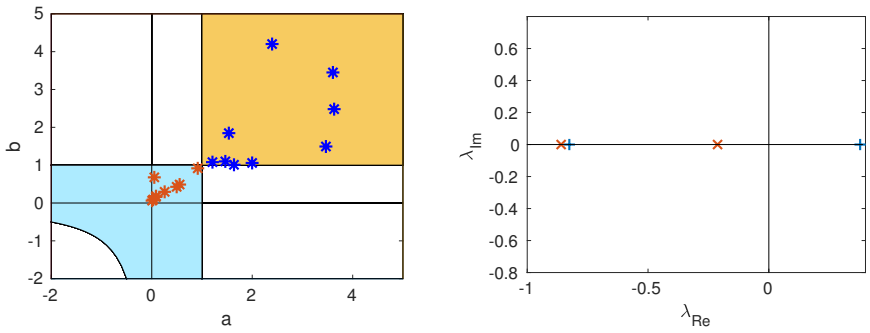
In the next Section we will begin by applying the developed method to a simple example, where just two species are made to evolve. This application bears pedagogical interest and it will pave the way to inspecting the general setting, on which we shall report in the subsequent Section.

4.3 Stabilizing a predator-prey dynamics

With the goal of gaining further insight on the control scheme developed above, we will here consider a simple setting where just two species are made to mutually interact. Hence, we will consider hereafter $N = 2$ and label with x_1 and x_2 the mean field concentrations of the interacting species. Equations (4.1) reduce therefore to:

$$\begin{cases} \dot{x}_1 = x_1(r_1 - s_1 x_1 + A_{12} x_2) \\ \dot{x}_2 = x_2(r_2 - s_2 x_2 + A_{21} x_1) \end{cases} \quad (4.15)$$

so as to describe a system composed by one predator and one prey. To proceed in the analysis we define $u_i \equiv (s_i/r_i)x_i$, $\tau \equiv r_1 t$, and eventually



(a) (b)

Figure 4.3: Panel (a): Existence and stability of the (non trivial) fixed point (u_1^*, u_2^*), as a function of the parameters a and b. In particular, the fixed point does not exist in the empty (no shading) areas (negative for $a > 1 \& b < 1$ and $a < 1 \& b > 1$, infinite for $a < 0 \& b < 0 \& b < 1/a$); it is unstable in the rectangular (orange) shaded region, entirely contained in the first quadrant; it is stable in the other (cyan) colored region, which extends in all quadrants. The blue stars refer to the initial choice of the parameters: the system is hence unstable. After the control procedure is applied, one obtains the red stars, which are distributed inside the region associated to linear stability. These latter symbols cluster in a limited portion of the plane, close to the threshold of instability. In this respect, the control method is minimal also in terms of the modification induced at the level of the dynamical parameters (and not only in relation to its effects in the complex plane where the spectrum of the Jacobian is depicted). Panel (b): original and modified spectrum are displayed for one representative case study among those depicted in panel (a).

get the following governing equation in the rescaled variables u_1 and u_2:

$$\begin{cases} \frac{du_1}{d\tau} = u_1(1 - u_1 + au_2) \equiv f_1(u_1, u_2) \\ \\ \frac{du_2}{d\tau} = \rho u_2(1 - u_2 + bu_1) \equiv f_2(u_1, u_2) \end{cases} \tag{4.16}$$

where $\rho = r_2/r_1$, $a \equiv -A_{12}\frac{r_2}{r_1 s_2}$ and $b \equiv -A_{21}\frac{r_1}{r_2 s_1}$. It is straightforward to prove that this system admits four equilibria: $(u_1, u_2) = (0, 0)$, $(u_1, u_2) = (0, 1)$, $(u_1, u_2) = (1, 0)$ and $(u_1, u_2) = (\frac{1-a}{1-ab}, \frac{1-b}{1-ab}) \equiv (u_1^*, u_2^*)$. We shall hereafter refer to the latter equilibrium, the only one to guarantee non trivial asymptotic concentrations for both species. Notice that this is admissible only if u_1^* and u_2^* are positive and finite, which, in turn, implies that the parameter space (a, b) is restricted to the colored region of Fig. 4.3, i.e.

- $a > 1$, $b > 1$
- $a < 1$, $b < 1$, $b < 1/a$

The stability of the selected fixed point is determined by the Jacobian matrix

$$J_{(u_1^*, u_2^*)} = \frac{1}{1 - ab}\begin{pmatrix} a - 1 & a(a - 1) \\ \rho b(b - 1) & \rho(b - 1) \end{pmatrix} \tag{4.17}$$

whose eigenvalues are

$$\lambda_\pm = \frac{1}{2(1 - ab)}\left[(b-1)(1+\rho) \pm \sqrt{(a - 1)^2(1 + \rho)^2 - 4\rho(1 - ab)(a - 1)(b - 1)}\right]. \tag{4.18}$$

The sign of $\lambda_\pm$ implies that the fixed point (u_1^*, u_2^*) is unstable for $a > 1$ and $b > 1$ and stable in the complementary domain, as depicted in Fig. 4.3.

Working in this simplified setting, it is therefore straightforward to implement, and graphically illustrate, the stabilization protocol, as addressed in the preceding Section. Starting from an unstable system corresponds to setting the parameters a and b in the (orange) shaded sub-portion of the first quadrant of Fig. 4.3. When implementing the control, panel (b) of Fig. 4.3, the rescaled parameters (a, b) are consequently moved to the other (cyan) shaded domain displayed in Fig. 4.3 (a), i.e. the region deputed to stability. Different symbols, as depicted in Fig. 4.3 (a), refer to distinct choices of the initial model parameters. In all cases, the stabilization is successfully produced and, more importantly, the modified parameters (a, b)

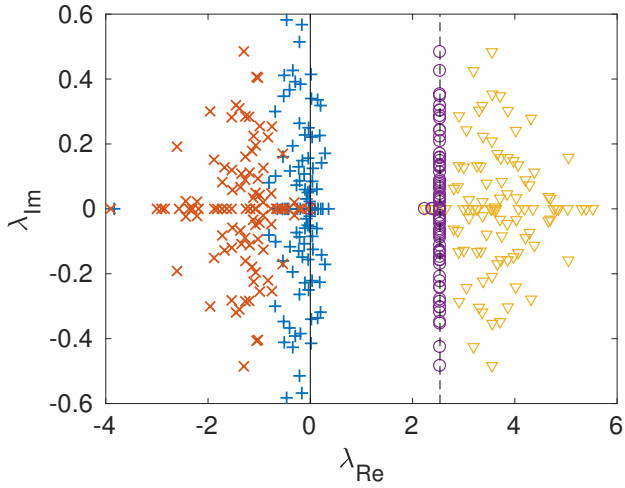


Figure 4.4: Eigenvalues in the complex plane, before, along and after the control procedure. The symbols are assigned as described in the caption of Fig. 4.1. The elements A_{ij} are assigned following the scheme discussed in the Appendix. Here, $N = 100$.

tend to cluster in a limited portion of the stability domain, close to the boundary of instability. This observation suggests again that the devised control acts by producing a somehow minimal perturbation to the original model. Following the alternative control recipe based on the Gershgorin theorem produces much more invasive changes (here not shown).

The controlled matrix of interaction, as well as the novel set of dynamical parameters, can be readily obtained from the modified quantities a' and b', following one of the interpretative scenarios discussed with reference to the N-dimensional case. Both strategies require altering the matrix of couplings $\boldsymbol{A}$ to eventually obtain its modified counterpart here denoted with $\boldsymbol{A'}$. More specifically:

- Changing the fixed point $\boldsymbol{x}^*$, the parameters $\boldsymbol{r}$ and $\boldsymbol{s}$ stay unchanged. Hence:

$$A'_{12} = -a' \frac{r_1 s_2}{r_2} \tag{4.19}$$

$$A'_{21} = -b' \frac{r_2 s_1}{r_1} \tag{4.20}$$

$$\tag{4.21}$$

The new fixed point $(x_1^*)'$, $(x_2^*)'$ is obtained by solving the self-consistent equations:

$$r_1 - s_1 (x_1^*)' + A'_{12} (x_2^*)' = 0 \tag{4.22}$$

$$r_2 - s_2 (x_2^*)' + A'_{21} (x_1^*)' = 0. \tag{4.23}$$

$$\tag{4.24}$$

- Modifying the parameters $\boldsymbol{s}$. The fixed point and $\boldsymbol{r}$ are not varied. One gets:

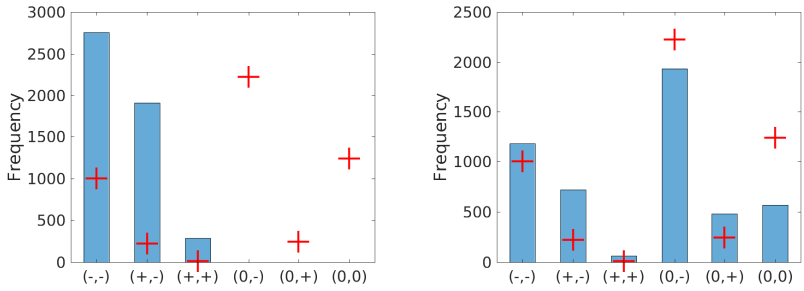
$$\begin{cases} a' = -A'_{12} \frac{r_2}{r_1 s'_2} \\ b' = -A'_{21} \frac{r_1}{r_2 s'_1} \end{cases} \tag{4.25}$$

which, together with the fixed point condition:

$$\begin{cases} r_1 - s'_1 x_1^* + A'_{12} x_2^* = 0 \\ r_2 - s'_2 x_2^* + A'_{21} x_1^* = 0. \end{cases} \tag{4.26}$$

allows one to obtain the entries of the matrix A' and the controlled parameters $\boldsymbol{s}$.

In the first case we have to make sure that the new fixed point is admissible, which amounts to meeting the conditions $(x_1^*)' > 0$ and $(x_2^*)' > 0$. In the second case, one has to impose $s'_1 > 0$ and $s'_2 > 0$. It is straightforward to prove that these constraints are always satisfied when a' and b' fall, as they do by definition, in the stability region ($1 - a' > 0$, $1 - b' > 0$ and $1 - a'b' > 0$). In the next Section we move on to considering the general setting by working with an arbitrarily large ecological system consisting of N interacting populations.



(a) (b)

Figure 4.5: Abundances of different types of couplings between species. (Red) plus symbols refer to the uncontrolled matrix of interaction (analytically computed in Appendix B.1). Panel (a): the (blue) bars report on the relative abundances of different classes, as obtained after the control has been applied. Results refer to just one realization of the process. Quantitatively similar conclusions are obtained for different realizations and/or averaging over a large ensemble of them. Panel (b): the frequency of appearance of different classes is plotted after a cut-off has been applied (see main text). Here the cut off is set to 0.005.

4.4 Generalization: a larger ecosystem

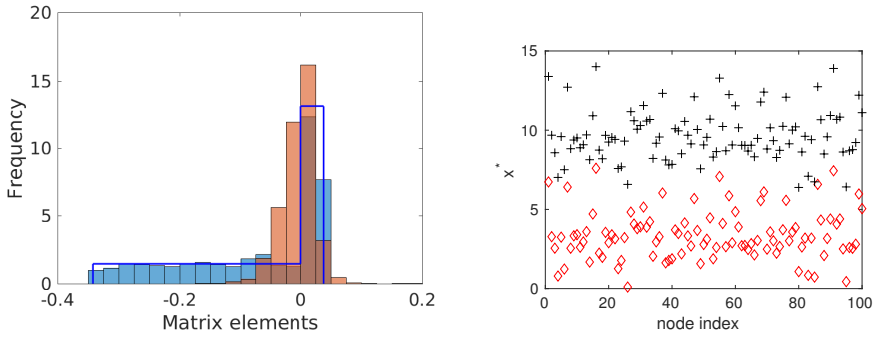
Let us now consider the general setting where an arbitrarily large number of species is made to interact. The initial coupling strengths which exemplify the interaction among distinct populations, as encoded in the elements of matrix $\boldsymbol{A}$, are initially assigned in such a way that a non trivial fixed point exists and is linearly unstable. This is achieved by generating the random entries A_{ij}, according to a specific distribution that we discussed in the Appendix and exploiting again the Gershgorin theorem. In the following, we will set $N = 100$ and illustrate the results which are obtained when allowing for the fixed point to be modified by the control procedure. The alternative control strategy, which leaves the original fixed point unchanged, is also analyzed and the results reported in the Appendix. As an interesting outcome, we will show that predator-prey interactions exert a stabilizing effect, as already pointed out in [6].

In Fig. 4.4, the original and modified spectrum are displayed in the complex plane, for one representative realization of matrix $\boldsymbol{A}$. The sym-

bols are chosen following the same convention adopted in Fig. 4.1. The stabilization produced by the control is clearly demonstrated.

To help visualizing how the control shapes the examined ecosystem, we extract from the matrix $\boldsymbol{A}'$ the number of pairs that belong to the different classes, categorized in five six large classes, here recalled for the sake of completeness: exploitation $(+,-)$, competition $(-,-)$, cooperation $(+,+)$, commensalism $(+,0)$, amensalism $(-,0)$ or null interaction $(0,0)$. The frequency of appearance of different classes is investigated in Fig. 4.5. The red plus symbols point to the uncontrolled setting, and reflect the specific rule chosen for generating matrix $\boldsymbol{A}$ and its associated, unstable fixed point (see Appendix B.1). For the controlled matrix, only three bins are populated, see Fig. 4.5 (a): interaction modalities that envisage a one directional coupling, or stated differently, a zero entry in matrix $\boldsymbol{A}$, are absent in the controlled scenario. The topological control activates in fact all pairwise connections, albeit often by a tiny amount. It is then interesting to silence, a posteriori, in the controlled adjacency matrix $\boldsymbol{A}'$, the links that are associated to a weight (in absolute value) smaller than a given cutoff. For a sufficiently small cutoff the stability of the controlled system is preserved: the number of newly added links can hence be considerably reduced, by eradicating from the collection those that bear no relevance in light of the modest exerted coupling. The effect of the cut off is visible in Fig. 4.5 (b): the final distribution of pairs resembles very closely the one generated at the beginning. Remarkably, the number of predator-prey interactions grows at the detriment of the last column of the histogram, implying that the new interactions that are to be established for stability to hold belong to this class, in qualitative agreement with the analysis by Allesina et al. [6].

A different view on the effect of the control method can be gained by looking at Fig. 4.6(a). Here, the distribution of the weights associated to predator-prey interactions is plotted before (blue bars and continuum profile) and after (orange histogram) application of the control. While considering the control procedure which allows the fixed point to change, it is clear that reducing the strength of the couplings proves beneficial for the system stabilization, in agreement with [47]. The same observation holds for the other classes of interactions (data not shown). As a final check, we show in Fig. 4.6(b) the new fixed points as obtained with the stabilized matrix $\boldsymbol{A}'$, compared to the trivial fixed point obtained when



(a)

Figure 4.6: (a): The distribution of the predator-prey interactions is displayed: blue bars refers to one realization of the initial system (averaging over many realizations yields the analytic profile represented by the blue line and obtained after equations (B.2)). Red bars photograph the distribution of couplings obtained once the control has been applied. (b): The equilibrium solution of the controlled system (red diamonds) is different from the trivial (uncoupled) fixed point (black plus symbols) obtained by setting $x_i^{*\prime} = r_i/s_i$. The index reported in the horizontal axis identifies the species.

the matrix of couplings is switched off, i.e. assuming that each species is bound to evolve on a isolated niche. The method devised here yields a genuinely complex stable equilibrium, which appears to be shaped by the couplings established among interacting units, despite the global tendency to reduce their associated strengths, as revealed in Fig. 4.6(a).

Summing up, as a generalization of the model analyzed in the previous chapters we have here considered a system composed of different families, each one represented by its concentration variable and associated to a node of an interaction network. Such variable evolves in time according to a reaction term and is enhanced or reduced by the interaction with the other species. Working in this context, we here contributed with a novel approach, mathematically grounded on first principles, to help identifying the topological features that should be possessed by a generic ecological network so as to ensure stability, hence resilience to external perturbations. To this end we have considered a system made up of N interacting species

and assumed the reaction dynamics to be logistic. We then generated a network of interactions, encoded in a weighted adjacency matrix, which is prone to instability. By rewiring the assigned links, and their associated weights, we have shown that it is possible to drive the system stable, via two alternative strategies which preserve, or not, the initial fixed point. The technique here developed implies minimal modifications to the spectrum of the Jacobian matrix which is responsible of the stability of the underlying equilibrium and traces these changes back to species-species interactions by means of the network modifications analysis.

Chapter 5

Control by adding one node to the network

The generalised stabilisation process described in the previous chapter constitutes an important tool for assessing the ecosystem stability. It presents however an important weakness: it is global. In other words, almost all the interactions between units must be modified in order to convert an unstable into a stable system. This, when dealing with real ecological systems, results clearly impracticable. The above described technique consequently represents a significant theoretical tool when the aim is to discover the general characteristics to be satisfied by a stable ecosystem, but fails from the point of view of practical realizations. In this chapter a different technique is presented that has the advantage of being local. Indeed, one species is added to the pool of interacting families and used as a dynamical controller to induce novel stable equilibria. The method is again based on the spectral properties of a stability matrix and, in addition, use is made of the root locus method to shape the needed control, in terms of intrinsic reactivity and adopted protocol of injection. The proposed method is tested on both synthetic and real data, thus enabling to demonstrate its robustness and versatility.

To set the reference frame we will hereafter consider a system consisting of N species (nodes) whose activities $\boldsymbol{x} = (x_1, x_2, \cdots, x_N)^T$ obey the

coupled nonlinear differential equations [47, 70, 81]:

$$\dot{x}_i = f_i(x_i) + \sum_j A_{ij} g_i(x_i, x_j) \quad i = 1, \ldots, N. \tag{5.1}$$

The first term on the right-hand side specifies the self-dynamics of species i while the second term stems from the interactions of species i with the other species. The nonlinear functions $f_i(x_i)$ and $g_i(x_i, x_j)$ encode the dynamical laws that govern the system's components, while the weighted connectivity matrix $\boldsymbol{A}$ captures the interactions between nodes. The elements A_{ij} can be positive or negative, depending on the specific nature of the interaction, i.e. cooperative or competitive. Notice that system (5.1) is assumed in [70] as a reference model to analyze resilience patterns in complex networks. Differently from [70], A_{ij} can here take positive and negative values (see also [191]).

The interaction shape is often epitomized by a quadratic response function $g_i(x_i, x_j)$ [81], as it is considered in the previous chapter and the reaction drive $f_i(x_i)$ to which each species is subjected represents a logistic growth with a prescribed carrying capacity [70].

Animals displaying competitive predator-prey interactions or, alternatively, subjected to a symbiotic dependence, such as in plant-pollinator relationships, are among the systems that fall within the aforementioned scenario [187]. Furthermore, the complex community of micro-organisms that live in the digestive tracts of humans and other animals, including insects, can be rooted on similar descriptive grounds [47].

Moreover, the systems which can be described by this generalized version of temporal equations are not limited to natural ecosystems, but include several other applications.

For genetic regulatory networks, the dynamical variables x_i represent the level of activity of a gene or the concentration of the associated proteins [22]. Species specific reaction terms $f_i(x_i)$ account for, e.g., degradation or dimerization. The pattern of activation could be effectively modeled by sigmoidal Hill-like functions [140], as follows the classical Michaelis-Menten scheme [100], which incorporates the known map of gene interactions. On a more general perspective, understanding the emerging dynamics in social communities [38], grasping the essence of the learning organization in the brain [150], and implementing efficient protocols for robot navigation in networked swarms [170] are among the very many ap-

plications that can be traced back to one of the variants of equations (5.1), with a suitable choice of the nonlinear functions $f_i(x_i)$ and $g_i(x_i, x_j)$.

5.1 Adding a species to enforce stable equilibria in a multidimensional system

Starting from the above illustrated setting, we will here discuss a suitable control scheme to drive system (5.1) towards a desired equilibrium $\boldsymbol{x}^* = (x_1^*, x_2^*, \cdots, x_N^*)^T$, which is linearly stable to externally imposed perturbations. To reach this goal we shall introduce one additional species, the $(N+1)$-th component of the collection, suitably designed to yield the sought effect. To set the notation, we indicate by u the component (e.g., concentration, activation level) assigned to the controller and write:

$$\begin{cases} \dot{x}_i = f_i(x_i) + \sum_j A_{ij} g_i(x_i, x_j) + \alpha_i h_i(x_i, u) \\ \dot{u} = -(u - u^*) - \rho \sum_j \beta_j (x_j - x_j^*). \end{cases} \tag{5.2}$$

The controller u can exert a direct influence on every component x_i, as specified by newly added terms $\alpha_i h_i(x_i, u)$ that modify the original system (5.1). $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_N)^T$ is a vector of N constant parameters, to be self-consistently adjusted following the scheme depicted below. $h_i(x_i, u)$ is a generic, in principle nonlinear, function of the components x_i and u that reflects the modality of interaction between the controller and the existing species. The equation for the dynamical evolution of the controller u displays two distinct contributions. The first represents a self-reaction term, assumed to be linear just for ease of presentation. The nonlinear self-dynamics of the controller u can be readily considered, with no further technical complication. The rate of change of u is assumed to be contextually driven by a global forcing that senses the relative distance of x_i from its deputed equilibrium x_i^*. The parameters $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_N)^T$ and ρ will prove central in enforcing the stabilization of the prescribed fixed point. A few comments are mandatory to fully appreciate the generality of the proposed framework, beyond the specific choices made for purely demonstrative purposes. Let us begin by remarking that the controller u can represent an artificially engineered component or, equivalently, belong to an extended pool of interacting populations. In the scheme here imagined, it is assumed that the values of u and x_i $\forall i$, are accessible to direct

measurement at any time and that this information can be processed to set the controller dynamics. This is largely reasonable for experiments that run under protected conditions like, e.g., the study of microbial dynamics in laboratory reactors, but certainly less realistic for applications that aim at in vivo multidimensional systems, think for instance to genetic regulatory circuits. The dynamical equation for u can, however, be amended to a large extent and with a great deal of flexibility, depending on the target application and the structural specificity of the employed controller, while still allowing for an analogous methodological treatment[1]. The dynamics of the original, unsolicited, components and the functional form that specifies the controller feedback bear unequivocal universality traits [70].

The global fixed point $(\boldsymbol{x}^*, u^*)$ of the controlled system (5.2) should match the following constraints

$$f_i(x_i^*) + \sum_j A_{ij} g_i(x_i^*, x_j^*) + \alpha_i h_i(x_i^*, u^*) = 0 \quad i = 1, .., N \tag{5.3}$$

which, provided the x_i^* and u^* are assigned, ultimately set the values of the parameters α_i. Conversely, as we shall illustrate in the following, one could assume the parameters α_i as a priori known and infer via equations (5.3) the fixed point(s) to be eventually stabilized. The next step in the analysis aims at ensuring the stability of the selected fixed point. This will be achieved by acting on the residual free parameters $\boldsymbol{\beta}$ and ρ. As routinely done, we perturb the equilibrium solution as $x_i = x_i^* + v_i$, $u = u^* + w$ and Taylor expand equations (5.2) assuming the imposed disturbances $\boldsymbol{\eta} = (\boldsymbol{v}, w)$ small in magnitude. At the linear order of approximation one obtains:

$$\dot{\boldsymbol{\eta}} = \begin{pmatrix} \boldsymbol{G} & \boldsymbol{q} \\ -\rho \boldsymbol{\beta}^T & -1 \end{pmatrix} \boldsymbol{\eta} \equiv \boldsymbol{J} \boldsymbol{\eta} \tag{5.4}$$

where $\boldsymbol{q}$ is a N-dimensional column vector of components $q_i = \alpha_i \frac{\partial h_i}{\partial u}(x_i^*, u^*)$. The $N \times N$ matrix $\boldsymbol{G}$ is defined as:

[1] As a matter of fact, we can equivalently assume a generalized equation for the controller of the type $\dot{u} = f_u(u) - \rho g_u(\boldsymbol{x}, u, \boldsymbol{\beta}) + b$ where $f_u(u^*) = 0$ and $b = \rho g_u(\boldsymbol{x}^*, u^*, \boldsymbol{\beta})$.

$$
\begin{aligned}
G_{ii} &= \frac{\partial f_i}{\partial x_i}(x_i^*) + \sum_k A_{ik} \frac{\partial g_i}{\partial x_i}(x_i^*, x_k^*) + \alpha_i \frac{\partial h}{\partial x_i}(x_i^*, u^*) \\
G_{ij} &= A_{ij} \frac{\partial g_i}{\partial x_j}(x_i^*, x_j^*).
\end{aligned}
$$

The fixed point $(\boldsymbol{x}^*, u^*)$ is linearly stable if all eigenvalues of the Jacobian matrix $\boldsymbol{J}$ have negative real parts. The associated characteristic polynomial $P(\lambda) = \det(\boldsymbol{J} - \lambda \boldsymbol{I})$ can be cast in the equivalent, affine in the ρ-parameter, form:

$$
\begin{aligned}
P(\lambda) &= -(1+\lambda)\det(\boldsymbol{G} - \lambda \boldsymbol{I}) + \rho \sum_{i,j=1}^{N} \beta_j [\mathrm{adj}(\boldsymbol{G} - \lambda \boldsymbol{I})]_{ji} q_i \\
&\equiv \mathcal{D}(\lambda) + \rho \mathcal{N}(\lambda)
\end{aligned}
$$

that is reminescent of the celebrated root locus method [61]. Here, $[\mathrm{adj}(\boldsymbol{Z})]_{ji} = (-1)^{i+j} \det[(\boldsymbol{Z})_{(i,j)}]$ denotes the adjugate of matrix $\boldsymbol{Z}$, $(\boldsymbol{Z})_{(i,j)}$ being the minor of $\boldsymbol{Z}$ obtained by removing the i-th row and the j-th column. The polynomials $\mathcal{D}(\lambda) = (-1)^{N+1} \prod_{k=1}^{N+1} (\lambda - p_k)$, and $\mathcal{N}(\lambda) = (-1)^{N+1} \prod_{k=1}^{N-1} (\lambda - z_k)$ have respectively degrees $N+1$ and $N-1$, without loss of generality $\boldsymbol{\beta}$ can be chosen so that $\mathcal{N}(\lambda)$ assumes this form. With a slight abuse of language we will refer to as *poles* the roots p_k of the polynomial $\mathcal{D}(\lambda)$ and *zeros* the roots z_k of $\mathcal{N}(\lambda)$. Notice that for $\rho = 0$ the eigenvalues of the Jacobian $\boldsymbol{J}$ correspond to the $N+1$ poles p_k. These latter quantities are uniquely determined, once the fixed point $(\boldsymbol{x}^*, u^*)$ has been assigned. In particular it cannot a priori be ensured that the real parts of all p_k are negative, as stability would require. In other words, when $\rho = 0$, we can enforce the desired fixed point into the system but cannot guarantee its stability. On the other hand, for $\rho \to \pm\infty$, $N-1$ eigenvalues of $\boldsymbol{J}$ tend to the zeros z_k, which depend self-consistently on the free parameters $\boldsymbol{\beta}$. As we shall show hereafter, it is in principle possible to assign the β_i to force the real parts of all z_k to be negative. The two remaining eigenvalues of matrix $\boldsymbol{J}$, in the limit of large $|\rho|$, diverge to infinity in the complex plane. More precisely, they travel along opposite directions following a vertical (resp. horizontal) asymptote, if ρ is bound to the positive (resp. negative) semiaxis. To confer stability in the limiting case $\rho \to \infty$ where $N-1$ eigenvalues of $\boldsymbol{J}$ coincide with the roots

of $\mathcal{N}(\lambda)$, it is therefore sufficient to (i) operate a supervised choice of $\boldsymbol{\beta}$ and (ii) impose the condition that yields a vertical asymptote ($\rho \to +\infty$), while, at the same time, requiring that this latter intersects the negative side of the real axis. In this respect, it is important to remark that the intersection occurs in the point of abscissa $\lambda_0 = \frac{1}{2}\big(\sum_{k=1}^{N+1} p_k - \sum_{k=1}^{N-1} z_k\big)$. Hence, the idea is to interpolate between the two limiting cases $\rho = 0$ and $\rho \to \infty$ by determining the minimal value ρ_c of ρ beyond which the desired fixed point becomes stable. The existence of the threshold ρ_c that makes the imposed fixed point attractive for any $\rho > \rho_c$ is obvious, being stability already assured in the limiting setting[2] $\rho \to +\infty$. For the sake of clarity we reiterate that this amounts to selecting $Re(z_k) < 0$ for all k and further imposing $\lambda_0 < 0$, by properly assigning the free parameters $\boldsymbol{\beta}$.

In order to study the assignability of the zeros z_k by means of $\boldsymbol{\beta}$, let us recall that for a generic square matrix $\boldsymbol{Z}$, we have that $\mathrm{adj}(\boldsymbol{Z} - \lambda \boldsymbol{I}) = -\sum_{m=0}^{N-1}\sum_{l=0}^{N-m-1} c_{l+m+2}\boldsymbol{Z}^m\lambda^l$ where c_k stands for the coefficients of the characteristic polynomial of $\boldsymbol{Z}$, namely $\det(\boldsymbol{Z} - t\boldsymbol{I}) = \sum_{l=0}^{N} c_{l+1}t^l$. The polynomial $\mathcal{N}(\lambda)$ can be consequently written as:

$$\mathcal{N}(\lambda) = -\sum_{m=0}^{N-1}\sum_{l=0}^{N-m-1} c_{l+m+2}[\boldsymbol{\beta}^T\boldsymbol{G}^m\boldsymbol{q}]\lambda^l \equiv \sum_{n=0}^{N-1} d_{n+1}\lambda^n \tag{5.5}$$

It is hence straightforward to establish a direct relation between the parameters $\boldsymbol{\beta}$ and the vector of coefficients $\boldsymbol{d}$:

$$d_n = -\sum_{k=0}^{N-n} c_{k+n+1}[\boldsymbol{\beta}^T\boldsymbol{G}^k\boldsymbol{q}] \tag{5.6}$$

that can also be equivalently stated as:

$$\boldsymbol{d} = \boldsymbol{H}\boldsymbol{\beta} \tag{5.7}$$

where $\boldsymbol{H}$ is the matrix defined by:

$$H_{nm} = -\sum_{k=0}^{N-n} c_{k+n+1}(\boldsymbol{G}^k\boldsymbol{q})_m. \tag{5.8}$$

[2]In principle, more than one value of ρ_c can exist for which the eigenvalues cross the imaginary axis, making stable an unstable fixed point. The intersections are found imposing $\lambda = i\omega$ in equation $\mathcal{D}(\lambda)+\rho_c\mathcal{N}(\lambda) = 0$, which yields a system of two equations, for respectively the real and imaginary parts. This system can then be solved for the two unknowns ρ_c and ω.

The suited vector $\boldsymbol{\beta}$ is thus obtained[3] from (5.7), provided matrix $\boldsymbol{H}$ is invertible. This latter requirement defines the condition of *controllability* for the scheme that we have implemented (see Appendix C for a discussion that aims at positioning this observation in the context of standard control theory [103]). Summing up, the devised strategy consists of the following steps. First, the fixed point is selected and the parameters $\boldsymbol{\alpha}$ frozen to their respective values as specified by (5.3). Then the complex roots z_k are chosen so that $Re(z_k) < 0$ for all k while, at the same time, matching the condition that makes the vertical asymptote cross the horizontal axis with a negative intercept. As we will clarify when discussing the applications, the z_k can be chosen to coincide with the poles p_k, except for point modifications whenever $Re(p_k) \geq 0$. Notice however that z_k should be real or come in conjugate pairs, as the coefficients d_k are, by definition, real. Once the roots z_k have been fixed, one can readily compute the associated polynomial coefficients d_k, and hence proceed with the determination of $\boldsymbol{\beta}$ via (5.7), provided that the controllability condition holds. Finally, by selecting $\rho > \rho_c > 0$ we obtain a linearly stable fixed point $(\boldsymbol{x}^*, u^*)$ for the controlled dynamics (5.2).

5.2 Testing the control method: from synthetic gene network to real microbiota dataset

As a first application of the above technique, we will study the dynamics of an artificial gene network [2, 59, 88, 98]. In our example, the network of connections is a regular tree with branching ratio $r = 4$. It is further assumed that the genetic activation between nodes i and j is described in terms of a Hill function, with cooperation coefficient equal to 2. In formulae, $A_{ij} = 1$ and $g_i(x_i, x_j) \equiv g(x_j) = x_j^2/(1 + x_j^2)$. Negative regulation loops are also accommodated for. These latter could, in principle, be modeled by assuming paired interactions of the type $1 - g(x_j)$, while still setting to one the relative entry of the connection matrix. As

[3]For obvious consistency reasons $\boldsymbol{\beta}$ must have real entries. This follows naturally if one chooses the zeros z_k to be real or complex conjugate in pairs, which implies that the coefficients d_n of the polynomial $\mathcal{N}(\lambda)$ (see (5.5)) are real. All other quantities involved are real by definition.

