



Reconstruction, mobility, and synchronization in complex networks

Luce Prignano



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial 3.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial 3.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial 3.0. Spain License.**

RECONSTRUCTION, MOBILITY, AND SYNCHRONIZATION IN COMPLEX NETWORKS

LUCE PRIGNANO

Departament de Física Fonamental

Facultat de Física

Universitat de Barcelona

Martí i Franqués 1

08028 Barcelona



Programa de Doctorat en Física
2009-2012

MEMÒRIA PRESENTADA PER LA LUCE PRIGNANO PER OPTAR AL TÍTOL DE
DOCTOR EN FÍSICA PER LA UNIVERSITAT DE BARCELONA

Certifico que la present tesi doctoral ha estat realitzada sota la meva direcció.
Barcelona, mayo 2012.

Albert Díaz-Guilera
Catedrático de Física de la Materia Condensada
Departament de Física Fonamental
Universitat de Barcelona

*To my new-coalesced
extended family*

Acknowledgement

There are so many people I would like to thank for their contribution in some way or another to this thesis. So, lets start with the easy ones and then try to remember all of them. But this is probably bound to lead to someone getting left out, and unfairly at that.

First and foremost, the person that deserves thanks more than anyone is Albert. He has been there from the very beginning, helping me to get a grant, till the very end, and beyond, during the hard unexpected heavy tail of this work. Without his guidance and experience, this thesis would quite simply not exist. And without his support and patience in recent weeks, I am pretty sure it would have remained in the archives to rot for a long time.

A big thank you also to Conrad, who was beside me the first days helping me surviving to the dangers of a foreign bureaucratic jungle, and who was also there during these three years to discuss almost every kind of issue.

An important debt of gratitude must go to my fellow Ula, whose arrival made the job truly cooperative and hence more interesting, and who made me discover the wonderful world of *Inkscape*. Without him, the cover of this thesis would be very different – much more boring.

I am also grateful to the people that have directly added to this work: Pablo Gleiser, for his suggestions, his confidence and his friendship. And Yamir Moreno, with whom we walked through unknown networks. Also I'd like to mention all the people of the PhysComp2 group: Marian, Alex, Sergi, Pol, Pau and Clara, for the friendly environment and for the last fantastic *calçotada*.

Special thank to Prof. Jurgen Kurths who made my stay at the Potsdam Institute for Climatology possible. And also to Prof. A. Vespignani for hosting me at the ISI Foundation in Turin.

A big thanks also to Anna, who was my guide in Berlin and Potsdam, and to Alla who opened the door of her house to an unknown student of a friend of the advisor of a friend of her... thank you, brave girl!

Moreover, I would like to express my sincere gratitude to my geek friend Sandro, who tries to appear as a regular guy, but he really is a generous mine of geek tricks. And also to Javi and Gorka, for useful discussion. I have learned so much from them. Now it is time to say “thank you, Jesus”, for listening to me and illuminating my way with your suggestions.

Many thanks also to the staff of both the secretary of the faculty and the secretary of the

department, and in particular to Bea, who is helping my in this very last bureaucratic rush.

A special mention goes to, in order of appearance: Oriol, Raul, Hicham, Dani, and Mario. We have shared not just an office, but also coffee, tea, pastries, chocolate, scissors, staplers, glue, comics... and, *ça va sans dire*, many useful information. An other one to Joaquin, Carlos, and Raquel, who have always considered me as a fellow, even if I was just a cheeky gatecrasher at their institute. And, last but not least, to Fakhteh, a fellow student and a pen-friend as those of yore. Scientific and human interchange with her has been stimulating and fruitful, both in the differences and in the similarities.

I would also like to say thank you to the people of *La Pantera RoSSa* in Zaragoza, a bookshop and social center which has adopted me and taught me how to resist the drowsiness in some sultry summer afternoons by administering to me generous doses of a very tasted caffeine-rich cola.

Looking ahead, I want to thanks Bernardo, Marco, and the whole *Simulpast Consolider*, for showing me new fascinating field where the skills I am supposed to have acquired during these years can be applied. The melting pot of heterogeneous researchers they have been able to build is really amazing.

Looking backward, at the first steps I took in Barcelona, I have to recognize that it is thanks to my dear friend and first flatmate Virginia that I did not get lost during my first days in Barcelona, if I have a NIE, a social security number, and many many other essentials stuffs. She is a real coach.

Of course, I cannot forget to mention all my flatmates, Vane, David, Sergio, and Nela, with whom I have shared not just a flat (actually two), but a true home. We have been enjoying nice moments in Barcelona together.

My life here would not have been so enjoyable without Marco, Paolo, Eva, and Narayan. When I found myself in trouble – all kinds of troubles, either practical or emotional, silly or very serious – their help was invaluable.

Big thanks and hugs to the *dispora crew*, which from Rome spread everywhere in Europe, but that was there again – both physically and electronically – when I needed them, in this alienating post-apocalyptic final stage. Francesca, Isa, Gab, Guido, Villa, Jack, Gino, Flavio, Carlo, Renzils, Nicco, Calta. You gave me the final proof that we have definitively grown up, and in the end, that's not so bad. On this side of the Tyrrhenian Sea, the same is for Francesco, Nathalie, Arianna, Ricci, Vane, and Nela, who waited me so warmly. But also for the good jaws, the dinners, the parties...

Now it's time to tanks who really was on my side since the very beginning. My sister, who visited me as often as possible. In the end, I know she loves me, although I am a "scientist", and some other disagreeable things. Although, sometimes, I said stupid things. My mother, who have shared my hopes, my dreams and my disappointments. She deserves the warmest, sweetest hug. My father, who made me forget months of mental stress with few days of physical

stress, while visiting very special places. He never gave up, never thought that what I was doing was too difficult or that I was too weird to be understood. There was nothing too difficult or too weird, almost nothing. He is going to miss the conclusion of this stage of my life, and many many other things. He took away something of us and of our future. But he also left us a lot of work to do, and the awareness that it is necessary and urgent, that we should not waste time. That is, a great motivation.

Finally, beyond a doubt, my greatest thanks to Emanuele, for the last twenty months. And maybe even more for these last few weeks. *Macchettelodicoaffa*'?

Contents

| | |
|--|------------|
| Acknowledgement | i |
| List of Figures | vii |
| List of Tables | xi |
| 1 Introduction | 1 |
| 2 Extracting topological features from dynamical measures in networks of Kuramoto oscillators | 8 |
| 2.1 Synchronization and phase transition | 10 |
| 2.2 Incoherent state | 13 |
| 2.3 Critical frequency and local topology | 15 |
| 2.4 Slightly above the critical point | 18 |
| 2.4.1 Hierarchical organization | 19 |
| 2.4.2 Recovering network topology | 21 |
| 2.5 Far from the critical point | 23 |
| 2.6 Conclusions | 28 |
| 3 Exploring complex networks by means of adaptive walkers | 30 |
| 3.1 Baseline model of walkers | 31 |
| 3.2 Characterizing the performance of the walkers | 33 |
| 3.3 Sequential steps and return time | 36 |
| 3.4 The adaptive strategy | 38 |
| 3.5 Recovering the degree distribution. | 42 |
| 3.6 Conclusions | 44 |
| 4 Synchronization of moving integrate and fire oscillators | 46 |
| 4.1 The geometric model | 48 |
| 4.1.1 Synchronization Mechanisms | 50 |
| 4.2 Minimal topological model | 54 |

| | | |
|----------|---|------------|
| 4.3 | Time to synchronize | 56 |
| 4.4 | Characterizing the system behavior | 57 |
| 4.5 | Fast limit | 61 |
| 4.5.1 | Assessment of the energy cost | 61 |
| 4.6 | Slow regime | 64 |
| 4.6.1 | Checking the hypothesis. | 66 |
| 4.6.2 | Empirical proposal of a scaling relation | 68 |
| 4.6.3 | From local coherence to global synchrony | 69 |
| 4.7 | Wasting time. The no-synchronization region. | 71 |
| 4.8 | Synchronizing faster, synchronizing cheaper | 72 |
| 4.9 | Conclusions | 74 |
| 5 | Conclusions and perspectives | 76 |
| 5.1 | Conclusions | 76 |
| 5.2 | Perspectives | 78 |
| A | Pacemakers in a Cayley tree of Kuramoto oscillators | 80 |
| A.1 | The model | 80 |
| A.2 | Characterizing the global behavior | 81 |
| A.3 | Results | 83 |
| A.4 | Conclusions | 88 |
| B | Estimations and micro-characterization for the minimal IFOS model | 90 |
| B.1 | Average size of connected components | 90 |
| B.1.1 | Significance in the case of moving agents | 91 |
| B.2 | Synchronization time in a fully connected IFOS system | 93 |
| B.3 | Local cluster synchronization time | 94 |
| B.3.1 | Local synchrony time scale | 95 |
| B.4 | Estimation of the (out-)neighbor change time | 97 |
| B.4.1 | The fast limit case. Heuristic derivation of V_f | 101 |
| B.4.2 | The in-neighbor change time | 102 |
| | Resumen | 103 |
| | Introducción | 103 |
| | Extrayendo características topológicas de medidas dinámicas en redes de osciladores de Kuramoto | 110 |
| | Explorando redes complejas mediante caminantes adaptativos | 112 |
| | Sincronización de osciladores móviles | 112 |
| | Perspectivas | 115 |

| | |
|--------------------------|------------|
| Bibliography | 117 |
| Publications list | 124 |

List of Figures

| | | |
|------|---|----|
| 1.1 | My thesis in a diagram. | 5 |
| 2.1 | Hierarchic network that will be used as a benchmark. | 11 |
| 2.2 | Order parameter (2.1.7) as a function of the natural frequency of the pacemaker. | 13 |
| 2.3 | Effective frequencies above the critical point as functions of time. | 14 |
| 2.4 | Average effective frequencies as a function of the natural frequency of the pacemaker ω | 14 |
| 2.5 | Critical frequency of a pacemaker as a function of its degree for a set of networks. | 18 |
| 2.6 | The network of Fig. 2.1 and its corresponding dendrogram. | 19 |
| 2.7 | Dendrogram of the cortical brain network of a cat. | 20 |
| 2.8 | Effective frequencies as function of time far above the critical point ($\omega = 20\omega_p^c$). | 24 |
| 2.9 | Normalized correlations of the cortical brain network of the cat for pacemaker on node 1 ($k_p = 10$). | 25 |
| 2.10 | Average number of reconstructed links as a function of the number of nodes we considered as pacemakers (number of trials). | 27 |
| 3.1 | Example of the motion of an agent: 4 snapshots taken at 4 sequential time steps. | 32 |
| 3.2 | Information \bar{I} (averaged over 3000 realizations) gathered by a walker during a searching time $T = 10000$ as a function of the q parameter. | 34 |
| 3.3 | Mean Information $\langle \bar{I} \rangle$ gathered by a walker performing its search starting from any home-node during a time lag of T steps, as a function of the q parameter. | 34 |
| 3.4 | Mean maximum number of consecutive steps forward $\langle \bar{S}_F \rangle$ (black line) and backward $\langle \bar{S}_B \rangle$ (red line) taken by a walker during a time lag of $T = 10000$ steps. | 37 |
| 3.5 | Mean time $\langle \bar{T}_R \rangle$ that a walker starting from any node in the network needs to come back to its home-node as a function of q | 38 |
| 3.6 | The figure represents, in a <i>flowchart</i> a possible implementation of the algorithm described in the main text for the adaptive strategy. | 40 |
| 3.7 | Left panel: the total efficiency $\langle \bar{E} \rangle = \langle \bar{V} \rangle / T$ as a function of the searching time T in the case of three searching strategies. Right panel: mean information $\langle \bar{I} \rangle$ as a function of T , again for the three best searching strategies. | 41 |
| 3.8 | Number of nodes of degree k as measured by a walker searching during $T = 2000$ (void circles) or $T = 10000$ (filled circles) time steps, in a single realization of the process. | 42 |

| | | |
|------|---|----|
| 3.9 | Kullback Divergence (D_{KL}) of the measured (normalized) degree distribution with respect to the real one. | 43 |
| 4.1 | The model of interaction between oscillators, based on geometrical constraints. Only those within a distance R are affected by the firing of the central one. | 49 |
| 4.2 | On the left: Heat map of the synchronization time as a function of r and V ; in the top picture a large region of the parameters space has been considered, while the bottom figure is a zoom on the most critical region ($r < 0.08$, $V < 0.25$). On the right: Profiles of the efficiency of the system rT_{sync} against r , for several values of V | 50 |
| 4.3 | Panel a): η against χ for several values of r and V . Panel b): the difference between the two order parameters ($\eta - \chi$) as a function of rT | 51 |
| 4.4 | On the left: the average value of χ at the synchronization time (χ_{sync}) as a function of r for several values of V . On the right: χ_{sync} as a function of rT_{sync} for several values of V . All the means have been performed over 200 realizations. | 53 |
| 4.5 | On the left: The cost rT_{sync} as a function of r and V . The heat map of χ_{sync} has been superimposed to the surface $rT_{sync}(r, V)$. On the right: the same quantities, zoomed in the most sensitive region of the parameter space. | 54 |
| 4.6 | Snapshot of the system in the incoherent state. Each disk is an oscillator, gray dashed arrows stand for first-neighbor relationships, while the continuous black one represents a firing event that is taking place at that precise instant. | 55 |
| 4.7 | The average synchronization time T_{sync} as a function of V , for $L = 400$, $N = 20$, $\epsilon = 0.1$ | 57 |
| 4.8 | Final ($T = T_{sync}$) network of the interactions mediated by a single oscillator (labeled "0"), in the fast limit, at $V = 2V_f$ (panel A) and at $V = V_m$ (panel B) respectively. | 58 |
| 4.9 | Final ($T = T_{sync}$) total interaction networks, in the fast limit (panel A) and at $V = V_m$ (panel B) respectively. | 59 |
| 4.10 | Global (η , blue line, left axis) and Individual (λ , red line, left axis) order parameters are plotted together with the Individual Mixing (m , black line, right axis) as a function of time, from $T = 0$ to the synchronization time. | 60 |
| 4.11 | Synchronization time dependence on the velocity in the right region. Panel A: each curve corresponds to a different value of ϵ , while the size of the system is kept fixed at $N = 20$. Panel B: each curve corresponds to a different value of N , while the coupling strength is $\epsilon = 0.1$ | 62 |
| 4.12 | Panel A: synchronization time at $V = V_f$ (red filled circles) and $V = L/T$ (blue empty squares) as a function of ϵ , for $N = 20$. Panel B: synchronization time at $V = L/T$ against $x = \epsilon^\beta/N^a$, with $a = (0.38 \pm 0.01)$, for several value of N | 62 |
| 4.13 | T_{sync} dependence on V in the slow regime for several values of N | 64 |
| 4.14 | Average synchronization time for a group of $N = 3$ static oscillators dependence on the velocity. The black line stands for a fit function $f(\epsilon) = \kappa/\epsilon$ | 66 |
| 4.15 | Rescaled synchronization time $T_{sync}(r) = (v/r)T_{sync}(v)$ against r in the slow regime for several values of N and ϵ | 67 |
| 4.16 | Coefficient c dependence on ϵ for several values of N | 68 |
| 4.17 | Rescaled synchronization time $T_{sync}(r) = (v/r)T_{sync}(v)$ times the empirical scaling function $\epsilon^{\Lambda(N)}$ against r for several values of N and ϵ | 69 |

| | |
|---|----|
| 4.18 Rescaled synchronization time $T_{sync}(r) = (v/r)T_{sync}(v)$ times the empirical scaling function $N \log(N)\epsilon^{\Lambda(N)}$ against r for several values of N and ϵ | 70 |
| 4.19 Efficiency dependence on the velocity for $N = 20$, $\epsilon = 0.1$, and $\lambda = 1$ (red circles) and $\lambda = 10$ (blue squares). | 73 |
| A.1 Cayley tree with coordination number $q = 3$ and radius $R = 5$. The pacemaker is located at the root of the tree (filled circle). | 81 |
| A.2 (Left) Effective Frequency Dispersion Order Parameter Δ_ω (A.2.2) vs. time. (Right) Normalized Frequency Dispersion Order Parameter r_ω (A.2.3) vs. natural frequency of the pacemaker ω | 82 |
| A.3 Values taken in the complex plane by parameter (A.3.9) for a Cayley tree for radius $R = 4$ and coordination number $q = 3$, during time windows of increasing duration. The pacemaker natural frequency is $\omega = 1.01\omega_c$ | 85 |
| A.4 Curves described in the complex plane by the order parameter (A.3.10) for a star of degree $q = 3$ and different value of the pacemaker natural frequency. That corresponding to the critical value is analytically computed while the other ones are obtained by numerical simulations. | 86 |
| A.5 Time evolution of the effective frequencies of the pacemaker (up) and its neighbors (down) in star of coordination number $q = 3$ for different value of the pacemaker natural frequency $\omega > \omega_c$: $\omega = 1.01\omega_c$ (black line), $\omega = 1.05\omega_c$ (red line), $\omega = 1.20\omega_c$ (blue line). | 87 |
| A.6 Comparison of the time evolution of the effective frequencies of the pacemaker (up) and its neighbors (down) in Cayley trees of coordination number $q = 3$ and different radius. The radii are respectively $R = 1$ (black line, analytically computed), $R = 2$ (red line), $R = 3$ (blue line) and $R = 8$ (pink line). | 87 |
| A.7 Comparison of the effective frequencies time evolution in a Cayley tree of radius $R = 4$ and coordination number $q = 3$. The red line (left panel) corresponds to the pacemaker, the blue one to the nodes of the first shell, the other to the second one (green) and the third one (black). In the right panel we have removed the pacemaker effective frequency zooming on the other ones in order to show the hierarchical organization of the amplitudes of the frequencies oscillation. The pacemaker natural frequency was 20 times larger than its critical value. | 88 |
| B.1 The average symmetry $\langle s \rangle$ (empty circles, left axis) and the global final mixing χ (filled circles, right axis) are plotted against the velocity V . The grey rectangle marks the no-synchronization region where no empirical data is available. | 93 |
| B.2 Fully connected topology. Synchronization time dependence on ϵ for the same value of N | 94 |
| B.3 Connected cluster of size $N = 2, 3, 4$ | 95 |
| B.4 Synchronization time as a function of ϵ for a 3-chain (green), a triangle (blue), a 4-chain (orange) and a 4-star (black). In the inset below, the synchronization time dependence on ϵ for a pair of reciprocal neighbors. | 96 |
| B.5 Oscillator i exiting from a square of side l . The green circles of radius V and center in i marks the possible positions that the oscillator will take at time $T + 1$ | 98 |

| | | |
|-----|--|-----|
| B.6 | The average out-neighbor time T_{out} (filled symbols, left axis) and the average in-neighbor time T_{in} (empty symbols, right axis) are plotted against the quantity $V\sqrt{N-1}/L$, for some values of N (red: $N = 10$, black: $N = 20$, blue: $N = 30$) and L (circles: $L = 200$, squares: $L = 400$, triangles: $L = 800$). | 99 |
| i | Mi tesis en un diagrama. | 108 |
| ii | Ejemplo de un dendograma representante la estructura jerárquica de una red, obtenido mediante medidas de correlaciones dinámica justo por encima del punto crítico. | 111 |
| iii | Diagrama de flujo del algoritmo adaptativo. | 113 |
| iv | Tiempo de sincronización en función de la velocidad para el segundo modelo considerado (interacción a primeros vecinos). | 115 |

List of Tables

2.1 Results of the reconstruction on several networks. The columns give the size of the system N , the total number of links in the original network L , the total error in the estimation of the degrees $K_{err} = \sum_{i=1}^N |k_i - k_i^*|$, the total number of links in the reconstructed network before the removal of exceeding links L' , the number of false positive F_p and false negative F_n links in this network, the same for the final reduced network L'_r , F_p/F_n and the final total error $err(\%)$. From the first row, the networks are our usual benchmark [1], a ring of 6 cliques of 3 nodes [2], a hierarchical network of 25 nodes and 2 levels [3], Zachary club social network [4], ring of 16 cliques of 3 nodes [2], the cortical brain network of a cat [5], a hierarchical network of 125 nodes and 3 levels [3], a network of 4 communities of 32 nodes each [4], and two networks of 3 levels of community structure [6]. 23

Chapter 1

Introduction

For centuries, scientific knowledge has been identified with the possibility to predict the future evolution of all phenomena starting from the knowledge of the laws governing them, and the objective of every scientific effort was considered to be the discovery of those laws. This program has been deeply influenced by the great success that Physics has had since the 16th century. Physics, an experimental discipline in which theoretical predictions are compared to experiments, differs from other sciences of nature in the crucial role of mathematics. Physicists describe the phenomena by using mathematical models and their predictions are obtained by mathematical reasoning [7]. The trajectory of a bullet, the velocity of a free falling ball, the time necessary to cover a given orbit, have been the first quantities that was possible to calculate with arbitrary precision, by applying the laws of dynamics and universal gravitation formulated by Newton. The empirical success of this method has progressively strengthen the belief that all physical phenomena, despite their heterogeneous appearance, could be explained in terms of *simple* and *universal* laws, that allow to predict, by means of calculations, the sequence of events that are going to occur in a given system, whenever the initial conditions are known [8].

At the beginning of the 19th Century, Laplace was writing that an infinitely intelligent mathematician would be able to predict with certainty the future by observing the state of the universe now, and using its knowledge of the laws of motion [7]. According to this view, in order to understand a system, it is necessary to decompose it into its minimal components. On the contrary, a detailed representation is useless and misleading since accidents do not change the *nature* of the phenomenon and, therefore, its evolution. Once performed this *cleaning* process, everything becomes simple and predictable. This view, generally labelled as *reductionist*, ruled the scientific thought during the most part of the science history and it reached its peak during the 20th century with the appearance of molecular biology and particle physics. Its role in the development of modern science is undeniable. Thanks to this approach, modern Physics has achieved highest precision level in the mathematical description of the reality, basing its description on an appropriate representation of the matter in terms of atomic and sub-atomic structure. Moreover, the same method allowed to get to the structure of biological

macromolecules and the logic of their functions. In those years, large part of the scientific community shared the view of Francis Crick, who thought that knowing “in all their details all the chemical steps that take place in the cell during the cellular cycle” was just a matter of time, and then “there will be nothing more to know about the cell itself, and mechanism of its living will be completely decoded”.

However, in the 1960s things started to change. Whereas among the great majority of active scientists it was accepted that “the workings of our minds and bodies, and of all the animate or inanimate matter of which we have any detailed knowledge, are [...] controlled by the same set of fundamental laws” [9], it was also becoming clear that understanding phenomena at higher scale starting from a lower one is often impossible. As stated by P. W. Anderson in 1972 [9],

“the reductionist hypothesis does not by any means imply a *constructionist* one: the ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe. [...] The constructionist hypothesis breaks down when confronted with the twin difficulties of scale and complexity. The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead, at each level of complexity entirely new properties appear, and the understanding of the new behaviors requires research which I think is as fundamental in its nature as any other [... and will show] how the whole becomes not only more than the sum of but very different from the sum of the parts.”

Systems displaying such features are generally known as complex systems, i.e. systems made up by many interacting elements. While, traditionally, fundamental science explores the very small and the very large, both of which lie beyond man’s everyday perception, the uniqueness of complex systems is that they have to do with a class of phenomena of fundamental importance in which the system and the observer may evolve on comparable time and space scales [10]. Examples of complex systems for which complexity models have been developed include ant colonies, human economies and social structures, climate, nervous systems, cells and living things, including human beings, as well as modern energy or telecommunication infrastructures. It is understood that such systems show emergent dynamical properties which are inherently related to the topology of the underlying pattern of connections among the constituent parts [11]. In the words of the biologist Brian Goodwin [12],

“The important properties of these complex systems are found less in what they are made of than in the way the parts are related to one another and the dynamic organization of the whole - their relational order [...] To understand these complex nonlinear dynamical systems it is necessary to study both the whole and its parts, and to be prepared for surprises due to the emergence of unexpected behavior.”

Hence, on the one hand, it is not possible to explain an organizational level only in terms of the lower ones. On the other hand, it seems clear that, more than details about the *nature* of the components, what we need is a map of the interactions. Therefore, complex systems are in

general suitably described through their networks of contacts, that is, in terms of nodes (representing the system's components) and edges (standing for their interactions), which allows to catch their essential features in a simple and general representation.

In 1998 Watts and Strogatz [13] presented the first evidence of what they called complex networks. At the time, networks of coupled dynamical systems had already been used to model biological oscillators, Josephson junction arrays, neural networks, genetic control networks and many other self-organizing systems. But, whereas ordinarily the connection topology was assumed to be either completely regular or completely random, they realized that “many biological, technological and social networks lie somewhere between these two extremes.” This was the case, for instance, of the neural network of the worm *Caenorhabditis elegans*, the power grid of the western United States, and the collaboration graph of film actors. All these patterns of connections showed common features that could not be explained by simple mathematical models. In particular, the presence of short-cuts made these networks extremely small in terms of the mean distance between nodes, while at the same time kept track of their ordered origin through a larger than expected clustering coefficient (related to the existence of triangles in the network) [11]. These two phenomena could not be explained simultaneously by the models used until the time (completely ordered or completely disordered). Slightly later, Barabasi and Albert [14], showed that the distribution in the number of connections coming out from a given node (e.g. computers in the Internet, pages in the world-wide-web, or number of papers authored by scientists) is skewed, which means that, whereas the large majority of the elements have very few connections, there exist some actors in the networks that are highly connected [11]. Since then, complex networks [15, 16, 17, 18] have become an important, largely used, framework for the understanding of both dynamical and topological aspects of systems such as the brain [19], protein-protein interaction networks [20], Internet and the WWW [21].

At the beginning, when complex networks was a new born scientific field, almost all the research effort was directed to map new systems and to analyze topological features of the resulting graphs. It was so possible to identify important characteristics shared by systems belonging to the same class. Increasing interest on this approach, thanks also to a favorable technological progress, led to the accumulation of an increasing amount of data. Usual network datasets encode the information about which pairs of nodes are connected and which are not, sometimes with some additional specification about the strength of the connections or the *importance* of the units. Nowadays, there is a huge amount of such data available. This situation has allowed the arising of new questions and, therefore, the diversification of the scientific work. Among them, we can point out three general issues that have been receiving a lot of interest: (i) is the available information always reliable and complete? (ii) how does a complex interaction pattern affect the emergence of collective behavior in complex systems? And (iii) which is the role of mobility within the framework of complex networks?

- i)** For what concerns question (i), nowadays it is well acknowledged that many of the mentioned networks are only partially known. Think, for instance, in artificial networks like on-line social networks or Internet itself, which are made up of millions of heterogeneous and non-identical nodes. In such large networks, a complete map is hardly available and difficult to get [22]. Thereby, providing efficient tools for their exploration has become a crucial challenge. On the other hand, in many situations, connections among elements are not directly accessible, in the sense that they are not physical objects like a cable, a hyper-link or an axon. Many natural networks, as genetic, brain or ecological networks, have to be understood as representations of the relations existing between different parts of the considered system. In all these cases, in order to map out a connectivity pattern, edges have to be inferred by measuring correlations of quantities related to the behavior of the units. Obviously, different methods and experiments can give rise to different results. This is the case, for instance, of gene regulatory networks. In order to construct connection patterns more and more complete, it is usually necessary to combine data obtained by means of different experimental procedures, not always coherent among them. Therefore, in recent years, the problem of the inference of network topologies from dynamical measures has been subject of intense investigation, both from an experimental and a theoretical perspective. Recently, it has also been presented a general mathematical and computational framework to deal with the problem of data reliability in complex networks [23]. In particular, this approach claims to reliably identify both missing and spurious interactions in noisy network observations.
- ii)** About question (ii) we have to say that the fact that the topology of the connection pattern affects the dynamical evolution of the elements is precisely what allows the reconstruction of many networks. Hence, it is straightforward that collective behaviors, that is, behaviors which characterize the system as a whole and which exist thanks to the interactions among the parts, may change depending on the features of the connection topology. Currently, much scientific effort is devoted to understanding how global dynamical properties are related with the units dynamics and the interactions between them [24]. One of the first studied examples of this relation was the transition between order and chaos that Random Boolean networks [25] display when increasing the of inputs (connections) to each node. More recently, the impact of the network topology has been object of much scientific interest in connection with, just to mention one among many other phenomena, the spreading of diseases [26]. However, many other important examples can be provided.
- iii)** There are two types of mobility that can be considered in the framework of complex networks. On the one hand, we can have different kinds of walkers that move throughout the network performing a large number of possible tasks. Agent-based models have been used to study phenomena such as epidemics spreading, opinion formation or Inter-

net traffic. On the other hand, in some cases the network is generated by the interactions between moving agents that are themselves the nodes of the connectivity pattern and the network topology evolves as an effect of the mobility. This situation is particularly interesting since it is common to many artificial as well as natural systems and applications range from robotic [27, 28] to ecology [29].

This thesis has been developed along these three lines, which are strictly interrelated. We expand on three case-studies, each one of which deals with two the above mentioned macro-issues. We consider the issue of the incompleteness of the available information both in the case of natural (Chapter 2) and artificial (Chapter 3) networks. As a paradigmatic emergent behavior, we focus on the synchronization of coupled phase oscillators (Chapter 2 and Chapter 4), deeply investigating how different patterns of connections can affect the achievement of a globally coherent state. Finally, we include moving agents in two different frameworks, using them as explorers of unknown networks (Chapter 3) and considering them as interacting units able to establish connections with their neighbors (Chapter 4).