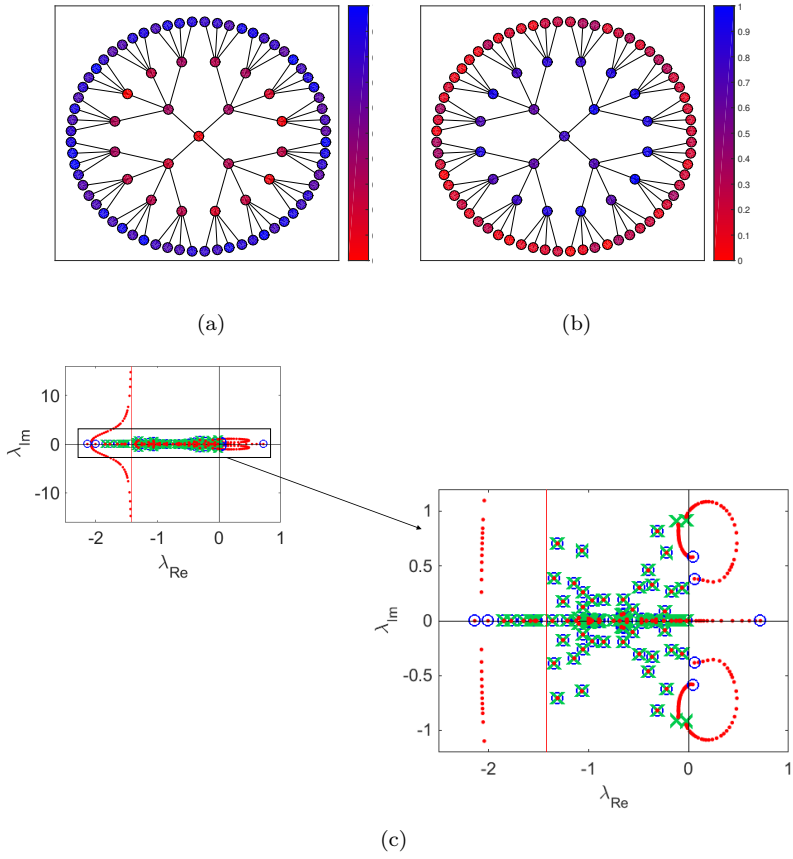


Figure 5.1: Panel (a): the control is modulated so as to enhance the activity of the peripheral nodes of the tree, as compared to the inner ones. Panel (b): the control makes now the bulk nodes more active as compared to the peripheral ones. Panel (c): the root locus diagram relative to the situation displayed in panel (b) is plotted. Blue circles stand for the position of the complex eigenvalues when $\rho = 0$, while green crosses identify the eigenvalues obtained for $\rho \to \infty$. The vertical red line represents the asymptote that attracts two of the modified eigenvalues, when $\rho \to \infty$. The red dots show the computed spectrum, calculated when increasing ρ. In this case the matrix $\boldsymbol{A}$ contains an identical number of ± 1 entries. These are randomly assigned and kept unchanged for all tests performed. The figure on the right is a zoom of the plot displayed on the left.

described in the Appendix C, we can equivalently set $A_{ij} = -1$, while assuming interactions to be modulated by $g(x_j)$ as indicated above. At the same time, the reaction part should be modified with an additional term, η_i, counting the number of negative loops that affects node i. More specifically, $f(x_i) = -\gamma_i x_i + \eta_i$, where the first term mimics constitutive degradation. In our tests, matrix $\boldsymbol{A}$ contains an identical number of randomly assigned ± 1. The parameters γ_i are random variables uniformly distributed over the interval $[0, 1]$. Working in this setting our aim is to control the equilibrum state of the system and thus shape the pattern of asymptotic activity. For this initial application we choose to operate with a simple linear control and set $h_i(x_i, u) \equiv h(u) = u$, for all i. In this case, u could e.g. represent the density of a suitable retroviral vector used to infect specific cell lines [176]. To provide an immediate graphical illustration of the power of the method, we set to stabilize two distinct fixed points. In the first example, see Fig. 5.1(a), the control is designated so as to enhance the degree of activity of the peripheral nodes of the tree. These latter are characterized by a similar value of the activity, apart for slight randomly superposed fluctuations. Similarly, the nodes that define the bulk of the tree display a shared degree (except for tiny stochastic modulation) of residual activity. In Fig. 5.1(b), the dual pattern is instead obtained and stabilized: the peripheral nodes are now being silenced and the activity concerns the nodes that fall in the center of the tree. In Fig. 5.1(c) the root locus diagram relative to the situation reported in Fig. 5.1(b) is displayed. By properly tuning ρ above a critical threshold ρ_c, we can enforce the stability of the obtained fixed point. Two eigenvalues diverge to $\pm\infty$ following a vertical asymptote in the complex plane. For each chosen fixed point that is being stabilized the zeros z_k can be selected so as to make the asymptote intercept the horizontal axis in the left-half of the plane.

As a second application of the proposed control strategy, we set to study the dynamics of the gut microbiota [47]. The intestinal microbiota is a microbial ecosystem of paramount importance to human health [174]. Efforts are currently aimed at understanding the microbiota ability to resist to enteric pathogens and assess the response to antibiotics cure of intestinal infections. Recent advances in DNA sequencing and metagenomics make it possible to quantitatively characterize the networks of interactions that rule the dynamics of the microbiota ecosystem. This was

for instance achieved in [180] by analyzing available data on mice [35] with an innovative approach which combines the classical Lotka-Volterra model and regression techniques. In the following we shall apply the method here developed to control the dynamics of the whole microbioma [180] or a limited sub-portion of it.

In this specific application, the self-dynamics is assumed to be logistic, namely $f_i(x_i) = x_i(r_i - s_i x_i)$, while $g(x_i, x_j) = x_i x_j$. Finally, $h_i(x_i, u) = ux_i$. In [180] the model has been applied to a relatively small (mice gut) microbiota system made up of 11 distinct populations. More precisely, the ten most abundant species have been identified: all together they account for the vast majority ($\sim 90\%$) of the total populations found in the mice gut. The remaining populations are grouped into a unique (non-homogeneous) category referred to as "Other". The authors of [180] provided a quantitative characterization of the coefficients that enter the definition of the relevant quantities $\boldsymbol{r}$, $\boldsymbol{s}$ and $\boldsymbol{A}$. These latter are reported in table 5.1 together with the names of the involved species:

Table 5.1: Populations identification: P1. *Barnesiella*; P2. undefined genus of Lachnospiraceae; P3. unclassified Lachnospiraceae; P4. Other; P5. *Blautia*; P6. undefined genus of unclassified Mollicutes; P7. *Akkermansia*; P8. *Coprobacillus*; P9. *Clostridium difficile*; P10. *Enterococcus*; P11. undefined genus of Enterobacteriaceae.

Species	$\boldsymbol{r}$	$\boldsymbol{s}$	$\boldsymbol{A}$									
P1	0.37	0.20	0	0.10	0.17	-0.16	-0.14	0.02	-0.51	-0.39	0.35	0.01	-0.27
P2	0.31	0.10	0.06	0	-0.04	-0.15	-0.19	0.03	-0.46	-0.41	0.30	0.02	-0.20
P3	0.36	0.10	0.14	-0.19	0	-0.14	-0.16	0.01	-0.50	-0.77	0.29	-0.01	-0.21
P4	0.54	0.83	0.22	0.14	4.6×10^{-4}	0	-0.22	0.22	-0.20	-1.01	0.67	-0.04	-0.40
P5	0.71	0.71	-0.18	-0.05	-5.0×10^{-5}	-0.054	0	0.02	-0.51	0.55	0.16	0.22	0.11
P6	0.47	0.42	-0.11	-0.037	-0.043	0.041	0.26	0	-0.18	-0.43	0.16	-0.061	-0.26
P7	0.23	1.2	-0.13	-0.19	-0.12	0.38	0.40	-0.16	0	1.39	-0.38	0.19	-0.096
P8	0.83	4.3	-0.071	6.0×10^{-4}	0.080	-0.45	-0.50	0.17	-0.56	0	0.44	-0.22	-0.21
P9	0.39	0.056	-0.037	-0.033	-0.050	-0.090	-0.10	0.032	-0.18	-0.30	0	0.014	-0.0077
P10	0.29	0.19	-0.042	-0.013	0.024	-0.12	-0.33	0.021	0.055	-2.1	0.11	0	0.024
P11	0.32	0.38	-0.37	0.28	0.25	-0.17	0.084	0.034	-0.23	-0.39	0.31	-0.039	0

The concentration of the species are measured in 10^{11} rRNAcopies/cm^3, the coefficients $\boldsymbol{r}$ have the dimension of the inverse of time (measured in days), and $\boldsymbol{s}$ is expressed as the inverse of the product of a time for a concentration.

Notice that the set of considered species includes the spore-forming pathogen *Clostridium difficile*. To lower its concentration (and so diminish the probability of infection) is one of the goals of the implemented control. The

results of the analysis are organized under different headings that reflect the three distinct control strategies explored. The physical dimension of the inserted controller u is again 10^{11} rRNAcopies/cm^3. The parameters $\boldsymbol{\beta}$ have dimension of the inverse of a time, while ρ is a-dimensional.

5.2.1 Stabilizing an unstable fixed point by means of an external controller (Case A).

For illustrative purposes, we will restrict the analysis to all sub-systems that combine 5 out of the 11 species analyzed in [180] (more precisely *Barnesiella*, *Blautia*, und. Mollicutes, *Coprobacillus* and und. Enterobacteriaceae). All the fixed points for the 5 species systems are calculated and only those displaying positive concentrations are then retained for subsequent analysis. The stability of each selected fixed point is established upon evaluation of the spectrum of the Jacobian of the reduced dynamics. In Fig. 5.2(a) the histogram of $(\lambda_{Re})_{max}$, the largest real parts of the recorded eigenvalues, is plotted: several fixed points exist that correspond to unstable equilibria. Let us consider one of them, denoted by $\bar{\boldsymbol{x}}$ (see table (5.9)). *The maximum eigenvalue of the Jacobian matrix evaluated at* $\bar{\boldsymbol{x}}$ *is* $(\lambda_{Re})_{max} = 0.0148 > 0$*, thus implying instability.* Starting from this setting, we will introduce a suitably shaped controller, following the above discussed guidelines, in order to stabilize a slightly perturbed version of the originally unstable fixed point, see pie charts in Fig. 5.2(a). Following the exposed procedure one can readily calculate the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ which eventually stabilize the fixed point $\boldsymbol{x}^*$. The values obtained are reported in table (5.9). The spectrum of the Jacobian matrix obtained for $\rho = 0$ (blue circles in Fig. 5.2) protrudes into the right half-plane. More specifically, one eigenvalue exhibits a positive real part, so flagging the instability that one aims to control. At variance, the green crosses in Fig. 5.2(b) stand for the roots z_k of $\mathcal{N}(\lambda)$ and fall in the left side of the complex plane. The vertical (red, in Fig. 5.2(b)) line identifies the location of the two residual eigenvalues of the Jacobian matrix, when $\rho \to \infty$. By tuning the parameter ρ, one can continuously bridge the two above limiting settings, as graphically illustrated in Fig. 5.2(b). When $\rho > \rho_c \simeq 0.01$, the eigenvalues populate the left half-hand plane and stability is, therefore, gained. Direct simulations of the controlled system, as displayed in

Fig. 5.2, confirms that stability has been indeed achieved.

Populations	$\bar{\boldsymbol{x}}$	$\boldsymbol{x}^*$	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}\ \times 10^4$
Barnesiella	0.9736	0.9917	0.0186	0.1860
Blautia	0.8840	0.9093	0.0089	-0.0638
und. uncl. Mollicutes	1.2361	1.2396	0.0089	-0.0032
Coprobacillus	0.1169	0.1363	0.1005	0.0495
und. Enterobacteriaceae	0.0756	0.0894	0.0175	-2.2067

(5.9)

5.2.2 Acting with one species of the pool to damp the concentration of the pathogens (Case B).

Select now a stable fixed point $\boldsymbol{x}^*$ (table 5.10), mixture of five distinct species (*Barnesiella*, und. Lachnospiraceae, Other, *Blautia* and *C. difficile*). The last one is a species of Gram-positive spore-forming bacteria that may opportunistically dominate the gut flora, as an adverse effect of antibiotic therapy. As controller we shall here employ one of the other 6 species that compose the microbioma [67,181]. Vector $\boldsymbol{\alpha}$ therefore follows in this case from the interaction matrix in table 5.1. The aim is to drive the system towards another equilibrium, stable to linear perturbations, which displays a decreased pathogen concentration.
Denote by $\bar{\boldsymbol{A}}$ the reduced 5×5 matrix that specifies all paired interactions between the pool of populations involved in the initial fixed point. The equilibrium solution that can be attained by the controlled system cannot be in this case a priori chosen but it is determined as $\boldsymbol{x}^* = -\bar{\boldsymbol{A}}^{-1}(\boldsymbol{r} + \boldsymbol{\alpha})$. Since it clearly depends on the species used as controller, our freedom lies in the choice of this last one. Retaining only the meaningful cases (fixed points with all positive entries) in the example depicted in Fig. 5.3(c), we obtain three possible solutions: $\boldsymbol{x}_L^*$ where the added species is uncl. Lachnospiraceae, $\boldsymbol{x}_M^*$ adding as external control the und. uncl. Mollicutes and $\boldsymbol{x}_E^*$ adding und. Enterobacteriaceae (see table 5.10). From inspection of the obtained solutions, one can appreciate the impact of the different employed controllers: in the latter case the concentration of *C. difficile* stays almost constant, in the second example it increases, while in the first case it is reduced. The pie charts in Fig. 5.3(c) represent, respectively, the initial fixed point and the final stationary equilibrium, as shaped by the

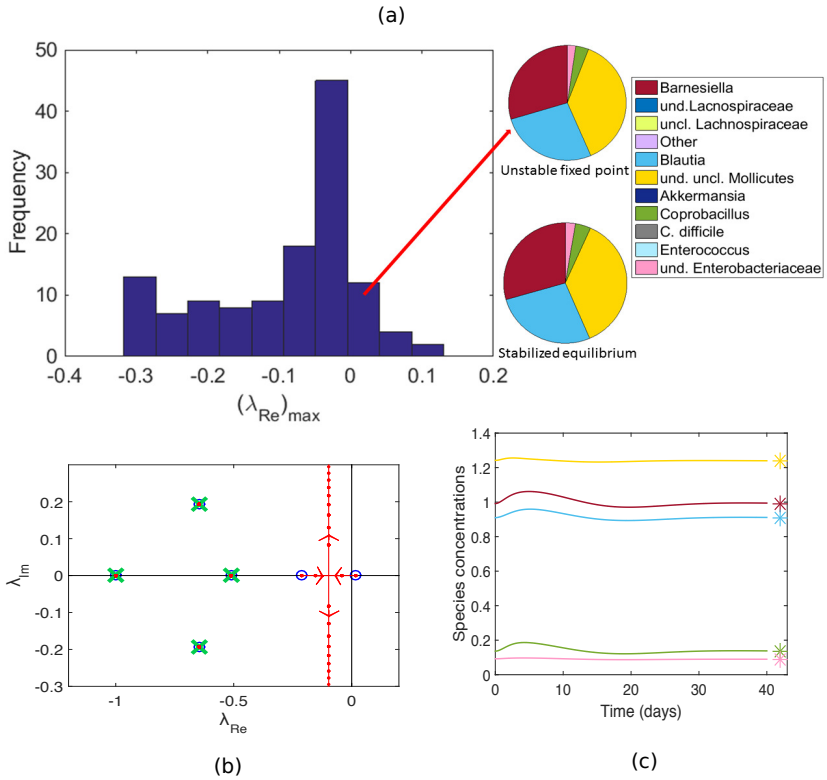


Figure 5.2: Case A. Panel (a): a reduced 5 species subsystem of the microbiota is considered (case A) and all possible fixed points computed. Only those displaying non-negative concentrations are retained and their stability assessed. In the main figure, the histogram of $(\lambda_{Re})_{max}$, the largest real parts of the eigenvalues obtained after the linear stability analysis, is depicted. The two pie charts refer to the initially unstable fixed point (upper chart) and the stabilized equilibrium (lower chart). Panel (b): the root locusf diagram relative to the case discussed in panel (a), is shown. Blue circles identify the position of the complex eigenvalues when $\rho = 0$, while green crosses stand for the eigenvalues obtained in the limit $\rho \to \infty$. The vertical red line is the asymptote that eventually attracts the two residual eigenvalues. The red dots show the computed spectrum, when progressively increasing ρ. Panel (c): numerical integration of the controlled system (see equations (5.2)). The equilibrium state stabilized upon injection of the controller (stars) is a slight modified version of the initially unstable fixed point, see table 5.9. The system is initialized out of equilibrium and, after a transient, converges to $\boldsymbol{x}^*$.

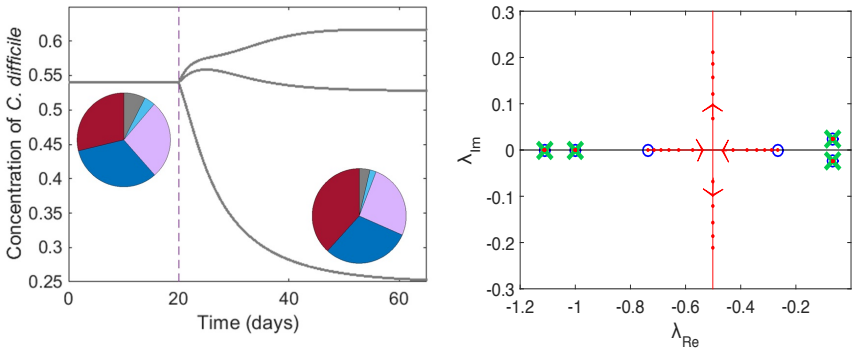


Figure 5.3: Case B. The goal is here to reduce the concentration of the pathogen species, *C. difficile*, by employing as controller one of the species that compose the microbioma. The concentration of *C. difficile* is monitored over time for three different control strategies, turning on the control at the same time ($t = 20$ days). The insertion of the species of uncl. Lachnospiraceae implies a substantial reduction (50 %) of the pathogen concentration, as also displayed by the enclosed pie charts (for interpreting the color-code refer to Fig. 5.2). Right panel: root locus diagram relative to the Case B where, starting from the pool of 5 populations the control is performed adding the species of uncl. Lachnospiraceae. Blue circles correspond to the position in the complex plane of the roots of $\mathcal{D}(\lambda)$ (eigenvalues of $\boldsymbol{J}$ when $\rho = 0$) while green crosses indicate the eigenvalues of the Jacobian in the limit $\rho \to \infty$. Paths followed by the eigenvalues when progressively increasing ρ are shown by the red lines, while the red dots represent the solutions for discrete values of ρ, scanning the interval from 0 to 0.1.

control in the most beneficial case, i.e., when the concentration of *C. difficile* is seen to shrink. The root locus plot obtained for this specific case is reported in Fig. 5.3.

Populations	$\boldsymbol{x}^*$	$\boldsymbol{x}^*_L$	$\boldsymbol{x}^*_M$	$\boldsymbol{x}^*_E$
Barnesiella	2.0745	2.5166	2.0480	2.2542
und. Lachnospiraceae	2.3607	1.9876	2.2674	2.7007
Other	1.9608	1.6929	1.9252	2.0842
Blautia	0.2724	0.1422	0.6171	0.1345
C. difficile	0.5402	0.2435	0.6194	0.5259

(5.10)

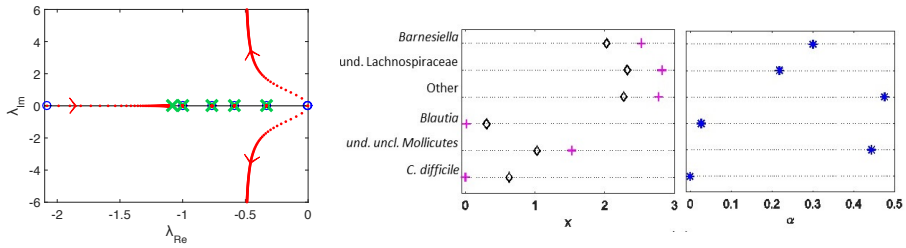


Figure 5.4: Case C. We now modify a stable fixed point, by driving to extinction one of the existing populations, the pathogen *C. difficile* (here species 6), with an indirect control strategy. The obtained concentrations are reported in the left-hand plot of right panel (pluses) and compared with the initial unperturbed solution (diamonds). As anticipated, $x_6^* \simeq 0$. The components of $\boldsymbol{\alpha}$ are plotted in the right-hand plot. Notice in particular that $\alpha_6 = 0$. Left panel: root locus diagram relative to this case.

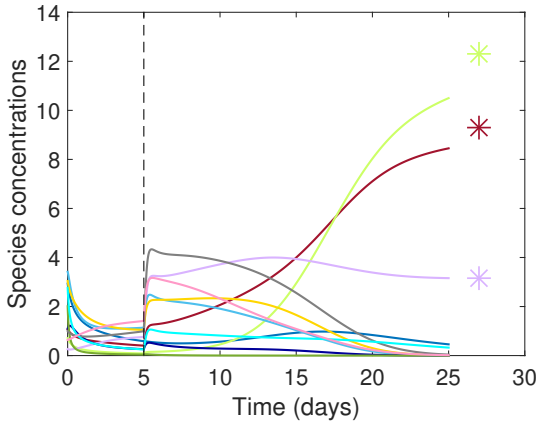
5.2.3 Driving to extinction one species, the other being the target of the control (Case C).

As an additional example, we wish to modify a stable fixed point of the dynamics, by silencing one of the existing populations with an indirect control. In other words we shall introduce and stabilize a novel fixed point, that displays a negligible residual concentration of the undesired species, by acting on the other species of the collection. This is for instance relevant when aiming at, e.g., eradicating a harmful infection that proves resistant to direct therapy. With this in mind, we consider a reduced ecosystem consisting of 6 species, selected among the 11 that define the microbiota. A stable fixed point exists (black diamonds in Fig. 5.4, right panel) which displays a significant concentration of *C. difficile*, the pathogen species. Assign to this latter species the index 6. We now insert a controller which cannot directly interfere with *C. difficile*. This amounts, in turn, to setting to zero the corresponding component of vector $\boldsymbol{\alpha}$ ($\alpha_6 = 0$). We then require the concentration of *C. difficile* to be small, i.e., $x_6^* = \varepsilon << 1$. This latter condition translates into a constraint that should be matched by the other 5 species, namely $\sum_{j\neq 6} \bar{A}_{6j} x_j^* = (s_6 - \bar{A}_{66})\varepsilon - r_6$. Given x_k^*, the components α_k, with $k \neq 6$, are chosen so as to match the constraint $\alpha_k = (-r_k + s_k x_k^* - \sum_{j\neq 6} \bar{A}_{kj} x_j^* - \bar{A}_{k6}\varepsilon)/u^*$. A possible solution of the problem is reported in Fig. 5.4 (right panel): in the left graph (plus

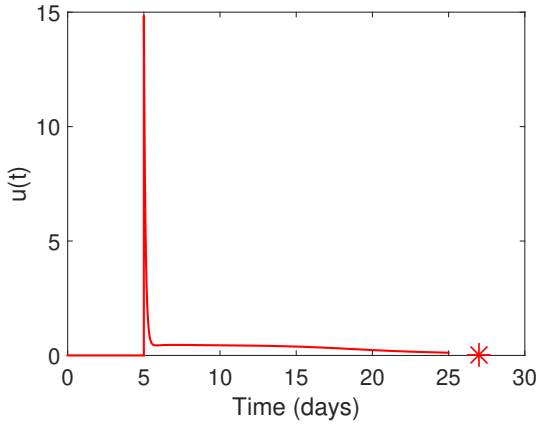
symbols) the components of the fixed point stabilized by the control are shown. As anticipated, the concentration of *C. difficile* is small. The right-hand plot of the same figure shows the components of the vector $\boldsymbol{\alpha}$ that specify the characteristics of the introduced controller. Notice that $\alpha_6 = 0$ so that the controller is not directly influencing the rate of production of *C. difficile*. The root locus diagram is reported in Fig. 5.4.

5.2.4 Controlling the Microbiota network: exploiting a transient control to drive the system towards an existing stable fixed point (Case D)

The control scheme here developed could be in principle exploited to drive the system towards a stable fixed point of the unperturbed dynamics, starting from out-of-equilibrium initial conditions. To achieve this goal u^* needs to be set to zero, thus requiring that the controller is turned off at equilibrium. Here, x_j^* is an equilibrium solution of the uncontrolled dynamics, which proves linearly stable to external perturbations. In this case, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are not subjected to specific constraints, as the existence and stability of the desired equilibrium are a priori granted. Such parameters could hence be chosen so as to reflect the specificity of the target system. The parameter γ can be tuned as desired so as to help the convergence towards x_j^* without falling in the basin of attraction of other existing fixed points. As a proof of principle of the method, we choose a stable fixed point of the global microbiota ecosystem, i.e. including the complete pool of 11 populations. This is characterized by $x_1^* = 9.299$, $x_3^* = 12.3085$, $x_4^* = 3.1627$ and $x_j^* = 0$ for $j \neq 1, 3, 4$. The largest real part of the eigenvalues of the associated 11×11 Jacobian matrix turns out to be $(\lambda_{Re}) = -0.1306 < 0$, thus implying stability of the aforementioned equilibrium. Imagine to initialize the system out of equilibrium with all species, including the pathogen *C. difficile*, being assigned a random concentration $x_j(0) \neq 0$. The system is let evolve for a while and then, at time t^*, the control is injected. Here, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are assigned as random, uniformly distributed over [0,1], parameters. As clearly displayed in Fig. 5.5, the system is steadily moved towards the equilibrium x_j^* (stars), while the control converges to zero after an abrupt jump. In other words, after a transient whose duration depends on the chosen parameters, the system achieves its asymptotic



(a)



(b)

Figure 5.5: Driving the system towards a stable fixed point of the unperturbed dynamics. The system is initiated out-of-equilibrium: the concentration of all species, including the pathogen *C. difficile*, is set to values different from zero. At $t^* = 5$ days the control u is injected. After a transient the system converges to its stable equilibrium characterized by $x_1^* = 9.299$, $x_3^* = 12.3085$, $x_4^* = 3.1627$ and $x_j^* = 0$ for $j \neq 1, 3, 4$, while the control u is turned to zero. Here $\gamma = 10$ and $u(0) = 15$.

(pathogen free) equilibrium and the control can be safely disconnected.

5.3 Generalized control

In the above implementation, the system is described by generic equations (5.1), however the controller u does not benefit of the same generality, see (5.2). The equation that governs its time evolution and interactions indeed assumes a specific form ensuring the success of the stabilization process. Moreover, we assume that the state of the system is always accessible to direct measurement so that the controller evolution is adjustable according to the distances $\boldsymbol{x} - \boldsymbol{x}^*$ and $u - u^*$ appearing in (5.2). Relaxing this working hypothesis is the main goal of this section, where the controller evolution will be governed by a different equation. We will in particular assume that the controller dynamics is ruled by an equation similar to that characterising the other N species. The controller will therefore configure as an effective $(N+1)$-th species of the pool and the only quantity to be externally controlled will consist of its equilibrium concentration u^*.
In their most generic form eqs. (5.2) become:

$$\begin{cases} \dot{x}_i = f_i(x_i) + \sum_j A_{ij} g_i(x_i, x_j) + \alpha_i h_i(x_i, u) \qquad i = 1, ..., N \\ \dot{u} = l(u) - \rho \sum_j \beta_j m_j(x_j, u) \end{cases} \tag{5.11}$$

with $f(\cdot)$, $g(\cdot,\cdot)$, $h(\cdot,\cdot)$, $l(\cdot)$ and $m(\cdot,\cdot)$ generic non-linear functions. We will however restrict our analysis to the simple case where the self-dynamics of each species, controller included, is logistic and the interaction functions are quadratic (product of the densities of the interacting species), as follows:

$$\begin{cases} \dot{x}_i = x_i(r_i - s_i x_i) + \sum_j A_{ij} x_i x_j + \alpha_i x_i u \qquad i = 1, ..., N \\ \dot{u} = u(r_{N+1} - s_{N+1} u) - \rho \sum_j \beta_j x_j u. \end{cases} \tag{5.12}$$

The fixed points $(\boldsymbol{x}^*, u^*)$ of this system are given by imposing $\dot{x}_i = 0$ and $\dot{u} = 0$. The only solution where each component is different from zero,

namely where none of the species becomes extinct [4], is given by:

$$\begin{cases} r_i - s_i x_i^* + \sum_j A_{ij} x_j^* + \alpha_i u^* = 0 & i = 1, ..., N \\ r_{N+1} - s_{N+1} u^* - \rho \sum_j \beta_j x_j^* = 0. \end{cases} \tag{5.13}$$

The stability of such equilibrium is assessed by the eigenvalues of the Jacobian matrix $\tilde{J}$ in the proximity of the fixed point. In order to compute $\tilde{J}$ we move from the equilibrium with a small perturbation $(\boldsymbol{v}, w)$, thus obtaining

$$\begin{pmatrix} \dot{\boldsymbol{v}} \\ \dot{w} \end{pmatrix} = \tilde{J} \begin{pmatrix} \boldsymbol{v} \\ w \end{pmatrix} = \begin{pmatrix} \tilde{G} & \boldsymbol{q} \\ -\rho \boldsymbol{\beta}^T u^* & -s_{N+1} u^* \end{pmatrix} \begin{pmatrix} \boldsymbol{v} \\ w \end{pmatrix} \tag{5.14}$$

where, coherently with the above sections, $\boldsymbol{q}$ is a N-dimensional column vector of components $q_i = \alpha_i x_i^*$, while the matrix $\tilde{G}$ is defined by

$$\begin{aligned} \tilde{G}_{ii} &= -s_i x_i^* \\ \tilde{G}_{ij} &= A_{ij} x_i^*. \end{aligned}$$

The eigenvalues of $\tilde{J}$ are the roots of the characteristic polynomial $\mathcal{P}(\lambda) = \det(\tilde{J} - \lambda \mathbb{1}_{N+1})$, which is explicitly given by:

$$\begin{aligned} \mathcal{P}(\lambda) = (-r_{N+1} - \lambda) \det(\tilde{G} - \lambda \mathbb{1}) + \rho \Big\{ \sum_{ij} \alpha_i \beta_j x_i^* u^* [\mathrm{adj}(\tilde{G} - \lambda \mathbb{1})]_{ji} + \\ + \sum_k \beta_k x_k^* \det(\tilde{G} - \lambda \mathbb{1}) \Big\} \equiv \tilde{\mathcal{D}}(\lambda) + \rho \tilde{\mathcal{N}}(\lambda). \end{aligned} \tag{5.15}$$

In the last equation we have defined the polynomials $\tilde{\mathcal{D}}(\lambda)$ and $\tilde{\mathcal{N}}(\lambda)$, respectively of degree $N+1$ and N. If we want the system to stabilise in an a priori chosen fixed point $\boldsymbol{x}^*$, we will need (5.13) to be satisfied for such equilibrium state and the roots of $P(\lambda)$ to have negative real parts. Following the above guidelines, we aim to stabilise the system by controlling the polynomial $\tilde{\mathcal{N}}(\lambda)$, in particular by exploiting its dependence

[4] While imposing $\dot{x}_i = 0$ we obtain $x_i(r_i - s_i x_i^* + \sum_j A_{ij} x_j^* + \alpha_i u^*) = 0$ which admits multiple solutions. However, if we impose $x_i \neq 0 \ \forall i$ (thus excluding the possibility of extinctions), we can divide by x_i. This leaves us with a linear equation corresponding to a unique solution given by (5.13).

on the free parameters $\boldsymbol{\beta}$. Whereas the polynomial $\tilde{\mathcal{D}}(\lambda)$, analogously to $\mathcal{D}(\lambda)$ in the previous sections, is fixed by the system constants and by the chosen fixed point of the dynamics. The idea, again, is to choose the vector $\boldsymbol{\beta}$ such that the roots of $\tilde{\mathcal{N}}(\lambda)$ have negative real part. This will ensure that for a large enough value of ρ, the roots of $\mathcal{P}$ are negative, even in case the roots of $\tilde{\mathcal{D}}(\lambda)$ happen to fall in the half-plane corresponding to positive real part. The other free parameters are represented by the entries of vector $\boldsymbol{\alpha}$, which can be determined by (5.13), together with the concentration of the controller u^*, once the vector $\boldsymbol{\beta}$ has been assigned. More in detail, u^* is obtained by the last equation of (5.13) and substituted in the other N equations to obtain the N components of $\boldsymbol{\alpha}$.

At this point all that remains to do is to devise optimal strategies to find good values for the components of the free parameter vector $\boldsymbol{\beta}$.

5.3.1 Strategy 1: single $\boldsymbol{\beta}$ approach

The first approach is very simple and also allows us to control the system with a partially local interaction.
Let us observe that by exploiting the fixed point equation (5.13) for u^* the polynomial $\tilde{\mathcal{N}}(\lambda)$ becomes

$$\begin{aligned}\tilde{\mathcal{N}}(\lambda) &= \\ &= \sum_k \beta_k \Big\{ \sum_i \big[x_i^* [\mathrm{adj}(\tilde{G} - \lambda \mathbb{1})]_{ki} - r_i + s_i x_i^* - \sum_j A_{ij} x_j^* \big] + x_k^* \det(\tilde{G} - \lambda \mathbb{1}) \Big\} \\ &\equiv \sum_k \beta_k \tilde{\mathcal{N}}_k(\lambda)\end{aligned} \tag{5.16}$$

thus defining a set of N polynomials $\tilde{\mathcal{N}}_k(\lambda)$ which are independent on the free parameters $\boldsymbol{\beta}$ and consequently a priori established. Now, let us suppose that all the roots of one of these polynomials, e.g. the one corresponding to the index $k = \bar{k}$, are completely negative. Then there exists a simple and convenient choice for the free parameters: $\beta_{\bar{k}} \neq 0$ and $\beta_k = 0$ for all $k \neq \bar{k}$. This solution therefore implies that the controller species is only affected by one of the species of the original set. Finding the possible solutions where the controller only interacts with some of the pre-existing species is one of the goals of this analysis. A local control is indeed easier to realize than a global one, especially when the system of

interest consists of a large number of interacting families. In order to make local also the action of the controller on the other families (not only the back-action of them on the controller) we should also reduce the number of non-zero components in vector $\boldsymbol{\alpha}$. A strategy to accomplish this goal will be explained at the end of this chapter.

5.3.2 Strategy 2: optimization

The second procedure that we expose allows us to find an optimal solution for $\boldsymbol{\beta}$ in the event that the previous approach is not applicable. This will eventually result in a stabilizability test more than in a real stabilization strategy. Proceeding in analogy with (5.5), we can write the polynomial $\tilde{\mathcal{N}}(\lambda)$ as:

$$\tilde{\mathcal{N}}(\lambda) = \sum_{n=1}^{N+1} \tilde{d}_n \lambda^{n-1} = \sum_{n=1}^{N+1} \left(\sum_{l=1}^{N} \tilde{H}_{nl} \beta_l \right) \lambda^{n-1} \tag{5.17}$$

where $\tilde{H}$ is a matrix with components

$$\tilde{H}_{nl} = -\sum_{k=0}^{N-n} \tilde{c}_{n+k+1} \sum_{i=1}^{N} \tilde{G}_{li}^{k} \tilde{q}_i + \tilde{c}_n x_l^* \tag{5.18}$$

where the first term is analogous to equation (C.3) in Appendix, being $\tilde{\boldsymbol{c}}$ the coefficients of the characteristic polynomial for $\tilde{G}$. In this generalised case it is however not possible to proceed following the previous approach, i.e imposing the coefficient $\tilde{d}$ in order to have negative roots and then obtaining $\boldsymbol{\beta}$ by the inversion of $\tilde{H}$, because this last one, differently from the matrix H, is not a square matrix and has dimension $(N+1) \times N$. In other words, the system $\tilde{d}_n = \sum_{l=1}^{N} \tilde{H}_{nl} \beta_l$ with $n = 1, ..., N+1$ is composed by $N+1$ equations with only N free parameters given by the elements of $\boldsymbol{\beta}$. We can therefore decompose such system in the N-dimensional system

$$\begin{aligned} \tilde{\boldsymbol{d}}^S &= \tilde{H}^S \boldsymbol{\beta} \\ \tilde{d}_{N+1} &= \sum_{l=1}^{N} \tilde{H}_{N+1,l} \beta_l \end{aligned} \tag{5.19}$$

where $\tilde{H}^S$ and $\tilde{\boldsymbol{d}}^S$ take the first N rows of respectively $\tilde{H}$ and $\tilde{\boldsymbol{d}}$ and those are subjected to the linear constraint defined by the last row. Here, we

exploit a method exposed by Blondel et al. in the article [26] to find the explicit solutions for root optimization of a polynomial family with one affine constraint. Let us first assume that $\tilde{\mathcal{N}}(\lambda)$ is monic ($\tilde{d}_{N+1} = 1$) without loss of generality, the coefficient multiplying λ^N can be indeed reabsorbed in the definition of parameter ρ. Consider the affine family of monic polynomials

$$P \equiv \{\sum_{n=1}^{N} \tilde{d}_n \lambda^{n-1} + \lambda^N \big| \sum_{l=1}^{N} \tilde{H}_{N+1,l}\beta_l - 1 = 0, \tilde{d}_n \in \mathcal{R}\} \tag{5.20}$$

where the constraint can be reshaped in the form $B_0 + \sum_{j=1}^{N} B_j \tilde{d}_j^S = 0$, with $B_0 = -1$ and $B_j = \sum_{l=1}^{N} \tilde{H}_{N+1,l}(\tilde{H}^S)^{-1}_{lj}$.
The stabilisation problem can be formulated as the problem of finding the scalars $\tilde{d}_1, ..., \tilde{d}_N$ for the polynomial family P such that the roots are as negative as possible. In detail, let us denote with $m(p) = max\{Re(\lambda)|p(\lambda) = 0, \lambda \in \mathbb{C}\}$ the *root abscissa* of a polynomial p of the family P. The optimization problem

$$\bar{m} \equiv \inf_{p \in P} m(p)$$

is solved by $\bar{m} = min\{z \in \mathbb{R}|h^{(i)}(-z) = 0 \text{ for some } i \in (0, ..., k-1)\}$, where $k = max\{j : B_j \neq 0\}$ and $h^{(i)}$ is the i-th derivative of the polynomial $h(z) \equiv B_N z^N + B_{N-1}\binom{N}{N-1} z^{N-1} + ... + B_1\binom{N}{1} z + B_0$. In short, with this method we obtain the minimum value $\bar{m}$ attainable by the root with maximum real part. The procedure is consequently successful if and only if $\bar{m}$ is negative and in this case we can obtain the coefficients $\tilde{d}_1, ..., \tilde{d}_{N+1}$ of the corresponding polynomial p. This method therefore provides a possible solution for defining polynomial $\tilde{\mathcal{N}}$. However this tipically corresponds to the case where all the roots coincide and take the real value given by $\bar{m}$. For our goal, stability, it is sufficient that the eigenvalues are negative, however a linearized system characterised by a Jacobian matrix with multiplicity different from one, implies an early amplification of the signal given by the non-exponential terms in (1.13) during the transient (before reaching equilibrium). In a multistable phase space such behaviour can sometimes bring the system out of the basin of attraction heading towards other equilibria, so mimicking instability. In order to avoid this inconvenience, we will use the optimization approach just to assess the possibility of stabilizing the system given by the sign of $\bar{m}$, but we will devolve to different strategies the task of finding the coefficients $\tilde{d}_n$, as explained in

the next sections.