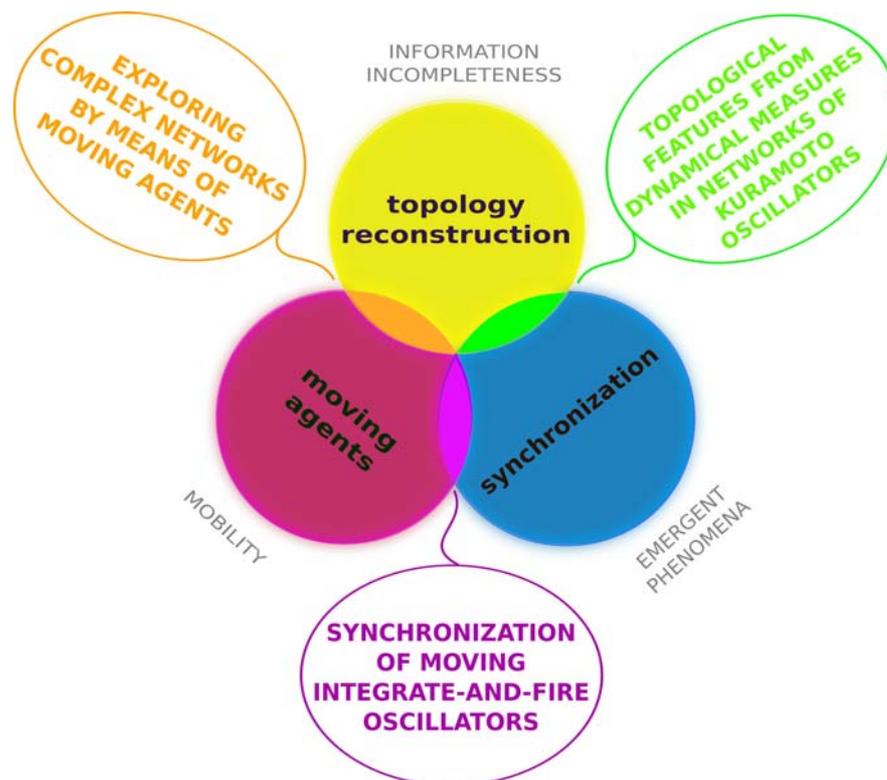


Figure 1.1: My thesis in a diagram.

If emergence refers to “the arising of novel and coherent structures, patterns and properties during the process of self-organization in complex systems”, and “emergent phenomena are conceptualized as occurring on the macro level, in contrast to the micro-level components

and processes out of which they arise” [30], then synchronization can be regarded as one the most interesting, abundant and well defined emergent behavior. First recognized in 1665 by Christiaan Huygens, synchronization phenomena take place in nature, engineering and social life. “Synchronous variation of cell nuclei, synchronous firing of neurons, adjustment of heart rate with respiratory and/or locomotory rhythms, different forms of cooperative behavior of insects, animals and even humans”, are universal phenomena and can be understood within the common framework based on modern nonlinear dynamics [31]. Our surrounding are full of oscillating objects. Systems as diverse as an applauding audience and a radio communication equipment share the common feature of producing rhythms. Usually these systems are not closed and interact with other objects. “This interaction can be very weak [...] but nevertheless it often causes a qualitative transition: an object adjusts its rhythm in conformity with the rhythms of other objects. [...] This adjustment of rhythms due to an interaction is the essence of synchronization” [31].

In this thesis we have considered coupled oscillators with fixed amplitudes that interact via mutual adjustment of their phases. In Chapter 2, we study the problem of the reconstruction of an unknown interaction network, whose nodes are Kuramoto oscillators. The Kuramoto model, first proposed by Yoshiki Kuramoto [32], provides a paradigmatic example of non-equilibrium transitions between an incoherent and a synchronized state. Its formulation was motivated by the behavior of systems of chemical and biological oscillators, and it has found widespread applications such as in neuroscience. The model makes several assumptions, including that there is weak coupling, that the oscillators are identical or nearly identical, and that interactions depend sinusoidally on the phase difference between each pair of objects. We analyze populations of almost identical oscillators in arbitrary interaction networks. Our aim is to extract topological features of the connectivity pattern from purely dynamical measures, based on the fact that in a heterogeneous network the global dynamics is not only affected by the distribution of the natural frequencies but also by the location of the different values. In order to perform a quantitative study we focused on a very simple frequency distribution considering that all the frequencies are equal but one, that of the pacemaker node. We then analyze the dynamical behavior of the system at the transition point and slightly above it as well as very far from the critical point, when it is in a highly incoherent state. The gathered topological information ranges from local features, such as the single node connectivity, to the hierarchical structure of functional clusters, and even to the entire adjacency matrix.

In Chapter 4, instead, we present a model of integrate and fire oscillators that are moving agents, freely displacing on a plane. The phase of the oscillators evolves linearly in time and when it reaches a threshold value they fire at their neighbors choosing them according to a certain interaction range (Sec. 4.1) or simply selecting the one at the minimal distance (Sec. 4.2 and following ones). In this way, the interaction network is a dynamical object by itself since it is re-created at each time step by the motion of the units. Depending on the velocity of

the motion, the average number of neighbors, the coupling strength and the size of the agents population, we identify different regimes. We characterize these regimes in terms of the time the system needs in order to reach the coherent state, that in some cases we are able to predict, also providing a detailed even if qualitative description of the mechanisms that enhance the synchronization.

Moving agents are employed also in Chapter 3 where they play the role of explorers of unknown artificial networks, having the mission to recover information about their structures. Usually, the problem of network reconstruction has been address by means of algorithms based on search and traffic routing [33, 34, 35], in many cases performed by moving “agents” which explore the topological space. We propose a model in which random walkers with previously assigned home nodes navigate through the network during a fixed amount of time. We consider that the exploration is successful if the walker gets the information gathered back home, otherwise, no data is retrieved. Consequently, at each time step, the walkers, with some probability, have the choice to either go backward approaching their home or go farther away. We show that there is an optimal solution to this problem in terms of the average information retrieved and the degree of the home nodes and design an adaptive strategy based on the behavior of the random walker. Finally, we compare different strategies that emerge from the model in the context of network reconstruction.

Chapter 2

Extracting topological features from dynamical measures in networks of Kuramoto oscillators

Currently, it is widely acknowledged that complex patterns of interaction are as ubiquitous in nature as in society [14]. Nonetheless, further research is required to completely understand how the topology affects the system dynamics [15, 16], particularly how global dynamical properties are related with the units dynamics and the interactions between them. A unique answer cannot be provided since complex networks respond differently depending on the dynamical processes that take place within them [24].

One of the most interesting of these macroscopically defined dynamical processes is synchronization, an emerging phenomenon in which populations of interacting units display a common periodic behavior [31, 36].

Understanding the role of connectivity in synchronization has been the subject of intense research in recent years [37]. On the one hand, much work has focused on the generic properties of dynamical systems, mainly looking for necessary and sufficient conditions that would grant that a population of units under a set of simple dynamical rules is able to synchronize [38]. On the other hand, much progress has been made by studying precise models of phase oscillators, one of the most paradigmatic being the model proposed by Kuramoto [32, 39], where the interaction between the units is proportional to the sine of the phase difference.

Here, we will continue along this line and analyze a population of Kuramoto oscillators with a precise distribution of frequencies. The original work by Kuramoto and many subsequent studies considered that the oscillators, each coupled equally to all the others, had natural frequencies taken from a given distribution. The non-zero width of those distributions made the units follow different trajectories, whereas the interaction term made their phases approach. In fact, depending on the width of the frequency distribution, there is a critical value of the interac-

tion strength above which the units tend to entrain their phases and hence leave the incoherent regime. If the natural frequencies of the oscillators are identical, a unique outcome is possible as the only attractor of the dynamics is a completely synchronized state in which all the oscillators end up in a common phase. And this occurs for any initial conditions and for any (connected) topology ¹.

In systems with regular patterns of connectivity (including all to all) the only complexity comes from the frequency distribution, whereas in more realistic (non homogeneous) patterns, not only the frequency values matter but also the precise location as well [41, 16].

We will focus on a particular frequency distribution, one that is just one step away from the homogeneous case. Such a distribution has identical frequencies for all oscillators except one. This singular oscillator, with a higher frequency than the rest, has received the name *pacemaker*, and its effect in populations of Kuramoto oscillators has been analyzed [42, 43]. In [42], Kori and Mikhailov consider a special case where the pacemaker affects its neighbors but is not affected by them; under these conditions they find numerically that the range of frequencies of the pacemaker for which the system can attain global synchronization depends on the “depth” of the network, where the depth is defined as the maximum distance from the pacemaker to peripheral nodes. Radicchi and Meyer-Ortmanns [43] consider regular structures for which the conditions to synchronize can be analytically computed.

We use several properties of the heterogeneity induced by the existence of the pacemaker to find useful relations between topology and dynamics. On one hand, by knowing the topology one should be able to infer the dynamical properties of the network. On the other hand, by performing dynamical measures some structural properties can be inferred, and this will be our purpose.

First, we use a similar procedure to the one used in [42] and [43], showing that there is a critical value for the frequency of the pacemaker above which the (frequency) synchronized state cannot exist. This is related with the existence of a synchronized solution (also exploited in [44]) that applies to any subset of oscillators. We find, however, that from a practical point of view the most restrictive condition is usually for the equation of the pacemaker that involves its connectivity, and hence there is a clear relationship between the critical frequency and the pacemaker connectivity that can be used as an experimental measure of the degree.

In order to get more details on the network structure we analyze the system above the critical value where correlations between dynamical evolution of the nodes appear. Such correlations enable us to reveal the hierarchical organization and to recover the network connectivity.

First, in Sec. 2.1 we characterize the coherent state and the transition to the incoherent one by means of a proper definition of the order parameter. Then, in Sec. 2.2 we qualitatively analyze the behavior of the system when it is not in the frequency-locked state. Sec. 2.3 is

¹There are, however, some limitations to these results; for instance when the units are placed in regular lattices, such as one-dimensional rings, where other attractors different from the synchronized (equal phases) state, may arise [40].

devoted to studying the relation between local connectivity and the ability of the system to reach a synchronized (frequency-locked) state. In Sec. 2.4 we focus on the system slightly above the transition toward the incoherent state. We show that it is possible to perform some hierarchical analysis concerning the connectivity network. Finally, in Sec. 2.5 we study the system far above the critical point, in a regime characterized by short range correlations where it becomes easy to identify the nodes directly connected to the pacemaker. Thus the reconstruction of the whole connectivity pattern is accurate and fast.

2.1 Synchronization and phase transition

In the original Kuramoto model [32, 39], the phases of the oscillators evolve according to the following equation

$$\dot{\varphi}_i = \omega_i + \sigma \sum_{j=1}^N \sin(\varphi_j - \varphi_i), \quad (2.1.1)$$

where N is the total number of units of the system, ω_i is the natural frequency of unit i , taken from a distribution, and σ stands for the coupling strength. This case corresponds to a fully connected topology; i.e., each unit interacts with all the other ones. The ability of the system to reach a coherent state, for a given coupling strength, depends only on the width of the distribution of natural frequencies.

Here we want to consider arbitrary connectivity patterns. In this situation, the behavior of the system can no longer be understood in terms of the ratio between the distribution width and the coupling strength only. Where the natural frequencies values are located is also relevant, since on a generic interaction network nodes are no longer equivalent.

From now on we are using the two-level hierarchical network of nine nodes represented in Fig. 2.1 as a benchmark, and, when not otherwise stated, all the figures refer to that connection pattern. This network has been presented in [1] as a very simple example of the class of deterministic scale-free hierarchical networks proposed by Ravasz and Barabasi in [3]. We choose this small regular connectivity pattern as a simple paradigmatic example showing general properties of the studied systems since it makes it easy to recognize the role of each node.

Let us rewrite the equation for the evolution of the phases including a symmetric connectivity matrix a_{ij} that takes a value of 1(0) if node i and j are connected (disconnected):

$$\dot{\varphi}_i = \omega_i + \sum_{j=1}^N a_{ij} \sin(\varphi_j - \varphi_i), \quad (2.1.2)$$

where we have rescaled time by setting $\sigma = 1$. Now we consider all the oscillators to have the same natural frequency (0, without loss of generality), except one of them, called the pacemaker, whose frequency is $\omega \neq 0$. It is precisely this extremely simple choice of frequencies that enables us to study the roles played by individual oscillators.

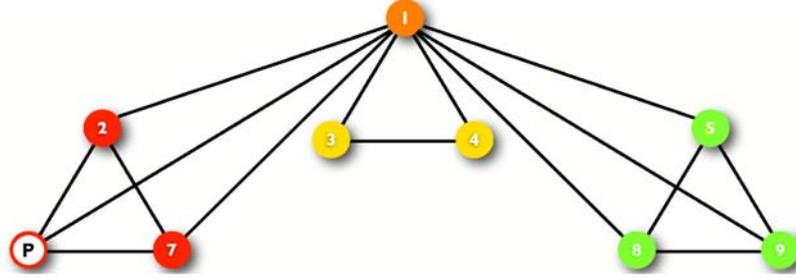


Figure 2.1: Hierarchic network that will be used as a benchmark. In this particular setting the pacemaker is located on a peripheral node (marked as P) of degree $k_p = 3$. The other nodes are grouped into sets using different colors. The elements of each set are topologically equivalent if we look at the network from the point of view of the pacemaker. Consequently their dynamical evolution is identical.

If a stationary state exists, then all the effective frequencies will take constant values and the following conditions have to be satisfied:

$$\sum_{j=1}^N a_{ij} \sin(\varphi_j - \varphi_i) = \Omega_i \quad \forall i \neq p \quad (2.1.3)$$

$$\omega + \sum_{j=1}^N a_{pj} \sin(\varphi_j - \varphi_p) = \Omega_p \quad (2.1.4)$$

where $\{\Omega_i\}$ are the effective frequencies of the oscillators. Notice that summing up Eqs. (2.1.3) and (2.1.4) the coupling terms cancel because of the symmetry of the interaction and it results in

$$\sum_{i=1}^N \Omega_i = \omega. \quad (2.1.5)$$

Looking at Eqs. (2.1.3) and (2.1.4) it is easy to recognize that there is an interplay between two effects. On the one hand the width of the frequencies distribution (in our present case this role is played by ω itself) tends to keep the evolution of the oscillators apart since each one follows its natural frequency. On the other hand, the interaction term makes them approach their phases as well as their effective frequencies. Then we conclude that if the pacemaker's natural frequency is small enough, the interaction term dominates and, after a transient time, all effective frequencies Ω_i will be identical:

$$\Omega_i = \omega/N \quad \forall i, \quad (2.1.6)$$

including the pacemaker. In this case we can say that the system is in a frequency-locked state since all oscillators have the same frequency, although the phases are not equal because there is a coupling term (that of the pacemaker) that cannot vanish.

When increasing the pacemaker frequency ω , some oscillators cannot keep the phase difference, and the frequency-locked state is broken. The left-hand side of Eq. (2.1.3) is indeed bounded because of the sine terms, whereas the right side increases as the pacemaker frequency is increased. A similar conclusion can be deduced from Eq. (2.1.4). Consequently, there will be a transition from a synchronized to an incoherent state. Thus we can define the critical value ω_p^c as the maximum value of the natural frequency of the pacemaker for which the system can attain global synchronization.

Such a transition for a population of phase oscillators is typically characterized by an order parameter R , defined through the equation $R e^{i\Psi} = \sum_j e^{i\varphi_j}$, where Ψ is a global phase (not constant) [45].

Here, following [41, 46], we adopt another order parameter that is a normalized measure of the effective frequency dispersion (standard deviation):

$$r_\omega = \sqrt{\frac{\sum_{i=1}^N [\dot{\varphi}_i / \langle \omega \rangle - 1]^2}{N - 1}}, \quad (2.1.7)$$

where $\langle \omega \rangle$ is the average effective frequency of the oscillators population, a constant quantity always equal to ω/N . According to its definition, r_ω takes values in the interval $[0, 1]$ (see Fig. 2.2). It should be noticed that, since above the critical frequency the system is not able to reach a steady state any longer, calculation of the order parameter (2.1.7) requires performing averages over an appropriate time window. Anyway, the value of $\langle \omega \rangle$ does not change because what we found in (2.1.5) is a general result, even for instantaneous values of the effective frequencies.

To find the precise value of the critical frequency we apply the Newton-Raphson method (NR) and check, as a function of the frequency ω , whether the synchronized solution of Eqs. (2.1.3) and (2.1.4) exists. To simulate the dynamics of the system in the incoherent state ($\omega > \omega_p^c$) we take as initial phases $\{\varphi_i(0)\}$ the stationary values of the differences provided by the NR solution for $\omega = \omega_p^c$. The system of equations (2.1.2) is numerically integrated with Euler's method (first order), unless otherwise stated, at fixed time step $\delta t = 10^{-2}$.

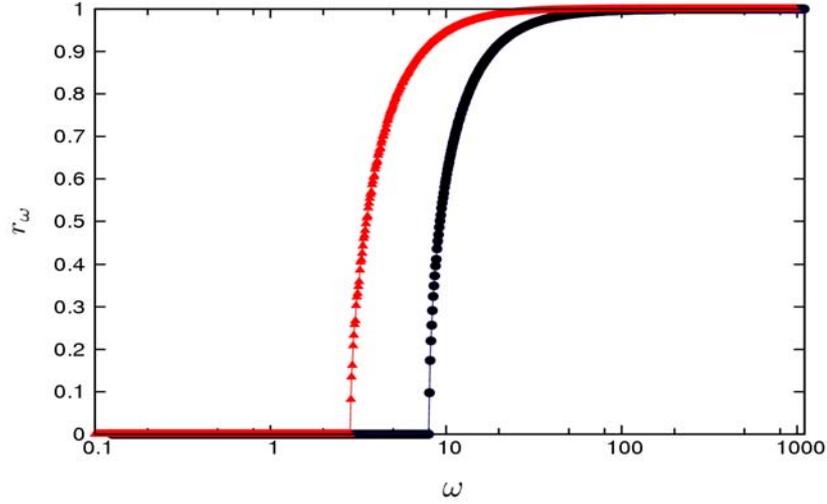


Figure 2.2: Order parameter (2.1.7) as a function of the natural frequency of the pacemaker. Different curves correspond to different settings: circles refer to the pacemaker located on node 1 in Fig. 2.1 ($k_p = 8$), triangles refer to the pacemaker on node 2 ($k_p = 3$). The average value $\langle r_\omega \rangle_t$ for $\omega > \omega_p^c$ was calculated on a time window $\Delta t = 100$.

2.2 Incoherent state

Above the critical frequency ω_p^c the system is no longer in a stationary state, and hence the effective frequencies are no longer constant.

Numerical simulations show that, after a transient time, the system enters into a “periodic” state (see Fig 2.3). The features of this periodic state are not affected by the initial conditions, and they only depend on the pacemaker frequency and location. It is precisely this fact that enables us to infer topological properties from dynamical measurements.

Figure 2.4 summarizes what we have learned up to now, shedding light on some interesting details. The time average of the effective frequency of the pacemaker $\langle \dot{\varphi}_p \rangle_t$ and that of one of its neighbor $\langle \dot{\varphi}_j \rangle_t$ are plotted as functions of the pacemaker natural frequency. These quantities are calculated from numerical simulations, taking into account appropriate time windows.

Starting from small values of ω , the picture shows how all the effective frequencies increase together linearly, following the reference line $\Omega_i = \omega/N$ defined by Eq. (2.1.6). Then, when ω reaches the critical value ω_p^c , they do separate. Initially, the average effective frequency of the pacemaker goes through a more than linear increasing, while the others start decreasing, keeping their (average) values very close to each other. For even larger values, when $\omega \gg \omega_p^c$, Fig. 2.4 shows how the average effective frequency $\langle \dot{\varphi}_p \rangle_t$ tends to ω , asymptotically increasing along a new reference line with a slope equal to 1. At the same time, $\langle \dot{\varphi}_i \rangle_t$ for $i \neq p$ goes to zero, as required by the conservation law (2.1.5).

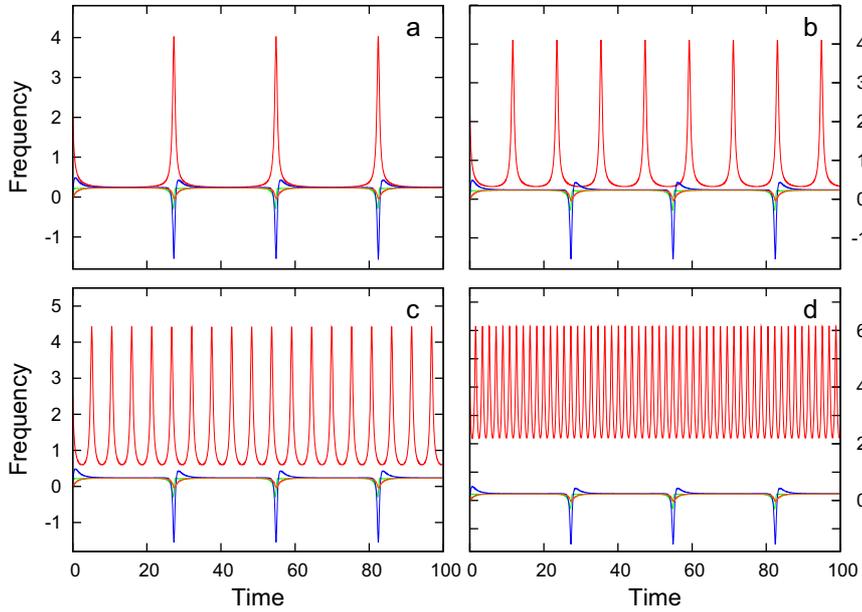


Figure 2.3: Effective frequencies above the critical point as functions of time. That of the pacemaker (red top curve), in this particular setting located on node 3 in Fig. 2.1, is on average much larger than the others (lower curves). The plots show a pacemaker natural frequency value that is (a) 1.01, (b) 1.05, (c) 1.2, and (d) 2 times its critical value. Time starts after a transient lag $T_s = 10$.

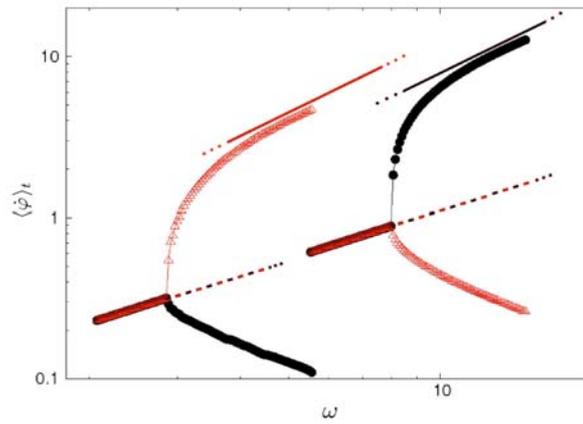


Figure 2.4: Average effective frequencies as a function of the natural frequency of the pacemaker ω . The behavior of two oscillators is shown: node 1 (black circles) and node 2 (red triangles). On the right side the pacemaker is node 1, and on the left one it is node 2. Initially the frequencies are synchronized and they increase linearly with a slope of $1/N$ (dashed line), as expected from Eq. (2.1.6). Then, when ω reaches the critical value, which is different for different location of the pacemaker, they grow apart. Far above the critical values, the average frequency of the pacemakers approach asymptotically a new reference line with a slope of 1 (solid line). The time averages were performed on a time window $\Delta t = 100$.

2.3 Critical frequency and local topology

In this section we explore the relation between the topology of the network and the value of the critical natural frequency of the pacemaker depending on the node where it is located.

Let us begin by writing the equation for the pacemaker in the synchronized state. As a consequence of Eq. (2.1.6), we have

$$\omega + \sum_{j=1}^N a_{jp} \sin(\varphi_j - \varphi_p) = \omega/N. \quad (2.3.1)$$

This equation links the natural frequency of the pacemaker to the constant values of the phase differences between it and its neighbors, when all the units are oscillating with the same effective frequency. Since the number of non-null terms a_{jp} in the previous expression is given by the number of nodes connected to the pacemaker and $\sin(\varphi_j - \varphi_i) \in [-1, 1]$, the degree (or connectivity) of the pacemaker is a bound for the absolute value of the sum in Eq. (2.3.1).

Thus there is an upper bound for the critical frequency:

$$\omega_p^c \leq k_p \frac{N}{N-1}, \quad (2.3.2)$$

where k_p is the degree of the pacemaker. Indeed, any value larger than the right-hand term in inequality (2.3.2) is surely unable to satisfy Eq. (2.3.1), and hence the system is unable to be frequency synchronized.

Notice that we have obtained this bound by taking into account a single equation, that of the pacemaker. We can write for any oscillator the analog of Eq. (2.3.1) as follows

$$\sum_{j=1}^N a_{ji} \sin(\varphi_j - \varphi_i) = \omega/N, \quad \forall i \neq p. \quad (2.3.3)$$

It is easy to verify that no stricter condition can arise from any of these equations². However, stronger bounds could exist due to the combination of Eq. (2.3.1) and some of Eqs. (2.3.3).

Let us consider a set of $(n+1)$ connected nodes, among which the pacemaker is included³. Labeling them by an increasing index $i = 1, 2, \dots, n+1 = p$ and summing up their equations, we obtain

$$(n+1) \frac{\omega}{N} = \omega + \sum_{i=1}^{n+1} \sum_{j=1}^N a_{ij} \sin(\varphi_j - \varphi_i). \quad (2.3.4)$$

If two nodes in the considered group are neighbors their respective interaction terms cancel

²Applying the same argument to the eq. of the i -th node, we obtain $\omega_p^c \leq k_i N$, whose smallest possible value is N , that is the largest possible value for the bound (2.3.2).

³It is not necessary to take into account the groups that do not include the pacemaker, since the bound we obtain for ω_p^c summing up $n+1$ equations including the pacemaker or the remaining $N-n+1$ (not including the pacemaker), is the same.

each other. So the number of remaining terms of the sums in Eq. (2.3.4) is given by

$$K_{out} = \sum_{i=1}^{n+1} k_i - \sum_{i,j=1}^{n+1} a_{ij} \quad (2.3.5)$$

where k_i is the degree of the i th node and K_{out} is equal to the number of links connecting the nodes of the considered set to external ones.

Consequently, Eq. (2.3.4) can be rewritten as

$$(n+1)\frac{\omega}{N} = \omega + \sum_{l=1}^{K_{out}} \sin(\phi_l), \quad (2.3.6)$$

where $\phi_l = \varphi_j - \varphi_i$, is the phase difference between two connected nodes i and j which are, respectively, inside and outside the group.

We are now able to write the expression of the upper bound for the critical frequency ω_p^c in a generalized form:

$$\omega_p^c \leq K_{out} \frac{N}{N - (n+1)} = N \frac{K_{out}}{N_{out}}, \quad (2.3.7)$$

where N_{out} stands for the number of nodes not belonging to the considered set. Eq. (2.3.7) reduces to the previous upper bound if one chooses $n = 0$.

In this way we can write a very large number of conditions, that is the number of the connected sets of nodes that include the pacemaker and whose size ranges from 1 to $N - 1$. Among these, the strongest one is that for which the ratio K_{out}/N_{out} takes its minimum value. This is a combinatorial problem, which is in principle very simple, but is hard from a computational point of view since the number of conditions grows at least exponentially with the network size.

Minimizing the ratio K_{out}/N_{out} we find the strictest condition on ω_c^p that can be expressed in the form of a single equation. No other equation obtained as a linear combination of Eqs. (2.1.3) and (2.1.4) may provide a stronger bound. This condition is analogous to the necessary condition for global synchronization concerning the surface (here K_{out}) of any subset of nodes derived in [44] for randomly distributed natural frequencies and generic oscillators. However, these conditions are not sufficient. In our case, it is not certain that the K_{out} remaining sine terms of Eq. (2.3.6) are allowed to take their minimal values simultaneously. This kind of problem directly involves the sine function arguments that may not be independent since they are differences between pairs of phases and we are dealing with a system of N coupled equations. It may happen that two or more phases are tied to each other by a certain set of equations of the kind $f_i(\varphi_i, \{\varphi_{i_j}\}) = 0$ (where nodes $\{i_j\}$ are neighbors of node i). Consequently, we cannot minimize the sum of sine terms on a hypercube $[0; 2\pi]^{K_{out}}$, but we have to restrict ourselves to a hypersurface of dimension $K_{out} - m$, where m is the number of constraints. A system may experience this kind of difficulty (which we can regard as a kind of angle frustra-

tion) only if cycles are present and there is some anisotropy and only when $1 < k_p < N - 1$. Therefore, for a good number of regular connectivity patterns, such as those analyzed in [43], there is not such a problem, and it is possible to analytically calculate the entire set of values $\{\omega_c^{(p)}\}$, $p = 1, \dots, N$.