5.3.3 Strategy 3: least squares

Another possible strategy to deal with the problem of $N+1$ equations with N variables is a least squares approximation. The approach consists of imposing the negative sign to the roots of $\mathcal{N}$ by forcing in advance their positions in the complex plane. This implies to fix the polynomial $\mathcal{N}$ and consequently the coefficients $\tilde{d}_n$. The procedure to obtain $\boldsymbol{\beta}$ is therefore based on the *pseudoinverse* of $\tilde{H}$.
The pseudoinverse, or Moore-Penrose inverse, A^+ of a matrix A is a generalization of the inverse matrix, which can be computed also for non-square matrices. It satisfies the criteria:

- $AA^+A = A$
- $A^+AA^+ = A^+$
- $(AA^+)^* = AA^+$
- $(A^+A)^* = A^+A$.

In our case, we have to deal with the system $\tilde{H}\boldsymbol{\beta} = \tilde{\boldsymbol{d}}$ from which we should obtain $\boldsymbol{\beta}$. Since the system is characterized by $N+1$ equations and N unknown, in general no solution exist. However, the least squares approach allows us to find a solution for $\boldsymbol{\beta}$ which minimizes the Euclidean norm $||\tilde{H}\boldsymbol{\beta} - \tilde{\boldsymbol{d}}||_2$. We can prove that $||\tilde{H}\boldsymbol{\beta} - \tilde{\boldsymbol{d}}||_2 \geq ||\tilde{H}\boldsymbol{z} - \tilde{\boldsymbol{d}}||_2$ where $\boldsymbol{z} = \tilde{H}^+\tilde{\boldsymbol{d}}$. The equality holds for $\boldsymbol{\beta} = \boldsymbol{\beta}_m \equiv \tilde{H}^+\tilde{\boldsymbol{d}} + (\mathbb{1} - \tilde{H}^+\tilde{H})\boldsymbol{w}$ for any vector $\boldsymbol{w}$. The family of vectors $\boldsymbol{\beta}_m$ therefore minimizes the above Euclidean norm and it is the closest we can go to the exact solution. This implies that if we follow this approach and compute the polynomial $\mathcal{N}(\lambda)$ by using β_m thus obtained, its roots will not exactly correspond to the set of negative roots previously chosen, but they will be close to them. Since we focus on the sign of these roots but not on the exact values of them, this proves sufficient for our goal, provided that that they did not cross the imaginary axis becoming positive.
Table 5.23 and Fig. 5.6 (left graph) depict a realization of this method on a subsystem composed by 6 species of the microbiota dataset of the paper [180]. The desired fixed point is stabilized by means of an external controller defined by the parameters reported in the table.

5.3.4 Strategy 4: LASSO regularization

The parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ define the intensity of, respectively, the action of the controller on the other species of the pool and the back-action of those on the added node. Ideally, one would like to control the dynamics of the system by coupling the external control only to a few dynamical variables. We have already seen, in section 5.3.1, that it is possible to find a stable system where the free parameter vector $\boldsymbol{\beta}$ is characterised by a unique component different from zero. This implies that only one species is affecting the concentration u. It is possible to also reduce the action exerted by the controller on the pre-existing families, so as to directly control just one or a few species. This can be done by using the LASSO regularization that we introduce here after. Such procedure will however require to relax the constraint on the choice of the fixed point $\boldsymbol{x}^*$. The system will then stabilize to an equilibrium where the species concentrations will be slightly different from the originally desired one.
The LASSO (Least Absolute Shrinkage and Selection Operator) regularization is a process which allows to solve ill-posed problems and was invented to improve the prediction accuracy and interpretability of regression models [189]. In detail, once the fixed point $\boldsymbol{x}^*$ that we aim to stabilize has been chosen there is only one possible choice for the product $\hat{\boldsymbol{\alpha}} \equiv \boldsymbol{\alpha} u^*$, given by the equations (5.13) for the existence of the fixed point. The relation between $\boldsymbol{x}^*$ and $\hat{\boldsymbol{\alpha}}$ is set by

$$\boldsymbol{x}^* = -M^{-1}(\hat{\boldsymbol{\alpha}} + \boldsymbol{r}), \tag{5.21}$$

where we defined matrix M as $M_{ij} = A_{ij} - s_i \delta_{ij}$. However, by allowing some flexibility on x^*, the set of options on the free parameters is extended and so the optimal value for $\hat{\boldsymbol{\alpha}}$, that we call $\hat{\boldsymbol{\alpha}}_L$, can be obtained from

$$\hat{\boldsymbol{\alpha}}_L = argmin_{\hat{\boldsymbol{\alpha}}}\{||\boldsymbol{x}^* + M^{-1}(\hat{\boldsymbol{\alpha}} + \boldsymbol{r})||_2^2 + \lambda_\alpha ||\hat{\boldsymbol{\alpha}}||_1\} \tag{5.22}$$

where $||\cdot||_2$ is the square norm defined by $||\boldsymbol{y}||_2 = \left(\sum_i |y_i|^2\right)^{1/2}$ and $||\cdot||_1$ denotes the 1-norm. We are thus looking for a vector $\hat{\boldsymbol{\alpha}}_L$ with minimal norm which minimizes the quantity inside the square norm. The regularization parameter λ_α has the role of weighting the importance of the constraint.
This operation returns multiple minimizing solutions for $\hat{\boldsymbol{\alpha}}$, all of them corresponding to a different fixed point. In choosing one of these solutions

the goal is twofold: minimizing the number of non-zero components α_i and at the same time minimizing the proximity of the corresponding fixed point to the originally chosen one. Indeed, with this method the exact agreement between $\hat{\boldsymbol{\alpha}}$ and $\boldsymbol{x}^*$ is lost, but we obtain a new vector of concentrations $\boldsymbol{x}^{*\prime}$ to be stabilized with the new, LASSO-originated, choice of the free parameters. This latter is given by $\boldsymbol{x}^{*\prime} = -M^{-1}(\hat{\boldsymbol{\alpha}}_L + \boldsymbol{r})$.
Two examples are reported in the second and third column of table (5.23) and in Fig. 5.6 (b), (c), respectively corresponding to $\lambda_\alpha = 0.2$ and $\lambda_\alpha = 0.7$, where $\boldsymbol{x}^{*\prime}$ denote the fixed point that we are able to stabilize. In the first case (second column) we gave more importance to the choice of the fixed point, the stabilized equilibrium is indeed close to the original fixed point, but the added node needs to have various interactions with the other nodes. Whereas in the second case (third column) we preferred to reduce the elements α_i and the added node indeed presents only three interactions, however, as a consequence, the stabilized fixed point is quite different from the originally desired one (first column).
The redefinition of the fixed point precedes the selection of the remaining free parameters $\boldsymbol{\beta}$ and ρ, which can be obtained by making use of one the methods reported in the previous sections. The controller concentration and $\boldsymbol{\alpha}$ are subsequently obtained from $u^* = (r_{N+1} - \rho \sum_j \beta_j x_j^*)/s_{N+1}*$ and $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}/u^*$. Moreover, if we choose to obtain $\boldsymbol{\beta}$ following the procedure of section 5.3.1 we end up with a really local controller, being it directly affected by only one of the original species and directly affecting only a few of them.

	least squares	LASSO ($\lambda_\alpha = 0.2$) + single $\boldsymbol{\beta}$	LASSO ($\lambda_\alpha = 0.7$) + single $\boldsymbol{\beta}$
$\boldsymbol{x}^{*\prime}$	0.22	0.24	1.15
	0.94	0.94	0.53
	0.48	0.11	0.09
	0.54	0.40	1.12
	0.22	0.005	0.98
	0.096	0.47	2.17
u^*	0.47	3.27	2.67
$\boldsymbol{\alpha}$	-0.19	-0.037	-10^{-4}
	0.82	0.075	0
	3.24	0	0
	-0.27	-0.073	-0.079
	1.73	0	0
	-0.34	-0.040	0
$\boldsymbol{\beta}$	1.66	1	1
	0.11	0	0
	-1.23	0	0
	0.23	0	0
	4.97	0	0
	-1.07	0	0
ρ	12	1	1

(5.23)

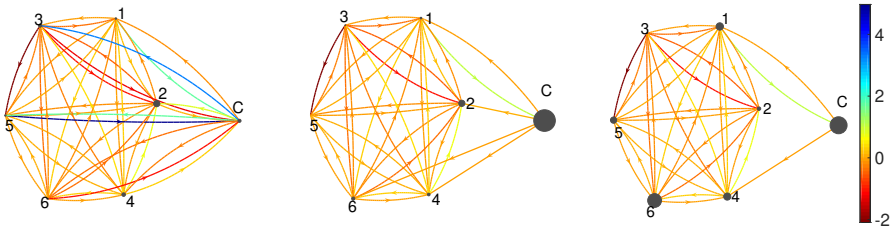


Figure 5.6: Graphic representation of a 6 species interaction network where an additional node is inserted. The control is performed by following the least squares method for the graph on the left and the LASSO regularization approach for the other two, where we used $\lambda_\alpha = 0.2$ for the central one and $\lambda_\alpha = 0.7$ for the graph on the right. The edges have been colored according to the intensity of the interactions (which can be positive or negative) and the nodes size is proportional to the stable fixed point concentration of each species.

Summing up, we have here proposed and tested a method to control the dynamics of multispecies systems on a complex graph. The original system is made up of N interacting populations obeying a set of general equations, which bear attributes of universality. One additional species, here referred to as the controller, is inserted and made interact with the existing constellation of species. By tuning the strength of the couplings (or equivalently the composition of the inserted controller), we can drive the system towards a desired equilibrium. The stability of the achieved solution is enforced by adjusting the parameters that ultimately govern the rate of change of the controller. Methodologically, we make use of the root locus method which can be naturally invoked once the control problem is suitably formulated. The tests that we have performed, both synthetic and drawn from real life applications, demonstrate the versatility and robustness of the proposed scheme. This latter therefore configures as a viable and innovative tool to tackle a large plethora of inter-disciplinary systems, from life science to man-made applications, that should be stably driven towards a desired configuration.

Chapter 6

Random walk and competition between traps

In this chapter we will devise a further peculiar method to control the way a set of microscopical agents asymptotically distribute on the different nodes when they travel a complex network. In many applied situations we are not concerned about the entire set of nodes and our interest only focuses on the eventual filling of certain specific vertices of the network with respect to each other. This is important for instance when we wish to compare two or more nodes of a graph from the point of view of their "reachability" in a random walk process. One possibility is to assign to such vertices the role of *target nodes* in a searching game among agents in the network. This will result in a curious way of controlling a mobility process on complex graphs where we still act on the topology but, instead of modifying edges and nodes, we change the nature of some nodes and the role they play within the network.
The game is not far from reality. Indeed, in real world applications, individual constituents, be they molecules, animals or bits of information, stochastically diffusing on the embedding graph should often head to specific targets, located on selected nodes. A web surfer crawling on a chain of hyper-linked pages to reach a given topic of interest [55], a molecule hunting for the deputed reaction site in topologically tortuous nano-reactors [129, 201] or porous media [24], exciton and electron hole recombination or trapping relevant to photonics and solar-energy science [112]; these are all examples that witness the widespread significance

of devising optimized searching schemes for a stochastic walker on complex geometries assimilated to networks [110, 113, 118, 123, 124, 127, 153, 154]. Moreover, in many cases of interest, the asymmetry in traveling the edges is a fundamental characteristic of the underlying network, which makes more complicated and intriguing the agent mobility. In our model we will then consider random walkers hopping from node to node by following specific imposed directions [64, 131].
Even more importantly, multiple target sites might coexist and mutually interfere with each other, by screening the flux of incoming particles. This will allow to compare different nodes at the level of strategic nature of their position. In particular, it will be possible to elaborate on viable strategies that could yield the most advantageous positioning of a set of target loci (in terms of their associated capturing ability), given a preexisting population of homologous destination sites.

To address this topic, we shall consider the stochastic dynamics of a walker bound to explore a directed graph, modified with the inclusion of *absorbing traps*. As we shall argue, the intransitivity of the examined process is a key ingredient to the forthcoming analysis. For pedagogical reasons, we will specialize on a simplified setting where just two traps are considered, although the analysis will extend straightforwardly to graph endowed with an arbitrary number of absorbing sinks.

Let us assume a first trap a priori set on a specific node of the hosting network and suppose we position a second trap on another random node. All the agents in the network will conclude their path in one of the two absorbing sinks. However these last ones will correspond to different basins of attraction, thus finally gathering different numbers of agents. So, in general, different choices of the second node will correspond to different eventual filling of the first absorbing trap. We then get to the main questions that we shall answer in this chapter:

is it possible to strategically position the second trap so as to obscure as much as possible the first, and so limiting its capacity to absorb diffusing agents? At the same time, can one minimize the risk that the newly added trap gets in turn weakened by the successive insertion of further absorbing sinks? Inquiring on the aforementioned items implies addressing an optimization problem, that we shall solve analytically. As we shall argue, the sought optimum depends on the topological characteristics of

the scrutinized network, the relevant mathematical quantities depending on the eigenfunctions of the associated discrete Laplacian. The theory will be discussed with reference to a specific family of graphs, which displays the small world effect [200]. While the optimization problem is trivial and largely uninteresting on directed regular lattices of connectivity, it is definitely relevant for system hosted on a disordered graph, with long range jumps assigned with a prescribed probability of relocation. Surprisingly, a relatively modest density of long range jumps suffices to yield a meaningful solution to the optimal problem.

To clarify the potential interest of our conclusions, imagine two competitors that are willing to advertise their own products by flagging them on a node of a complex asymmetric network, e.g. a page on the web. The first makes his/her choice and promotes the activity on a specific site, which therefore configures as an absorbing trap for agents (clients) unawarely surfing around. Following our recipe, the second investor can place the second trap on a designated node which (i) limits the number of visitors that can reach the site flagged by the opponent and (ii) secure a strategic positioning to reduce the risk of being shaded by other competitors that might join the venture. On a different level, strategies for optimal integration of multiple reactive sites might have been at play, from biology to chemistry, to shape the world the way we know it.

As already mentioned, the optimization scheme to which we alluded above exploits a fundamental property which ultimately stems from having assumed an asymmetric, hence directed, spatial support. In the context of game theory this property is termed intransitivity [72]. Non-transitive games produce at least one loop of preferences: if strategy A is to be preferred over strategy B, and strategy B outperforms strategy C, then strategy A is not necessarily preferred over strategy C. This is for instance the case for the classical rock, paper, scissor game, which is deliberately constructed to yield a three steps loop. A more subtle implementation of non transitive game is provided by the so called Penney's game [165], a head and tail sequence generating game. The first player bets on a binary sequence of assigned length, and discloses it to the second player, who selects in turn another sequence of identical length. A string is produced by successive tossing of a fair coin, and the player whose sequence appears first, as consecutive readings of the toss outcomes, wins. Provided sequences of at least length three are used, and because of the emerging

intransitivity, the second player statistically wins over the starting player: for any given sequence of length three (or longer), another sequence can be always found that has higher probability of occurring first. Mathematically, the Penney game can be reformulated as a problem of random walk on a directed network, whose nodes are the different sequences of fixed length which can be assembled with a binary alphabet. Remarkably, the non-transitivity relates to the asymmetry of the underlying graph, but the two concepts are to some extent different, as we shall argue in the following. As a matter of fact, the analytical treatment proposed here will materialize into a macroscopic indicator to quantify the global intransitivity of the examined asymmetric graph, enabling one to establish a priori if one contender can outperform the other or, equivalently, if optimal strategies can be played.

6.1 Competition between traps on asymmetric regular lattices

Assume a walker to hop randomly on a one-dimensional directed regular lattice made of N nodes and subject to periodic boundary conditions. Each node is solely connected to its adjacent nearest neighbors. Denote by a (resp. b) the probability of jumping towards the right (resp. left). Here, a and b are positive real numbers chosen to match the condition $a+b \leq 1$. The stochastic $N \times N$ matrix $\mathbf{\Pi}$ which controls the diffusive process (Π_{ji} being the probability of moving from i to j) is therefore circulant, with entries specified by $\Pi_{i-1,i} = b$, $\Pi_{i,i} = 1 - a - b$, $\Pi_{i+1,i} = a$ in such a way that $\sum_j \Pi_{ji} = 1$, $\forall i$. Notice that in this preliminary example the edges among connected nodes are symmetric. The asymmetry that makes the graph directed comes from the probability which controls microscopic jumps. In the following Section we shall turn to consider graphs that are topologically asymmetric, namely graphs that present an heterogeneous distribution of links. In all cases, for the sake of simplicity, we shall refer to direct or, equivalently, asymmetric networks.

In the continuum limit, assuming that nodes are densely distributed on the circle, the probability $p(x,t)$ of seeing the walker in a specific spatial location (identified by the continuum variable x) at time t is governed by:

$$\partial_t p(x,t) = -v\partial_x p(x,t) + D\partial_x^2 p(x,t) \tag{6.1}$$

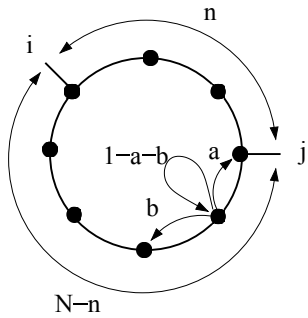


Figure 6.1: A schematic illustration of the diffusion process on an asymmetric lattice, subject to periodic boundary conditions is provided. The two competing traps are located on nodes i and j, respectively. The problem can be equivalently reformulated, by studying the flux of particles inside two adjacent intervals, composed by n and $N-n$ nodes and constrained to match absorbing boundary conditions at the edges.

where $v = b - a$ and $D = (a + b)/2$ respectively denotes the drift velocity and the diffusion constant. In performing the continuum limit we are implicitly setting both space and time elementary intervals to unit. Working in this context, and given any fixed pair of nodes, i and j, we aim at evaluating their relative scores in terms of visits of independent and mutually transparent random walkers. More specifically, we imagine i and j to act as fully absorbing traps. Starting from a uniform distribution (nodes are equally populated at time $t = 0$), we wish to estimate the number of paths that take a walker to i (without hitting j) and viceversa. This analysis translates into a scalar indicator V_{ij}, positive and smaller than one, if properly normalized, that weights the probability of i to *win over* j. Conversely, V_{ji} will measure the probability of j to prevail over i. Clearly, $V_{ij} + V_{ji} = 1$, as it follows from the obvious conservation of the total probability. In the end, a $N \times N$ matrix $\boldsymbol{V}$ can be obtained that quantifies the probability of every node to win against any other selected competitor site. The diagonal elements of $\boldsymbol{V}$ are arbitrarily set to zero.

To explicitly determine the matrix $\boldsymbol{V}$, we first position two fully absorbing traps in respectively nodes i and j. Then, we divide the circle (regular lattice under periodic boundary conditions) into two distinct domains, as schematically depicted in Fig. 6.1. One domain is constituted

by the n nodes encountered when circulating from i to j, clockwise. The ensemble made by the complementary $N-n$ nodes (from j to i, clockwise) defines the second set. To quantify the asymptotic density of walkers that fall on each trap, we need to solve the Fokker-Planck equation (6.1), inside both domains and subject to absorbing boundary conditions at their respective edges. More specifically, we will consider the general solution of the Fokker-Planck equation defined on a one dimensional segment $[0, L]$, with $p(0,t) = p(L,t) = 0$. Further, we shall begin by assuming as initial condition a Dirac delta centered in $x_0 = \alpha L$, with $0 < \alpha < L$. Following [63] (see Materials and Methods), the sought solution reads:

$$p(x,t) = \frac{\Gamma}{2L} e^{\frac{v}{2D}(x-\alpha L) - \frac{v^2}{4D}t} \tag{6.2}$$

where

$$\Gamma \equiv \left[\theta_3\left(\frac{\pi}{2}(\alpha - \frac{x}{L}), z(t)\right) - \theta_3\left(\frac{\pi}{2}(\alpha + \frac{x}{L}), z(t)\right)\right] \tag{6.3}$$

and:

$$\theta_3(r,q) \equiv 1 + 2\sum_{k=1}^{\infty} \cos(2rk) q^{k^2} \qquad z(t) \equiv e^{-\frac{\pi^2 D t}{L^2}}. \tag{6.4}$$

Here $\theta_3(\cdot,\cdot)$ stands for the Jacobi theta function. We are now in a position to evaluate the probability current $J(x,t) = -D\partial_x p(x,t) + vp(x,t)$, flowing to the boundaries, namely $J(L,t)$ e $J(0,t)$. As outlined in Materials and Methods, one eventually gets:

$$J_{\rightarrow}(t) \equiv J(L,t) = \frac{D\pi}{2L^2} e^{\frac{v}{2D}L(1-\alpha) - \frac{v^2}{4D}t} \theta_3'\left(\frac{\pi}{2}(\alpha+1), z(t)\right) \tag{6.5}$$

$$J_{\leftarrow}(t) \equiv -J(0,t) = -\frac{D\pi}{2L^2} e^{-\frac{v}{2D}\alpha L - \frac{v^2}{4D}t} \theta_3'\left(\frac{\pi}{2}\alpha, z(t)\right) \tag{6.6}$$

where $\theta_3'(\tilde{r}, q) = \partial_r \theta_3(r,q)|_{r=\tilde{r}}$, and where we have introduced the positive quantity $J_{\rightarrow}(t)$ (resp. $J_{\leftarrow}(t)$) to denote the current flowing from the right (resp. left) boundary. The probability that particles get absorbed to either right ($\mathcal{J}_{\rightarrow}$) or left ($\mathcal{J}_{\leftarrow}$) boundary is obtained by respectively integrating $J_{\rightarrow}(t)$ and $J_{\leftarrow}(t)$ to yield:

$$\mathcal{J}_{\rightarrow}(L) = \int_0^{\infty} J_{\rightarrow}(t) dt = \frac{1 - e^{-\frac{\alpha v L}{D}}}{1 - e^{-\frac{vL}{D}}} \tag{6.7}$$

$$\mathcal{J}_{\leftarrow}(L) = \int_0^{\infty} J_{\leftarrow}(t)dt = \frac{e^{-\frac{\alpha v L}{D}} - e^{-\frac{vL}{D}}}{1 - e^{-\frac{vL}{D}}} \tag{6.8}$$

To estimate the relative performance of the two traps i and j, we have to generalize the analysis to the case of an initial uniform distribution of the walkers on the lattice. For this reason, we shall integrate the above expressions over the allowed interval in α to yield:

$$\mathcal{V}_{\rightarrow} = \int_0^1 \mathcal{J}_{\rightarrow} d\alpha = \frac{1}{1 - e^{-\frac{vL}{D}}} - \frac{D}{vL} \tag{6.9}$$

$$\mathcal{V}_{\leftarrow} = 1 - \mathcal{V}_{\rightarrow} = 1 - \frac{1}{1 - e^{-\frac{vL}{D}}} + \frac{D}{vL}. \tag{6.10}$$

These preliminary relations will be used to assess the relative performance of the two traps i and j. We recall that the circular lattice that defines the spatial background of the model has been split into two distinct domains: the first formed by the n nodes, visited when going from i to j, clockwise. The second domain is constituted by the remaining $N - n$ nodes, encountered when circulating the ring clockwise from j to i. With reference to the former, $\mathcal{V}_{\rightarrow}$ stands for the flux of particles that eventually hits j, while $\mathcal{V}_{\leftarrow}$ refers to the particles that are eventually attracted towards i. For the other domain, the situation is clearly specular. Hence, the probability V_{ij} that an agent is eventually attracted to node i instead of node j (i.e. without passing from node j) is the sum of two terms: $\mathcal{V}_{\leftarrow}$ calculated for $L = n$ and $\mathcal{V}_{\rightarrow}$ for $L = N - n$. The first term should be weighted by a factor n/N to reflect the average over the initial uniform distribution, while the second needs to be multiplied by a factor $(N - n)/N$. In formulae:

$$\begin{aligned} V_{ij} &= \frac{n}{N}\mathcal{V}_{\leftarrow}(n) + \frac{N-n}{N}\mathcal{V}_{\rightarrow}(N-n) = \\ &= \frac{n}{N} - \frac{n}{N}\frac{1}{1 - e^{-\frac{v}{D}n}} + \frac{N-n}{N}\frac{1}{1 - e^{-\frac{v}{D}(N-n)}}. \end{aligned} \tag{6.11}$$

With analogous considerations one gets:

$$\begin{aligned} V_{ji} &= \frac{n}{N}V_{\rightarrow}(n) + \frac{N-n}{N}V_{\leftarrow}(N-n) = \\ &= \frac{N-n}{N} + \frac{n}{N}\frac{1}{1 - e^{-\frac{v}{D}n}} - \frac{N-n}{N}\frac{1}{1 - e^{-\frac{v}{D}(N-n)}}. \end{aligned} \tag{6.12}$$

For $v/D \simeq 0$ diffusion prevails over drift: on average, half of the particles are expected to fall on trap i and the remaining ones to get absorbed by trap j. This is in turn the limit of a symmetric adjacency matrix for the investigated stochastic dynamics, that yields $V_{ij} = V_{ji} = 1/2$. At variance, for large values of the ratio v/D, the exponential functions in equations (6.11) and (6.12) can be neglected and one eventually obtains $V_{ij} \simeq 1 - n/N$, and, obviously, $V_{ji} \simeq n/N$.

The intransitivity ultimately stems from the underlying network asymmetry, here implemented through unbalanced jumping rates. Although related, the concepts of asymmetry and intransitivity are however subtly different. When the adjacency matrix is asymmetric, it is not a priori guaranteed that, for any selected node i, at least another node j exists that wins over i. Stated differently, for a generic random walker hopping on a directed graph, the entries of a given column(s) of matrix $\boldsymbol{V}$ can be in principle smaller than $1/2$. In the following, we are interested in identifying a specific subclass of asymmetric networks, that we shall term globally intransitive. For these networks, any trap i can be always (statistically) outperformed, in terms of its ability to absorb, by at least another trap, positioned on a given node j. To formally classify the asymmetric network according to this scheme, we introduce an *index of global intransitivity*, η, calculated via the following procedure. We select the maximum from each column of $\boldsymbol{V}$, and then identify the global minimum among collected values. This latter quantity is then shifted by $-1/2$, to yield the index η, which is therefore bound to the interval $[-1/2, 1/2]$. If η is positive the system is globally intransitive, according to the definition evoked above. Classical measures of networks intransitivity rely on triad census. The transitivity coefficient of a network, often termed *clustering coefficient*, is the ratio of the number of loops of length three and the number of paths of length two [28]. In other words, the clustering coefficient quantifies the frequency of loops of length three in the network. The parameter η returns instead a more general estimate of the intransitivity degree, as it does require assuming a priori a specific size of underlying loops.

For the case under scrutiny of a regular asymmetric lattice, $\eta \to 0$, when the drift is virtually silenced ($v/D \to 0$) and the process approaches the symmetric limit. Conversely, for $v/D \neq 0$, $\eta > 0$ and it approaches the limiting value $\eta \to 1/2$ for $v/D \to \infty$. The process of asymmetric particles' hopping on a regular lattice with short ranged connections is

therefore globally intransitive for any $v/D \neq 0$.

Since the process is always globally intransitive, a secondly added trap can always be found that wins over the first, by attracting more walkers. Is it however possible to position the second competing trap (j) on a node that reduces as much as possible the risk of being obscured by yet another trap, the third of the sequence, while still performing better than the first, (i)? To answer this question, we take advantage of the composite information stocked inside matrix $\boldsymbol{V}$. We introduce in fact an additional indicator, called σ_{ij} and defined as follows:

$$\sigma_{ij} = \frac{\sum_{k \neq i} V_{jk}}{N-2} \tag{6.13}$$

The larger the value of σ_{ij} the less the average screening on trap j (selected after trap i), as exerted by an hypothetical third trap, installed in one of the remaining $N-2$ nodes of the network. Notice that, by definition, σ_{ij} stays in the interval $[0, 1]$.

Building on the above we are now in a position to define a supervised strategy for optimizing the selection of trap j, given the pre-defined location of trap i and for a random walk process, taking place on a globally intransitive graph. For fixed i, the key idea is to select j in such a way that it both maximizes V_{ij}, the factor that quantifies direct competition versus i, and σ_{ij}, a measure of competition against the residual bulk. In Fig. 6.2 V_{ij} is plotted versus σ_{ij} for regular one-dimensional asymmetric lattices, characterized by increasing values of v/D amount. Different symbols refer to different choices of v/D. In carrying out the analysis we considered all possible combination of i and j. As it can be clearly appreciated by visual inspection, the data align on an almost vertical line, the value of σ_{ij} being, for practical purposes, constant. The larger the value of v/D, the wider is the vertical band. Optimizing the selection of node j, given i, proves therefore a trivial exercise, when the system is made to diffuse on a regular directed lattice: the best choice is to select the node j which maximizes the V_{ij} score, irrespectively of the corresponding σ_{ij}. The newly introduced trap will be manifestly fragile, as concerns the successive intrusion of additional traps. As we shall see in the following section, complex asymmetric networks, that accommodate for directed long-range jumps to distant sites, yield however a definitely richer scenario and, consequently, more intriguing optimization protocols.

Before concluding this Section, we remark that, in the limit $v/D \to \infty$,

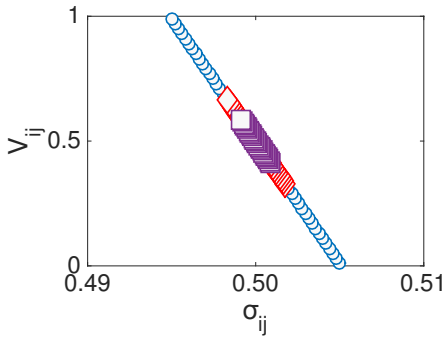


Figure 6.2: (V_{ij} vs. σ_{ij} as obtained for regular asymmetric lattices of the type schematized in Fig. 6.1. Violet squares refer to $v/D = 0.01$, red diamonds to $v/D = 0.02$ and blue circles to $v/D = 2$. As v/D gets larger the distribution of points stretches vertically. To favor visualization we have chosen to plot a subset of the full data point list, for each choice of v/D. Notice the horizontal scale: the points align almost vertically.

the matrix element $V_{ij} \simeq 1 - n/N$. Hence, σ_{ij} as defined in (6.13) can be calculated analytically to return $\sigma_{ij} = (N-3)/(N-2)/2 + n/(N-2)/N$, where n refers to the number of sites entrapped in between nodes i and j, see Fig. 6.1. This latter estimate accurately explains the peculiar distributions as seen in Fig. 6.2.

6.2 Traps on a directed disordered network and the optimization problem

We here aim at extending the above analysis to the case of a walker that is randomly moving on a generic directed network. The entries of the adjacency matrix are one, if two nodes are mutually connected, or zero otherwise. Hence, at variance with the case considered above, the asymmetry of the support is topological, namely related to the distribution of assigned edges, while the probability of individual jumps is constant and set to one, without loss of generality. In the following, we shall denote with $\mathbf{\Pi}$ the stochastic matrix obtained by dividing the columns of the adjacency matrix by the associated nodes connectivity, $\Pi_{ij} = A_{ij}/k_j$.