As a simple, analytically solvable network let us consider a Cayley tree with coordination number z , made up of S shells. For each node it is indeed possible to single out a connected “group” such that $K_{out} = 1$, taking in all the nodes on the branch starting from the considered pacemaker. In this way we are minimizing the ratio K_{out}/N_{out} so that we can consider the strictest equation among Eqs. (2.3.7). Moreover, since there is no cycle, there are no problems of angle frustration either. Therefore, the obtained expressions give the correct values, not just bounds. In this way we obtain for the critical frequency

$$\omega_c^{(s)} = N \frac{1}{N - \sum_{i=0}^{S-s} (z-1)^i},$$

where s is the shell of the pacemaker (see Application A for more analytical details about this case).

Even though in real complex networks it is not so easy to calculate $\{\omega_c^p\}$, we have empirically verified that only in a few cases is the critical frequency much smaller than its first upper bound (2.3.2). This can be clearly observed in Fig. 2.5, where we plotted the ratios between the real critical values and the corresponding upper bound for every choice of the pacemaker in several networks.

The accuracy of this estimation enables us to use it in the opposite direction, i.e., to get an estimation of the pacemaker degree from an experimental measure of the critical frequency. We can invert Eq. (2.3.2) obtaining

$$k_p \geq \omega_p^c \frac{N-1}{N}, \quad (2.3.8)$$

but, since the right term is not an integer, the smallest allowed value for k_p is

$$k_p^* = \left[\omega_p^c \frac{N-1}{N} + 1 \right], \quad (2.3.9)$$

where $[x]$ stands for the integer part of x . We can conclude that Eq. (2.3.9) gives the correct value of k_p whenever

$$\omega_p^c \in \left[(k_p - 1) \frac{N}{N-1}, k_p \frac{N}{N-1} \right].$$

This fact implies that the estimator (2.3.9) for the degree of the pacemaker is very reliable. Indeed, it only fails when the critical frequency is really smaller than its bound (2.3.2).

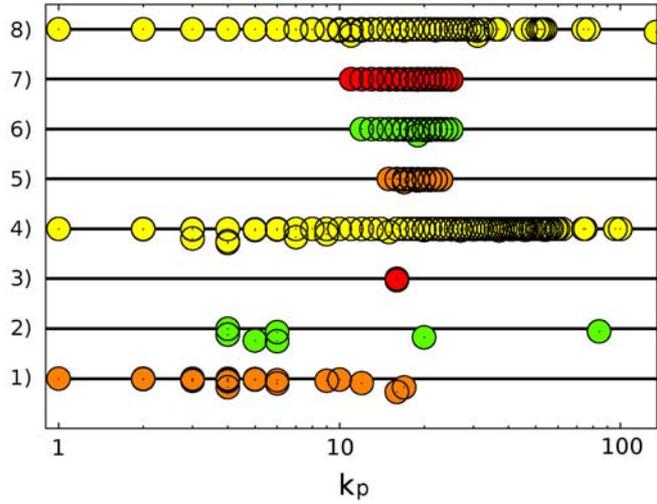


Figure 2.5: Critical frequency of a pacemaker as a function of its degree for a set of networks. We have divided the critical frequency by the degree and by $N/(N - 1)$ such that the bound given by Eq. (2.3.2) is 1. We have shifted the data for the different networks and the horizontal lines are the reference (equal to 1) for each case. From bottom to top the networks are (1) the Zachary club social network [4] used in community detection applications, (2) a hierarchical network of 125 nodes and 3 levels [3], (3) a network of 4 communities of 32 nodes each used as benchmark in community detection algorithms [4] where all the nodes have the same degree, (4) a network of jazz bands [47], (5-7) three networks of three levels of community structure used to relate topological and temporal scales in synchronization [6], and (8) the *Caenorhabditis elegans* neural network [48].

2.4 Slightly above the critical point

In this and in the next section we translate the rich dynamical information that the system provides in the incoherent state into useful topological information. Here we focus on the behavior of the system slightly above the critical point, while in Sec. 2.5 we will analyze the system when the natural frequency of the pacemaker is many time larger than its critical value.

We are interested in estimating how similar two nodes are from a global topological perspective. For this purpose we need to define an appropriate correlation function, able to relate the dynamical responses of pairs of oscillators.

Looking for the expression of a good correlation function, we get no help from the average values $\langle \dot{\varphi}_i \rangle_t = \int_0^\infty \varphi_i(t) dt$. Indeed, in this regime, all the oscillators, except the pacemaker, have the same average effective frequency. On the contrary, it can be useful to look at the difference between instantaneous values. We measure the frequency of every oscillator at each time, inside a suitable interval. In order to define a correlation, that is, a quantity that has to be non-negative and symmetric with respect to nodes indexes i and j , it is reasonable to start from a power of the absolute value of the difference $|\dot{\varphi}_i^{(p)}(t) - \dot{\varphi}_j^{(p)}(t)|$, where (p) stands for

the pacemaker that induces the considered dynamical evolution. Therefore, we propose

$$c_{ij}^p(t) = 1 - \sqrt{\frac{|\dot{\varphi}_i^{(p)}(t) - \dot{\varphi}_j^{(p)}(t)|}{\omega}}.$$

Dividing by ω makes the argument of the root less than 1 because, even if the frequency's may take negative values (see Fig. 2.3), the condition $|\dot{\varphi}_i(t)| \ll \omega$ always holds.

The period of the effective frequencies oscillation depends on which node is the pacemaker. Then, in order to compute averages in time that are really independent from the considered interval, we have to choose a time window many times larger than the oscillation period. Furthermore, since $|\dot{\varphi}_i - \dot{\varphi}_p| \gg |\dot{\varphi}_i - \dot{\varphi}_j|$ for any $i, j \neq p$, we decide to exclude these contributions, taking into account only terms of the kind $c_{i,j}^p$ where $i \neq p$ and $j \neq p$.

Finally, in order to remove the dependence from the index p we have to average all the possible pacemakers. Summarizing in a compact expression, our correlation function can be written as follows:

$$c_{ij} = 1 - \frac{1}{N-2} \sum_{p=1, p \neq i, j}^N \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \sqrt{\frac{|\dot{\varphi}_i^{(p)}(t) - \dot{\varphi}_j^{(p)}(t)|}{\omega}} dt \quad (2.4.1)$$

2.4.1 Hierarchical organization

Once we have obtained the correlation matrix we can proceed to some hierarchical analysis. In the present work we use the standard unweighted pair group method average (UPGMA) [49] algorithm to compute such diagrams. What we find out is a hierarchy of dynamical communities, whose meaning is immediately understandable in the case of small networks, such as our benchmark in Fig. 2.1 (see Fig. 2.6).

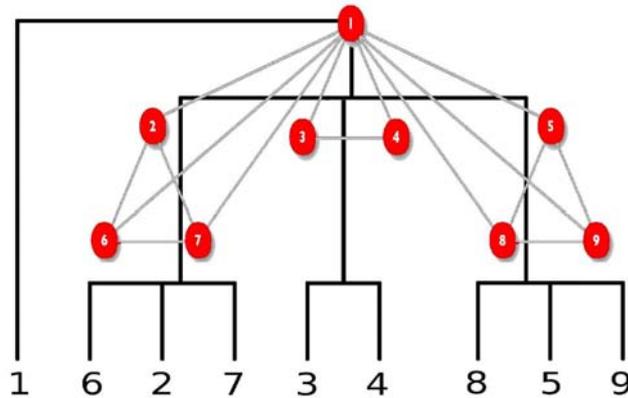


Figure 2.6: The network of Fig. 2.1 and its corresponding dendrogram. Correlations are calculated by averaging over a time window $\Delta t = 60$, after a transient lag $T_s = 10$, for $\omega = 1.1\omega_p^c$.

Obviously, this simple network does not need any analysis to obtain its hierarchical organization, but this methodology can be very useful when applied to functional hierarchical network.

As a paradigmatic example, let us consider the corticocortical network of a cat at the macroscopic level. We look at each cortical area as a basic unit, modeling it as a Kuramoto oscillator, finding similar results as in [50, 51].

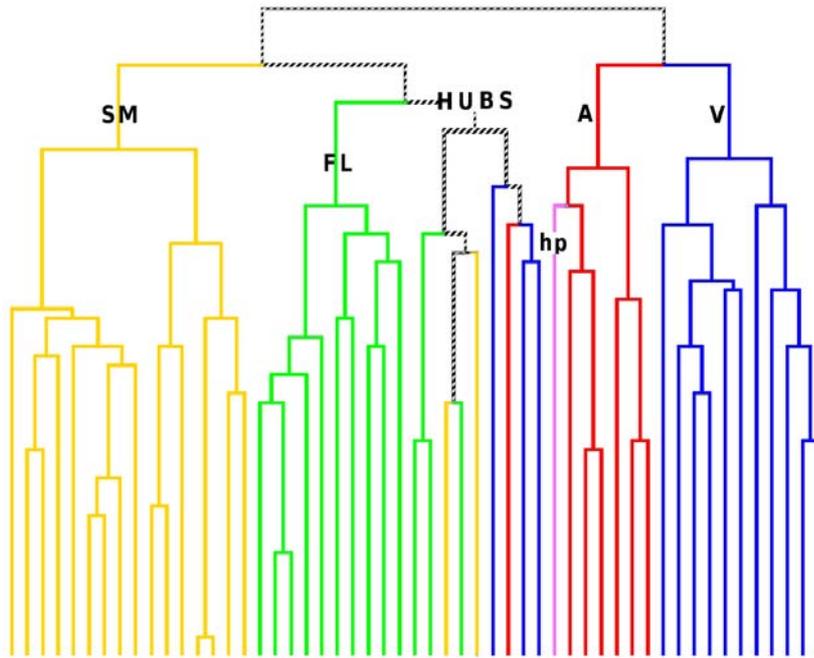


Figure 2.7: Dendrogram of the cortical brain network of a cat. Different colors correspond to different sub-systems: the fronto-limbic (FL), the somatosensory-motor (SM), the auditory (A), and the visual (V). The rich club is labeled with HUBS, while the branch indicated with the label hp is the area that belongs to the hippocampus and it is out of place. Correlations are calculated by averaging over a time window $\Delta t=100$, after a transient lag $T_s=10$, for $\omega=1.1\omega_p^c$.

In Fig. 2.7 we show that, going down along our dendrogram starting from the root, it is possible to recognize two communities clearly separated. Then, the right branch splits into two parts, and the left one undergoes two subsequent bifurcations, so that it is possible to identify three groups of nodes on it. At this level we have five communities. Four of them correspond to well known physiological subsystems: the fronto-limbic (FL), the somatosensory-motor (SM), the auditory (A) and the visual (V). The fifth one (HUBS) is composed, except for a single area⁴, by super-hubs, sometimes considered as a metacommunity (rich club) [50, 51].

The most relevant aspect of our hierarchical analysis is that there is no way to recognize this meta-community if the dendrogram is constructed by means of static methods. It cannot

⁴The cortical area that is not a super-hub is a border area that can be seen as a hub only joined with another one very similar to it, but anyway regarded as a super-hub in itself.

be obtained throughout correlation matrices generated from the adjacency matrix using, for instance, Pearson's coefficient [52] either. Nor these nodes emerge as a community when the modularity function is maximized. Indeed, maximizing the modularity we obtain as an optimal partition the same four groups corresponding to the four physiological sub-systems.

In general, complex networks can be organized, and thus analyzed, at different hierarchical levels. For social networks it is very important that a group is tight, so that the multiple connections within the group give rise to the concept of community. On the contrary, in biological networks the most crucial concept is function rather than connectivity *per se*. Therefore, methods that rely on the connections within groups and maximize modularity will not be enough to identify biological units based primarily on function [53, 54]. In this case, our method, which analyzes the dynamical correlation between units, provides a better approach to inferring functional relationships.

One of the known problems of the methods commonly used for detecting community structures in complex networks is the existence of the so called resolution limit, found by Fortunato and Barthelemy [2]. This issue is related to the impossibility for the methods based on modularity optimization to go beyond a certain resolution which is related to the community size and to the number of links between communities. The paradigmatic example of the problem is a network formed by "cliques" (small groups of totally connected nodes) which are very sparsely connected. We have checked such structures and found that, dynamically, the correlations are very strong within the cliques and not among nodes belonging to different modules, showing that our method for detecting the hierarchical organization is not affected by the resolution limit problem.

2.4.2 Recovering network topology

Let us now take a step backward and recover something we had previously discarded. In the sum of Eq.(2.4.1) we had excluded terms in which one of the indexes was equal to p since they were *heterogeneous*. So $c_{ij} = \frac{1}{N-2} \sum_{p \neq i,j} c_{ij}^p$. However, the set of elements c_{pj}^p , $p = 1, \dots, N$, also contains information. We may ask ourselves which oscillators are most strongly correlated to the pacemaker and if they share some topological property. The simplest hypothesis is that the set of k_p largest c_{pj}^p identifies the neighbors of the pacemaker. This is reasonable since, even if the pacemaker is very weakly correlated to the rest of the oscillators, coefficients c_{pj}^p are not uniform and the topological distance is the most immediate quantity that we may suppose this variability is related to. In Sec 2.3 we showed how to find out an estimator of the degree of each node from the critical frequencies. Thus if we are able to select the possible neighbors we would be, in principle, able to reconstruct the entire network.

The first problem we face in the attempt to validate this hypothesis is that our list of likely

neighbors gives us an asymmetric and weighted adjacency matrix, whose elements are

$$\begin{aligned} a'_{pji} &= c_{pji}^p \text{ for } i = 1, \dots, k_p^*, \\ a'_{pji} &= 0 \text{ for } i = k_p^* + 1, \dots, N, \end{aligned}$$

where k_p^* is the estimator for the degree of the pacemaker given by (2.3.9) and $c_{pji}^p > c_{pji}^p$ whenever $i \leq k_p^*$ and $l > k_p^*$.

Moreover $a'_{mn} \neq a'_{nm}$ since, generally speaking, $c_{mn}^m \neq c_{mn}^n$. Therefore we have to remove the weights and to symmetrize this matrix. Here we propose an algorithm to perform this task that is at the same time simple and efficient. It consists of four steps.

- (1) Symmetrize the matrix in the usual way: $a_{mn}^s = (a'_{mn} + a'_{nm})/2$.
- (2) Compute a list of temporary degree $k'_n \geq k_n^*$ as the number of non-null elements a_{nm}^s .
- (3) Order all the non-zero values a_{nm}^s in a list, from the smaller to the larger.
- (4) Check which ones among the corresponding likely links have to be removed, starting from the weakest one.

We proceed as follows: given a pair of nodes m and n whose link is the weakest one, if and only if $k'_m > k_m^*$ and $k'_n > k_n^*$, we remove that link, setting $a_{mn}^s = a_{nm}^s = 0$. In this case both k'_m and k'_n are reduce by one unit. Otherwise we go to the next link, going on along the entire list until reaching the strongest link.