As discussed above, we are interested in resolving the degree of mutual interference between two distinct traps, respectively located in i and j. Walkers can reach the absorbing traps, but cannot escape from them. To accommodate for this effect we replace the i-th and j-th columns of matrix $\mathbf{\Pi}$ with zeros, except for the diagonal elements which are instead set to one. The obtained matrix is hereafter referred to as to $\mathbf{\Pi}^{[i,j]}$. The master equation that governs the evolution of p_k, the probability of detecting a particle on node k, reads:

$$\dot{p}_k(t) = \sum_l \Pi_{kl}^{[i,j]} p_l(t) - \sum_l \Pi_{lk}^{[i,j]} p_k(t) = \sum_l L_{kl}^{[i,j]} p_l(t) \tag{6.14}$$

where $L_{kl}^{[i,j]} = \Pi_{kl}^{[i,j]} - \delta_{kl}$ is the random walk Laplacian operator modified for the presence of absorbing traps. Again we implicitly assume discrete time updates with $\Delta t = 1$. The last equality in (6.14) follows immediately from the normalization condition $\sum_l \Pi_{lk}^{[i,j]} = 1$. To solve the linear problem (6.14) we need to develop the time dependent probability on a proper basis, which diagonalizes the Laplacian operator. The associated eigenvalue problem takes the form $\sum_l L_{kl}^{[i,j]} \psi_l^{(\alpha)} = \lambda^{(\alpha)} \psi_k^{(\alpha)}$ where $\lambda^{(\alpha)}$ and $\boldsymbol{\psi}^{(\alpha)}$ define, respectively, the eigenvalue and its associated, N-dimensional, eigenvector. It can be proven that two eigenvalues of the discrete Laplacian operator exist which are identically equal to zero and that reflect the imposed absorbing traps. Importantly, and because of the asymmetry of the network, all remaining eigenvalues are complex, and bear a negative real part. The eigenvectors that correspond to null eigenvalues have a rather simple structure: all their components are zero, except for the entry identified by the trap index. This latter component is equal to one. The eigenvectors associated to the Laplacian operators are linearly independent, but they do not constitute an orthonormal basis, as it instead happens when the underlying graph structure is supposed to be symmetric. To solve the linear equation (6.14) we shall preliminarily define an appropriate orthonormal basis $\{\boldsymbol{u}^{(\beta)}\}$, expressed in terms of the original eigenvectors $\boldsymbol{\psi}^{(\alpha)}$ of the Laplacian operator. As we shall see, the request of dealing with an orthogonal basis is fundamental to carry out the forthcoming derivation. Mathematically, one can always find a linear transformation such that $u_k^{(\beta)} = \sum_\alpha C_{\alpha\beta} \psi_k^{(\alpha)}$, where $\boldsymbol{C}$ is the $N \times N$ matrix that specifies the change of basis. This can can be calculated via a straightforward implementation of the Gram-Schmidt orthogonalization

algorithm. We can, therefore, set:

$$p_k(t) = \sum_\beta \hat{p}_\beta(t) u_k^{(\beta)} \tag{6.15}$$

where $\hat{p}_\beta$ represents the coefficients of the expansion on the introduced orthonormal basis. Inserting the above ansatz into equation (6.14) and carrying out the calculation, that we discuss in some detail in the Materials and Methods Section, one eventually obtains the following explicit solution:

$$p_k(t) = \sum_l p_l(0) \sum_\beta (u_l^{(\beta)})^* \sum_\alpha C_{\alpha\beta} \psi_k^{(\alpha)} e^{\lambda^{(\alpha)} t}. \tag{6.16}$$

The asymptotic solution p_k^∞ that is relevant for our purposes can be readily obtained, by performing the limit for $t \to \infty$ in (6.16), and so yielding:

$$p_k^\infty = \sum_l p_l(0) \sum_\beta (u_l^{(\beta)})^* \left[C_{i\beta} \psi_k^{(i)} + C_{j\beta} \psi_k^{(j)} \right] \tag{6.17}$$

where use has been made of the fact that two eigenvalues (those associated to the traps) are zero and all the other have negative real parts (their contributions, stored in the exponential, fade away in the large time limit). In the above equation $(\cdot)^*$ stands for the complex conjugate. Recall now that $\psi_k^{(i)} = \delta_{ik}$ e $\psi_k^{(j)} = \delta_{jk}$, which allow to further simplify equation (6.17) as:

$$p_k^\infty = \sum_l p_l(0) \sum_\beta (u_l^{(\beta)})^* \left[C_{i\beta} \delta_{ki} + C_{j\beta} \delta_{kj} \right] \tag{6.18}$$

Eventually the walker has to land either on site i or j, where the absorbing sinks are located. The relative ability of i and j to trap stochastically walking entities is quantified through the following elements of matrix $\boldsymbol{V}$:

$$V_{ij} = \sum_l p_l(0) \sum_\beta (u_l^{(\beta)})^* C_{i\beta} \tag{6.19}$$

and

$$V_{ji} = \sum_k p_l(0) \sum_\beta (u_k^{(\beta)})^* C_{j\beta} \tag{6.20}$$

with $p_l(0) = 1/N$, in the relevant case where the initial condition is assumed to be uniform. Having determined the elements of the matrix $\boldsymbol{V}$,

we are in the position to estimate both the global intransitivity index η and the parameter σ_{ij} as defined in the preceding Section. As we will demonstrate in the following, the optimization problem discussed above admits a non trivial solution, when the hosting graph is heterogeneous, and as opposed to the simplified setting where the random walk occurs on a regular asymmetric lattice. Incidentally, we note that the above relations provide closed analytical solutions to the family of Penney's games.

To prove our claim we consider a family of directed graphs generated via a straightforward procedure which is adapted from the Watts-Strogatz recipe. Assign the desired number of nodes N and be K their assigned constant degree (connectivity). We then construct a K-regular ring lattice, by connecting each node to its K nearest neighbors, on one side only. Then, for every node i, we select all its associated edges and rewire them with a given probability $q \in [0, 1]$. Rewiring implies replacing the target node, with one of the other nodes, selected with a uniform probability from the ensemble of possible destination sites. The rewiring is directed and the outgoing connectivity is preserved. In Fig. 6.3 we report η as a function of the connectivity K, for (i) the K-ring lattice, and the corresponding disordered graph obtained by imposing a different probability of rewiring, respectively (ii) $q = 0.02$ and (iii) $q = 0.2$. In all cases the index η is positive, hence implying that the system is globally intransitive, a prerequisite condition for the optimization protocol to be applicable. Moreover η decreases as K is increased and, more importantly, as q gets larger. The more disordered the graph, the less intransitive the network appears at the global scale, a reasonable result as rewiring amounts to breaking the perfect asymmetry of the initial lattice and so enforcing a macroscopic symmetrization in the topology of the hosting support.

We now turn to consider the optimization process, following the approach illustrated in the preceding Section. We recall that the idea is to select the location of the second trap j, after having fixed the first one i, so as to maximize, at the same time, V_{ij}, the measure of direct competition versus i, and σ_{ij}, the quantity that controls the degree of competition against the remaining $N - 2$ nodes. The results of analysis are reported in Fig. 6.4, for two different choices of the parameter q (top and lower panels, respectively). For a given network realization, we select a generic node i, which identifies the location of the first trap. The $N - 1$ symbols scattered in the plane (σ_{ij}, V_{ij}) (panels in the left) gauge the performance

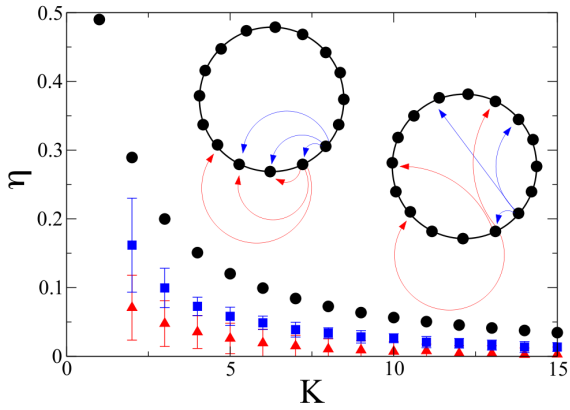


Figure 6.3: The index of global intransitivity η is plotted as a function of the (outgoing) connectivity K for the case of regular K-ring lattices (circles) and disordered graphs, with $q = 0.01$ (squares) and $q = 0.1$ (triangles), respectively. In all cases $N = 100$. The data obtained for the disordered graphs have been averaged over 100 independent realizations. The errors are the recorded standard deviations. The insets provide a pictorial illustration of a K-regular lattice (left) and a disordered network (right)

of the other nodes, imagined as the competitor trap j.

Performing the same analysis for the limiting case $q = 0$, returns a distribution that is substantially uninteresting in the perspective of devising a viable optimization strategy, consistently with the analysis carried out in the preceding Section (data not shown). For $q \neq 0$, instead, the distribution of points gets distorted and progressively elongates along the bisectrix, as clearly testified by visual inspection of Fig. 6.4. Remarkably, the more the global intransitivity index η gets reduced, the more the density points tend to populate the top-right portion of the parameter plane, which incidentally identifies the region of interest for the optimization method here addressed. To cast it differently, when the graph becomes disordered, while still being asymmetric, it is definitely possible to operate a supervised selection of an absorbing sink j, for *any* given choice of i, that outperforms the latter in terms of ability to attract and, still, minimizes

the risk of being in turn buried by successively added traps. The rightmost panels of Fig. 6.4 display the density plot obtained upon averaging over 20 independent realizations of the networks, generated with an assigned rewiring probability q. The trend agrees with the general conclusion illustrated above for a specific choice of i. It is remarkable that a modest amount of long-range edges suffice to yield a significant optimization scenario.

Before ending this Section, we wish to assess the effectiveness of the proposed method. With reference to the choice $q = 0.02$, we place the second trap in the optimal position, as identified in the top left panel of Fig. 6.4. We evolved numerically the stochastic dynamics of the system, starting from a uniform initial distribution, and found that trap j captures almost 75% of the agents, the remaining ones heading to i. But what is going to happen when a third trap is introduced into the competition? Averaging over the $N - 2$ possible locations of the third trap, we see that the second trap still has the lion's share with about 50 % visits out of the total. Conversely, when the second trap j is assigned to the sub-optimal position, as highlighted in Fig. 6.4, it is solely invested by 15 % of the total flux, the remaining amount being directed towards the other two competing sinks. These results are schematically summarized in the insets of the top-left panel of Fig. 6.4. For the case $q = 0.2$, a similar scenario holds: the optimal trap j scores 85 %, while trap i displays only 15 %. Inserting the third trap proves mainly at the detriment of the first sink, the second winning the competition with a final 60 % score. If trap j is assigned to its sub-optimal configuration, as depicted in Fig. 6.4 (lower-left panel), the final score is, as expected, very modest, of 15 %. The histograms inserted in the left-lower panel of Fig. 6.4 summarizes these results. Finally, we varied j, among those nodes that display similar V_{ij} entries, for i fixed. Reducing σ_{ij} is indeed beneficial, as anticipated by our interpretative scheme.

In this chapter we have specifically considered the stochastic dynamics of a walker moving on a directed graph, endowed with two absorbing sinks. Given the network, we imagined the location of the first trap to be assigned a priori. The position of the second trap is designated so as to obscure as much as possible the first, upon estimation of a quantitative indicator that characterizes the degree of pair-interference. At the same time, an optimal location of the second absorbing sink can be determined that allows

to minimize the average screening due to a newly added (third) trap. Analytical formulae are derived which implicitly depend on the topology of the scrutinized network and that enable us to tackle the above optimization process. For walkers exploring a regular lattice, and subject to a constant drift, the optimization protocol is largely ineffective: for any given trap i, a competitor sink j can always be found that absorbs a substantial amount of incoming flux of walkers, at the detriment of i. The average screening coming from an hypothetic third trap is however relevant and substantially independent on the specific location of the assigned trap j. As opposed to this conclusion, the proposed optimization strategy is definitely significant when the random walker is made to explore a disordered directed graph, with long range relocation edges. A modest degree of disorder suffices to yield an effective optimization scheme, as we demonstrated for a specific family of asymmetric complex networks of the Watts-Strogatz type. Explicit formulae are derived which materialize in a new class of indicators for the topological characteristic of complex random graphs. Interestingly, and as a side result, we also propose a global measure for the grade of intransitivity of a network, which does not require to identify closed loops of a given size, as customarily done. The analysis here carried out could be relevant for a large plethora of applications, where multiple reactive sites are concurrently at play. Smart positioning of fully or partially absorbing traps might also translate in innovative non-invasive strategies to control, and consequently shape, the response of a dynamical system bound to evolve on a complex network-like spatial support.

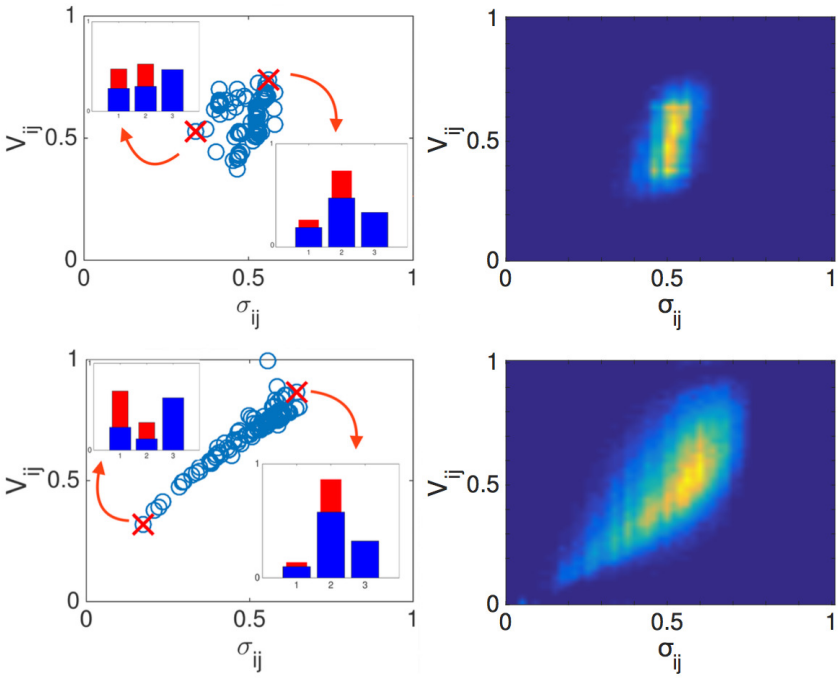


Figure 6.4: Top left panel: (V_{ij} vs. σ_{ij}) for a specific choice of i, and running j on the remaining $N-1$ nodes. Here, $q = 0.02$. Two different choices of j are evidenced which correspond to optimal (top-right) and un-optimal (lower-left) selections, according to the criteria illustrated in the main body of the paper. The histograms report on the performance of the traps $i = 1$ and $j = 2$, when these are the sole sinks present (red thin bars) and when they are competing with trap number 3 (blues tick bars). The data are calculated by averaging over the $N-2$ possible locations for the third trap. The lower histogram depicts the optimal scenario (rightmost red cross). The upper histogram corresponds to the sub-optimal condition (leftmost red cross). Top right panel: density plot for (V_{ij} vs. σ_{ij}), as obtained for a K-ring with $q = 0.02$, upon averaging over 20 independent realizations and allowing all possible selection for the first trap i. Lower panels: same as for the top panels, with $q = 0.2$. Here, N=100. As it can be clearly appreciated, the distribution of points gets progressively stretched along the diagonal, making the proposed optimization protocol gradually more effective for increasing value of the long-ranged distortion q. Notice that a modest probability of relocation q suffices to yield an interesting optimization, an intriguing observation which can be ascribed to the peculiarities of small world networks.

Chapter 7

Reactive random walk

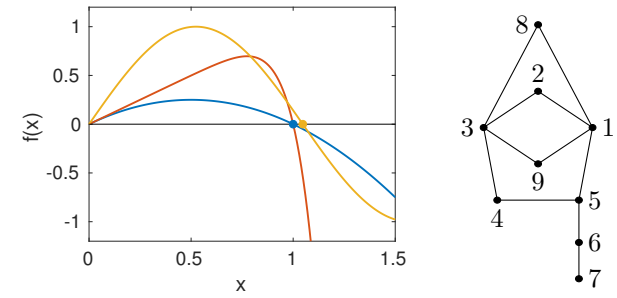
In the previous chapters we have studied reaction-diffusion systems and random walk models, especially focusing on the possible equilibria that the system can reach, and we analyzed their stability. We have then devised peculiar methods to control such systems in order to modify the above equilibria by acting on the topology of the underlying network.
In this last chapter before conclusions, we wonder about what might happen when mixing the ingredients of the previously analyzed models. We will obtain a novel process for which the equilibrium states are anything but trivial and by analyzing these we will be able to follow the inverse process: starting from the naturally obtained equilibrium we will try to get some hints on the structure of the underlying network.

When dealing with a dynamical system defined on a complex network, the coupling describes the nature of mobility, i.e. how the microscopical agents are displaced in time among the different nodes, be they physical space regions or interacting entities. The kind of coupling consequently plays a pivotal role in defining the stationary solutions for node filling. If the process solely involves mobility among nodes, be it described by random walk or diffusion, without a local reaction term, the asymptotically stable equilibrium is completely described by the Laplacian eigenvector corresponding to the maximum eigenvalue, $\Lambda^{(0)} = 0$. This last eigenvector is uniform (i. e. characterized by identical entries for each node) for the diffusion Laplacian L^D and proportional to the nodes degree $\boldsymbol{k}$ for L^{RW}. We have already seen that for reaction-diffusion systems the uniform state

is still a solution of the dynamical equations, even if there is no guarantee that it is stable, indeed the dynamical interplay between diffusion and reaction modifies the equilibrium conditions with respect to the case of pure diffusion. The question that we wish to answer in this last chapter is: what happens if we insert a local reaction term in a random walk system? We will immediately discover that in general the degree vector $\boldsymbol{k}$ is no longer a solution for the system. The fixed point should therefore be found by also taking into account the reaction function and in the following we will devise a strategy to analytically compute it.
Moreover, let us observe that the degree vector corresponds to a measure of centrality for networks which allows us to rank the nodes in order of importance. However when the system is endowed with a reaction the concept of "importance" is no longer merely related to the connectivity because also the local part carries some weight. For this reason the new reactive fixed point can also be thought as a novel measure of centrality. In studying this innovative dynamical model we will also discover its potential for uncovering the underlying network topology, so bringing to light the hidden nodes symmetries and the degree correlations between adjacent nodes.

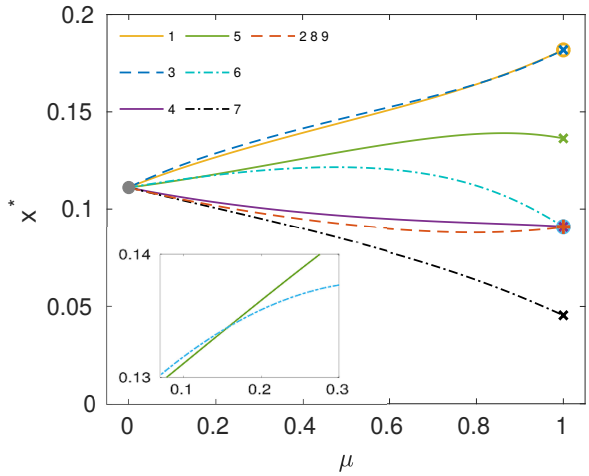
7.1 Model

Our model describes the dynamics of *reactive random walkers*, i.e. random walkers moving over the links of a complex network and interacting at its nodes. Let us consider an undirected and unweighted network with N nodes and K edges, described by a symmetric adjacency matrix $A = \{a_{ij}\}$, where $a_{ij} = 1$ if nodes i and j are linked, and $a_{ij} = 0$ otherwise. We denote as $x_i(t)$ the occupation density, at time t, of node i, with $i = 1, 2, \ldots, N$, so that the state of the entire network at time t is completely described by the vector $\boldsymbol{x}(t) = (x_1(t), x_2(t), \ldots, x_N(t))$. The occupation density $\boldsymbol{x}$ shall be normalised as $\sum_i x_i(t) = 1 \ \forall t$, so that it can be considered as an occupation probability. The law governing the time evolution of $\boldsymbol{x}(t)$ takes into account the network topology, i.e. the adjacency matrix A, and also the specific characteristics of each individual node through a set of local reaction functions. This is formally expressed by the following N



(a)

(b)



(c)

Figure 7.1: (a): Examples of possible reaction functions to be used in (7.1): $f(x) = x - x^2$ in blue, $f(x) = x - x^{10}$ in red and $f(x) = \sin(3x)$ in yellow. (b) A graph of $N = 9$ nodes, and (c) the fixed point $\boldsymbol{x}^*$ obtained when a reactive random walk model with $f(x) = x - x^2$ and different values of the mobility parameter μ is implemented on such a graph. The inset represents a zoom showing the inversion of x_5^* and x_6^* obtained by changing μ.

equations:

$$\dot{x}_i = (1-\mu) f(x_i) + \mu \sum_{j=1}^{N} l_{ij}^{\mathrm{RW}} x_j \qquad i = 1, ..., N \tag{7.1}$$

where μ is a tuning parameter, thereon referred to as the *mobility parameter*, which takes values in $[0, 1]$ and enables us to modulate the weight of two contributions. The first term on the right-hand side of (7.1) accounts for the local reaction at each node i, and is ruled by a function $f(x_i)$ of the occupation probability x_i. For simplicity we assume that the reaction function f is the same for all nodes. The second term takes into account the topology of the network and describes the mobility on it by means of the *random walk Laplacian* $L^{\mathrm{RW}} = \{l_{ij}^{\mathrm{RW}}\}$. This Laplacian is defined as:

$$l_{ij}^{\mathrm{RW}} = \pi_{ij} - \delta_{ij} \tag{7.2}$$

where $\Pi = \{\pi_{ij}\}$ is the transition matrix of a random walk. Entry π_{ji} of matrix Π represents the probability of the random walker to move from node i to node j (see Appendix A). Notice that $\sum_j \pi_{ji} = 1 \ \forall i$. In the simplest possible case we can assume that the random walk is unbiased. This means that the probability of leaving node i is equally distributed among all its adjacent nodes j, so that we can set $\pi_{ji} = a_{ij}/k_i$ for each j. Here, we consider instead a more general transition matrix in the form:

$$\pi_{ji} = \frac{a_{ij} k_j^{\alpha}}{\sum_l a_{il} k_l^{\alpha}} \tag{7.3}$$

which describes degree-biased random walks, i.e. random walkers whose motion also depends on the degree of the node j, and such a dependence can be tuned by changing the value of the exponent α [77]. Namely, for $\alpha > 0$, the walker at node i will preferentially move to neighbours with high degree while, for $\alpha < 0$, it will instead prefer low degree neighbours. Finally, for $\alpha = 0$, we recover the transition matrix $\pi_{ji} = a_{ij}/k_i$ of the standard unbiased random walk.

Summing up, the main ingredients and tuning parameters of the reactive random walkers model in (7.1) are: the network topology, encoded in the adjacency matrix A of the underlying mobility graph; the bias parameter $\alpha \in \mathbb{R}$, which allows to explore the graph in different ways; the local reaction functions ruling the interactions at nodes; and the mobility

parameter $\mu \in [0,1]$ to weight the relative strength of reaction and relocation. Notice that the model of reactive random walkers we have introduced recalls metapopulation models [122, 148, 157], for which the occupation probability of each node of the network wherein the population is allocated is governed by a random walk process, as well as by a local term accounting for birth and death on each environment. (7.1) are also similar to those describing reaction-diffusion processes, but where $x_i(t)$ represents the density at node i at time t, and the Laplacian matrix L^{RW} of (7.2) is replaced by the matrix $L^{\mathrm{Diff}} = \{l_{ij}^{\mathrm{Diff}}\}$ that stems from a purely diffusive process. For similarities and differences between the two definitions of Laplacian see Appendix A.

Limiting case $\mu = 0$. Let us begin the analysis of the reactive random walk model by considering its two limits, namely $\mu = 0$ and $\mu = 1$. In the first limit, the mobility is completely suppressed and the dynamics of each node is independent of the others. Since we have assumed that the function f is the same for each node, Eqs. (7.1) reduce to solve the 1-dimensional system $\dot{x} = f(x)$. In principle the reaction function f can be freely chosen among all the functions $f\colon \mathbb{R} \to \mathbb{R}$. However interesting cases are found when the variable $x(t)$ is bound to converge towards a stationary point, x^*, defined by $f(x^*) = 0$. The function f should then be chosen among the continuous functions and such that 0 is included in its image. Moreover, in order to have equilibrium stability, it is necessary that f is monotonically decreasing in, at least, one of the points where it vanishes, in order to ensure that there exists (at least) one stable fixed point x^*. Some possible examples of reaction functions are reported in Fig. 7.1(a).

Limiting case $\mu = 1$. In the opposite limit, when the mobility parameter takes its maximum value $\mu = 1$, equations (7.1) describe a pure random walk process. The stationary distribution $\boldsymbol{x}^* = \{x_1^*, x_2^*, \ldots, x_N^*\}$ of the dynamics in this limit is obtained by $L^{\mathrm{RW}}\boldsymbol{x}^* = 0$, which is equivalent to $\Pi\boldsymbol{x}^* = \boldsymbol{x}^*$. The Perron-Frobenius [68, 166] theorem ensures that, if the graph is connected and contains at least one odd cycle, the fixed point $\boldsymbol{x}^*$ always exists and is unique. In the case of degree-biased random walks we get [77]:

$$x_i^* = \frac{c_i k_i^\alpha}{\sum_l c_l k_l^\alpha} \text{ with } c_i = \sum_j a_{ij} k_j^\alpha. \tag{7.4}$$

Such an expression, for $\alpha = 0$, reduces to $x_i^* = k_i/2K$, meaning that the

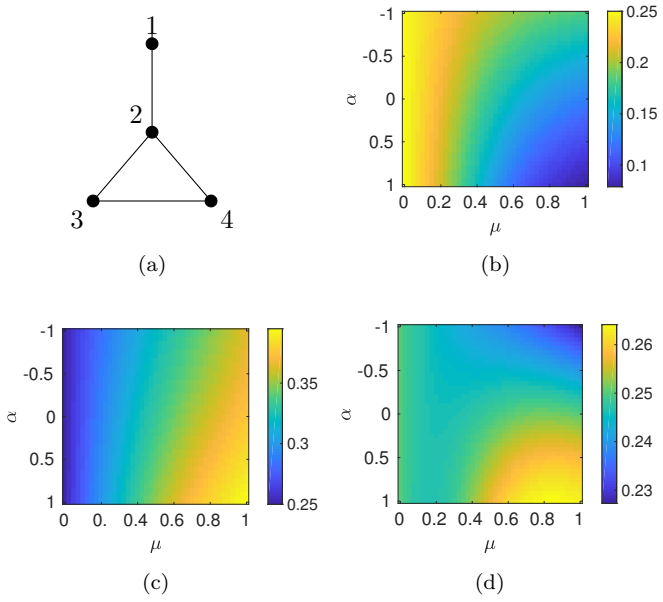


Figure 7.2: Stationary occupation probability of the nodes of graph in (a) is shown for a reactive random walk as a function of the mobility parameter μ and of the bias exponent α: node 1 (b), node 2 (c), nodes 3 and 4 (d). The latter two nodes yield identical patterns (consequently displayed in just one figure), being symmetric nodes. The same reaction function as in Fig. 7.1 has been chosen.

walker, after a long enough period of time, is found on a node i with a probability linearly proportional to the node degree k_i. In this case the asymptotic distribution is completely characterized by the degree $\boldsymbol{k}$ of the graph, with better-connected nodes having a larger probability of being visited by the walker.

The general expression for the asymptotic distribution at a node i, when $\alpha \neq 0$, depends instead not only on the degree k_i of node i, but also on the degrees of the first neighbours of node i, through the coefficient c_i, and such dependence can be tuned by changing the value of the exponent α. For instance, optimal values of the bias, which depend both on the degree distribution and on the degree-degree correlations of a network, can be found to obtain maximal-entropy random walks [36, 77, 177] or to induce

the emergence of synchronization [76].

The general case.

The most interesting dynamics of our model emerges at intermediate values of the mobility parameter μ, when interactions at nodes and random movements between nodes are entangled. In this case, the walkers move on the network jumping from node to node, so that the node occupation probability depends on the network connectivity because of the Laplacian contribution but, at the same time, it evolves at each node according to the reaction function. Reaction functions in turn depend on the occupation probability, so that we have different contributions for differently populated node. This leads to a stationary probability $\boldsymbol{x^*}$ reflecting the topology of the graph in a way that is non trivial and worth analysing. The stationary probability of the model can be obtained, for any value of μ in $[0, 1]$, by setting $\dot{x}_i = 0$ in (7.1) and solving numerically the following implicit equations:

$$x_i^* = \sum_j \frac{a_{ij}}{k_j} x_j^* + \frac{(1-\mu)}{\mu} f(x_i^*). \tag{7.5}$$

Notice however that, when $\mu \neq 1$, the state $x_i(t)$ of node i in (7.1) is not constrained between 0 and 1. This is an effect caused by the reaction term, which behaves as a source term at each node. If we want to interpret the state of the network as an occupation probability, we need then to further impose the normalization, for instance we can consider the vector $\boldsymbol{x}/\sum_i x_i$ instead of the vector $\boldsymbol{x}$.

In the following, we will consider a series of examples so as to get a first insight on the properties of the stationary distribution $\boldsymbol{x}^*$ for different network structures and for different values of the two main tuning parameters of the model, namely the mobility parameter μ and the bias exponent α.

In Fig. 7.1, as local interaction, we consider the logistic function $f(x) = x - x^2$ shown in panel (a), and we implement the model on the graph of $N = 9$ nodes displayed in panel (b). Panel (c) reports the obtained values of the components of the normalized fixed point $\frac{1}{\sum_i x_i^*}(x_1^*, x_2^*, ..., x_9^*)$ as functions of the mobility parameter μ, when α is fixed to zero. The numerical results are in agreement with the expected behaviours in the two limiting cases $\mu = 0$ and $\mu = 1$. In particular, we get $\boldsymbol{x}^* = \boldsymbol{k}/2K$ for $\mu = 1$, and $\boldsymbol{x}^* = \mathbf{1}/N$ for $\mu = 0$, where $\mathbf{1}$ denotes an N-dimensional

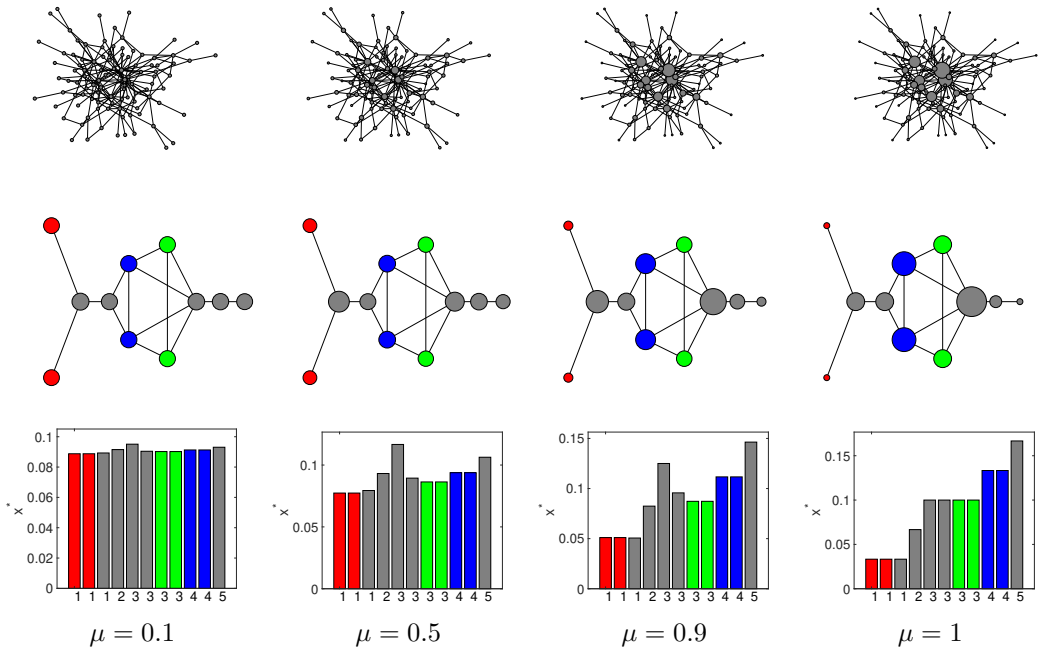


Figure 7.3: Reactive random walks on a scale-free network with $p_k \simeq k^{-\gamma}$ and $\gamma = 2.2$, $N = 100$ and mean degree $\langle k \rangle = 3.6$, and on a toy graph with only 11 nodes. The four columns represent four different values of the mobility parameter: $\mu = 0.1$, $\mu = 0.5$, $\mu = 0.9$ and $\mu = 1$, where the size of the nodes is proportional to the different components of $\boldsymbol{x}^*$. The reaction function selected are $f(x) = x - x^2$ for the first network, and $f(x) = \sin(3x)$ for the second one. In the second graph, the same colour has been used for nodes with the same symmetry, while colour gray has been used for all other (non-symmetric) nodes. The histograms report the stationary occupation at each node, while the node degree is indicated on the x axis.

vector with all entries being identically equal to 1. This means that all the curves in the figure start from the same point $\boldsymbol{x}^*$ at $\mu = 0$, while for $\mu = 1$ we observe four different points $\boldsymbol{x}^*$. The graph considered has in fact nodes with four different degrees, namely $k = 1, 2, 3$ and 4, and curves corresponding to nodes with same number of links will converge to the same point x^* for $\mu = 1$. However, at intermediate values of μ, even nodes with the same degree can exhibit different values of x^* (with the exception of some of them, see Section 7.4 for a discussion on symmetric nodes) going from their degree class at $\mu = 1$ towards $1/N$ at $\mu = 0$. In particular, the various curves of x^* as a function of μ can cluster in a different way when heading towards the limit $\mu = 0$. Let us focus for instance on the behaviour of the node 6 of the graph. Such a node belongs to the degree-2 class but, following the curve of its stationary state when it goes from $\mu = 1$ to $\mu = 0$, we notice that it separates from the curves of the other nodes of its class, approaching the curve of node 5, x_5^*, although the latter node is characterized by a larger degree ($k_5 = 3$). Moreover, node 6 even overcomes node 5 for small values of μ before both curve collapse towards the homogeneous solution. The crossing between the two curves is highlighted in the inset of Fig. 7.1(c).

In the most general case, in our model it is possible to tune both the local dynamics, by choosing different reaction functions $f(x)$, and the bias in the random walk, by considering values of the exponent $\alpha \neq 0$. An illustrative example is reported in Fig. 7.2 in the case of a smaller graph with only four nodes. The three coloured panels show the three different values of the fixed point at the nodes of the network as functions of the mobility parameter μ and the bias exponent α. Notice that node 3 and 4 have the same symmetry in the graph, so they reach the same fixed point (see Section 7.4 for a discussion of symmetries). In detail, while for μ and α equal to zero the four nodes exhibit the same value of the occupation probability, $x_i^* = 0.25\ \forall i$, when we increase the mobility parameter we observe a non-trivial behaviour of these values, which in general decrease for low-degree nodes and increase for high-degree nodes. The effect of introducing a degree-bias in the random walk by turning on and tuning the bias parameter is instead that the occupation probability of the most connected nodes (see nodes 2, 3 and 4) is enhanced for positive and decreased for negative values of α. The opposite happens for the less connected nodes (node 1).

Our third and last numerical example is reported in Fig. 7.3. In this case, we have considered two different topologies, namely a scale-free network with $N = 100$ nodes (first row panels) and a smaller network with $N = 11$ nodes (second row panels). Again, the stationary occupation probability at the nodes of the graphs is shown for various values of μ. For both networks, the size of the nodes in the graphs is proportional to $\boldsymbol{x}^*$, while the four different columns represent respectively the four values of the mobility parameter, $\mu = 0.1, 0.5, 0.9, 1$. While all the nodes have almost equal size for small values of μ, they clearly tend to differentiate when μ increases. Notice that for $\mu = 1$ the node size only reflects their degree, so that the nodes with the largest sizes are the hubs of the scale-free network in the first row. For intermediate values of the mobility parameter (see for instance $\mu = 0.9$), instead the nodes with the largest occupation probability are those connecting isolated vertices to the rest of the network, irrespective of their own degree. This is evident for the second graph in the second and third rows. For this graph, symmetric nodes are also highlighted in figure (see Section 7.4 for a formal definition of symmetric nodes) by adopting the same colours for pairs of nodes with the same symmetry, and reporting in gray nodes not having a symmetric counterpart.

7.2 Analytical derivation of the stationary state

The fixed point $\boldsymbol{x}^*$ of the reactive random walk model in Eqs. (7.1) is in general not easy to obtain analytically because of the interplay between random walk dynamics and local interactions. Approximate techniques can however be employed in the low-mobility limit $\mu \simeq 0$, when the local dynamics is only slightly modified by coupling between network nodes due to the movement. In this limit, it is possible to derive a perturbative estimate for $\boldsymbol{x}^*$: $\boldsymbol{x}^* = s^*\mathbf{1} + \sum_{n=1}^{\infty} \mu^n \boldsymbol{\delta x}^{(n)}$, where $\boldsymbol{\delta x}^{(n)}$ stands for the n-th correction to the uncoupled case. The first two corrections take the explicit form:

$$\delta x_i^{(1)} = -\frac{s^*}{f'(s^*)} \sum_j l_{ij}^{\mathrm{RW}} \tag{7.6}$$

and

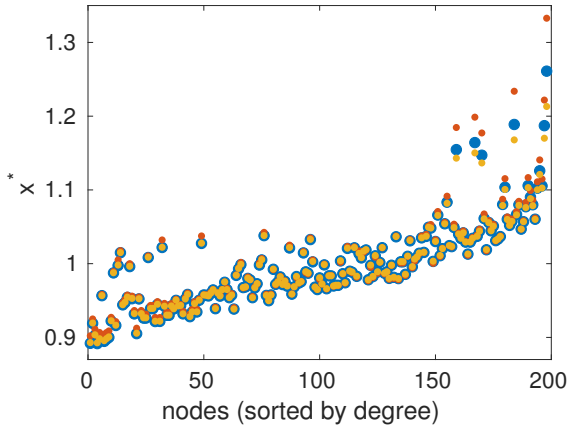
$$\begin{aligned} \delta x_i^{(2)} =& -\frac{(s^*)^2}{2}\frac{f''(s^*)}{f'(s^*)^3}\left(\sum_j l_{ij}^{\mathrm{RW}}\right)^2 - \frac{s^*}{f'(s^*)}\sum_j l_{ij}^{\mathrm{RW}} + \\ &+ \frac{s^*}{f'(s^*)^2}\sum_j l_{ij}^{\mathrm{RW}} \sum_k l_{jk}^{\mathrm{RW}} = \\ =& \delta x_i^{(1)} - \frac{f''(s^*)}{2f'(s^*)}(\delta x_i^{(1)})^2 - \frac{1}{f'(s^*)}\sum_j l_{ij}^{\mathrm{RW}} \delta x_j^{(1)} \end{aligned} \tag{7.7}$$

where s^* is the solution for $\mu = 0$, $f(s^*) = 0$. In Fig. 7.4 we show that the analytical predictions are in agreement with the numerical solution. In particular, we consider reactive random walkers with a mobility parameter $\mu = 0.1$ and a logistic function as local interaction term, and we implement the model on the graph of collaborations among jazz musicians [74].
If f is a C^∞ function, the perturbative terms can be computed for each order n. In this case the hypothesis of small μ can be relaxed and the analytical solution for the fixed point can be, in principle, exactly determined. In such a case, the generic n-th correction can be cast in the form:

$$\delta x_i^{(n)} = -\frac{1}{f'(s^*)}\Bigg\{\sum_{r=2}^{n}\frac{f^{(r)}}{r!}\Bigg[\sum_{m_1=1}^{n-r+1}\sum_{m_2=1}^{n-r-m_1+2}\sum_{m_3=1}^{n-r-m_1-m_2+3}\dots$$
$$\dots\sum_{m_{r-1}=1}^{n-\sum_{j=1}^{r-2}m_j-1}\delta x_i^{(m_1)}\delta x_i^{(m_2)}\dots\delta x_i^{(m_{r-1})}\delta x_i^{(n-\sum_{k=1}^{r-1}m_k)}\Bigg]-$$
$$-\sum_{r=2}^{n-1}\frac{f^{(r)}}{r!}\Bigg[\sum_{m_1=1}^{n-r}\sum_{m_2=1}^{n-r-m_1+1}\sum_{m_3=1}^{n-r-m_1-m_2+2}\dots\sum_{m_{r-1}=1}^{n-\sum_{j=1}^{r-2}m_j-2}\delta x_i^{(m_1)}$$
$$\delta x_i^{(m_2)}\dots\delta x_i^{(m_{r-1})}\delta x_i^{(n-1-\sum_{k=1}^{r-1}m_k)}\Bigg] - f'(s^*)\delta x_i^{(n-1)}+$$
$$+\sum_j l_{ij}^{\mathrm{RW}}\delta x_j^{(n-1)}\Bigg\} \tag{7.8}$$

where $f^{(r)}$ is the r-th derivative computed in s^*.

As expected, at different perturbative orders the local dynamics involves successive derivatives of f at s^*. In particular, the first correction $\boldsymbol{\delta x}^{(1)}$ is only sensitive to the the first derivative, while in $\boldsymbol{\delta x}^{(2)}$ the second derivative appears. In general, the n-th correction is characterized by all the derivatives of f until the n-th one. More interestingly it is worth noticing that $\delta x_i^{(1)}$ contains a term that, when the random walk is unbiased, is proportional to $\sum_j a_{ij}/k_j$, which essentially is a sum over all neighbours of node i of their inverse degree. This implies that the first correction to the generic i-th component of the uniform fixed point depends on the inverse degree of all the nodes of the graph that are adjacent to i. In the second order correction, we instead find the term $\sum_{jl} a_{ij}/k_j a_{jl}/k_l$. The fixed point computed at the second order in μ thus not only depends on the inverse degree of the nearest neighbours of node i, but also on the inverse degree of its second-nearest neighbours. By iterating forward this reasoning, the n-th correction will depend on the n-th nearest neighbours degrees: the term $\sum_j l_{ij}^{\mathrm{RW}}\delta x_j^{(n-1)}$ in eq. (7.8) takes implicitely into account all the nodes of the network that can be reached, in at most n time steps, when starting from node i. Obviously, when n goes to infinity all the nodes of the network contribute with their inverse degree.

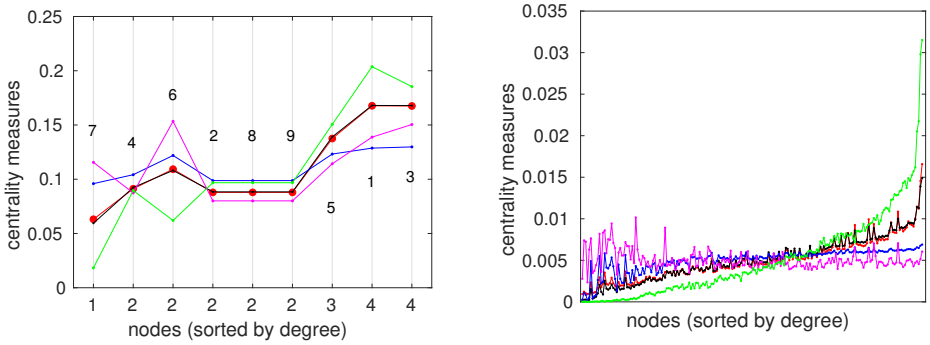


(a)

Figure 7.4: Comparison between analytical predictions (first order in red and second order in yellow) and numerical results (blue dots). The stationary occupation probability of different nodes (sorted by their degree) is shown for reactive random walkers with logistic growth $f(x) = x - x^2$ on the graph of collaborations among jazz musicians [74]. The mobility parameter μ has been set to 0.1.

It is also worth observing that the perturbative calculation can be readily extended to the general case of biased random walks. To this end one should consider the more general Laplacian form in the last term of each correction: $\sum_j l_{ij}^{\mathrm{RW}} \delta x_j^{(n-1)}$. In this case, the first correction $\delta x_i^{(1)}$ is not solely influenced by first neighbours of node i, but also depends on the second neighbours, being proportional to $k_i^\alpha \sum_j \frac{a_{ij}}{\sum_l a_{lj} k_l^\alpha}$. Analogously, for the second correction term, the biased random walks introduces a dependency on the neighbours of all nodes at distance two from each vertex, and so on. In general, considering a degree bias always has the effect of moving the set of involved nodes to further proximity level in the network, as already observed in [77] in the case of non-reactive random walks.

In the next three sections we will explore how the occupation probability of reactive random walkers can turn useful to define novel measures of functional centrality for the nodes of a network, to detect network symmetries, or to distinguish assortative from disassortative networks.



(a) (b)

Figure 7.5: Measures of centrality based on Eq. (7.5) and on different choices of f and α are compared to PRC (red curves) in the case of two networks, the graph of $N = 9$ nodes in Fig. 7.1(b) and the graph of $N = 198$ nodes representing the jazz musician network [74].

7.3 Measures of functional ranking

Centrality measures allow to rank the nodes according to their location in the network [117]. Originally employed in social network analysis to infer the influent actors in a social system, but soon adopted in many other fields, different centrality measures have been constructed to capture different aspects which make a node important, from the number and strength of its connections to its reachability. Commonly used centrality measures are the eigenvector centrality [68, 166], the α-centrality [29, 30], the betweenness centrality [65], the closeness centrality [66] and, of course the simplest one, the degree centrality. This latter corresponds to the fixed point of our model in the limit $\mu = 1$. In this case, the stationary occupation probability x_i^* is indeeed proportional to the degree of node i. However, in our model of reactive random walkers, when $\mu \neq 1$, the stationary state of the model will also depend on the choice of the local dynamics, resulting in a plethora of distinct configurations fostering different roles within the network. In other words, for a fixed value of the mobility parameter we can interpret our dynamical system as a *reaction-dependent centrality measure*. Moreover, we note that the form of Eq. (7.5) on which

this centrality measure is based, is reminiscent of other existing definitions of centralities such as a generalization of the Bonacich centrality [29] known as the α-centrality [30], and the PageRank centrality (PRC) [33]. For instance, the PageRank centrality x_i^{PR} of a graph node i is defined as [34, 73, 116]:

$$x_i^{\mathrm{PR}} = d \sum_j \frac{a_{ij}}{k_j} x_j^{\mathrm{PR}} + \frac{1-d}{N} \tag{7.9}$$

where $d \in (0, 1)$ is a parameter usually set equal to 0.85. PRC was originally proposed as a method to rank the pages of the World Wide Web. Indeed, it mimics the process of a typical user navigating through the World Wide Web as a special random walk with "teleportation" on the corresponding graph. Such a random walker with a probability d performs local moves on the graph (most of the times a user surfing on the Web randomly clicks one of the links in the page that is currently being visited), while with a probability $1 - d$ starts again the process at a node randomly chosen out of the N nodes of the graph (the surfer starts again from a new Web site). The latter action, the so-called "teleportation", is represented by the term $(1 - d)/N$ in Eq. (7.9). Notice that the value of $d = 0.85$ is estimated from the average frequency at which surfers recur to their browser's bookmark feature. The introduction of the teleportation term assigns a uniform non-zero weight to each vertex, and it is particularly useful to avoid pathological cases of nodes with null centrality, in the case the graph is not connected (or strongly connected if a directed graph). In some cases, however, the teleportation contribution is not uniform, but can be designed to gauge an intrinsinc importance of each node. This implies enforcing a dependence on the generic node index i in the second term in the right hand side of Eq. (7.9). The advantage of using Eq. (7.5) instead of Eq. (7.9) as a measure of centrality then consists in the possibility of freely choosing the reaction term. The adoption of function $f(x_i)$ in Eq. (7.5), assigning a different contribution to each node i that depends on x_i, finds a plausible justification in the fact that the importance of a node may also depend on other factors, not necessarily directly linked to the topology of the graph, such as the status or functionality of the node. In a social network, for instance, this factor could be related to the age, social status or income of an individual. Moreover, f can be chosen so as to take into account the time evolution of some features of the nodes of the network. Let us consider again the problem of ranking Web

pages. PageRank centrality in Eq. (7.9) can be modified by replacing the constant teleportation term with a variable contribution, due for instance to the number of visualizations of each page, which could be suitably described by a non-constant term proportional to $x_i(t)$, or more generally by a function $f(x_i)$ as in Eq. (7.5).

As a practical example let us come back to examining the graph in Fig. 7.1(b) and focus again on node 6. According to stardard centrality measures such a node would not result as a very central one, being in a peripheral part of the graph and having just two neighbours. However, one the neighbours is node 7, which is a graph leaf and this makes node 6 its only bridge towards the rest of the graph. This consideration highlights the importance of nodes bridging other nodes of the network and, depending on which characteristics we want to focus on, could be an extremely useful feature to take into account when devising a measure of node centrality. Increasing the importance of this class of nodes can be for instance obtained by an appropriate choice of function $f(x)$ in Eq. (7.5). This is clearly shown in Fig. 7.5(a), where the rankings of the graph nodes obtained for different reference reaction functions and also for different choices of the mobility and bias parameters are compared. The nodes are sorted according to their degree, which is explicitly indicated on the x-axis, while the other reported numbers correspond to node labels as in Fig. 7.1(b). Node 6, which bridges node 7 to the rest of the graph, appears to be more sensitive than the others to the changes, with a large variety of ranking positions, especially if compared to the other nodes with the same degree. The green and magenta symbols respectively refer to a positive ($\alpha = 1$) and a negative ($\alpha = -1$) bias with $f(x) = x - x^2$ and $\mu = 0.85$. We observe that it is also possible to reproduce the same trend of the PRC (red symbols) by again using the logistic function with the same value of the mobility parameter, but setting the bias to zero (black symbols). A different reaction is used for the blue curve: $f(x) = x - x^{10}$ with $\mu = 0.7$ and $\alpha = 0$.
The same types of functional ranking as in Fig. 7.5(a) have also been adopted in Fig. 7.5(b) for the nodes of the network of collaborations among jazz musicians, and the results are reported with the same color code. A similar general trend appears, with low degree nodes enhanced by a negative degree bias and vice versa hubs enhanced by a positive bias. In addition to this, we observe some fluctuations with peaks appearing in the

different ranking measures, most of them corresponding to nodes bridging one or more otherwise isolated nodes of the network.

In conclusion, the proposed measure of functional ranking can mimic other centrality measures, like PRC, in the limit of large μ, where diffusion is important and it is only slightly modified by the local interactions. In general, for every value of μ between 0 and 1, our model of reactive random walkers can be thought as a new way to measure centrality which accounts for the differences between nodes at a deeper level, with the focus on different time-varying characteristics of the nodes themselves.

7.4 Detecting network symmetries

Symmetries are ubiquitous in nature, and one of the main reasons by which humans have been long attempted to describe and model the world through the tools and the language of mathematics. In complex networks, despite the fact that symmetric nodes may appear as special cases, they are surprisingly numerous in real and artificial network structures [175].

In mathematical terms, network symmetries form a group, each element of which can be described by a permutation matrix that re-orders the nodes in a way that leaves the graph unchanged. More precisely, a graph G with N nodes described by the adjacency matrix A has a symmetry if there exists a permutation matrix P, i.e. a $N \times N$ matrix with each row and each column having exactly one entry equal to 1 and all others 0, such that P commutes with A: $PA = AP$. This is equivalent to say that $PAP^{-1} = A$, namely that PAP^{-1} performs a relabeling of the nodes of the original graph which preserves the adjacency matrix A. Therefore, two nodes of the graph are said *symmetric* if their swapping preserves the adjacency relation. This implies that two symmetric nodes are necessarily characterized by the same degree, but also that their neighbours must have the same degree, so as the neighbours of their neighbours, and so on.

While network symmetries may be easy to spot in small graphs like those considered in Fig. 7.1 and in Fig. 7.3, this is typically not the case for large graphs. Different techniques to reveal symmetries in networks have been developed, both numerical and analytical [150, 164, 175, 206]. As we will show below, reactive random walkers provide another method to detect symmetric nodes by looking at the value of the stationary occupation probability at different nodes. In fact, while in the case $\mu = 1$ of a

pure random walk process the fixed point $\boldsymbol{x}^*$ is solely determined by the node degrees, when $\mu \neq 1$ the dynamics is governed by the network as a whole and the value of the stationary occupation probability at a node will depend of its degree, but also on the properties of the second, third and so on neighbours. Hence, it is plausible to conclude that only perfectly symmetric nodes can assume the same asymptotic occupation probability, and to propose to detect symmetric nodes of a graph by looking at those having the same value of $\boldsymbol{x}^*$ for a reactive random walker model with $\mu \neq 1$ on the graph.

An analytical argument in support of this can be obtained from the perturbative derivation of the stationary state presented in Section 7.2. In the limit $\mu \simeq 0$, the expression for the first correction $\delta x_i^{(1)}$ to the uniform stationary state given in Eq. (7.6) contains a term proportional to $\sum_j a_{ij}/k_j$, which indicates the dependence of the stationary state on the degree of the neighbours of i. Analogously, the degree of the second nearest neighbours can be found in the second correction $\delta x_i^{(2)}$, while the degree of the n-th nearest neighbours appears in the n-th correction. The value of x_i^* of a node i will consequently depend on the degrees of all the nodes in the graph. Since two symmetric nodes share the same connectivity at each level of neighbourhood, we can then find symmetric nodes as those with exactly the same value of x^*.

Let us come back to the graphs considered in Figs. 7.1 and 7.3. In the first example the three nodes labeled as 2, 8 and 9 are symmetric, as can be seen directly from figure 7.1 (b), given that they share the same set of neighbours. The existence of such a symmetry is also revealed by looking at the behaviour of the occupation probability of different nodes when varying μ: Fig. 7.1(b) shows that the curves corresponding to these three nodes are indistinguishable. Another remarkable example is reported in Fig. 7.3, where the graph reported in the second row panels is taken as reference model to observe the variation in the occupation probability state for different values of μ. Here, nodes with the same symmetries are shown with the same colour, while the remaining nodes are in grey, and correspond to exactly the same value of $\boldsymbol{x}$, as reported in the third row panels of the same figure.

A more general argument that extends the results above from $\mu \simeq 0$ to the general case $\mu \neq 1$ can be obtained by proving that Eqs. (7.1) are equivariant under a permutation of symmetric nodes [164]. Such equations

can be rewritten in vectorial notation as:

$$\dot{\boldsymbol{x}} = AK^{-1}\boldsymbol{x} - \boldsymbol{x} + \mathcal{F}(\boldsymbol{x}) \tag{7.10}$$

where $K = \{k_{ij}\}$ is a diagonal matrix whose entries are defined as $k_{ij} = k_i\delta_{ij}$, and the functional $\mathcal{F} : \mathbb{R}^N \to \mathbb{R}^N$ is defined such that the generic i-th element of the image vector $[\mathcal{F}(\boldsymbol{x})]_i$ is equal to $f(x_i)$. Our goal is now to prove that Eq. (7.10) also holds for the permuted vector $P\boldsymbol{x}$. Left-multiplying the equation by matrix P we get:

$$\begin{aligned} P\dot{\boldsymbol{x}} &= PAK^{-1}\boldsymbol{x} - P\boldsymbol{x} + P\mathcal{F}(\boldsymbol{x}) = \\ &= AK^{-1}P\boldsymbol{x} - P\boldsymbol{x} + P\mathcal{F}(\boldsymbol{x}) \end{aligned} \tag{7.11}$$

where in the last equality we have used the fact that P commutes with A and, since symmetric nodes have the same degree, it also commutes with K and consequently with its inverse. Now we observe that the role of matrix P is to permute symmetric nodes leaving the others unchanged. The effect of P on a generic vector $\boldsymbol{v} \in \mathbb{R}^N$ is $[P\boldsymbol{v}]_i = v_{\tilde{i}}$ where $\tilde{i}$ denotes the node of the network which is the symmetric twin of i, if it exists, otherwise $\tilde{i} = i$. Consequently, when we apply P to $\mathcal{F}(\boldsymbol{x})$ we obtain a vector whose i-th component is:

$$[P\mathcal{F}(\boldsymbol{x})]_i = [\mathcal{F}(\boldsymbol{x})]_{\tilde{i}} = f(x_{\tilde{i}}) = f([P\boldsymbol{x}]_i) = [\mathcal{F}(P\boldsymbol{x})]_i \tag{7.12}$$

Making use of this result, Eq. (7.11) becomes the equivalent of Eq. (7.10) evaluated for $P\boldsymbol{x}$ instead of $\boldsymbol{x}$, which is what we wanted to prove.

7.5 Measuring degree correlations

A distinguishing feature of many real-world networks is the presence of non-trivial patterns of degree-degree correlations [145, 146, 159]. In the case of positive degree-degree correlation the network is said to be *assortative*: this is often the case for social networks, where hubs have a pronounced tendency to be linked to each other. Conversely, a network is said *disassortative* if the correlations are negative and connections between hubs and poorly connected nodes are favored. Well-known examples of disassortative networks are the Internet, and biological networks such as protein-protein interaction networks, where high degree nodes tend to avoid each other.

One possible way to reveal the presence of degree-degree correlations in a network is to compute the average degree of neighbours of nodes of degree k, and to look at how this quantity depends on the value of k. The average degree $k_{nn,i}$ of the neighbours of node i is defined as $k_{nn,i} = \frac{1}{k_i}\sum_j a_{ij}k_j$. To obtain the average degree of neighbours of nodes of degree k, we need to average the quantity $k_{nn,i}$ over all nodes i of degree k. Let us denote as $p_{k'|k}$ the conditional probability[1] that a link from a node of degree k is connected to a node of degree k'. Now, by expressing the sum over nodes as a sum over degree classes, the average degree of the nearest neighbours of nodes with a given degree k can be written as:

$$\langle k_{nn}\rangle_k = \sum_{k'} k' p_{k'|k}.$$

The function $\langle k_{nn}\rangle_k$ is a good indicator of the presence of degree correlations in a network. In fact, the quantity $\langle k_{nn}\rangle_k$ increases with k when the network has positive degree correlations, it is decreasing when the network has negative correlations, while it is constant and equal to $\langle k^2\rangle/\langle k\rangle$ for uncorrelated networks.

We will now show that the dynamics of the reactive random walker model of (7.1) is sensitive to the presence of correlations in the underlying network, and it is therefore possible to detect and measure the assortative or disassortative nature of a network from the asymptotic node occupation probability. To this end we need to return to the perturbative approach to obtain the equilibrium occupation probability discussed in Section 7.2. As already remarked, a full hierarchy of terms are found to appear as a byproduct of the calculation, which respectively relate to paths connecting nodes that are $1, 2, 3, \ldots$ steps away from any selected node. Let us focus on the first correction to the uniform state, namely the term $\delta x_i^{(1)}$, as specified in (7.6). Up to the multiplicative node-invariant factor $s^*/f'(s^*)$, $\delta x_i^{(1)}$ is equal to $\sum_j l_{ij}^{\mathrm{RW}} = \sum_j a_{ij}/k_j - 1$. Therefore, at the first order, the difference between the equilibrium distribution and the uniform state

[1] To construct the conditional probabilities $p_{k'|k}$ it is convenient to define a matrix E such that the entry $e_{kk'}$ is equal to the number of edges between nodes of degree k and nodes of degree k', for $k \neq k'$, while $e_{kk'}$ is twice the number of links connecting two nodes having both degree k. The conditional probability $p_{k'|k}$ can be then expressed as $p_{k'|k} = e_{kk'}/\sum_{k'} e_{kk'}$ [117]. By definition such a probability satisfies the normalization condition $\sum_{k'} p_{k'|k} = 1\ \forall k$.

is governed by the quantity:

$$w_i^{(1)} \equiv \sum_{j=1}^{N} \frac{a_{ij}}{k_j} = k_i \langle \frac{1}{k_{nn}} \rangle, \tag{7.13}$$

representing, for a generic node i, the sum of the inverse degrees of all its neighbours. The quantity $w_i^{(1)}$ is always non-negative, and gets larger when many nodes are adjacent to node i (large degree k_i, corresponding to many terms in the sum) and all such nodes display smaller degrees. In the particular case in which all the nodes connected to i have exactly degree equal to k_i, we get $w_i^{(1)} = 1$. When instead, the degree k_i of node i is smaller than the inverse of the mean inverse degree of the nodes adjacent to i , then we have $0 < w_i^{(1)} < 1$. In the extreme case of low degree nodes connected to hubs $w_i^{(1)}$ tends to zero[2].

Looking at the whole network, the vector $\boldsymbol{w}$ can be turned into an effective indicator for the presence of degree-degree correlations that relies on the harmonic mean of the degrees instead that on the standard mean.

For instance, we can consider the average value of $w_i^{(1)}$ for all nodes i of degree $k_i = k$. Such a quantity can be written in terms of the adjacency matrix of the graph as:

$$\langle w^{(1)} \rangle_k = \frac{1}{N_k} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{a_{ij}}{k_j} \delta_{k_i,k} \tag{7.14}$$

where $N_k = \sum_{i=1}^{N} \delta_{k_i,k}$ is the number of nodes of degree k. We can rewrite the previous equation by making use of the conditional probability $p_{k'|k}$, so that the sum over all neighbours j of i becomes a sum over the degrees k' of the nodes adjacent to those of degree k. We finally obtain:

$$\langle w^{(1)} \rangle_k = k \sum_{k'} \frac{1}{k'} p_{k'|k} = k \langle \frac{1}{k'} \rangle_k, \tag{7.15}$$

where the quantity $\langle 1/k' \rangle_k$ denotes the average of the inverse degree of the first neighbours of nodes of degree k. In absence of degree correlations the

[2]Using the definition given in section 7.3, the quantity $\boldsymbol{w}^{(1)}$ weights the role of i in bridging the gap between neighbours. In other words, it gauges how much node i is important in linking *isolated nodes* to the main bulk, so keeping the graph connected. We already mentioned the role of node 6 in the graph of Fig. 7.1(b) and how its intrinsic relevance stems from the stationary solution $\boldsymbol{x^*}$ (see Fig. 7.5(a)). The formal explanation of this phenomenon is indeed due to the presence of the quantity $\boldsymbol{w}$ in the first term of the perturbative expansion of $\boldsymbol{x}^*$.

conditional probability takes the form: $p^{\rm nc}_{k'|k} = k' p_{k'} / \langle k \rangle$ [117, 145], where p_k is the degree distribution of the network, $\langle k \rangle$ is the average degree, and "nc" stands for no correlations. Hence, in uncorrelated networks the quantity in (7.15) reduces to:

$$\langle w^{(1)}_{\rm nc} \rangle_k = \big(k \sum_{k'} \frac{1}{k'} p^{\rm nc}_{k'|k} \big) = \frac{k}{\langle k \rangle} \tag{7.16}$$

and is a linearly increasing function of k with slope equal to $1/\langle k \rangle$. Such a function represents the reference case to compare to when evaluating the quantity $\langle w^{(1)} \rangle_k$ for a given network.

In Fig. 7.6(a) we plot $\langle w^{(1)} \rangle_k$ as a function of k for three synthetic networks, respectively with positive, negative and no degree correlations. The uncorrelated network is an Erdős-Rényi random graph with $N = 1000$ nodes and $K = 10000$ edges, while the other two have been generated from the uncorrelated one by using an algorithm that swaps edges according to the degree of the corresponding nodes [31, 203], to produce respectively a disassortative graph with correlation coefficient $r = -0.94$ and an assortative graph with $r = 0.93$ and [145]. The algorithm preserves not only the average degree $\langle k \rangle$, but also the entire degree distribution, which is shown in lower-right inset. Consequently, the results for the three networks, shown respectively as purple, yellow and red pluses, can be directly compared to the same analytical prediction $\langle w^{(1)}_{\rm nc} \rangle_k$ (straight line), which is clearly well in agreement with the randomized network. In the disassortative graph we observe that the quantity $\langle w^{(1)} \rangle_k$ is larger than $\langle w_{\rm nc} \rangle_k$ for degree values $k > \langle k \rangle$. This is because the first neighbours of the hubs are typically poorly connected, i.e. $\langle 1/k' \rangle_k > 1/\langle k \rangle$ when $k > \langle k \rangle$. Conversely, $\langle w^{(1)} \rangle_k$ is smaller than $\langle w^{(1)}_{\rm nc} \rangle_k$ for poorly connected nodes, i.e. for $k < \langle k \rangle$. In the assortative graph, as expected, $\langle w \rangle_k \simeq 1$ for most of the degree classes. Deviations from perfect assortativity only occur at the two extremes of the degree distribution, i.e. for limit values of the degree: a sample node with low (high) degree is in fact linked to nodes whose degree is in average larger (lower) than its own.

In Fig. 7.6(b) and (c) we show the results obtained for two real-world networks with known mixing patterns, namely the collaboration networks of astrophysicists [144] and the Internet at the autonomous systems (AS) level [159]. The first network has $N = 17903$, an average degree $\langle k \rangle$ equal to 22.2 and is assortative with correlation coefficient $r = 0.23$, while the second one has $N = 11174$, $\langle k \rangle = 4.3$ and is disassortative with $r = -0.19$.

A logarithmic scale has been adopted in the two plots, as both networks exhibit long-tailed degree distributions. The plots show larger fluctuations that those observed for the artificially generated graphs. The general behaviour is however preserved and allows to identify the two different types of degree-degree correlations. In particular, the inversion of the trend, which occurs around the mean degree, is clearly preserved. For the assortative network of collaborations in astrophysics $\langle w^{(1)} \rangle$ is larger than the value expected for the uncorrelated case when $k < \langle k \rangle$, while it is smaller that this for almost all the larger values of k. The opposite behaviour is displayed by the Internet network, which is instead disassortative. A possible way to detect the sign and, at the same time, to quantify the entity of the correlations in a network from the study of the quantity $\langle w^{(1)} \rangle_k$ is to extract the slope of the curve $\langle w^{(1)} \rangle_k$ as a function of k at point $k = \langle k \rangle$, and compare it to the slope of $\langle w^{(1)}_{\rm nc} \rangle_k$ vs k for the corresponding randomized case. For instance, we can evaluate the difference $\mathcal{S}$ between the two slopes multiplied by $\langle k \rangle$:

$$\begin{aligned} \mathcal{S} &= \langle k \rangle \frac{d}{dk} \Big(\langle w^{(1)}_{\rm nc} \rangle_k - \langle w^{(1)} \rangle_k \Big) \Big|_{k=\langle k \rangle} = \\ &= 1 - \langle k \rangle \frac{d}{dk} \left(k \sum_{k'} \frac{1}{k'} p_{k'|k} \right) \Bigg|_{k=\langle k \rangle} \end{aligned} \tag{7.17}$$

that we name *slope variation*. The multiplying mean degree has the role of rescaling $\mathcal{S}$, which becomes a quantity of order 1 (instead of $1/\langle k \rangle$) and consequently a comparable measure for networks with different connectivity. Such a quantity has been computed for the networks analysed in Fig. 7.6. Results are reported in Table 7.1 and compared to the standard quantities usually adopted, namely the Pearson correlation coefficient r and the exponent ν governing the behaviour, $\langle k_{nn} \rangle_k \sim k^{\nu}$, of the average degree of first neighbours of nodes of degree k as a function of k. We notice that positive values of the slope variation $\mathcal{S}$ are associated to assortative networks, while negative slope differences indicate disassortative ones, in agreement with the standard indicators of degree-degree correlations.

Table 7.1 also reports the values of $\mathcal{S}$ obtained in a sample of other artificial and real-world networks, and shows that the proposed indicator agrees not only for the sign but also for the order of magnitude with the standard measures, when evaluated for networks with strong degree-degree correlations, namely the network of collaboration in Astrophysics,

Internet AS and Caida, as well as for artificial networks. The exceptional cases where the value of $\mathcal{S}$ results considerably different from r and ν are those where the degree correlation does not prove to be clearly defined, corresponding to a significant error $\Delta\nu$ obtained from the fit of $k_{nn}(k)$.

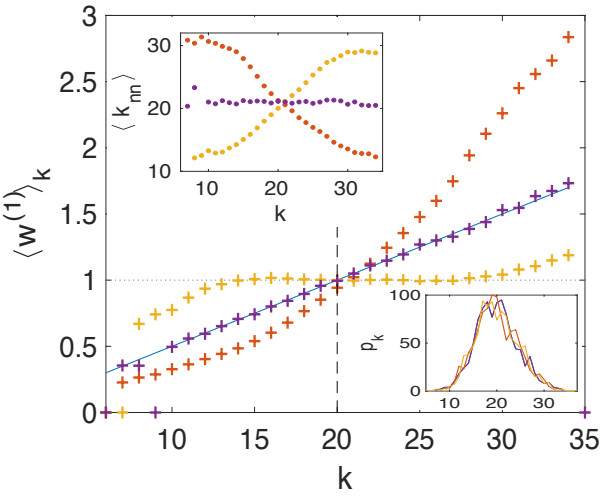
In summary the value of $\mathcal{S}$ provides indication on the presence of degree-degree correlations that are in all similar to r or ν. However, as the n-th term of the expression of $\boldsymbol{x}^*$ in Eq. (7.8) takes into account the degree correlations of a node to those which are n steps away, our indicator can be easily generalized and employed to detect higher order degree correlations. Let us consider for instance the second term $\delta x_i^{(2)}$ of the Taylor expansion in Eq.(7.7). A second order analogue of $w_i^{(1)}$ can be defined as $w_i^{(2)} \equiv \sum_{jl} \frac{a_{ij}}{k_j} \frac{a_{jl}}{k_l}$ to measure the inverse degree of the second neighbours of node i. Such a quantity represents a measure of the connectivity of node i compared to that of nodes which are two steps away from it. The degree k_j present at the denominator mitigates the impact of the number of nodes adjacent to i, so that the comparison only takes into account k_i and the degree of the second neighbours. Indeed, we have $w_i^{(2)} = 1$ when all the second neighbours l of i have degree $k_l = k_i$. As for the case of $w_i^{(1)}$, we can consider the average value of $w_i^{(2)}$ over all nodes i of degree $k_i = k$. Writing this as a summation over degree classes, we have:

$$\langle w^{(2)} \rangle_k = \frac{1}{N_k} \sum_{i=1}^{N} w_i^{(2)} \delta_{k_i,k} = k \sum_{k',k''} \frac{1}{k''} p_{k'|k} p_{k''|k'} \tag{7.18}$$

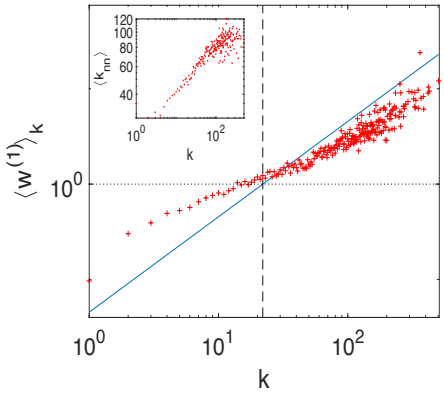
where k'' represents the degree of second neighbours. It is important to notice that the above introduced quantity does not measure genuine second order degree correlations in a network but rather how the effect of first order degree correlations reflects on nodes which are at distance of two steps. The generalization to higher orders follows naturally. Assessing the effectiveness of this latter quantity as compared to other possible generalization of standard degree correlation measures to higher order [5] is left as a challenge for future investigations.