This method is rooted in the hypothesis, empirically very well verified, that the matrix a_{mn}^s contains all the links of the real network, plus a number of *false positive* ones, i.e., that there is no *false negative* link. Thus we need just remove, never add, edges.

Moreover it works properly only if our estimators $\{k_n^*\}$ of the actual degrees $\{k_n\}$ are correct; otherwise, we may make additional errors. Fortunately it is a very infrequent problem. The sole hypothesis we make is that the probability for a link to be a “false”one is a monotonously decreasing function of the correlation between the nodes it joins.

Finally, the method does not ensure that in the final estimated network $k'_n = k_n^* \forall n$ because it is possible that even if $k'_n > k_n^*$, the n th oscillator has no possible neighbor whose temporary degree is larger than its estimated one. Sometimes this fact may cause new errors; other times it acts as a compensation of the underestimation of the real degrees.

In order to quantify how good a reconstruction is, we introduce the following error definition:

$$\text{err}(\%) = \frac{Fp + Fn}{L} \times 100,$$

where Fp and Fn are respectively the number of false positive (spurious) and false negative (missing) links in the reconstructed network, and L is the number of edges in the original connectivity pattern. Globally speaking, we can state that our method allows for a reconstruction of an arbitrary connectivity pattern with a good precision. Taking into account the networks in Table 2.1, on average we have $\text{err}(\%) = 6.5$.

Among these networks there are artificial as well as real connectivity patterns. They were selected to be representative of several classes of networks, including hierarchical as well as nonhierarchical, with and without community structure, regular and irregular. For this reason, the average error calculated on this set of benchmarks can be considered to be a good estimator of the accuracy of the proposed reconstruction method when applied on a given unknown connectivity pattern.

| N | L | K_{err} | L' | F_p/F_n | L'_r | F_p/F_n | err(%) |
|-----|------|-----------|------|-----------|--------|-----------|--------|
| 9 | 15 | 0 | 15 | 0/0 | 15 | 0/0 | 0 |
| 18 | 24 | 0 | 24 | 0/0 | 24 | 0/0 | 0 |
| 25 | 66 | 0 | 82 | 16/0 | 66 | 0/0 | 0 |
| 34 | 78 | 7 | 99 | 27/6 | 75 | 7/10 | 21.8 |
| 48 | 64 | 0 | 64 | 0/0 | 64 | 0/0 | 0 |
| 53 | 391 | 0 | 445 | 53/0 | 392 | 5/4 | 2.3 |
| 125 | 394 | 33 | 475 | 81/0 | 383 | 1/12 | 3.3 |
| 128 | 1024 | 0 | 1060 | 57/21 | 1026 | 36/34 | 6.8 |
| 256 | 2311 | 0 | 3223 | 1040/128 | 2324 | 259/246 | 21.8 |
| 256 | 2301 | 0 | 2851 | 607/57 | 2312 | 116/105 | 9.6 |

Table 2.1: Results of the reconstruction on several networks. The columns give the size of the system N , the total number of links in the original network L , the total error in the estimation of the degrees $K_{err} = \sum_{i=1}^N |k_i - k_i^*|$, the total number of links in the reconstructed network before the removal of exceeding links L' , the number of false positive F_p and false negative F_n links in this network, the same for the final reduced network L'_r , F_p/F_n and the final total error err(%). From the first row, the networks are our usual benchmark [1], a ring of 6 cliques of 3 nodes [2], a hierarchical network of 25 nodes and 2 levels [3], Zachary club social network [4], ring of 16 cliques of 3 nodes [2], the cortical brain network of a cat [5], a hierarchical network of 125 nodes and 3 levels [3], a network of 4 communities of 32 nodes each [4], and two networks of 3 levels of community structure [6].

2.5 Far from the critical point

Far above the critical point the system behaves quite differently. As clearly shown in Fig. 2.8 (left panels) all units are characterized by effective frequencies that, after a transient time, oscillate around precise values that are equal to their own natural frequency. From this point of view, by increasing the natural frequency of the pacemaker the coupling is less and less important. But, in any case, there are still remnants of the interactions since the amplitudes of the oscillations decay very quickly with the distance from the pacemaker. Indeed, the frequencies of the neighbors of the pacemaker oscillate with an amplitude that is roughly $A_{neigh} \simeq 2$, while all the other oscillators are almost at rest compared with them. These conditions allow us to recognize the neighbors of a given pacemaker even if we do not know how many there are. Therefore, we may define a simplified correlation function that better suits this situation and

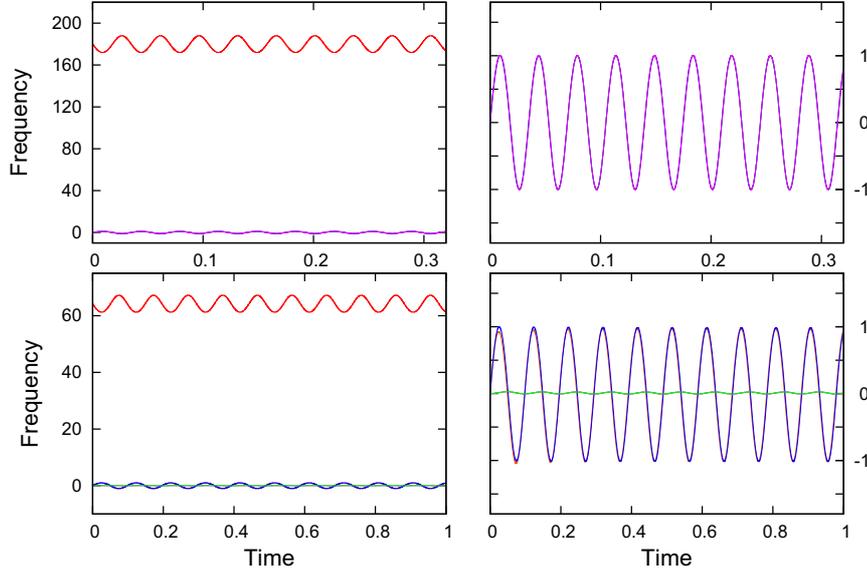


Figure 2.8: Effective frequencies as function of time far above the critical point ($\omega = 20\omega_p^c$). Plots refer to the same network used for the previous pictures, in the case of two different choices of the pacemaker: node 1 ($k_1 = 8$) above and node 2 ($k_2 = 3$) below. On the left hand side we plotted the frequency of all the nodes in the network. On the right side the scale has been changed and the pacemakers are left out. Notice how above, where all the nodes are neighbors of the pacemaker, we may observe a unique curve. On the contrary, below there are two different kinds of oscillations. The largest ones are those of the neighbors of the pacemaker; the others belong to the oscillator not directly connected to it. Time starts after a transient lag $T_s = 1$. The integration time step used is $\delta t = 10^{-4}$.

that only connects each pacemaker with its neighbors:

$$c_{pi}^F = \frac{\max_t[\dot{\varphi}_i(t)] - \min_t[\dot{\varphi}_i(t)]}{\max_t[\dot{\varphi}_p(t)] - \min_t[\dot{\varphi}_p(t)]} = \frac{A_i}{A_p}. \quad (2.5.1)$$

The above expression is the ratio between two positive terms (amplitudes) and it is equal to 1 for $i = p$.

On any connectivity pattern, the values c_{pi}^F are distributed along a set of stairs whose highest step is easy to identify even if we consider short time windows. The transient time, indeed, is always very short in this regime. We no longer need to completely reconstruct the entire connection topology.

All we have to do is to compute the values c_{pi}^F for each pacemaker. After finding out the maximum values $\max_{i \neq p} c_{pi}^F \forall p$, we choose an appropriate threshold, say 0.5. A node j will be a neighbor of the pacemaker p if $c_{pj}^F / (\max_{i \neq p} c_{pi}^F) \geq 0.5$. Now we are able to construct a connectivity matrix.

Let us notice that in this case there is no need for symmetrization since the adjacency matrix constructed in this way is already symmetric because this method is based on a reliable general

property that holds for any connectivity pattern. The use of a threshold is therefore, in principle,

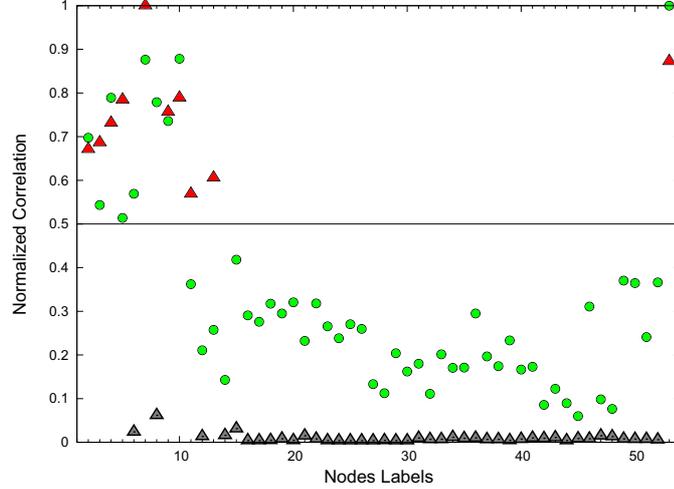


Figure 2.9: Normalized correlations of the cortical brain network of the cat for pacemaker on node 1 ($k_p = 10$). The circles are the correlation values calculated through Eq. (2.4.1) for $\omega = 1.2\omega_p^c$ ($\Delta t = 100, T_s = 20$). The triangles correspond to the correlations given by expression (2.5.1) when $\omega = 20\omega_p^c$, calculated on a time window $\Delta t = 1$ and waiting a transient time $T_s = 0.1$. All the values have been divided by the maximum on each set (excluding the autocorrelation). Notice that while in the first case there is an almost continuous spectrum of values, in the second one it is easy to identify a group of points (red triangles) above the line at 0.5 clearly separated from the rest. Those are the 10 neighbors of node 1.

unnecessary since all the neighbors have the same amplitude of the frequency oscillation, when the pacemaker natural frequency is above a certain value. But since this value is not known *a priori* and it may be very large if the distribution of the degrees among the neighbors of the pacemaker is very wide, it is useful from an empirical point of view. It is important to stress that, even if we are still in a regime where some degree of *heterogeneity* among the neighbors is conserved, there is no chance to make any errors in the recovered topology. Indeed, the amplitude of the frequency oscillations of the pacemakers neighbors is at least one order of magnitude larger than that of any other oscillator (see Figs. 2.8 and 2.9). By means of this method all the topologies considered in Table 2.1 are properly reconstructed, without errors.

In addition, not all nodes need to be considered as pacemakers. While the method discussed in Sec. 2.4.2 requires us to perform dynamical measures for every possible location of the pacemaker, for the current description this is not necessary. Indeed, we can look for the neighbors of a number $N' < N$ of pacemakers in order to get all the connections in the considered network. From an experimental point of view, adopting the conceptual framework proposed in [55], we may consider the choice of a certain pacemaker as the application of a drift on a given unit in a system of identical coupled oscillators. This means that it is possible to solve the problem with less than N experiments.

The criterion for choosing the ordered sequence of nodes on which we locate the pacemaker

can vary. We may operate a random extraction, or we may start from a randomly chosen node and then move to one of its neighbors along a random walk. Another option, which is much more convenient, especially in the case of scale-free networks, can be adopted if the critical frequencies associated with each oscillator are known. We can order the nodes according to decreasing critical frequency, starting from the highest one. In this way we proceed from larger to smaller (estimated) degrees, taking an important advantage if the degree distribution is not uniform and there are hubs. The hubs, indeed, provide information about a large number of links by means of very few experiments (Fig.2.10).

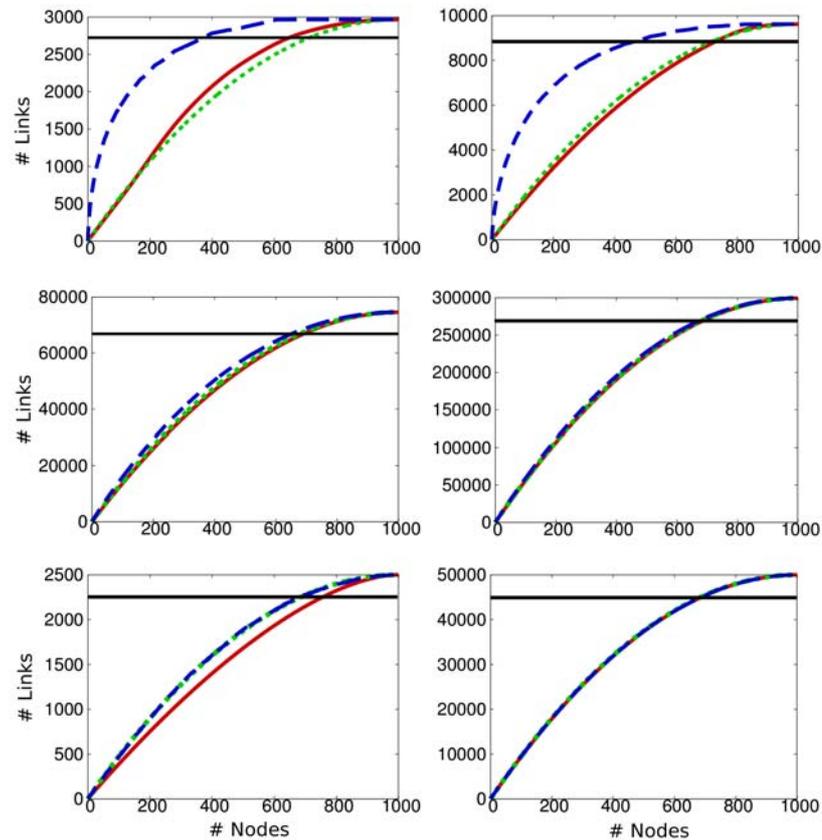


Figure 2.10: Average number of reconstructed links as a function of the number of nodes we considered as pacemakers (number of trials). From the top to the bottom, the considered networks are a pair of Barabasi-Albert networks, with parameter $k = 3$ (left) and $k = 10$ (right); a pair of Erdos-Reyni graphs with average degree equal to 15 (left) and 60 (right); and a pair of random regular graphs with degree 5 (left) and 100 (right). The size is $N = 1000$ for all of them. Different lines corresponds to different selection algorithms. Blue dashed lines stand for the ordered sequence on the basis of the critical frequency values; the red solid lines are for the random walk; the green dotted lines are for random extractions. Both the random walk and the random extractions are averaged over 1000 samples. The horizontal black line marks 90% of links: notice how in any case we never need more than 70% of the nodes in order to reconstruct 90% of the links, decreasing to 30% – 40% in the case of the ordered sequence for scale-free networks. Correlations are computed under the same conditions as those in Fig. 2.9.