Networks	N	$\langle k \rangle$	r	$\nu \pm \Delta\nu$	$\mathcal{S}$
Synthetic uncorrelated	1000	20	-0.003	-0.02 ± 0.01	-0.01
Synthetic assortative	1000	20	0.93	0.83 ± 0.08	1.02
	1000	20	0.71	0.61 ± 0.06	0.83
	1000	20	0.50	0.34 ± 0.05	0.59
	1000	20	0.30	0.19 ± 0.03	0.36
Synthetic disassortative	1000	20	-0.94	-0.89 ± 0.07	-0.86
	1000	20	-0.71	-0.66 ± 0.05	-0.74
	1000	20	-0.50	-0.35 ± 0.04	-0.52
	1000	20	-0.30	-0.23 ± 0.02	-0.32
Astrophysics collaboration [144]	17903	22.01	0.23	0.22 ± 0.02	0.41
Facebook [120]	4039	43.69	0.11	0.054± 0.051	0.40
Jazz collaboration [74]	198	27.70	0.03	0.11± 0.04	0.46
Email URV [84]	1134	9.61	0.078	0.05 ± 0.03	0.03
C. elegans frontal [104]	453	8.97	0.035	0.062 ± 0.050	0.28
Internet AS [159]	11174	4.19	-0.19	-0.52 ± 0.04	-0.33
Caida [119]	26475	4.03	-0.19	-0.52 ± 0.03	-0.38
US politics books [114]	105	8.42	-0.019	-0.13 ± 0.07	-0.045
US power grid [200]	4941	2.67	0.003	-0.035 ± 0.10	-0.18

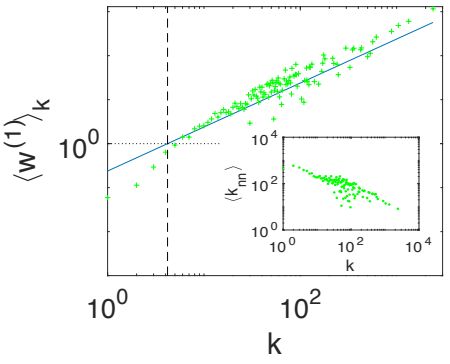
Table 7.1: Pearson correlation coefficient r, exponent ν and slope variation $\mathcal{S}$ for different synthetic and real-world networks with N nodes and average degree $\langle k \rangle$. The highlighted rows correspond to the three artificial networks analysed in Fig. 7.6(a).



(a)



(b)



(c)

Figure 7.6: The quantity $\langle w^{(1)} \rangle_k$ is reported as a function of k for (a) three synthetic graphs with the same number of nodes and links, and respectively disassortative ($r = -0.94$, red pluses), assortative ($r = 0.93$, yellow pluses) and uncorrelated ($r \sim 0$, purple pluses). The vertical dashed line identifies the mean degree $\langle k \rangle$. The mean degree $\langle k_{nn} \rangle_k$ of the nearest neighbours of nodes of degree k is displayed in the upper-left inset, while the degree distribution p_k is shown in the lower-right inset. Same quantities as in panel (a) for two real networks: (b) the network of collaborations in astrophysics and (c) Internet AS. Double-logarithmic scales have been used.

Conclusions

Complex webs of interactions between individual constituents are everywhere, from the Internet to epidemic spreading, via the molecular processes occurring inside a cell. Complex networks prove indeed highly versatile in describing numerous processes and they represent a powerful tool to quantitatively address the study of a wide gallery of complicated phenomena. Nevertheless, network science is a relatively new discipline, having reached its maximum development in the last twenty years. The interplay between couplings and the microscopic dynamics taking place on the various nodes, and how this gives rise to macroscopic complex behaviors, represents a fascinating and relatively new field of investigation.

In this thesis different scenarios have been explored, from agents mobility to ecological and reaction-diffusion systems, each chapter reporting on a different model. Methods have been proposed to unveil the structure of the network representing the underlying support of the considered dynamical system, since the topological details are not easy to identify in large graphs, even if they configure as fundamental ingredients for the observed behavior. In other chapters we have, instead, successfully made global or local adjustments to the structure of inter-nodes couplings with the ambitious and challenging goal of affecting the dynamical equilibria eventually reached by the system, in order to obtain a priori chosen asymptotic states. Modeling the form of the considered systems, i.e. the physical support represented by network structures, yields strategies to shape the ensuing dynamics, thus resulting in novel control methods.

Let us summarize the obtained results and outline possible future developments.
In Chapters 2 and 3 we analyzed the process of pattern formation in

reaction-diffusion systems. We have in particular shown how the spectral properties of the Laplacian operator affect the dynamical response of the network. Hence, by properly playing with eigenvalues and eigenvectors we succeeded in obtaining two important results. The first one involves network generation. It is indeed possible, in some cases, to identify a subnetwork or a set of nodes that can be acknowledged as practically irrelevant for pattern formation. It is then clear that distinct networks with differences solely localized on these particular substructures can, under apt conditions, generate the same irregular patterns. As a remarkable example of this situation, we have shown that scale-free networks, often taken as reference model to describe many real systems, can be, for what concerns reaction-diffusion systems, replaced by graphs characterized by highly different structures. This result provides a novel scheme for network generation, with a purely functional perspective: instead of building the specific network structure brick by brick, paying attention to local characteristics of nodes, the complex graph is globally established and constrained to yield an a priori chosen dynamical output. The most immediate application and possible future development of this method concerns neuronal systems; one of the main open questions in neuroscience today indeed involves structural neuronal network reconstruction starting from the observable response to specific stimuli.

The second important result obtained from the spectral analysis of reaction-diffusion systems deals instead with the possibility of dampening the irregular behavior of the reaction-diffusive system, thus allowing it to reach a synchronized stable equilibrium, even in regions of parameters where normally the synchronized solution would result unstable. This goal is achieved by modifying the Laplacian spectrum so as to drive the complete set of eigenmodes in the manifold identified as irrelevant for pattern formation. This work opens up a new perspective in developing methods to reach consensus. This latter is a useful concept in multi-agents systems when the ultimate goal is to enforce an agreement among microscopical entities on certain quantities of interest [125], like in smart power grids, swarming control, and group decision making.

A similar control technique targeted at stabilizing a predetermined equilibrium has been applied to ecosystems in Chapter 4, where the number of nodes reflects the biodiversity of the scrutinized sample [47, 81, 133] and each population is customarily identified in terms of its continuous

density. This latter evolves in time, as dictated by specific self-reaction stimuli, that generally bear nonlinear contributions. We choose for the sake of simplicity to operate under the deterministic approximation, by deliberately neglecting the role played by finite size fluctuations and disregarding spatial variability. In a complex and dynamical environment, species experience a large plethora of mutual interactions, notably pairwise exchanges. Cooperative and competitive interferences are simultaneously at play, and shape the ultimate fate of the system as a whole. The analysis carried out within this operating framework is principally useful in order to identify the mathematical constraints that have to be theoretically satisfied by a stable ecosystem. The model suggests that predator-prey interactions exert a stabilizing effects, in qualitative agreement with the conclusion reached in [6]. Furthermore, it has been found that a preponderance of weak interactions is beneficial to stability, as suggested by [47].

The above strategy is however of difficult application, qualifying as a control method that globally involves all the species at play. This is particularly true when dealing with natural ecosystems. For this reason, another chapter of the thesis is devoted to the fascinating subject of ecosystems stability, which many scientists have addressed, from Robert May [133] on. In Chapter 5 a different control method is devised, that consists in a local change of the interaction network. This ultimately implies adding one species (which configures as a further node of the network) to the pool of interacting families. The additional agent allows for customizing the eventual equilibrium attained by the system, by means of a properly engineered interaction either with the complete set or just with some of the preexisting species. As a future development it would be interesting to explore how to apply this technique to different systems, such as power grids, where it is important to identify strategic nodes of the network to act on, with a local control (and therefore at low cost) in such a way to stabilize a desired equilibrium.

Searching for strategic nodes within a large sample is crucial for many applications as e.g. mobility systems, on a physical space like a urban context, or on a virtual domain like the World Wide Web. In Chapter 6, for instance, we analyzed the case where a set of agents randomly explore a graph, chasing for specific target sites. These latter configure as absorbing traps which capture the walkers. This setting becomes interesting when

multiple sinks are simultaneously present on a given directed network and, by interfering with each other, they can mutually screen the flux of incoming agents. We have proposed an optimization strategy that results in the most advantageous positioning of an ensemble made of competing traps, so as to minimize their reciprocal hindrance and favor cooperation. It has also been considered that it is sometimes strategically advantageous to maximize the mutual competition between traps, so that the lastly added sink prevails over previously existing ones. The obtained results can contribute to explain a large plethora of natural phenomena, as shaped by the evolutionary pressure and devise novel efficient man-made solutions to specific technological problems. In the context of urban mobility, for instance, often roads can be trodden in one direction only, and this implies dealing with asymmetric edges between nodes. Agents can be assumed to stochastically diffuse across the embedding graph, even if they are individually heading towards specific targets nodes. For instance, parking spots, as well as shops selling a desired item, can act as veritable absorbing traps. When reaching the sought area, agents abandon the system and no longer belong to the population of hopping walkers.
On a more refined level of abstraction, traps could be made partially absorbing, by enhancing at will their relative degree of attractiveness. Playing with this additional variable amounts to studying a more complex mathematical problem. Interestingly, the hierarchy of traps as dictated by mere topological considerations can be completely subverted, an observation which opens up interesting perspectives for cheap decision making strategies built on existing infrastructures. Moreover, the analysis could also be extended to the relevant setting where finite volume effects limiting the capacity of each site, as e.g. competition for space in crowded traffic condition, can be accounted for.

Finally, the versatility of random walk processes has also been assessed in Chapter 7, where we have introduced a class of random walkers subject to node dependent reaction terms. Our model of reactive random walk is formulated in such a way that the relative contribution of the interaction term at the nodes and of the relocation term can be tuned at will, and this improves the sensitivity of the walkers to the structure of the network. In particular, the occupation probability of a given node is shaped by the non trivial interplay between the connectivity patterns and the local interaction functions. This was shown by determining analytically the

asymptotic occupation probability via a perturbative approach that takes a purely reactive dynamics as reference point. Exploiting the dependence of the occupation probabilities on the two tuning parameters of the model, namely the mobility parameter μ and the bias parameter α, and on the shape of the local reaction functions, we have shown that reactive random walkers can turn useful in many different ways. We have first discussed how, by properly adjusting the reaction contribution, one can emphasize nodes which bridge otherwise disconnected parts of the network, so that reactive random walkers can readily lead to generalized definitions of node centrality measures. Furthermore, with the help of general arguments and of a series of worked examples we have shown that, by making the random walkers reactive and inspecting their associated density distribution, one can easily detect the symmetries of a network. Finally, the specific form of the perturbative solution has inspired the introduction of a novel indicator for the presence, sign and entity of degree-degree correlations, which, differently from other standard measures, is based on harmonic averages. We have illustrated how reactive random walkers can distinguish assortative from disassortative networks. The approach can be in principle generalized so as to include next-to-leading correlations, and this defines an intriguing avenue for the investigation of higher-correlation in complex networks, which is left for future work.

As a final comment we emphasize that most of the procedures here outlined ultimately require determining the eigenvectors of a Laplacian matrix, with an associated algorithmic cost that scales as N^3. For large networks, this is a computationally demanding task and other suboptimal procedures might be devised, which would return approximate solutions to the examined problem. Relevant dynamical information stored in the stationary distribution can be for instance mathematically accessed by computing a limited set of eigenvectors, so resulting in principle in a significant reduction of computationally complexity. Alternatively, one could build simplified representations of the original network which exploit, where possible, its modular, community-like, structure. In doing so one could eventually define a large-scale backbone of the network for the optimization protocol to be implemented. The degree of imposed coarse graining, could reflect the available computational resources. Future investigations will be targeted to shed light onto these possibilities.

Appendix A

Spectral stabilization: supplementary details

A.1 Eigenmode randomization for directed networks

The procedure outlined in chapter 2 requires a special attention when applied to a system where the spatial support is a directed network. The main difference lies in the asymmetry of the Laplacian matrix, so that the eigenvalues and eigenvectors are, in general, complex. In this case, the unstable eigenmodes can be identified by observing how the eigenvalues are placed in the complex plane. An instability region, corresponding to the positive part of the dispersion relation, can be delineated following the procedure explained in the main text, as represented in Fig. A.1. The eigenvalues appearing in this region correspond to eigenmodes that compose the unstable manifold, which should be preserved to obtain similar patterns in a secondly generated network. All the other eigenvalues and eigenvectors can be modified so as to obtain a new Laplacian matrix. Let us recall that the eigenvalues $\Lambda^{(\alpha)}$, due to the fact that the original Laplacian matrix is real, either are real or are complex and come in conjugate pairs. The same happens for the components of the eigenvectors $\boldsymbol{\phi}^{(\alpha)}$. These characteristics have to be preserved during the randomization in order to obtain a meaningful spectrum for the new Laplacian.

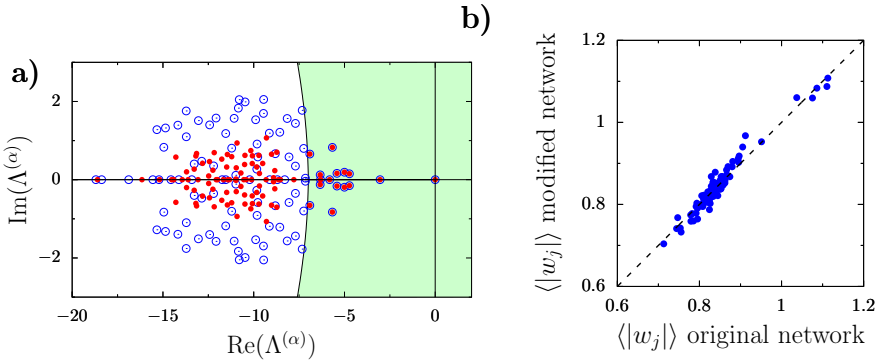


Figure A.1: (a): Original (red dots) and modified (blue circles) eigenvalues in the complex plane, the coloured area corresponds to the instability region. (b): patterns correlation, corresponding to $R^2 = 0.95$.

Let us first consider the case where the eigenvectors are not modified during the randomization procedure, i.e. the rotation matrix $\mathbf{R} = \mathbb{1}$. Following the guidelines provided in [39], we prove hereunder that the Laplacian entries are kept real also when dealing with a directed network. Let us bring into evidence the real and imaginary parts of every element of (2.2.1). It is immediate to see that the imaginary part of $L^{D'}_{il}$ reads:

$$\begin{aligned}
&(L^{D'}{}_{Im})_{il} = \\
&= \sum_j (D'_{Im})_{jj}[(\Phi_{Re})_{ij}(\Phi^{-1}_{Re})_{jl} - (\Phi_{Im})_{ij}(\Phi^{-1}_{Im})_{jl}] + \\
&+ \sum_j (D'_{Re})_{jj}[(\Phi_{Re})_{ij}(\Phi^{-1}_{Im})_{jl} + (\Phi_{Im})_{ij}(\Phi^{-1}_{Re})_{jl}].
\end{aligned} \tag{A.1}$$

To match condition $(L^{D'}{}_{Im})_{il} = 0$, both terms on the right hand side of (B.11) should be zero. To prove this fact, let us begin by recalling that the eigenvalues of a real asymmetric matrix either are real or are complex and come in conjugate pairs. Consider first the latter case and label with α and β the generic pair of conjugate eigenvalues. By definition $\mathbf{L}^D\boldsymbol{\phi}^{(\alpha)} = \Lambda^{(\alpha)}\boldsymbol{\phi}^{(\alpha)}$. Taking the complex conjugate yields $\mathbf{L}^D(\boldsymbol{\phi}^{(\alpha)})^* = (\Lambda^{(\alpha)})^*(\boldsymbol{\phi}^{(\alpha)})^*$ where $(\cdot)^*$ stands for the complex conjugate and where use has been made of the condition $\mathbf{L}^D = \mathbf{L}^{\mathbf{D}^*}$. Recalling that $(\Lambda^{(\alpha)})^* = \Lambda^{(\beta)}$ we can immediately conclude that $(\boldsymbol{\phi}^{(\alpha)})^*$ is an eigenvector of $\mathbf{L}^D$ relative

to the eigenvalue $\Lambda^{(\beta)}$ and thus $\boldsymbol{\phi}^{(\beta)} = (\boldsymbol{\phi}^{(\alpha)})^*$. Hence

$$\begin{aligned}(\Phi_{Re})_{i\alpha} &= (\Phi_{Re})_{i\beta} \\ (\Phi_{Im})_{i\alpha} &= -(\Phi_{Im})_{i\beta}\end{aligned} \tag{A.2}$$

for every i and (α, β). Consider now the equation

$$(\boldsymbol{\phi}^{-1})^{(\alpha)}\mathbf{L}^D = \Lambda^{(\alpha)}(\boldsymbol{\phi}^{-1})^{(\alpha)} \tag{A.3}$$

with $(\boldsymbol{\phi}^{-1})^{(\alpha)}$ α-th row of $\boldsymbol{\Phi}^{-1}$. Proceeding in analogy with the above, one gets:

$$\begin{aligned}(\Phi^{-1}_{Re})_{\alpha l} &= (\Phi^{-1}_{Re})_{\beta l} \\ (\Phi^{-1}_{Im})_{\alpha l} &= -(\Phi^{-1}_{Im})_{\beta l}.\end{aligned} \tag{A.4}$$

Let us go back to (B.11). Performing the summation on $j = \alpha$ and $j = \beta$, using (B.12) and (B.14) and the fact that the corrections $D'_{\alpha\alpha}$ and $D'_{\beta\beta}$ are complex conjugated as the original eigenvalues $\Lambda^{(\alpha)}$ and $\Lambda^{(\beta)}$ are, we finally conclude that the terms of the sums in (B.11) cancel in pairs. Consider now the case of a real eigenvalue $\Lambda^{(k)}$. Hence, by definition, $(D'_{Im})_{kk} = 0$, since, in this case, the stabilization can be solely achieved by acting on the real part of the eigenvalue (see next Section). To prove that $(L^D_{Im})_{il} = 0$ we need therefore to focus on the second term of (B.11), with $j = k$. Without loss of generality (up to a constant scaling factor) $\boldsymbol{\phi}^{(k)}_{Im} = 0$: the eigenvector associated to $\Lambda^{(k)}$ is hence real. The ik entries of matrix $\boldsymbol{\Phi}$ are indeed the elements of $\boldsymbol{\phi}^{(k)}$ and, for this reason, $(\Phi_{Im})_{ik} = 0$ $\forall i$. To conclude the reasoning and eventually prove that $(L^{D\ \prime}_{Im})_{il} = 0$, one needs to show that $(\Phi_{Re})_{ik}(\Phi^{-1}_{Im})_{kl} = 0$. This is in fact the case: $(\boldsymbol{\phi}^{-1})^{(k)}$ is the left eigenvector of matrix $\mathbf{L^D}$, relative to the real eigenvalue $\Lambda^{(k)}$. Reasoning as above, one can take $(\boldsymbol{\phi}^{-1})^{(k)}$ to be real and thus $(\Phi^{-1}_{Im})_{kl} = 0$. Then, summing up, $(L^{D\ \prime}_{Im})_{il} = 0$ $\forall i, l$.

Let us consider now the case where the eigenvectors are modified by using a rotation matrix. The new eigenvectors $\tilde{\boldsymbol{\phi}}^{\alpha}$ will not, in general, preserve the complex conjugate couples, which happen to be randomly rotated in the $(N - n)$-dimensional space. This ultimately implies that the imaginary parts do not mutually cancel out and that the entries of a Laplacian obtained from

$$\tilde{\boldsymbol{\Delta}} = \tilde{\boldsymbol{\Phi}}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\Phi}}^{-1}$$

would be complex. Let us however observe that cutting away the imaginary part of $\tilde{\boldsymbol{\Delta}}$, we obtain a third version of the Laplacian, $\hat{\boldsymbol{\Delta}} = \mathrm{Re}(\tilde{\boldsymbol{\Delta}})$,

whose eigenmodes can still serve to our aim. In fact, the new eigenvectors and eigenvalues either are real or complex conjugate and in particular, the n eigenmodes of the original Laplacian to be left invariant are still preserved. This can be proven by observing that the Laplacian $\tilde{\mathbf{\Delta}}$, even if it is complex, is built so as to correctly maintain the unstable manifold, and its imaginary part (which is eliminated in defining $\hat{\mathbf{\Delta}}$) only involves the stable eigenmodes. Therefore, cutting out the imaginary part of $\hat{\mathbf{\Delta}}$ translates into a second randomization of the stable eigenmodes, that does not affect the unstable manifold where the pattern information is eventually stored. In order to prove this statement we should consider the imaginary part of a generic (i,l) Laplacian entry:

$$\begin{aligned}(\tilde{\Delta}_{\mathrm{Im}})_{il} = &\sum_{\alpha=1}^{N} \tilde{\Lambda}^{\alpha}_{\mathrm{Im}}[(\tilde{\Phi}_{\mathrm{Re}})_{i\alpha}(\tilde{\Phi}^{-1}_{\mathrm{Re}})_{\alpha l} - (\tilde{\Phi}_{\mathrm{Im}})_{i\alpha}(\tilde{\Phi}^{-1}_{\mathrm{Im}})_{\alpha l}]+ \\ &+ \tilde{\Lambda}^{\alpha}_{\mathrm{Re}}[(\tilde{\Phi}_{\mathrm{Re}})_{i\alpha}(\tilde{\Phi}^{-1}_{\mathrm{Im}})_{\alpha l} + (\tilde{\Phi}_{\mathrm{Im}})_{i\alpha}(\tilde{\Phi}^{-1}_{\mathrm{Re}})_{\alpha l}]\end{aligned} \tag{A.5}$$

where $\tilde{\mathbf{\Phi}}_{\mathrm{Re}}$ and $\tilde{\mathbf{\Phi}}_{\mathrm{Im}}$ are the real and the imaginary part of the eigenvectors. The sum over α can be separated into two sums: one over $\alpha = 1, ..., n$, the other over $\alpha' = n+1, ..., N$, the first containing all the unmodified eigenmodes, the second the randomized ones. We shall prove that the first sum is equal to zero. Let us begin by observing that for index α up to n, the eigenvalues $\tilde{\Lambda}$ and eigenvectors $\tilde{\phi}$ appearing in the sum can be replaced by the original Λ and ϕ. Considering a conjugate couple labeled with β and γ, we have $\Lambda^{(\gamma)} = (\Lambda^{(\beta)})^*$, with

$$\begin{aligned}(\phi_{\mathrm{Re}})_{i\gamma} &= (\phi_{\mathrm{Re}})_{i\beta} \\ (\phi_{\mathrm{Im}})_{i\gamma} &= -(\phi_{\mathrm{Im}})_{i\beta}\end{aligned} \tag{A.6}$$

and

$$\begin{aligned}(\phi^{-1}_{\mathrm{Re}})_{\gamma l} &= (\phi^{-1}_{\mathrm{Re}})_{\beta l} \\ (\phi^{-1}_{\mathrm{Im}})_{\gamma l} &= -(\phi^{-1}_{\mathrm{Im}})_{\beta l}.\end{aligned} \tag{A.7}$$

This implies the cancellation in pairs of the terms corresponding to complex conjugate couples. For what concerns the remaining terms, corresponding to real eigenvalues and eigenvectors, it is clear that all the terms in (B.11) are automatically zero (remember that $\boldsymbol{\phi}^{-1}_{jl} = \boldsymbol{\phi}_{lj}$). This proves the claim.

The proofs (ii) and (iii) of chapter 2, respectively corresponding to the zero-stochasticity and symmetry preservation of the Laplacian are trivially

extended to the directed case without further remarks. We here report however an important complement concerning the balanced networks. A balanced network is defined by equal incoming and outgoing connectivity, $k_i^{out} = k_i^{in} \equiv k_i$, characteristic which is preserved during the randomization process, as it is hereunder proven:

If $\mathbf{L}^D$ is balanced, then also $\mathbf{L}_c^D$ is. The balanced Laplacian is characterized by zero-row-sum [1]. It is then sufficient to prove $\sum_l {L^D}'_{il} = 0$. First of all, recall that the columns of matrix $\mathbf{\Phi}$ are the right eigenvectors of $\mathbf{L}^D$, while the rows of $\mathbf{\Phi}^{-1}$ are the left eigenvectors, namely:

$$\begin{aligned} \mathbf{L}^D\mathbf{\Phi} &= \mathbf{\Phi}\mathbf{D} \\ \mathbf{\Phi}^{-1}\mathbf{L}^D &= \mathbf{D}\mathbf{\Phi}^{-1} \end{aligned} \tag{A.8}$$

By definition of Laplacian, the uniform vector $\mathbf{1}$ is the left eigenvector of $\mathbf{L}^D$ corresponding to $\Lambda^{(1)} = 0$:

$$\mathbf{1}^T\mathbf{L}^D = 0 \quad \Rightarrow \quad \sum_i L_{ij}^D = 0. \tag{A.9}$$

If $\mathbf{L}^D$ is balanced then also the right eigenvector corresponding to $\Lambda^{(1)} = 0$ is equal to $\mathbf{1}$, hence:

$$\mathbf{L}^D\mathbf{1} = 0 \quad \Rightarrow \quad \sum_j L_{ij}^D = 0. \tag{A.10}$$

By controlling the network of connections we modify the eigenvalues of the Laplacian operator, while preserving the uniform right and left eigenvector corresponding to 0, thus:

$$\begin{aligned} \mathbf{L^D}_c\mathbf{1} = 0 \quad &\Rightarrow \quad \textstyle\sum_j ({L^D}')_{ij} = 0 \\ \mathbf{1}^T\mathbf{L^{D'}} = 0 \quad &\Rightarrow \quad \textstyle\sum_i ({L^D}')_{ij} = 0 \end{aligned} \tag{A.11}$$

which proves the claim.

In Fig. A.1(a) the eigenvalues of an original random network with $N = 100$ and $\langle k \rangle = 10.2$ are displayed in the complex plane together with those of the modified spectrum. The instability region has been drawn by again using the CGL equation as self-dynamics. Fig. A.1(b) shows the correlation between time-average modulus patterns obtained from the original and the network generated from the Laplacian $\hat{\mathbf{\Delta}}$.

[1] The network is balanced if and only if $\sum_j L_{ij}^D = 0$. Indeed we have: $\sum_j L_{ij}^D = \sum_j A_{ij} - k_i^{out} = k_i^{in} - k_i^{out}$. So $\sum_j L_{ij}^D = 0$ if and only if $k_i^{in} = k_i^{out}$.

A.2 Local rewiring: acceptance threshold

The local rewiring method for isodynamic network generation is based on an iterative scheme. Each move consists of removing or adding a link between two random nodes and is accepted only if the unstable manifold is preserved in the new network. The degree of similarity between new and old eigenmodes is established by means of the error function:

$$E = NE_l + E_q;$$

where

$$E_l = \frac{1}{n}\sum_{\alpha=1}^{n} |\tilde{\Lambda}^{(\alpha)} - \Lambda^{(\alpha)}|^2$$

and

$$E_q = \frac{1}{n}\sum_{\alpha=1}^{n} |\langle \tilde{\boldsymbol{\phi}}^{(\alpha)}, \boldsymbol{\phi}^{(\alpha)} \rangle - 1|^2.$$

The factor N is used to ensure that both terms in the expression of E are of the same order. The change is accepted only if the error function E is smaller than a chosen threshold parameter τ.

Appendix B

Mathematical details on ecosystem stability

B.1 Couplings in the original system

We shall here briefly discuss the algorithm that we have employed to generate the initial matrix of interactions $\boldsymbol{A}$. For each pair i, j (with i and j running from 1 to N) we draw a random number p from a uniform distribution in the interval $[0, 1]$. If $p > p_0$ a link that goes from j to i is established. Here $0 < p_0 < 1$ is a free parameter. The strength of the link, namely the element A_{ij} of matrix $\boldsymbol{A}$, is assigned as follows: we extract a random number from a uniform distribution defined in the compact interval $[-c_0, 1-c_0]$, with $c_0 > 0$, and multiply the selected number by a scalar amplitude factor a. It is immediate to prove, that increasing c_0 makes the system progressively more unstable. This follows a direct application of the Gershgorin theorem, mentioned in chapter 4. The relative abundance of the pairs can be on average computed and shown to be related to the choices of p_0 and c_0. The percentage of null entries (no links) of $\mathbf{A}$ will be $1 - p_0$, the percentage of negative entries $p_0 c_0$, while the percentage of positive entries is $p_0(1 - c_0)$. Building on this observation, we obtain the following estimate for $f_{(\cdot,\cdot)}$, the frequencies of occurrence of the different

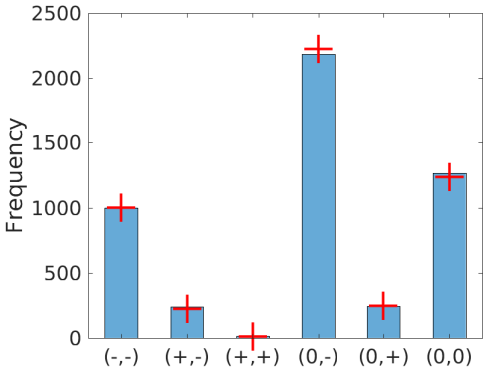


Figure B.1: The amount of pairs of the different kinds are shown. Red crosses are analytically computed after equations (B.2). Blue bars are numerically obtained following one individual realization of the scheme discussed above for generating $\boldsymbol{A}$. Here, $N = 100$, $a = 0.38$, $p_0 = 0.5$, $c_0 = 0.9$. The agreement between the analytic calculation and the numerical implementation is already satisfying (for just one realization) for $N = 100$. This observation justifies assuming the analytic abundances as a reference stand for the controlled matrices of interactions (of the same dimension $N = 100$) to be compared with.

classes:

$$\begin{aligned} f_{(0,0)} &= (1-p_0)^2 \\ f_{(0,-)} &= 2p_0(1-p_0)c_0 \\ f_{(0,+)} &= 2p_0(1-p_0)(1-c_0) \\ f_{(+,-)} &= 2p_0^2c_0(1-c_0) \\ f_{(+,+)} &= p_0^2(1-c_0)^2 \\ f_{(-,-)} &= p_0^2c_0^2. \end{aligned} \tag{B.1}$$

Each quantity is to be multiplied for a factor $(N^2 - N)/2$ to obtain the abundances of elements, which are numerically tested in Figure B.1. The red crosses depicted in the Figures enclosed in chapter 4 refer to (B.2).

B.2 Extending the analysis to account for a generalized nonlinear reaction term

The control scheme developed in this paper can also be applied to a more general reaction model. Assume that the logistic dynamics is replaced by a generic nonlinear function $f(x_i, \boldsymbol{r}_i)$:

$$\dot{x_i} = f(x_i, \boldsymbol{r}_i) + x_i \sum_{j \neq i} A_{ij} x_j \tag{B.2}$$

where $\boldsymbol{r}_i$ identifies an arbitrary set of constant parameters.
The associated fixed point is obtained by solving:

$$f(x_i, \boldsymbol{r}_i) + x_i \sum_{j \neq i} A_{ij} x_j = 0. \tag{B.3}$$

By performing the change of variables $y_i \equiv x_i / x_i^*$ and $B_{ij} = A_{ij} x_j^*$ we obtain:

$$\dot{y_i} = \frac{1}{x_i^*} f(y_i x_i^*, \boldsymbol{r}_i) + y_i \sum_{j \neq i} B_{ij} y_j \tag{B.4}$$

In the new variables, the fixed point equation takes the form:

$$\frac{1}{x_i^*} f(x_i^*, \boldsymbol{r}_i) + \sum_{j \neq i} B_{ij} = 0 \tag{B.5}$$

where use has been made of the condition $y_i^* = 1$.
The local stability analysis requires introducing a modest perturbation v_i around the fixed point $y_i^* = 1$, which amounts to writing $y_i = 1 + v_i$. The inhomogeneous perturbation v_i will evolve according to:

$$\dot{v_i} = v_i \left[\frac{1}{x_i^*} \frac{\partial f}{\partial y_i} (y_i x_i^*, \boldsymbol{r}_i) \bigg|_{y_i = 1} + \sum_{j \neq i} B_{ij} \right] + \sum_{j \neq i} B_{ij} v_j \equiv \sum_j C_{ij} v_j. \tag{B.6}$$

where the matrix $\boldsymbol{C}$ is defined as

$$C_{ij} \equiv \begin{cases} B_{ij} & \text{if } i \neq j \\ \sum_{k \neq i} B_{ik} + \frac{1}{x_i^*} \frac{\partial f}{\partial y_i} (y_i x_i^*, \boldsymbol{r}_i) \Big|_{y_i = 1} & \text{if } i = j, \end{cases} \tag{B.7}$$

and the stability of the fixed point is ultimately controlled by the sign of the real part of the eigenvalues of $\boldsymbol{C}$. In analogy with the procedure

described in chapter 4, we then define $\boldsymbol{D}$ as

$$D_{ij} \equiv \begin{cases} B_{ij} & \text{if } i \neq j \\ \frac{1}{x_i^*} f(x_i^*, \boldsymbol{r}_i) & \text{if } i = j, \end{cases} \tag{B.8}$$

which is a zero-row-sum matrix, because of the fixed point condition, $\sum_j D_{ij} = \sum_{j \neq i} B_{ij} + \frac{1}{x_i^*} f(x_i^*, \boldsymbol{r}_i) = 0$.
We can then modify the matrix $\boldsymbol{D}$, which hence transforms into $\boldsymbol{D'}$, so as to enforce the desired stability, following the recipe outlined in chapter 4 . The implemented changes can be interpreted as follows:

- for the off-diagonal entries ($i \neq j$) we impose $D'_{ij} = B'_{ij} = A'_{ij} x_j^{*\prime}$ which enables to calculate the elements of the controlled interaction matrix. As usual, we can decide to freeze the fixed point to its original value or modify it consistently, while assuming constant the parameters of the model.
- for the diagonal entries ($i = j$) we impose $D'_{ii} = \frac{1}{x_i^{*\prime}} f(x_i^{*\prime}, \boldsymbol{r}'_i)$ where $\boldsymbol{r}'_i$ is the new vector of parameters of the stabilized system. This latter can be readily obtained by inverting the above equation (f needs therefore to be invertible, with respect to $\boldsymbol{r}$). Notice that, in general, only a subset of the elements of $\boldsymbol{r}_i$, need to be adjusted. As remarked above, it is alternatively possible to leave the parameters $\boldsymbol{r}_i$ unchanged, and modify $\boldsymbol{x^*}$

B.3 The controlled matrix $\boldsymbol{D'}$ is real and zero-row-sum as $\boldsymbol{D}$ is

Recall the definition of $\boldsymbol{D'}$:

$$\boldsymbol{D'} = \boldsymbol{D} + \boldsymbol{\Phi}(\boldsymbol{\delta\Lambda})\boldsymbol{\Phi}^{-1} \equiv \boldsymbol{D} + \boldsymbol{\delta D} \tag{B.9}$$

where $\boldsymbol{\Phi}$ is the matrix whose columns are the eigenvectors $(\boldsymbol{\phi^{(1)}}, ..., \boldsymbol{\phi^{(N)}})$ of $\boldsymbol{D}$. Observe that since $\boldsymbol{D}$ is real and zero-row-sum, it is sufficient for our purposes to prove that $\boldsymbol{\delta D}$ exhibits the same properties.

The elements of $\boldsymbol{D'}$ are real. Consider the generic entries of $\boldsymbol{\delta D}$:

$$(\delta D)_{il} = \sum_j \Phi_{ij} \delta\Lambda^{(j)} (\Phi^{-1})_{jl}. \tag{B.10}$$

Isolate the real and imaginary parts of every element in (B.10). It is immediate to see that, being $\delta\Lambda^{(j)}$ real $\forall j$, the imaginary part of $(\delta D)_{il}$ reads:

$$(\delta D_{Im})_{il} = \sum_j \delta\Lambda^{(j)}_{Re}[(\Phi_{Re})_{ij}(\Phi^{-1}_{Im})_{jl} + (\Phi_{Im})_{ij}(\Phi^{-1}_{Re})_{jl}]. \tag{B.11}$$

To match condition $(\delta D_{Im})_{il} = 0$, the term on the right hand side of (B.11) should be zero. To prove this, let us first recall that the eigenvalues of a real matrix are either real or complex and come in conjugate pairs. Consider a complex eigenvalue $\Lambda^{(\alpha)}$, by definition $\boldsymbol{D}\boldsymbol{\phi}^{(\alpha)} = \Lambda^{(\alpha)}\boldsymbol{\phi}^{(\alpha)}$. Taking the complex conjugate yields $\boldsymbol{D}(\boldsymbol{\phi}^{(\alpha)})^* = (\Lambda^{(\alpha)})^*(\boldsymbol{\phi}^{(\alpha)})^*$ where $(\cdot)^*$ stands for the complex conjugate and where we have used the condition $\boldsymbol{D} = \boldsymbol{D}^*$. Let the indexes α and β to be defined by the relation $(\Lambda^{(\alpha)})^* = \Lambda^{(\beta)}$. We can immediately conclude that $\boldsymbol{\phi}^{(\beta)} = (\boldsymbol{\phi}^{(\alpha)})^*$. Hence

$$\begin{aligned}(\Phi_{Re})_{i\alpha} &= (\Phi_{Re})_{i\beta} \\ (\Phi_{Im})_{i\alpha} &= -(\Phi_{Im})_{i\beta}\end{aligned} \tag{B.12}$$

for each node index i and every pair (α, β) of complex conjugate eigenvalues. Consider now the equation

$$(\boldsymbol{\phi}^{-1})^{(\alpha)}\boldsymbol{D} = \Lambda^{(\alpha)}(\boldsymbol{\phi}^{-1})^{(\alpha)} \tag{B.13}$$

with $(\boldsymbol{\phi}^{-1})^{(\alpha)}$ left eigenvector of $\boldsymbol{D}$, corresponding to the α-th row of $\boldsymbol{\Phi}^{-1}$. Proceeding in analogy with the above, one eventually gets:

$$\begin{aligned}(\Phi^{-1}_{Re})_{\alpha l} &= (\Phi^{-1}_{Re})_{\beta l} \\ (\Phi^{-1}_{Im})_{\alpha l} &= -(\Phi^{-1}_{Im})_{\beta l}.\end{aligned} \tag{B.14}$$

Return now to (B.11). By performing the summation on $j = \alpha$ and $j = \beta$, using (B.12) and (B.14) and the fact that the corrections $\delta\Lambda^{(\alpha)}$ and $\delta\Lambda^{(\beta)}$ are complex conjugated like the original eigenvalues $\Lambda^{(\alpha)}$ and $\Lambda^{(\beta)}$, we finally conclude that the terms of the sums in (B.11) cancel in pairs relative to complex conjugate eigenvalues.

The remaining terms in the summation correspond to real eigenvalues. Without loss of generality, we can always consider the relative left and right eigenvectors to be real, up to a constant scaling factor. In detail, if $\Lambda^{(\gamma)} \in \mathbb{R}$, $(\Phi_{Im})_{i\gamma}$ and $(\Phi^{-1}_{Im})_{\gamma l}$ can always be set equal to zero $\forall i, l$. This implies that all the remaining terms in equation (B.11) disappear, one by

one. Then, summing up, $(\delta D_{Im})_{il} = 0 \ \forall i, l$.

***D′* is a zero-row-sum matrix.** Matrix $\boldsymbol{D}$ zero-row-sum. Hence, the vector $\mathbf{1}$ (with all entries equal to one) is the right eigenvector of $\boldsymbol{D}$ corresponding to $\Lambda^{(1)} = 0$:

$$\sum_j D_{ij} = 0 \quad \Longleftrightarrow \quad \boldsymbol{D}\mathbf{1} = 0. \tag{B.15}$$

Recall that the proposed approach implies a modification of the eigenvalues of matrix $\boldsymbol{D}$, while keeping the eigenvectors unchanged. As a consequence, vector $\mathbf{1}$ is still solution of the eigenvalue problem. Moreover, the zero eigenvalue is not responsible for the instability and is therefore preserved upon application of the control scheme. Hence:

$$\boldsymbol{D'}\mathbf{1} = 0 \quad \Longleftrightarrow \quad \sum_j (D')_{ij} = 0 \tag{B.16}$$

which proves the claim.

B.4 On the conditions of controllability

As explained in chapter 4, the proposed control method is based on shifting the eigenvalues of a zero-row-sum matrix in the complex plane by applying to their values a real and negative correction (which is identically equal to zero, for the subset of eigenvalues which should be preserved). The goal of this Appendix is to single out the conditions which allow for the control procedure to be effectively implemented.

Denote by $\mathcal{M}$ the set of indices corresponding to the eigenvalues which are responsible for the instability and which should be modified by the controller. Denote by $\delta\Lambda_i$, $i \in \mathcal{M}$, the (real) shift imposed to the selected ensemble of eigenvalues for stabilization. In the sequel, we will assume that the translation takes the interested eigenvalues to a constant value, R, smaller than r_{min} (other strategies can clearly be adopted, consequently altering the analysis reported below):

$$(\Lambda_{Re})_i + \delta\Lambda_i = R < r_{min} \tag{B.17}$$

In choosing the constant R we must ensure that the applicability constraint (4.11) is satisfied:

$$\tilde{s}_i - \sum_j \Phi_{ij}\delta\Lambda_j\Phi_{ji}^{-1} \geq 0 \ \forall i \tag{B.18}$$

Imposing condition (B.17), one gets:

$$\tilde{s}_i - R\sum_{j\in\mathcal{M}}\Phi_{ij}\Phi_{ji}^{-1} + \sum_{j\in\mathcal{M}}\Phi_{ij}(\Lambda_{Re})_j\Phi_{ji}^{-1} \equiv k_i - R\sum_{j\in\mathcal{M}}\Phi_{ij}\Phi_{ji}^{-1} \geq 0 \ \forall i. \tag{B.19}$$

where we defined $k_i = \tilde{s}_i + \sum_{j\in\mathcal{M}}\Phi_{ij}(\Lambda_{Re})_j\Phi_{ji}^{-1}$.

If $k_i - r_{min}\sum_{j\in\mathcal{M}}\Phi_{ij}\Phi_{ji}^{-1}$ is positive, for each index $i \in [1, N]$, then it is sufficient to set $R = r_{min}$ for achieving stabilization. The complementary situation, in which the above quantity turns out to be negative for some i, is more intricate, as we shall clarify hereafter.

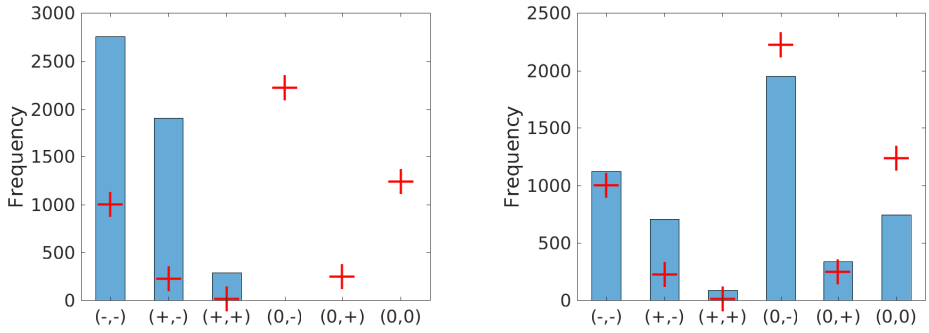
First, it is convenient to sort the node indices in such a way that for indices $i \in [1, n]$, with $n < N$, the quantity $\sum_{j\in\mathcal{M}}\Phi_{ij}\Phi_{ji}^{-1}$ is positive, while it takes negative values for the remaining indices of the collection, $i \in [n+1, N]$. Recalling that R is bound to be smaller than r_{min}, the following inequalities hold:

$$k_i - R\sum_{j\in\mathcal{M}}\Phi_{ij}\Phi_{ji}^{-1} \leq k_i - r_{min}\sum_{j\in\mathcal{M}}\Phi_{ij}\Phi_{ji}^{-1} \quad \forall i \in [n+1, N] \tag{B.20}$$

$$k_i - R\sum_{j\in\mathcal{M}}\Phi_{ij}\Phi_{ji}^{-1} \geq k_i - r_{min}\sum_{j\in\mathcal{M}}\Phi_{ij}\Phi_{ji}^{-1} \quad \forall i \in [1, n] \tag{B.21}$$

where R is chosen so as to make the term on the left hand side positive. The inequality (B.20) is obviously violated if the expression on the right hand side is negative for at least one i in the interval $[n+1, N]$. In this case, the system cannot be controlled, using the recipe here discussed (which amounts, among the other specificities, to select a constant R). For what concerns the other inequality (B.21), the righthand term is instead allowed to be negative. Suppose that this happens for indices $i \in [1, \tilde{n}]$ with $\tilde{n} \leq n$ (the indices are imagined to be properly sorted). Then, the constant R must be smaller than r_{min}, let us say $R = r_{min} - \epsilon$, with $\epsilon > 0$. Substituting it into (B.19) we obtain a lower bound for ϵ:

$$\epsilon \geq \max_{i\in[1,\tilde{n}]}\left(r_{min} - \frac{k_i}{\sum_{j\in\mathcal{M}}\Phi_{ij}\Phi_{ji}^{-1}}\right). \tag{B.22}$$



(a) (b)

Figure B.2: Abundances of different types of couplings between species, as explained in caption of Figure 4.5. Here the blue bars refer to the system controlled upon application of the strategy which leaves the fixed point to its original value, while the parameters $\boldsymbol{s}$ are modified. The Figure in panel (b) reports the coupling abundances after a cut-off of 0.05 has been applied.

Recall again that R is (our choice) constant. Assuming $R = r_{min} - \epsilon$, one could eventually loose the controllability condition for indices $i \in [n+1, N]$. The following additional condition needs therefore to be considered:

$$k_i - r_{min} \sum_{j \in \mathcal{M}} \Phi_{ij} \Phi_{ji}^{-1} + \epsilon_{min} \sum_{j \in \mathcal{M}} \Phi_{ij} \Phi_{ji}^{-1} \geq 0 \tag{B.23}$$

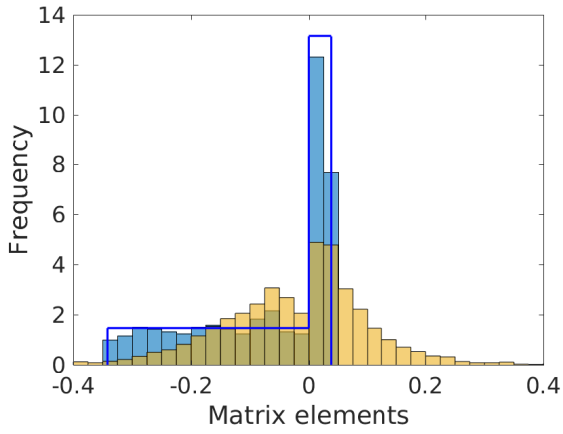
thus resulting in an upper bound for ϵ:

$$\epsilon \leq \min_{i \in [n+1, N]} \left(r_{min} - \frac{k_i}{\sum_{j \in \mathcal{M}} \Phi_{ij} \Phi_{ji}^{-1}} \right). \tag{B.24}$$

Another necessary condition for controllability is then found by imposing that the upper bound for ϵ is larger than the lower bound (both positive):

$$0 < \max_{i \in [1, \tilde{n}]} \left(r_{min} - \frac{k_i}{\sum_{j \in \mathcal{M}} \Phi_{ij} \Phi_{ji}^{-1}} \right) \leq \min_{i \in [n+1, N]} \left(r_{min} - \frac{k_i}{\sum_{j \in \mathcal{M}} \Phi_{ij} \Phi_{ji}^{-1}} \right) \tag{B.25}$$

which coincides with the equations reported in chapter 4.



(a)

Figure B.3: Distribution of predator-prey interactions: the blue bars and the blue line refer to one realization of the initial system, as explained in the caption of Figure 4.6(a). Yellow bars represent the different couplings in matrix $\boldsymbol{A'}$ as obtained from controlling the system leaving the fixed point unchanged. In this case, strong predator-prey links are generated and the distribution of positive vs. negative weights symmetrized

B.5 Controlling without modifying the fixed point

The Figures appearing in this Appendix illustrate the results obtained for the control strategy where the fixed point is kept unchanged while the inverse carrying capacity vector $\boldsymbol{s}$ is modified together with the weights of the interaction matrix. Figures B.3 and B.2 represent the analogue of Figures, respectively 4.5 and 4.6(a), already discussed in chapter 4, which are obtained with the alternative control strategy.

Appendix C

Control adding one node: inside the model and the controllability

C.1 Genetic network model

We shall here justify the model of genetic regulatory network analyzed in chapter 5. Consider first a small regulatory network consisting of one gene (whose activity is labelled x) and one protein (associated to the continuous concentration y). A positive regulation loop can be modeled as:

$$\begin{aligned} \dot{x} &= k_1 g(y) - \gamma_1 x \\ \dot{y} &= k_2 x - \gamma_2 y \end{aligned}$$

where:

$$g(y) = \frac{y^n}{K + y^n}. \tag{C.1}$$

In the following, we will make the choice $n = 2$ and $K = 1$. Similarly, a negative regulation loop can be modeled as:

$$\begin{aligned} \dot{x} &= k_1 (1 - g(y)) - \gamma_1 x \\ \dot{y} &= k_2 x - \gamma_2 y \end{aligned}$$

As it should be, the concentration of proteins grows with the level of gene activity. As a first step approximation, and to avoid dealing with two distinct families of mutually interlinked constituents (genes and proteins), we can replace y with x in the argument of $g(\cdot)$, via adiabatic elimination (apart from redefinition of the involved constants). Building on the above we model the extended regulatory network as:

$$\dot{x}_i = k_i \sum_j A_{ij} g(x_j) - \gamma_i x_i + k_i \eta_i \tag{C.2}$$

where x_i stands for the activity of gene i. For *positive* feedbacks between species i and j, $A_{ij} = 1$, while, for *negative* loops $A_{ij} = -1$. The parameter η_i stands for the number of negative loops (number of negative entries of the i-th row of $\boldsymbol{A}$) that are associated to node i. Finally, to keep the structure as simple as possible we set $k_i = 1$ for all i. The matrix $\boldsymbol{A}$ employed in the example reported in chapter 5 has a simple structure and it has been chosen for purely illustrative purposes: it represents a regular tree network with branching ratio $r = 4$.

C.2 On the controllability condition

Observe that the control procedure here discussed requires computing the vector $\boldsymbol{\beta}$. Determining this latter implies inverting a matrix, an operation that imposes a mathematical constraint that we shall hereafter analyze more in depth. The matrix to be inverted, as defined in chapter 5, reads:

$$H_{nm} = -\sum_{k=0}^{N-n} c_{k+n+1} (\boldsymbol{G}^k \boldsymbol{q})_m \tag{C.3}$$

where c stands for the coefficients[1] of the characteristic polynomial of $\boldsymbol{G}$, namely $\det(\boldsymbol{G}-\lambda\boldsymbol{I}) = \sum_{l=0}^{N} c_{l+1}\lambda^l$. The invertibility is ensured if $\det(\boldsymbol{H}) \neq 0$, in formulae:

$$\begin{aligned}\det(\boldsymbol{H}) &= \sum_{i_1=1}^{N}\cdots\sum_{i_N=1}^{N}\epsilon_{i_1,\ldots i_N}H_{1i_1}H_{2i_2}\ldots H_{N-1i_{N-1}}H_{Ni_N} = \\ &= \sum_{i_1=1}^{N}\cdots\sum_{i_N=1}^{N}\epsilon_{i_1,\ldots i_N}\left[-\sum_{k_1=0}^{N-1}c_{k_1+2}(\boldsymbol{G}^{k_1}\boldsymbol{q})_{i_1}\right]\left[-\sum_{k_2=0}^{N-2}c_{k_2+3}(\boldsymbol{G}^{k_2}\boldsymbol{q})_{i_2}\right] \\ &\ldots\left[-\sum_{k_{N-1}=0}^{1}c_{k_{N-1}+N}(\boldsymbol{G}^{k_{N-1}}\boldsymbol{q})_{i_{N-1}}\right][-c_{N+1}q_{i_N}] \neq 0\end{aligned} \tag{C.5}$$

where $\epsilon_{i_1,\ldots i_N}$ is the Levi-Civita tensor. This complicated expression can be heavily simplified. The Levi-Civita symbol is in fact totally antisymmetric in the permutation of its indices. As a consequence, all the terms multiplied by $\epsilon_{i_1,\ldots i_N}$ which are symmetric in the permutations, cancel out. It follows that the terms containing the product of two or more factors $(\boldsymbol{G}^k\boldsymbol{q})$ with the same power k are identically equal to zero. The only terms which survive are those obtained by just retaining the largest possible value of k in each summation ($k_1 = N-1$, $k_2 = N-2$,..., $k_{N-1} = 1$).

[1]

$$\begin{aligned}c_1 &= \det(\boldsymbol{G}) \\ c_k &= \frac{(-1)^{k-1}}{(k-1)!}\sum_{\substack{j_1\ldots j_{k-1}=1 \\ j_1\neq\ldots\neq j_{k-1}}}^{N}\det\boldsymbol{G}_{(j_1j_1)(j_2j_2)\ldots(j_{k-1}j_{k-1})} \text{ for } k=2,\ldots,N-1 \\ c_N &= (-1)^{N-1}Tr(\boldsymbol{G}) \\ c_{N+1} &= (-1)^N\end{aligned} \tag{C.4}$$

where with $\boldsymbol{G}_{(j_1j_1)(j_2j_2)\ldots(j_{k-1}j_{k-1})}$ we identify the minor obtained from matrix $\boldsymbol{G}$ by removing the j_1-th, j_2-th,...,j_{k-1}-th rows and columns.

In formulae:

$$\begin{aligned}\det(\boldsymbol{H}) = (-1)^N \sum_{i_1=1}^{N} \cdots \sum_{i_N=1}^{N} \epsilon_{i_1,\ldots i_N} c_{N+1} (\boldsymbol{G}^{N-1}\boldsymbol{q})_{i_1} c_{N+1} (\boldsymbol{G}^{N-2}\boldsymbol{q})_{i_2} \ldots \\ \ldots c_{1+N} (\boldsymbol{G}\boldsymbol{q})_{i_{N-1}} c_{N+1} q_{i_N} = \\ = (-1)^N (-1)^{N^2} \sum_{i_1=1}^{N} \cdots \sum_{i_N=1}^{N} \epsilon_{i_1,\ldots i_N} (\boldsymbol{G}^{N-1}\boldsymbol{q})_{i_1} (\boldsymbol{G}^{N-2}\boldsymbol{q})_{i_2} \ldots \\ \ldots (\boldsymbol{G}\boldsymbol{q})_{i_{N-1}} q_{i_N}.\end{aligned} \tag{C.6}$$

where use has been made of the fact that $c_{N+1} = (-1)^N$ (see equation (C.4) in the footnote 1).

Drawing on this preliminary observations it is possible to re-interpret the above controllability constraint, making contact with standard control theory. The controllability condition amounts to require that the matrix

$$\boldsymbol{C} \equiv [\boldsymbol{q}, \boldsymbol{G}\boldsymbol{q}, ..., \boldsymbol{G}^{N-1}\boldsymbol{q}] = \begin{pmatrix} q_1 & (\boldsymbol{G}\boldsymbol{q})_1 & (\boldsymbol{G}^2\boldsymbol{q})_1 & \ldots & (\boldsymbol{G}^{N-1}\boldsymbol{q})_1 \\ q_2 & (\boldsymbol{G}\boldsymbol{q})_2 & (\boldsymbol{G}^2\boldsymbol{q})_2 & \ldots & (\boldsymbol{G}^{N-1}\boldsymbol{q})_2 \\ \vdots & & & & \\ q_N & (\boldsymbol{G}\boldsymbol{q})_N & (\boldsymbol{G}^2\boldsymbol{q})_N & \ldots & (\boldsymbol{G}^{N-1}\boldsymbol{q})_N \end{pmatrix} \tag{C.7}$$

has maximum rank. Here, with the notation $(\boldsymbol{G}^l\boldsymbol{q})_k$ we identify the k-th entry of the vector obtained from the product of the matrix $\boldsymbol{G}$ to the power of l with vector $\boldsymbol{q}$. In system theory the matrix $\boldsymbol{C}$ is called the *controllability* matrix of the pair $(\boldsymbol{G}, \boldsymbol{q})$. Since $\boldsymbol{C}$ is a square matrix, the maximum rank condition is equivalent to require $\det(\boldsymbol{C}) \neq 0$, in formulae:

$$\begin{aligned}\det(\boldsymbol{C}) &= \sum_{\sigma_1=1}^{N} \cdots \sum_{\sigma_N=1}^{N} \epsilon_{\sigma_1,\ldots,\sigma_N} C_{\sigma_1 1} C_{\sigma_2 2} \ldots C_{\sigma_{N-1} N-1} C_{\sigma_N N} = \\ &= \sum_{\sigma_1=1}^{N} \cdots \sum_{\sigma_N=1}^{N} \epsilon_{\sigma_1,\ldots,\sigma_N} (\boldsymbol{G}^0 \boldsymbol{q})_{\sigma_1} (\boldsymbol{G}^1 \boldsymbol{q})_{\sigma_2} \ldots \\ &\quad \ldots (\boldsymbol{G}^{N-2} \boldsymbol{q})_{\sigma_{N-1}} (\boldsymbol{G}^{N-1} \boldsymbol{q})_{\sigma_N} = \\ &\quad = (-1)^N \sum_{\sigma_1=1}^{N} \cdots \sum_{\sigma_N=1}^{N} \epsilon_{\sigma_1,\ldots,\sigma_N} (\boldsymbol{G}^{N-1} \boldsymbol{q})_{\sigma_1} (\boldsymbol{G}^{N-2} \boldsymbol{q})_{\sigma_2} \ldots \\ &\quad \ldots (\boldsymbol{G}\boldsymbol{q})_{\sigma_{N-1}} (\boldsymbol{q})_{\sigma_N} \neq 0 \end{aligned} \tag{C.8}$$

where use has been made of the definition of $\boldsymbol{C}$ (C.7), namely $C_{ij} = (\boldsymbol{G}^{j-1}\boldsymbol{q})_i$.

Expressions (C.6) and (C.8) are identical (except for the sign) and consequently the two conditions, $\det(\boldsymbol{H}) \neq 0$ e $\det(\boldsymbol{C}) \neq 0$, prove equivalent. Stated differently, the condition of invertibility of matrix $\boldsymbol{H}$, obtained as a self-consistent constraint for the introduced control scheme, coincides with the standard controllability condition, as known in control theory.

As a final remark we recall (see chapter 5) that our goal is not to arbitrarily assign the polynomial $\mathcal{N}(\lambda)$ but rather to locate its roots z_k within the open left-hand plane. In this respect, the necessary and sufficient system-theoretic condition is the so-called *stabilizability* of the pair $(\boldsymbol{G}, \boldsymbol{q})$, which results from the Popov-Belevitch-Hautus rank condition reported in [103, 205].

Appendix D

Traps competition: supplementary details

D.1 Explicit solution of the 1D Fokker-Planck with two absorbing boundaries

The solution of the Fokker-Planck equation employed in chapter 5 follows the derivation by [63], that we shortly review in the following. Introduce the operator $F = D\partial_x^2 - v\partial_x$ which is defined on $D_F = \{\phi(x) \mid \phi(0) = \phi(L) = 0\} \cap D^2$ where D^2 identifies the set of twice differentiable functions. If $\phi_\lambda \in D_F$ is an eigenfunction of F associated to the eigenvalue λ, then $e^{\lambda t}\phi_\lambda$ is a particular solution of the Fokker-Planck equation (6.1). The operator F can be made Hermitian by defining the scalar product as:

$$\langle \phi_1 | \phi_2 \rangle \equiv \int_0^L e^{-\frac{v}{D}x} \phi_1^*(x)\phi_2(x)dx, \tag{D.1}$$

We can therefore find a orthonormal basis formed by the eigenfunctions of F which read:

$$\phi_k(x) = \sqrt{\frac{2}{L}} e^{\frac{v}{2D}x} \sin\big(\frac{k\pi}{L}x\big) \tag{D.2}$$

The associated eigenvalues are $\lambda_k = -\frac{v^2}{4D} - D(\frac{k\pi}{L})^2$.
Hence, by denoting with $\phi^0(x)$ the initial probability distribution, one can

cast the solution of the Fokker-Planck equation (6.1) in the explicit form:

$$p(x,t) = \sum_{k=1}^{\infty} \langle \phi_k | \phi^0 \rangle \phi_k(x) e^{\lambda_k t}. \tag{D.3}$$

Assuming the initial distribution to be a delta function centered in $x_0 = \alpha L$ $(0 < \alpha < 1)$ yields

$$\begin{aligned} \langle \phi_k | \phi^0 \rangle &= \langle \phi_k | \delta(x - \alpha L) \rangle = \\ &= \sqrt{\frac{2}{L}} \int_0^L e^{-\frac{v}{D}x} \delta(x - \alpha L) e^{\frac{v}{2D}x} \sin\left(\frac{k\pi}{L}x\right) = \\ &= \sqrt{\frac{2}{L}} e^{-\frac{v}{2D}x} \sin\left(\frac{k\pi}{L}x\right) \end{aligned} \tag{D.4}$$

from which equation (6.2) immediately follows.

D.2 The case of a generic network: details of the calculation.

By inserting ansatz (6.15) in equation (6.14), one gets:

$$\sum_{\beta} \dot{\hat{p}}_\beta(t) u_k^{(\beta)} = \sum_l L_{kl}^{[i,j]} \sum_\beta \hat{p}_\beta(t) u_l^{(\beta)}. \tag{D.5}$$

To proceed, we make explicit the dependence on the eigenvectors:

$$\begin{aligned} \sum_{\beta} \dot{\hat{p}}_\beta(t) \sum_\alpha C_{\alpha\beta} \psi_k^{(\alpha)} &= \sum_l L_{kl}^{[i,j]} \sum_\beta \hat{p}_\beta(t) \sum_\alpha C_{\alpha\beta} \psi_l^{(\alpha)} = \\ &= \sum_\beta \hat{\rho}_\beta(t) \sum_\alpha C_{\alpha\beta} \lambda^{(\alpha)} \psi_k^{(\alpha)} \end{aligned} \tag{D.6}$$

Since $\psi^{(\alpha)}$ are linearly independent, one gets:

$$\sum_\beta \dot{\hat{p}}_\beta(t) C_{\alpha\beta} = \sum_\beta \hat{p}_\beta(t) C_{\alpha\beta} \lambda^{(\alpha)}. \tag{D.7}$$

which yields:

$$\sum_\beta \hat{p}_\beta(t) C_{\alpha\beta} = \sum_\beta \hat{p}_\beta(0) C_{\alpha\beta} e^{\lambda^{(\alpha)} t}. \tag{D.8}$$

Making use of the above relations, we obtain:

$$\begin{aligned} p_k(t) &= \sum_\beta \hat{p}_\beta(t) \sum_\alpha C_{\alpha\beta}\psi_k^{(\alpha)} = \\ &= \sum_\alpha \psi_k^{(\alpha)} \sum_\beta \hat{p}_\beta(0) C_{\alpha\beta} e^{\lambda^{(\alpha)}t}. \end{aligned} \tag{D.9}$$

The only quantity that we have to determine is $\hat{p}_\beta(0)$, that we wish to express as a function of the initial condition $p_k(0)$. To this end, we make use of the inverse of (6.15):

$$\hat{p}_\beta(t) = \sum_l p_l(t)(u_l^{(\beta)})^*, \tag{D.10}$$

that, introduced into equation (D.9), results in the general solution reported in chapter 5. Finally, let us verify that equation (D.10) is indeed the inverse of (6.15). In formulae:

$$\begin{aligned} p_k(t) &= \sum_\beta \hat{p}_\beta(t) u_k^{(\beta)} = \sum_\beta \sum_l p_l(t)(u_l^{(\beta)})^* u_k^{(\beta)} \\ &= \sum_l p_l(t)\delta_{kl} = p_k(t) \end{aligned}$$

where use has been made of the condition $\sum_\beta (u_l^{(\beta)})^* u_k^{(\beta)} = \delta_{kl}$. It is hence clear the importance of dealing with an orthonormal basis in order to carry out the calculation.

Appendix E

Publications

This research activity has led to several publications in international journals. These are summarized below.[1]

International Journals

1. **G. Cencetti**, F. Bagnoli, F. Di Patti, D. Fanelli. "The second will be first: competition on directed networks", *Scientific Reports*, **6**, 27116, 2015.
2. **G. Cencetti**,F. Bagnoli, G. Battistelli, L. Chisci, F. Di Patti, D. Fanelli. "Topological stabilization for synchronized dynamics on networks", *Eur. Phys. J. B*, **90**, 9, 2017.
3. **G. Cencetti**, F. Bagnoli, G. Battistelli, L. Chisci, D. Fanelli. "Control of multidimensional systems on complex network", *Plos One*, **12**, 9, 2017.
4. **G. Cencetti**, F. Bagnoli, G. Battistelli, L. Chisci, D. Fanelli. "Spectral control for ecological stability", *Eur. Phys. J. B*, **91**, 264, 2018.

Submitted

1. **G. Cencetti**, P. Clusella, D. Fanelli. "Pattern invariance for reaction-diffusion systems on complex networks", *Scientifc Reports*, 2018. (Accepted after major revision)
2. **G. Cencetti**, F. Battiston, D. Fanelli, V. Latora. "Reactive random walkers on complex networks", *Phys. Rev. E*, 2018. (Accepted after minor revision)

[1]The author's bibliometric indices are the following: H-index = 2, total number of citations = 23 (source: Google Scholar on Month 10, 2018).

Bibliography

[1] L. A. Adamic, "The small world web," in *International Conference on Theory and Practice of Digital Libraries*. Springer, 1999, pp. 443–452.

[2] R. M. Albert J., "Dynamic modeling of gene expression in prokaryotes: application to glucose-lactose diauxie in escherichia coli," *Systems and Synthetic Biology*, vol. 5, no. 1, pp. 33–43, 2011. [Online]. Available: http://dx.doi.org/10.1007/s11693-011-9079-2

[3] D. Aldous and J. A. Fill, "Reversible markov chains and random walks on graphs," 2002, unfinished monograph, recompiled 2014, available at http://www.stat.berkeley.edu/$\sim$aldous/RWG/book.html.

[4] D. Aldous and J. Fill, "Reversible markov chains and random walks on graphs," *unfinished monograph*, 2002.

[5] A. Allen-Perkins, J. M. Pastor, and E. Estrada, "Two-walks degree assortativity in graphs and networks," *Appl. Math. Comput.*, vol. 311, no. C, pp. 262–271, Oct. 2017. [Online]. Available: https://doi.org/10.1016/j.amc.2017.05.025

[6] S. Allesina and S. Tang, "Stability criteria for complex ecosystems," *Nature*, vol. 483, no. 7388, pp. 205–208, 2012.

[7] R. M. Anderson and R. May, "Infectious diseases ofhumans," *Dynamics and Control. Oxford, England: Oxford University Press*, 1991.

[8] D. Andow, P. M. Kareiva, S. A. Levin, and A. Okubo, "Spread of invading organisms," *Landscape Ecology*, vol. 4, no. 2-3, pp. 177–188, 1990.

[9] E. Andreotti, D. Remondini, and A. Bazzani, "Equivalence between spectral properties of graphs with and without loops," *arXiv preprint arXiv:1804.05988*, 2018.

[10] C. N. Angstmann, I. C. Donnelly, and B. I. Henry, "Pattern formation on networks with reactions: A continuous-time random-walk approach," *Physical Review E*, vol. 87, no. 3, p. 032804, 2013.

[11] I. S. Aranson and L. Kramer, "The world of the complex ginzburg-landau equation," *Rev. Mod. Phys.*, vol. 74, pp. 99–143, Feb 2002. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.74.99

[12] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhoug, "Synchronization in complex networks," *Phys. Rep.*, vol. 469, pp. 93–153, 2008.

[13] M. Asllani, T. Biancalani, D. Fanelli, and A. J. McKane, "The linear noise approximation for reaction-diffusion systems on networks," *The European Physical Journal B*, vol. 86, no. 11, pp. 1–10, 2013.

[14] M. Asllani, T. Carletti, F. Di Patti, D. Fanelli, and F. Piazza, "Hopping in the crowd to unveil network topology," *Phys. Rev. Lett.*, vol. 120, p. 158301, Apr 2018. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.120.158301

[15] M. Asllani, T. Carletti, and D. Fanelli, "Tune the topology to create or destroy patterns," *arXiv preprint arXiv:1604.07046*, 2016.

[16] M. Asllani, J. D. Challenger, F. S. Pavone, L. Sacconi, and D. Fanelli, "The theory of pattern formation on directed networks," *Nature Communications*, vol. 5, 2014.

[17] M. Asllani, F. Di Patti, and D. Fanelli, "Stochastic turing patterns on a network," *Physical Review E*, vol. 86, no. 4, p. 046105, 2012.

[18] E. M. Azoff, *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc., 1994.

[19] J. R. Banavar, A. Maritan, and A. Rinaldo, "Size and form in efficient transportation networks," *Nature*, vol. 399, no. 6732, p. 130, 1999.

[20] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999. [Online]. Available: http://science.sciencemag.org/content/286/5439/509

[21] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical processes on complex networks*. Cambridge University Press, 2008.

[22] A. Becskei and L. Serrano, "Engineering stability in gene networks by autoregulation," *Nature*, vol. 405, no. 6786, pp. 590–593, 2000.

[23] H. E. Bell, "Gershgorin's theorem and the zeros of polynomials," *The American Mathematical Monthly*, vol. 72, no. 3, pp. 292–295, 1965.

[24] O. Bénichou, C. Chevalier, J. Klafter, B. Meyer, and R. Voituriez, "Geometry-controlled kinetics," *Nature Chemistry*, vol. 2, no. 6, p. 472, 2010.

[25] T. B. Benjamin and J. Feir, "The disintegration of wave trains on deep water part 1. theory," *Journal of Fluid Mechanics*, vol. 27, no. 03, pp. 417–430, 1967.

[26] V. D. Blondel, M. Gurbuzbalaban, A. Megretski, and M. L. Overton, "Explicit solutions for root optimization of a polynomial family with one affine constraint," *IEEE Transactions on Automatic Control*, vol. 57, no. 12, pp. 3078–3089, 2012.