2.6 Conclusions

Systems of non-identical Kuramoto oscillators have been recently shown to display a degree of synchronization that depends strongly on the topology of the underlying complex network. Here, these dynamical properties, particularly by setting different types of correlations between the dynamical evolution of the oscillators, have been used to gather information on the connectivity patterns. Remarkably, this is the case for most experimental situations, where the *a priori* unknown connectivity of a particular network is inferred from purely dynamical measurements.

When the oscillators are identical (all of them having the same natural frequency) any topological configuration has a unique attractor, which is the complete synchronized state, meaning that the oscillators end up in such a state that all effective frequencies and phases are identical. This state does not offer any information about the topology. We perturb this setting by allowing one of the oscillators to have a different natural frequency than the rest. This unit is called the pacemaker of the network. Such a perturbation causes the final state to no longer be phase synchronized. But if the natural frequency of the pacemaker is not very different from the value of the rest of the population, the system still will retain a certain degree of synchronization since the whole system can evolve with the same effective frequency. However, if the frequency difference becomes larger, the system will be unable to find any kind of synchronization. The threshold between the former case and this latter case is a well defined value, which is strictly dependent on the location of the pacemaker in the network. In this context, we can use the correlations between the effective frequencies of the oscillators in such an incoherent state to reproduce the network connectivity.

Moreover, we show that the dynamical correlations in different situations, whether close to or far from the critical point, provide complementary information on the network.

1. Working around the critical point we are able to estimate the degree of each pacemaker merely by its critical frequency.
2. Slightly above the transition point the hierarchical structure of the whole network (related to functional modules) can be obtained from the correlations between effective frequencies. A further refinement enables us to recover the whole connection network with a good degree of accuracy.
3. Far above the critical point it is possible to recognize which oscillators are directly connected to an individual pacemaker from a very short measurement of the time evolution of the effective frequencies. In this way we can recover the connectivity pattern, and this method turns out to be much more precise and more efficient than the previous one.

In summary, this chapter deals with different approaches relating dynamical properties of individual nodes to the topology of the network. The topological properties inferred from dynamics can be local (the existence of a link between two nodes) as well as global (hierarchical

organization of the nodes in the functional network). In particular, for a scale-free network and if the node degrees are known (or have been estimated from the critical frequencies), considering 30% of the possible pacemakers, always selecting the most connected nodes, will be enough to reconstruct approximately 90% of the links.

Other works have considered the reconstruction of the network from dynamical information. Similar to our proposal with specific targets, Tegner *et al.* [56] analyzed the dynamical response of a gene-regulatory network by changing expression levels of particular genes. On the other hand, di Bernardo *et al.* [57] considered the global effect of different types of perturbations to infer the network topology. This approach has been followed recently also by Gorur Shandilya and Timme [58], who assumed that there is some information about the dynamical evolution of the isolated units and about the coupling. Our method, based on the change of the frequency of a single unit and how it enhances correlations among the nodes, can be more effective in oscillatory systems. In any case, for practical purposes the method chosen will depend on the specific details of the experimental setup and even a combination of different ones can be the most appropriate.

Chapter 3

Exploring complex networks by means of adaptive walkers

Finding efficient algorithms to explore large networks with the aim of recovering information about their structure is an open problem. In general, network features are discovered by means of algorithms based on search and traffic routing [33, 34, 35]. In many cases, the latter can be performed by means of moving “agents”, which explore the topological space and recover information. Nonetheless, it is still a key issue the investigation and characterization of the efficiency of different strategies [59, 60, 61] as far as the quality and quantity of information gathered is concerned.

Here, we investigate this challenge by proposing a model in which random walkers with previously assigned home nodes navigate through the network during a fixed amount of time. We consider that the exploration is successful if the walker gets the information gathered back home, otherwise, no data is retrieved. Consequently, at each time step, the walkers, with some probability, have the choice to either go backward approaching their home or go farther away. Our aim is to find out an optimal strategy to maximize both the number of visited nodes and the chance to meet again the starting point, independently of which is the choice for the latter. To this end, we consider an arbitrary (heterogeneous) network of N nodes and a single agent (explorer or walker) initially located on a given node (home-node), and let it move during a time frame T , the walker’s lifetime. Every time the agent comes back to the starting point, all the nodes it has visited until that moment are marked as *visited* and the total information gathered is updated. Obviously, it could also be possible to send several agents at once, but it has been demonstrated for several similar situations [62] that increasing the number of walkers (and reducing their lifetime proportionally) does not produce better results. Consequently, we focus on the performance of single agents. The most important novelty of our proposal is that the agents are not markovian random walkers, nor a modified version of random walks’ dynamics in which additional rules (for instance, preferential or self-avoiding random walks [60, 63]) are introduced. Indeed, we introduce a parameter q which governs how likely it is for a walker,

at each time step, to go forward or backward (with respect to the walker's home). Thus, by changing the value of this parameter, the two probabilities can be tuned and hence different strategies are defined. In one limiting case, the walkers will tend to move back home, whereas in the other limiting setting, they will tend to move away from home. In between these two asymptotic behaviors, we recover a classical random walk, for which all directions are equally probable. We explore different strategies and their dependencies with both the degree of the home nodes and the walkers' lifetimes. Moreover, we show that it is possible to built up an adaptive algorithm whose efficiency in terms of the information gathered and the quality of the reconstructed network is, in general, the best.

The rest of the chapter is organized as follows. Section 3.1 introduces the model which is characterized in Sections 3.2-3.3. Our proposal for an adaptive strategy is presented in Section 3.4. In Section 3.5 we present the application of the algorithms previously discussed to the reconstruction of the degree distribution. Finally, the last section (Sec. 3.6) is devoted to briefly discuss the potentialities and possible applications of our model.

3.1 Baseline model of walkers

Let us first discuss a baseline model in which a given set of walkers explore the network starting from a home node. As previously discussed, in order to collect the results of walkers' exploration, they should go back home. Therefore, we introduce two probabilities when the walker is at a given node, provided it has tracked the information about the path followed from the home-node to the current position. These two probabilities correspond to the forward (F) and backwards (B) motion along the already tracked path and read, respectively, as

$$P_F(k_i) = q^2(k_i - 1)/[1 + q^2(k_i - 1)], \quad (3.1.1)$$

$$P_B(k_i) = 1/[1 + q^2(k_i - 1)], \quad (3.1.2)$$

where the label i indicates the node that the explorer is going to leave and k_i is its degree. These equations stand for every step whenever the agent is not in the starting node – the home, h –. In the latter case, i.e., while at home, it can only go forward, thus at that position we have $P_F^h = 1$ and $P_B^h = 0$.

From Eqs. (3.1.1)-(3.1.2), we recover the pure random walk (without any bias, i.e., all possible directions are equally probable) for $q = 1$. For very large values of the parameter q , no backward step is allowed. Consequently the explorers can get back to their starting node only by chance, through a different path, not being *aware* that they are coming back, but being able to recognize where they are (at home). Conversely, when q goes to 0, after the first move, no more steps forward are allowed. Therefore, only the first neighbors of the starting node

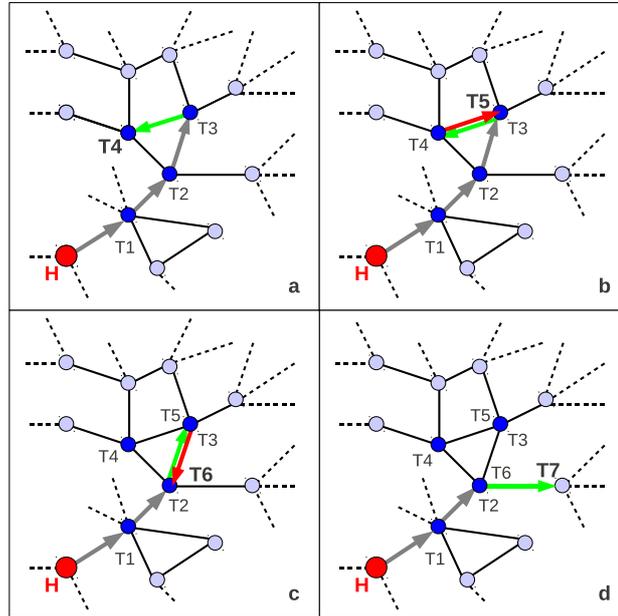


Figure 3.1: Example of the motion of an agent: 4 snapshots taken at 4 sequential time steps. The red node labeled with “H” is the home-node and the nodes that have been visited are colored in blue. Panel (a): T4. Grey arrows stand for previous steps forward, while the green arrow stands for the last one. Panel (b): T5. The agent takes a step backward (red arrow) reverting and removing from its “memory” the last step forward. Panel (c): T6. At this point it is the step taken at time instant T3 that has to be regarded as the last step forward (green arrow); the walker takes another step backward (red arrow) reverting it. Panel (d): T7. The walker takes a step forward toward a new node.

can be explored. We also consider that the walker’s lifetime is T steps, which represents the time allowed for the network exploration before the dynamics stops. We define the information gathered as the fraction of nodes marked as *visited* after T time steps: $I = V/N$, where V is the number of visited nodes and N is the size of the network. Moreover, if the agent is not at home at time T , the new nodes visited after its last return to the home node are not computed in V (i.e., we consider that only the information brought counts).

We first discuss the expected behavior of I at the two limiting values of q (very high or very small). On one hand, for very low q values only the nearest neighbors are visited and hence I will be small independently of T . On the other hand, for very large values of q the walkers only return to home by chance, being the search also inefficient provided the exploration time is not very large (see next section). Then, if we fix the total number of steps we can expect that the information collected will have a maximum as a function of q . Therefore, there should exist, for any given network, a precise value $q^*(T)$ such that, if we average over all the possible choices of the home-node and over many realizations of the dynamical exploration, the mean information $\langle \bar{I}(q^*) \rangle$ is maximal. In other words, there is no other value q' for which $\langle \bar{I}(q') \rangle > \langle \bar{I}(q^*) \rangle$, where “ $\langle \cdot \rangle$ ” stands for the mean performed over all the nodes in the network and “ $\bar{\cdot}$ ”

for the average over many realizations.

The previous analysis indicates that the best efficiency in terms of maximal recovery of information can only be obtained for two values of q^* . In the next section, we explore the dependency of I on the network properties (as given by the degree of the home node) and walkers' lifetime. Admittedly, when this time is very long ($T \gg N$) we should expect to recover most information by setting $q^* \rightarrow \infty$. However, even if this is the best choice on average, it might not be the case when the home of the walker is at a low degree node. On the other hand, for shorter searching times, a value of $q = q^* < 1$ gives almost the same performance for I , but this time the results are independent of the degree of the home node and $\langle \bar{I}(q^*) \rangle$ is a global maximum – the caveat is that q^* cannot be known *a priori*.

3.2 Characterizing the performance of the walkers

Hereafter we will use as a benchmark a scale free network of $N = 10^4$ nodes and mean degree $\langle k \rangle = 10$ generated by the uncorrelated configuration model [64]. We however note that all results reported are valid for any network with a power-law degree distribution provided that it does not have a tree-like topology. Actually, the only relevant difference in the case of a tree-like network is that we will observe a different behavior for large values of q . This is because leaves would make very difficult for a walker to come back through a different path making their performance very poor, even for very large values of T and for very large degrees of the home nodes.

In Fig. 3.2 the information $\langle I \rangle$ is plotted as a function of q for several home-nodes and a searching duration of $T = N = 10000$ steps. As it is clearly shown, starting from small values of the parameter q , $\langle I \rangle$ initially increases but soon afterwards there is an abrupt decay to give way to a new increase as q grows further. For very large values of q , the information gathered saturates to an asymptotic value. Interestingly enough, as seen in the figure, the amount of information gathered for both very small values of q and when $q \gg 1$, as well as the size of the abrupt decay, depend on the degree of the node from which the walker started the exploration. However, there exists a universal value of $q = q_p$ at which almost all curves corresponding to different degrees of the home node collapse; i.e., there is a local maximum which is roughly independent of the connectivity of the home node. Nevertheless, whether this point is also a global maximum for $I(q)$ or just a local one depends on the degree of the initial node. Indeed, when the home-node is highly connected, for this searching duration, an agent performs better for $q \rightarrow \infty$, but if this is not the case, q_p gives the optimum efficiency.

In Fig. 3.3 we plot the same quantity as in Fig. 3.2 but averaged over all the possible home-nodes (then the dependency with the degree washes out) and considering different lifetimes T . The figure makes it more clear that at $q = q_p$ the value of $\langle \bar{I} \rangle$ is a global maximum unless T is many times larger than the network size N . This definitively means that if we are interested

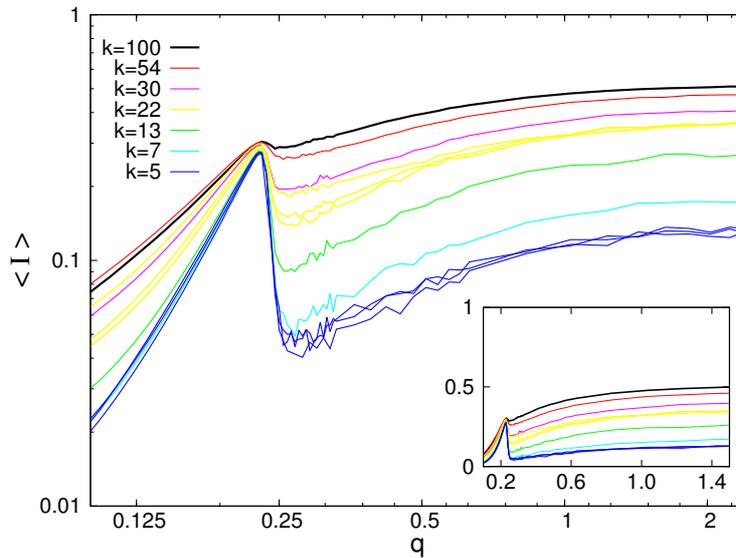


Figure 3.2: Information \bar{I} (averaged over 3000 realizations) gathered by a walker during a searching time $T = 10000$ as a function of the q parameter. Each curve refers to a different home-node and different colors refer to different degrees of the starting node: $k = 100$ (black), $k = 54$ (red), $k = 30$ (purple), $k = 22$ (yellow), $k = 13$ (green), $k = 7$ (light blue), $k = 5$ (blue). In the inset: the same quantity in a linear scale.