[27] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, "The structure and dynamics of multilayer networks," *Physics Reports*, vol. 544, no. 1, pp. 1–122, 2014.

[28] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Physics Reports*, vol. 424, pp. 175–308, 2006.

[29] P. Bonacich, "Factoring and weighting approaches to status scores and clique identification," *Journal of Mathematical Sociology*, vol. 2, no. 1, pp. 113–120, 1972.

[30] P. Bonacich and P. Lloyd, "Eigenvector-like measures of centrality for asymmetric relations," *Social networks*, vol. 23, no. 3, pp. 191–201, 2001.

[31] M. Bonaventura, V. Nicosia, and V. Latora, "Characteristic times of biased random walks on complex networks," *Physical Review E*, vol. 89, no. 1, p. 012803, 2014.

[32] A. Brechtel, P. Gramlich, D. Ritterskamp, B. Drossel, and T. Gross, "Master stability functions reveal diffusion-driven pattern formation in networks," *Phys. Rev. E*, vol. 97, p. 032307, Mar 2018. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.97.032307

[33] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer Networks and ISDN Systems*, vol. 30, no. 1-7, pp. 107–117, 1998.

[34] K. Bryan and T. Leise, "The $25,000,000,000 eigenvector: The linear algebra behind google," *SIAM Review*, vol. 48, no. 3, pp. 569–581, 2006.

[35] C. G. Buffie, I. Jarchum, M. Equinda, L. Lipuma, A. Gobourne, A. Viale, C. Ubeda, J. Xavier, and E. G. Pamer, "Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to clostridium difficile-induced colitis," *Infection and Immunity*, vol. 80, no. 1, pp. 62–73, 2012.

[36] Z. Burda, J. Duda, J.-M. Luck, and B. Waclaw, "Localization of the maximal entropy random walk," *Physical Review Letters*, vol. 102, no. 16, p. 160602, 2009.

[37] G. Caldarelli and A. Chessa, *Data Science and Complex Networks: Real Case Studies with Python*. Oxford University Press, 2016.

[38] M. Cavallaro, D. Asprone, V. Latora, G. Manfredi, and V. Nicosia, "Assessment of urban ecosystem resilience through hybrid social-physical complex networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 29, no. 8, pp. 608–625, 2014. [Online]. Available: http://dx.doi.org/10.1111/mice.12080

[39] G. Cencetti, F. Bagnoli, G. Battistelli, L. Chisci, F. Di Patti, and D. Fanelli, "Topological stabilization for synchronized dynamics on networks," *The European Physical Journal B*, vol. 90, no. 1, p. 9, 2017.

[40] J. D. Challenger, R. Burioni, and D. Fanelli, "Turing-like instabilities from a limit cycle," *Physical Review E*, vol. 92, no. 2, p. 022818, 2015.

[41] X.-Q. Cheng and H.-W. Shen, "Uncovering the community structure associated with the diffusion dynamics on networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2010, no. 04, p. P04024, 2010.

[42] C. Clopath, T. Bonhoeffer, M. Hübener, and T. Rose, "Variance and invariance of neuronal long-term representations," *Phil. Trans. R. Soc. B*, vol. 372, no. 1715, p. 20160161, 2017.

[43] E. A. Codling, M. J. Plank, and S. Benhamou, "Random walk models in biology," *Journal of the Royal Society Interface*, vol. 5, no. 25, pp. 813–834, 2008.

[44] V. Colizza, R. Pastor-Satorras, and A. Vespignani, "Reaction–diffusion processes and metapopulation models in heterogeneous networks," *Nature Physics*, vol. 3, no. 4, p. 276, 2007.

[45] J. J. Collins and I. N. Stewart, "Coupled nonlinear oscillators and the symmetries of animal gaits," *Journal of Nonlinear Science*, vol. 3, no. 1, pp. 349–392, 1993.

[46] S. Contemori, F. Di Patti, D. Fanelli, and F. Miele, "Multiple-scale theory of topology-driven patterns on directed networks," *Physical Review E*, vol. 93, no. 3, p. 032317, 2016.

[47] K. Z. Coyte, J. Schluter, and K. R. Foster, "The ecology of the microbiome: Networks, competition, and stability," *Science*, vol. 350, no. 6261, pp. 663–666, 2015. [Online]. Available: http://science.sciencemag.org/content/350/6261/663

[48] M. C. Cross and P. C. Hohenberg, "Pattern formation outside of equilibrium," *Reviews of modern physics*, vol. 65, no. 3, p. 851, 1993.

[49] H. Daido, "Critical conditions of macroscopic mutual entrainment in uniformly coupled limit-cycle oscillators," *Prog. Theor. Phys.*, vol. 89, no. 4, pp. 929–934, 1993.

[50] ——, "A solvable model of coupled limit-cycle oscillators exhibiting perfect synchrony and novel frequency spectra," *Physica D*, vol. 69, pp. 394–403, 1993.

[51] ——, "Onset of cooperative entrainment in limit-cycle oscillators with uniform all-to-all interactions: Bifurcation of the order function," *Physica D*, vol. 91, pp. 24–66, 1996.

[52] M. De Domenico, "Diffusion geometry unravels the emergence of functional clusters in collective phenomena," *Physical Review Letters*, vol. 118, no. 16, p. 168301, 2017.

[53] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[54] F. Di Patti, D. Fanelli, F. Miele, and T. Carletti, "Benjamin-feir instabilities on directed networks," *arXiv preprint arXiv:1607.06301*, 2016.

[55] F. Di Patti, D. Fanelli, and F. Piazza, "Optimal search strategies on complex multi-linked networks," *Scientific Reports*, vol. 5, p. 9869, 2015.

[56] F. Dorfler and F. Bullo, "Synchronization and transient stability in power networks and nonuniform kuramoto oscillators," *SIAM Journal on Control and Optimization*, vol. 50, no. 3, pp. 1616–1642, 2012.

[57] F. Dörfler, M. Chertkov, and F. Bullo, "Synchronization in complex oscillator networks and smart grids," *Proceedings of the National Academy of Sciences*, vol. 110, no. 6, pp. 2005–2010, 2013.

[58] P. G. Drazin and W. H. Reid, *Hydrodynamic stability*. Cambridge University Press, 2004.

[59] M. B. Elowitz and S. Leibler, "A synthetic oscillatory network of transcriptional regulators," *Nature*, vol. 403, no. 6767, pp. 335–338, 2000.

[60] P. Erdos and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, no. 1, pp. 17–60, 1960.

[61] W. R. Evans, "Graphical analysis of control systems," *Transactions of the American Institute of Electrical Engineers*, vol. 67, no. 1, pp. 547–551, 1948.

[62] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *ACM SIGCOMM computer communication review*, vol. 29, no. 4. ACM, 1999, pp. 251–262.

[63] Z. Farkas and T. Fülöp, "One-dimensional drift-diffusion between two absorbing boundaries: application to granular segregation," *Journal of Physics A: Mathematical and General*, vol. 34, no. 15, p. 3191, 2001.

[64] S. Fortunato and A. Flammini, "Random walks on directed networks: the case of pagerank," *International Journal of Bifurcation and Chaos*, vol. 17, no. 07, pp. 2343–2353, 2007.

[65] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, pp. 35–41, 1977.

[66] ——, "Centrality in social networks conceptual clarification," *Social networks*, vol. 1, no. 3, pp. 215–239, 1978.

[67] S. Freilich, R. Zarecki, O. Eilam, E. S. Segal, C. S. Henry, M. Kupiec, U. Gophna, R. Sharan, and E. Ruppin, "Competitive and cooperative metabolic interactions in bacterial communities," *Nature Communications*, vol. 2, p. 589, 2011.

[68] G. F. Frobenius, F. G. Frobenius, F. G. Frobenius, F. G. Frobenius, and G. Mathematician, *Über Matrizen aus nicht negativen Elementen*. Königliche Akademie der Wissenschaften, 1912.

[69] Y. Gandica, R. Lambiotte, T. Carletti, F. S. dos Aidos, and J. Carvalho, "Circadian patterns on wikipedia edits," in *Complex Networks VII*. Springer, 2016, pp. 293–300.

[70] J. Gao, B. Barzel, and A.-L. Barabási, "Universal resilience patterns in complex networks," *Nature*, vol. 530, no. 7590, pp. 307–312, 2016.

[71] V. García-Morales and K. Krischer, "The complex ginzburg–landau equation: an introduction," *Contemporary Physics*, vol. 53, no. 2, pp. 79–95, 2012.

[72] M. Gardner, "Paradoxical situations that arise from nontransitive relations," *Scientific American*, vol. 231, no. 4, pp. 120–125, 1974.

[73] D. F. Gleich, "Pagerank beyond the web," *SIAM Review*, vol. 57, no. 3, pp. 321–363, 2015.

[74] P. M. Gleiser and L. Danon, "Community structure in jazz," *Advances in complex systems*, vol. 6, no. 04, pp. 565–573, 2003.

[75] A. Goldbeter *et al.*, "Biochemical oscillations and cellular rhythms," *Biochemical Oscillations and Cellular Rhythms, by Albert Goldbeter, Foreword by MJ Berridge, Cambridge, UK: Cambridge University Press, 1997*, vol. 1, 1997.

[76] J. Gómez-Gardeñes, V. Nicosia, R. Sinatra, and V. Latora, "Motion-induced synchronization in metapopulations of mobile agents," *Phys. Rev. E*, vol. 87, p. 032814, Mar 2013. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.87.032814

[77] J. Gómez-Gardeñes and V. Latora, "Entropy rate of diffusion processes on complex networks," *Physical Review E*, vol. 78, no. 6, p. 065102, 2008.

[78] S. Goss, S. Aron, J.-L. Deneubourg, and J. M. Pasteels, "Self-organized shortcuts in the argentine ant," *Naturwissenschaften*, vol. 76, no. 12, pp. 579–581, 1989.

[79] G. A. Gottwald, "Model reduction for networks of coupled oscillators," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 5, p. 053111, 2015. [Online]. Available: https://doi.org/10.1063/1.4921295

[80] J. Grilli, M. Adorisio, S. Suweis, G. Barabas, J. R. Banavar, S. Allesina, and A. Maritan, "Feasibility and coexistence of large ecological communities," *Nature Communications*, vol. 8, 2017.

[81] J. Grilli, T. Rogers, and S. Allesina, "Modularity and stability in ecological communities," *Nature Communications*, vol. 7, 2016.

[82] R. Grimshaw, *Nonlinear ordinary differential equations*. Routledge, 2017.

[83] J. Guare, *Six degrees of separation: A play*. Vintage, 1990.

[84] R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas, "Self-similar community structure in a network of human interactions," *Physical review E*, vol. 68, no. 6, p. 065103, 2003.

[85] E. J. Hancock and G. A. Gottwald, "Model reduction for kuramoto models with complex topologies," *arXiv preprint arXiv:1804.07444*, 2018.

[86] I. Hanski, "Metapopulation dynamics," *Nature*, vol. 396, no. 6706, p. 41, 1998.

[87] L. G. Harrison, *Kinetic theory of living pattern*. Cambridge University Press, 1993, vol. 28.

[88] J. Hasty, D. McMillen, F. Isaacs, and J. J. Collins, "Computational studies of gene regulatory networks: in numero molecular biology," *Nature Reviews Genetics*, vol. 2, no. 4, pp. 268–279, 2001.

[89] S. Hata and H. Nakao, "Localization of laplacian eigenvectors on random networks," *Scientific Reports*, vol. 7, no. 1, p. 1121, 2017. [Online]. Available: https://doi.org/10.1038/s41598-017-01010-0

[90] S. Hata, H. Nakao, and A. S. Mikhailov, "Global feedback control of turing patterns in network-organized activator-inhibitor systems," *EPL (Europhysics Letters)*, vol. 98, no. 6, p. 64004, 2012.

[91] T. Hoffmann, M. A. Porter, and R. Lambiotte, "Generalized master equations for non-poisson dynamics on networks," *Phys. Rev. E*, vol. 86, p. 046102, Oct 2012. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.86.046102

[92] E. E. Holmes, M. A. Lewis, J. E. Banks, and R. R. Veit, "Partial differential equations in ecology: Spatial interactions and population

dynamics," *Ecology*, vol. 75, no. 1, pp. 17–29, 1994. [Online]. Available: http://www.jstor.org/stable/1939378

[93] R. B. Hoyle, *Pattern formation: an introduction to methods*. Cambridge University Press, 2006.

[94] L. Huang, Q. Chen, Y.-C. Lai, and L. M. Pecora, "Generic behavior of master-stability functions in coupled nonlinear dynamical systems," *Physical Review E*, vol. 80, no. 3, p. 036204, 2009.

[95] L. Hufnagel, D. Brockmann, and T. Geisel, "Forecast and control of epidemics in a globalized world," *Proceedings of the National Academy of Sciences*, vol. 101, no. 42, pp. 15 124–15 129, 2004.

[96] I. Iacopini, S. Milojević, and V. Latora, "Network dynamics of innovation processes," *Physical Review Letters*, vol. 120, no. 4, p. 048301, 2018.

[97] F. Iannelli, A. Koher, D. Brockmann, P. Hövel, and I. M. Sokolov, "Effective distances for epidemics spreading on complex networks," *Physical Review E*, vol. 95, no. 1, p. 012313, 2017.

[98] F. J. Isaacs, J. Hasty, C. R. Cantor, and J. J. Collins, "Prediction and measurement of an autoregulatory genetic module," *Proceedings of the National Academy of Sciences*, vol. 100, no. 13, pp. 7714–7719, 2003.

[99] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, "The large-scale organization of metabolic networks," *Nature*, vol. 407, no. 6804, p. 651, 2000.

[100] K. A. Johnson and R. S. Goody, "The original michaelis constant: translation of the 1913 michaelis–menten paper," *Biochemistry*, vol. 50, no. 39, pp. 8264–8269, 2011.

[101] S. Johnson, V. Domínguez-García, L. Donetti, and M. A. Muñoz, "Trophic coherence determines food-web stability," *Proceedings of the National Academy of Sciences*, vol. 111, no. 50, pp. 17 923–17 928, 2014.

[102] D. Jung and S. Kettemann, "Long-range response in ac electricity grids," *Phys. Rev. E*, vol. 94, p. 012307, Jul 2016. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevE.94.012307

[103] T. Kailah, *Linear Systems*. Prentice-Hall Englewood Cliffs, NJ, 1980.

[104] M. Kaiser and C. C. Hilgetag, "Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems," *PLoS Computational Biology*, vol. 2, no. 7, p. e95, 2006.

[105] J. E. Kamienkowski, H. Pashler, S. Dehaene, and M. Sigman, "Effects of practice on task architecture: Combined evidence from interference experiments and random-walk models of decision making," *Cognition*, vol. 119, no. 1, pp. 81–95, 2011.

[106] E. Kandel, J. Schwartz, and T. Jessell, *Principles of Neural Science*, 4th ed. McGraw-Hill, 2000.

[107] M. Kang and Z. Petrášek, "Random graphs: Theory and applications from nature to society to the brain," *Internationale mathematische Nachrichten= International mathematical news*, vol. 227, pp. 1–24, 2014.

[108] S. A. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic nets," *Journal of theoretical biology*, vol. 22, no. 3, pp. 437–467, 1969.

[109] D. H. Kelley and N. T. Ouellette, "Emergent dynamics of laboratory insect swarms," *Scientific Reports*, vol. 3, p. 1073, 2013.

[110] J. M. Kleinberg, "Navigation in a small world," *Nature*, vol. 406, no. 6798, p. 845, 2000.

[111] B. Kogut and G. Walker, "The small world of germany and the durability of national networks," *American sociological review*, pp. 317–335, 2001.

[112] R. Kopelman, "Reaction kinetics: Catalysis without a catalyst," *Nature Chemistry*, vol. 2, no. 6, p. 430, 2010.

[113] K. Kosmidis, S. Havlin, and A. Bunde, "Structural properties of spatially embedded networks," *EPL (Europhysics Letters)*, vol. 82, no. 4, p. 48005, 2008.

[114] V. Krebs, *unpublished*, http://www.orgnet.com/.

[115] Y. Kuramoto, *Chemical oscillations, waves, and turbulence*. Springer Science & Business Media, 2012, vol. 19.

[116] A. N. Langville and C. D. Meyer, "Deeper inside pagerank," *Internet Mathematics*, vol. 1, no. 3, pp. 335–380, 2004.

[117] V. Latora, V. Nicosia, and G. Russo, *Complex networks: principles, methods and applications*. Cambridge University Press, 2017.

[118] S. H. Lee and P. Holme, "Exploring maps with greedy navigators," *Physical Review Letters*, vol. 108, no. 12, p. 128701, 2012.

[119] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graph evolution: Densification and shrinking diameters," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 1, no. 1, p. 2, 2007.

[120] J. Leskovec and J. J. Mcauley, "Learning to discover social circles in ego networks," in *Advances in neural information processing systems*, 2012, pp. 539–547.

[121] T. Leung Yee, A. Cheng Hou Tsang, B. Malomed, and K. Wing Chow, "Exact solutions for domain walls in coupled complex ginzburg–landau equations," *Journal of the Physical Society of Japan*, vol. 80, no. 6, p. 064001, 2011.

[122] R. Levins, "Some demographic and genetic consequences of environmental heterogeneity for biological control," *American Entomologist*, vol. 15, no. 3, pp. 237–240, 1969.

[123] G. Li, S. Reis, A. Moreira, S. Havlin, H. Stanley, and J. Andrade Jr, "Towards design principles for optimal transport networks," *Physical Review Letters*, vol. 104, no. 1, p. 018701, 2010.

[124] ——, "Optimal transport exponent in spatially embedded networks," *Physical Review E*, vol. 87, no. 4, p. 042810, 2013.

[125] Z. Li, Z. Duan, G. Chen, and L. Huang, "Consensus of multiagent systems and synchronization of complex networks: A unified viewpoint," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 1, pp. 213–224, 2010.

[126] D. H. Lim, J. M. LeDue, M. H. Mohajerani, and T. H. Murphy, "Optogenetic mapping after stroke reveals network-wide scaling of functional connections and heterogeneous recovery of the peri-infarct," *Journal of Neuroscience*, vol. 34, no. 49, pp. 16 455–16 466, 2014.

[127] Y. Lin and Z. Zhang, "Mean first-passage time for maximal-entropy random walks in complex networks," *Scientific Reports*, vol. 4, p. 5365, 2014.

[128] H. Lodish, A. Berk, S. L. Z. , P. Matsudaira, D. Baltimore, and J. Darnell, *Molecular Cell Biology*, 4th ed. New York: W. H. Freeman, 2000.

[129] Y. Lu and M. Ballauff, "Thermosensitive core–shell microgels: from colloidal model systems to nanoreactors," *Progress in Polymer Science*, vol. 36, no. 6, pp. 767–792, 2011.

[130] M. Lucas, D. Fanelli, T. Carletti, and J. Petit, "Desynchronization induced by time-varying network," *EPL (Europhysics Letters)*, vol. 121, no. 5, p. 50008, 2018.

[131] J. Malmros, N. Masuda, and T. Britton, "Random walks on directed networks: Inference and respondent-driven sampling," 2016.

[132] N. Masuda, M. A. Porter, and R. Lambiotte, "Random walks and diffusion on networks," *Physics Reports*, 2017.

[133] R. M. May, "Will a large complex system be stable?" *Nature*, vol. 238, pp. 413–414, 1972.

[134] H. Meinhardt, *Models of biological pattern formation.* Academic Press, 1982.

[135] R. E. Mirollo and S. H. Strogatz, "Synchronization of pulse-coupled biological oscillators," *SIAM Journal on Applied Mathematics*, vol. 50, no. 6, pp. 1645–1662, 1990.

[136] J. M. Morales, P. R. Moorcroft, J. Matthiopoulos, J. L. Frair, J. G. Kie, R. A. Powell, E. H. Merrill, and D. T. Haydon, "Building the bridge between animal movement and population dynamics," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 365, no. 1550, pp. 2289–2301, 2010.

[137] R. Muolo, M. Asllani, D. Fanelli, P. K. Maini, and T. Carletti, "Patterns of non-normality in networked systems," *arXiv preprint arXiv:1812.02514*, 2018.

[138] J. D. Murray, E. A. Stanley, and D. L. Brown, "On the spatial spread of rabies among foxes," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 229, no. 1255, pp. 111–150, 1986. [Online]. Available: http://rspb.royalsocietypublishing.org/content/229/1255/111

[139] J. D. Murray, "Mathematical biology i: an introduction, vol. 17 of interdisciplinary applied mathematics," 2002.

[140] J. Murray, "Mathematical biology ii: Spatial models and biochemical applications, volume ii," 2003.

[141] H. Nakao, "Complex ginzburg-landau equation on networks and its non-uniform dynamics," *The European Physical Journal Special Topics*, vol. 223, no. 12, pp. 2411–2421, 2014.

[142] H. Nakao and A. S. Mikhailov, "Turing patterns in network-organized activator-inhibitor systems," *Nature Physics*, vol. 6, no. 7, pp. 544–550, 2010.

[143] M. E. J. Newman, *Networks: An Introduction*. Oxford: Oxford University Press, 2010.

[144] M. E. Newman, "The structure of scientific collaboration networks," *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 404–409, 2001.

[145] ——, "Assortative mixing in networks," *Physical Review Letters*, vol. 89, no. 20, p. 208701, 2002.

[146] ——, "Mixing patterns in networks," *Physical Review E*, vol. 67, no. 2, p. 026126, 2003.

[147] M. E. Newman and D. J. Watts, "Scaling and percolation in the small-world network model," *Physical Review E*, vol. 60, no. 6, p. 7332, 1999.

[148] V. Nicosia, F. Bagnoli, and V. Latora, "Impact of network structure on a model of diffusion and competitive interaction," *EPL (Europhysics Letters)*, vol. 94, no. 6, p. 68009, 2011.

[149] V. Nicosia, R. Criado, M. Romance, G. Russo, and V. Latora, "Controlling centrality in complex networks," *Scientific Reports*, vol. 2, 2012.

[150] V. Nicosia, M. Valencia, M. Chavez, A. Díaz-Guilera, and V. Latora, "Remote synchronization reveals network symmetries and functional modules," *Physical Review Letters*, vol. 110, no. 17, p. 174102, 2013.

[151] J. D. Noh and H. Rieger, "Random walks on complex networks," *Physical Review Letters*, vol. 92, no. 11, p. 118701, 2004.

[152] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[153] C. L. Oliveira, P. A. Morais, A. A. Moreira, and J. S. Andrade Jr, "Enhanced flow in small-world networks," *Physical Review Letters*, vol. 112, no. 14, p. 148701, 2014.

[154] G. Oshanin, K. Lindenberg, H. S. Wio, and S. Burlatsky, "Efficient search by optimized intermittent random walks," *Journal of Physics A: Mathematical and Theoretical*, vol. 42, no. 43, p. 434008, 2009.

[155] H. G. Othmer and L. Scriven, "Instability and dynamic pattern in cellular networks," *Journal of Theoretical Biology*, vol. 32, no. 3, pp. 507–537, 1971.

[156] ——, "Non-linear aspects of dynamic pattern in cellular networks," *Journal of Theoretical Biology*, vol. 43, no. 1, pp. 83–112, 1974.

[157] O. Ovaskainen and I. Hanski, "Spatially structured metapopulation models: Global and local assessment of metapopulation capacity," *Theoretical Population Biology*, vol. 60, no. 4, pp. 281 – 302, 2001. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0040580901915485

[158] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Reviews of modern physics*, vol. 87, no. 3, p. 925, 2015.

[159] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, "Dynamical and correlation properties of the internet," *Phys. Rev. Lett.*, vol. 87, p. 258701, Nov 2001. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.87.258701

[160] R. Pastor-Satorras and A. Vespignani, *Evolution and structure of the Internet: A statistical physics approach*. Cambridge University Press, 2007.

[161] K. Pearson, "The problem of the random walk," *Nature*, vol. 72, no. 1867, p. 342, 1905.

[162] L. M. Pecora and T. L. Carroll, "Master stability functions for synchronized coupled systems," *Phys. Rev. Lett.*, vol. 80, pp. 2109–2112, Mar 1998. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.80.2109

[163] L. M. Pecora, T. L. Carroll, G. A. Johnson, D. J. Mar, and J. F. Heagy, "Fundamentals of synchronization in chaotic systems, concepts, and applications," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 7, no. 4, pp. 520–543, 1997.

[164] L. M. Pecora, F. Sorrentino, A. M. Hagerstrom, T. E. Murphy, and R. Roy, "Cluster synchronization and isolated desynchronization in complex networks with symmetries," *Nature Communications*, vol. 5, p. 4079, 2014.

[165] W. Penney, *Journal of Recreational Mathematics*, vol. 2, p. 241, 1969.

[166] O. Perron, "Uber matrizen," *Math. Annalen*, vol. 64, pp. 248–263, 1907.

[167] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: a universal concept in nonlinear sciences*. Cambridge University Press, 2003, vol. 12.

[168] L. M. Pismen, *Patterns and interfaces in dissipative dynamics*. Springer Science & Business Media, 2006.

[169] M. Rohden, A. Sorge, M. Timme, and D. Witthaut, "Self-organized synchronization in decentralized power grids," *Phys. Rev. Lett.*, vol. 109, p. 064101, 2012.

[170] M. Rubenstein, A. Cornejo, and R. Nagpal, "Programmable self-assembly in a thousand-robot swarm," *Science*, vol. 345, no. 6198, pp. 795–799, 2014.

[171] A. Samukhin, S. Dorogovtsev, and J. Mendes, "Laplacian spectra of, and random walks on, complex networks: Are scale-free architectures really important?" *Physical Review E*, vol. 77, no. 3, p. 036115, 2008.

[172] P. Schippers, J. Verboom, J. Knaapen, and R. v. Apeldoorn, "Dispersal and habitat connectivity in complex heterogeneous landscapes: an analysis with a gis-based random walk model," *Ecography*, vol. 19, no. 2, pp. 97–106, 1996.

[173] J. Sethna, *Statistical mechanics: entropy, order parameters, and complexity*. Oxford University Press, 2006, vol. 14.

[174] T.-C. D. Shen, L. Albenberg, K. Bittinger, C. Chehoud, Y.-Y. Chen, C. A. Judge, L. Chau, J. Ni, M. Sheng, A. Lin, B. J. Wilkins, E. L. Buza, J. D. Lewis, Y. Daikhin, I. Nissim, M. Yudkoff, F. D. Bushman, and G. D. Wu, "Engineering the gut microbiota to treat hyperammonemia," *The Journal of Clinical Investigation*, vol. 125, no. 7, pp. 2841–2850, 7 2015. [Online]. Available: https://www.jci.org/articles/view/79214

[175] A. B. Siddique, L. M. Pecora, and F. Sorrentino, "Symmetries in the time-averaged dynamics of networks: reducing unnecessary complexity through minimal network models," *arXiv preprint arXiv:1710.05251*, 2017.

[176] E. Signaroldi, P. Laise, S. Cristofanon, A. Brancaccio, E. Reisoli, S. Atashpaz, M. R. Terreni, C. Doglioni, G. Pruneri, P. Malatesta *et al.*, "Polycomb dysregulation in gliomagenesis targets a zfp423-dependent differentiation network," *Nature Communications*, vol. 7, 2016.

[177] R. Sinatra, J. Gómez-Gardenes, R. Lambiotte, V. Nicosia, and V. Latora, "Maximal-entropy random walks in complex networks with limited information," *Physical Review E*, vol. 83, no. 3, p. 030103, 2011.

[178] P. S. Skardal and A. Arenas, "On controlling networks of limit-cycle oscillators," *arXiv preprint arXiv:1603.00842*, 2016.

[179] D. L. Stein, "Lectures in the sciences of complexity the proceeding of the 1988 complex systems summer school, held june-july, 1988 in santa fe, new mexico," 1989.

[180] R. R. Stein, V. Bucci, N. C. Toussaint, C. G. Buffie, G. Rätsch, E. G. Pamer, C. Sander, and J. B. Xavier, "Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota," *PLoS Computational Biology*, vol. 9, no. 12, p. e1003388, 2013.

[181] S. N. Steinway, M. B. Biggs, T. P. Loughran Jr, J. A. Papin, and R. Albert, "Inference of network dynamics and metabolic interactions in the gut microbiome," *PLoS Comput Biol*, vol. 11, no. 6, p. e1004338, 2015.

[182] S. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering.* Westview Press, 2001, vol. 272.

[183] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, p. 268, 2001.

[184] J. T. Stuart and R. C. DiPrima, "The eckhaus and benjamin-feir resonance mechanisms," *Proc. R. Soc. Lond. A*, vol. 362, no. 1708, pp. 27–41, 1978.

[185] S. Suweis, J. Grilli, J. R. Banavar, S. Allesina, and A. Maritan, "Effect of localization on the stability of mutualistic ecological networks," *Nature Communications*, vol. 6, 2015.

[186] S. Suweis, J. Grilli, and A. Maritan, "Disentangling the effect of hybrid interactions and of the constant effort hypothesis on ecological community stability," *Oikos*, vol. 123, no. 5, pp. 525–532, 2014. [Online]. Available: http://dx.doi.org/10.1111/j.1600-0706.2013.00822.x

[187] S. Suweis, F. Simini, J. R. Banavar, and A. Maritan, "Emergence of structural and dynamical properties of ecological mutualistic networks," *Nature*, vol. 500, no. 7463, pp. 449–452, 2013.

[188] E. Thébault and C. Fontaine, "Stability of ecological communities and the architecture of mutualistic and trophic networks," *Science*,

vol. 329, no. 5993, pp. 853–856, 2010. [Online]. Available: http://science.sciencemag.org/content/329/5993/853

[189] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[190] J. Travers and S. Milgram, "The small world problem," *Phychology Today*, vol. 1, no. 1, pp. 61–67, 1967.

[191] C. Tu, J. Grilli, and S. Suweis, "Universality of resilience patterns in generalized lotka volterra dynamics and beyond," *arXiv preprint arXiv:1606.09630*, 2016.

[192] A. M. Turing, "The chemical basis of morphogenesis," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 237, no. 641, pp. 37–72, 1952.

[193] D. Urban and T. Keitt, "Landscape connectivity: a graph-theoretic perspective," *Ecology*, vol. 82, no. 5, pp. 1205–1218, 2001.

[194] T. Valente, "Network models of the diffusion of innovations (hampton, cresskill, nj)," 1995.

[195] A. van Harten, "On the validity of the ginzburg-landau equation," *Journal of Nonlinear Science*, vol. 1, no. 4, pp. 397–422, 1991.

[196] D. P. Vázquez, C. J. Melián, N. M. Williams, N. Blüthgen, B. R. Krasnov, and R. Poulin, "Species abundance and asymmetric interaction strength in ecological networks," *Oikos*, vol. 116, no. 7, pp. 1120–1127, 2007.

[197] A. Vespignani, "Modelling dynamical processes in complex socio-technical systems," *Nature physics*, vol. 8, no. 1, p. 32, 2012.

[198] A. Wagner and D. A. Fell, "The small world inside large metabolic networks," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 268, no. 1478, pp. 1803–1810, 2001.

[199] X.-J. Wang, "Probabilistic decision making by slow reverberation in cortical circuits," *Neuron*, vol. 36, no. 5, pp. 955–968, 2002.

[200] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[201] N. Welsch, A. Wittemann, and M. Ballauff, "Enhanced activity of enzymes immobilized in thermoresponsive core-shell microgels," *The Journal of Physical Chemistry B*, vol. 113, no. 49, pp. 16 039–16 045, 2009.

[202] J. Wyller, P. Blomquist, and G. T. Einevoll, "Turing instability and pattern formation in a two-population neuronal network model," *Physica D: Nonlinear Phenomena*, vol. 225, no. 1, pp. 75–93, 2007.

[203] R. Xulvi-Brunet and I. M. Sokolov, "Changing correlations in networks: assortativity and dissortativity," *Acta Physica Polonica B*, vol. 36, p. 1431, 2005.

[204] S.-J. Yang, "Exploring complex networks by walking on them," *Physical Review E*, vol. 71, no. 1, p. 016107, 2005.

[205] Z. Yuan, C. Zhao, Z. Di, W.-X. Wang, and Y.-C. Lai, "Exact controllability of complex networks," *Nature Communications*, vol. 4, 2013.

[206] L. Zhang, A. E. Motter, and T. Nishikawa, "Incoherence-mediated remote synchronization," *Physical Review Letters*, vol. 118, no. 17, p. 174102, 2017.

Acknowledgments

The end of the thesis, the end of an era. As many other natural phenomena, this thesis represents the finally emerging result, stemming from the interplay between all the microscopic dynamics that have composed the complex system that was my PhD. Microscopic dynamics which, if analysed in details, are composed by daily interactions, discussions, considerations, ideas and communications with all my nearest neighbouring nodes in the last three years.

All of them being essential, the most important node has surely been Duccio who, with his dedication and constant incentive, has given me all the equipment which, today, allows me to call myself a researcher. I must thank, not only for all the acquired skills, but also for all the passion that he always put in his work and in our collaboration. The afternoons spent in front of the blackboard were the best moments of my entire PhD! Thanks also for making me feel part of a team. Team which, under the flag of CSDC, has represented an amazing and incredible experience for me. Other milestone of the group, of my thesis and of my PhD, with his incredible experience, personality and care, who I must thank from the bottom of my heart, is definitely Franco. Thanks to him I discovered the fascinating world of complex systems, and beyond, because with Franco you always learn something and are enchanted by something.
Thanks to Giorgio, for sharing his smartness and thanks to Luigi for the confidence placed in me.
Thanks also, obviously, to Giovanna, the most central node of the CSDC, who, with her friendship and her sunshine, enlightened most of my working days. Thanks to Clément, for all the laughs and all his affection. Thanks to Maxime, Pau, Leo, Nicco, Sara, Lorenzo, Lorenzo, Ihusan, Francesca, Pierfrancesco and everyone who I happened to share the office, the joys, the laughs, the satisfactions and the frustrations with.

Thanks also to the other labs that I had the chance to visit. Thanks to Vito first of all for all the interesting discussions and for introducing me to the family of Queen Mary, my first big group of friends-collegues with whom I learned how much PhD life can be funny and stimulating, falling even more in love with the academic world. A special thank you to Fede, a great mate and friend, for all the proposals, the incentives, the encouragements and the coaching. And, of course, for Budapest and the Central European University. Thanks to QMUL and CEU for hosting me.

Thanks to Nicco for giving me a new perspective and for making me realize how useful and interesting my work can be.

Thanks to Dani, a person from whom there's always more to learn, for sharing with me joy and sorrow of research life, and for being such an inexhaustible source of good advice and strategies.

Thanks to Cate, la mia socia, my perfect mate and match for the entire pathway, for holding my hand since the dawn of time, giving me the strength to walk uphill. A rising which nowadays still doesn't stop. Such a vertigo to imagine where this launch pad will take us!

A final word of thanks goes to my family, for supporting and encouraging me, almost always. And thanks to all the others who were part of the last three years and who, even if not strictly linked to my PhD, represented crucial connections, everyone with his own weight, and made possible all of this to emerge.