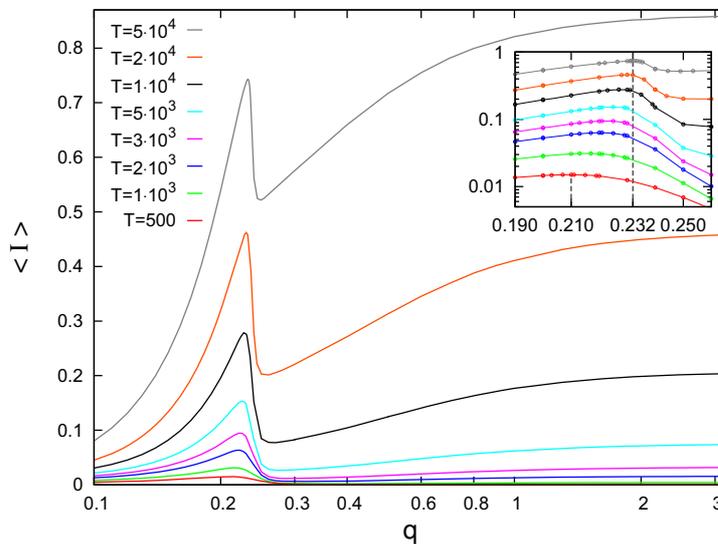


Figure 3.3: Mean Information $\langle \bar{I} \rangle$ gathered by a walker performing its search starting from any home-node during a time lag of T steps, as a function of the q parameter. The mean is performed over all the nodes in the network, and for each of them averaging over 100 realizations. Different colors refer to different durations of the searching. From the bottom to the top: $T = 500, 1000, 2000, 3000, 5000, 10000, 20000$ and 50000 . In the inset: zoom of the peak. Notice that q_p displays a small shift increasing T , up to $T = 10000$ when it reaches an asymptotic value.

in the information an agent may gather for a very long searching time, what we have to do is to set $q \gg 1$. Otherwise, if we are interested in more realistic situations where there can be limitations on the duration of the exploration (for instance, due to energy constraints), the best choice would be to set $q = q_p$. The latter option has a caveat, however: the precise value of q_p depends in an unknown way on the topological features of the underlying network. Nevertheless, one can obtain useful insights into the problem by inspecting how the behavior of a walker changes when q varies.

Looking more carefully to the results plotted in Fig. 3.3, one can distinguish three regions that qualitatively correspond to the three distinct behaviors of the walker. In the first one, for $q < q_p$, $\langle \bar{l} \rangle$ monotonously increases as a function of q ; in the second one, $\langle \bar{l} \rangle$ experiences an abrupt decay; whereas the third region shows that $\langle \bar{l} \rangle$ starts to increase again, until it saturates to a value that depends on T . It is easy to realize that the first increase corresponds to small enough values of q . In this region, the walker moves just a few hops away home and consequently it takes only a few steps to get back home. The larger the value of q is, the longer the mean path covered by the walker will be. Since for very small values of q the exploration is local, the relevance of the home-node degree is very high (see Fig. 3.2). Then, increasing q , we are allowing the walker to explore farther nodes, that is to collect new information, and the initial differences due to the degree of the home-node become progressively smaller. At $q = q_p$ they have almost vanished.

In the second region, for q slightly larger than q_p , the walker often gets lost and its performance is, on average, less efficient. In other words, the explorer wastes an important fraction of the lifetime T gathering information that it will not be able to bring back home before the time is over. The precise value at which this start to occurs is slightly affected by the duration of the exploration, as shown in the inset of Fig. 3.3. This can be explained as a combination of two factors. On the one hand, to increase q means to increase the number of nodes visited, but also the risk to get “lost”. Indeed, if an agent is performing a long trip and it is going to bring a lot of information back home, when the searching time is suddenly over, the loss is big. On the other hand, the very first trips are those that provides the largest fraction of new information since the majority of nodes are being visited for the first time. Thus, getting lost after a couples of returns causes a much worse loss than if the same happens after a few round trips. Again it is a matter of balance and the optimum value q_p is smaller when the lifetime is shorter. The second region ends at a value of q for which the previous balance is the worst possible one, thus giving raise to another increase, which marks the start of the third region. Here, for even larger values of q , it begins to be quite frequent that, wandering across the network almost randomly, the explorer returns to its home-node through a different path just by chance. This new behavior entails a new increasing of $\langle \bar{l} \rangle$ due to the fact that this kind of random returns start to balance the inefficiency of the walkers that get lost. The likelihood of these events increases with q and it is maximum when $q \rightarrow \infty$, that is, when $P_B = 0$ at each time step.

The previous dependency of $\langle \bar{I} \rangle$ on the walker's lifetime T defines two optimal values for q , either $\langle \bar{I}(q) \rangle$ takes its maximum value at $q^* = q_p$ or at $q^* = \infty$. However, we stress again that for $q \gg 1$, the walker gets back home by chance (recall that for these values of q the backward probability $P_B = 0$). Consequently the asymptotic values of $\langle \bar{I} \rangle$ in the $q = \infty$ limit strongly depends on the degree of the home nodes (see Fig. 3.2). Therefore, setting $q^* = q_p$ could be a better choice even when T is large enough. In order to be able to take advantage of the agents' behavior at q_p , we need to characterize deeper the transition that occurs for that value of the parameter. To this end, in the next section we focus on the behavior of some dynamical quantities which display a relevant change around q_p .

3.3 Sequential steps and return time

In Fig. 3.4 we plot the average maximum number of sequential steps backward ($\langle \bar{S}_B \rangle$) and forward ($\langle \bar{S}_F \rangle$) that a walker takes in a time lag $T = 10000$ as a function of q . These two quantities, estimated by averaging over many realizations and over all the possible home-nodes, give a useful picture of the transition between the first and the second regimes previously described. They initially increase together, then $\langle \bar{S}_B \rangle$ start increasing slower than $\langle \bar{S}_F \rangle$, it reaches a maximum and start decreasing, asymptotically going to zero. Notice that for small q the value of $\langle \bar{S}_F \rangle$ is small. Consequently, $\langle \bar{S}_B \rangle$ is bounded (even if $P_B(k) \sim 1 \forall k$) since, when an agent is back to its home, no more steps backward can be taken. The value of q for which $\langle \bar{S}_B \rangle$ and $\langle \bar{S}_F \rangle$ take the maximum value before getting apart roughly corresponds to q_p . It is when the walker goes as far as possible from its starting point, being still able to come back on its own steps. Increasing q a little bit further provokes that the number of steps forward exceeds that of steps backward and the home-node is not recovered any more, so that the searching efficiency rapidly decreases. This phenomenology helps us to find out an heuristic definition for the peak. It is indeed possible to state that q_p is the precise value of q for which a walker is allowed to take enough steps forward to be able to visit a large region of the network, but at the same time it is also allowed to take enough steps backward so as to return to its home not by chance.

Admittedly, it is possible to translate this heuristic statement into a quantitative condition starting from one simple observation. There exists, for any k , a value of q such that $P_F(k) = P_B(k)$ and from Eqs. (3.1.1)-(3.1.2) we know that this value is $q(k) = 1/\sqrt{k-1}$. If $q = q(k_{max})$ it is guaranteed that $P_F \leq P_B \forall k$, so the mean path is short and the explorer will come back to home very often. If $q = q(k_{min})$, the situation is the opposite, $P_F \geq P_B \forall k$, so for the agent it is very difficult to recover its home. Therefore, the conclusion is that the peak lies between these two extremal values. A reasonable estimation could be obtained using the mean degree $\langle k \rangle$. However, this actually results in an overestimation of the actual value. Indeed, if the walker reaches a highly connected node and $q = q(k_{max})$, due to the fact that in

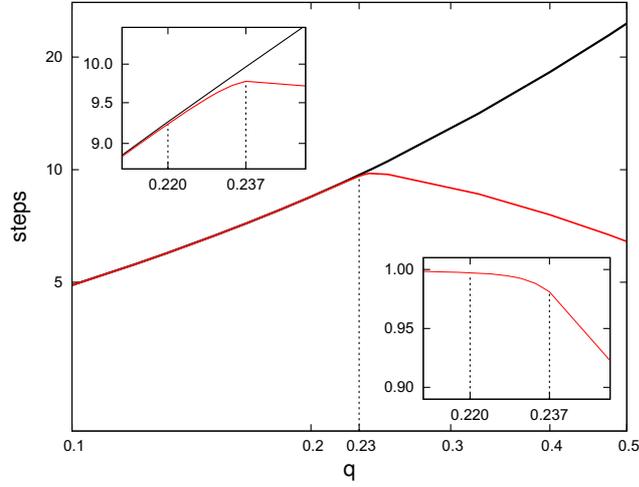


Figure 3.4: Mean maximum number of consecutive steps forward $\langle \bar{S}_F \rangle$ (black line) and backward $\langle \bar{S}_B \rangle$ (red line) taken by a walker during a time lag of $T = 10000$ steps. The mean is performed over all the nodes and averaging over 2000 realizations for each of them. In the first inset (above): zoom around the value of q at which the two curves get apart. In the second one (below): the ratio $\langle \bar{S}_B \rangle / \langle \bar{S}_F \rangle$ in the same range. According to the arguments discussed in this section, the peak should lie between the two values indicated with the dashed lines ($q_p \in [0.220 ; 0.237]$) and this is in a good agreement with what we can observe in Fig. 3.3.

a scale-free network $k_{max} \gg \langle k \rangle$, it has a very small probability to take a step backward and consequently the walker will likely get lost. Moreover, the probability to meet any of the nodes with a high degree is larger than their relative frequency (it is proportional to $kP(k)$). Hence, an improvement of the first estimation consists of replacing $\langle k \rangle$ with the quadratic mean degree $\langle k^2 \rangle$. Although, this value is still too large and also overestimates q_p , it does constitute a better upper bound for q_p , that is

$$q_p \leq \frac{1}{\sqrt{(\langle k^2 \rangle)^{1/2} - 1}}. \quad (3.3.1)$$

In order to complete this phenomenological picture it can be useful to look at another quantity strictly related with what we have said in the previous paragraphs. In Fig. 3.5 we plot the mean time that a walker needs to come back to its home-node ($\langle \bar{T}_R \rangle$) as a function of q . In this case we did not set a lifetime. We then let N_w walkers wander through the network starting from a given node. Each time an agent recovers its home it is not allowed to leave it anymore and the duration of the trip is recorded. We wait until every walker has come back and calculate the average return time $\bar{T}_R^{(i)}$, where $i \in [1, N]$ stands for the considered home-node. Finally we average over all the possible starting nodes. What we obtain is a curve that closely resembles that of the order parameter in a second order phase transition, but with the critical point located slightly above q_p . Furthermore, if we look at the dispersion of the values of $T_R^{(i)}$ we recover a behavior quite similar to that of the susceptibility, i.e., a divergence at the critical point [65]. Actually, the divergence takes place very close to q_p , but even closer to the value

of q for which the average number of consecutive backward steps is maximum (see the inset in Fig. 3.4). Again, this scenario corroborates the intuition that q_p is the maximum value of q able to guarantee that the walker will not get lost.

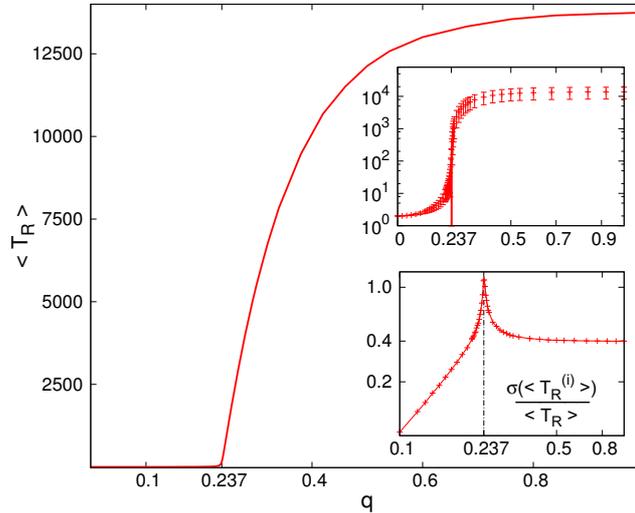


Figure 3.5: Mean time $\langle \bar{T}_R \rangle$ that a walker starting from any node in the network needs to come back to its home-node as a function of q . The average is performed over all the nodes, and for each of them over 5000 realizations. In the inset above: the same quantity represented in a log scale with error bars (standard deviation among home-nodes); in the inset below: the relative standard deviation of the values $\bar{T}_R^{(i)}$.

3.4 The adaptive strategy

With the information gained up to this point, one can set up an adaptive algorithm aimed at optimizing the performance of an agent exploring a heterogeneous network in a number of steps T that is equal to (or less than) the number of nodes N . The basic idea is simple. We have a walker and a value of q associated to it. We let it wander and when it is at home again we evaluate the contribution of this last round trip to the information gathered until that moment and, if necessary, the value of q is modified. In order to build up such an algorithm, three main elements are needed. The first one is an appropriate quantitative way to evaluate the performance of the agents. The second one is a criterion to decide whether or not q would be modified. Finally, the adaptive rule applies whenever the choice is to change the value of q . This third element is an algorithm able to connect what the agent has learned about the network until its last return, the efficiency of its performance and the current value of q in order to provide a new, more suitable, value for the parameter.

Let us start with the first element. Since the aim of the exploration is to collect the maximum amount of information in a fixed time frame, to be *efficient* means to visit as many new

nodes as possible per unit of time (step). The final efficiency of a searching process can thus be defined as $E = I/T$. This definition can be expressed as a function of the number of round trips. If we indicate with t_r the time of the r -th return of the explorer ($0 < t_1 < t_2 < \dots < T$), we have $E(t_r) = V(t_r)/t_r$, where $V(t_r)$ stands for the number of visited nodes after t_r steps ($V(t_r)N = I(t_r)$). It is also possible to measure the efficiency of a single trip as $e_r = [V(t_r) - V(t_{r-1})]/(t_r - t_{r-1})$, but this is not a very useful procedure as e_r is very noisy. Therefore, in order to compare the performance at time t_r with that at time t_{r-1} it is better to consider the efficiency variation $\Delta E(t_r) = E(t_r) - E(t_{r-1})$. Hence, a good criterion to decide whether a change of q is needed is $\Delta E(t_r) < 0$.

Notice that if we start with a small value of q , the number of steps forward and backward will be the same (see Fig. 3.4) and the explorer will pass on each visited node at least two times. Therefore the first return time t_1 will be twice the number of steps ahead that the walker was allowed to take. The maximum number of different nodes that the agent may have visited during its trip is therefore equal to the number of steps it took forward. This happens whenever the walker does not cross each link more than twice (forward and backward). Thus, the efficiency has an upper bound, $E \leq 1/2$, that can be easily reached for any small value of q when the explorer performs its first trip. In particular, for $q = 0$ we surely have $E(t_1) = 1/2$ since only one step forward is allowed and the agent will visit one node in two time steps. Consequently, we expect $E(t_r)$ to start from a value very close to $1/2$ and then necessarily decreases. Hence, changing q has the effect of decelerating the decay of $E(t)$, or at most, to make $E(t)$ reach a stationary value.

When q is varied, we should also take special care in not letting the agent to get lost. With this aim, one needs to fix an upper bound for q . This bound can be provided by Eq. (3.3.1) if we use the degrees of the visited nodes to evaluate the effective value of $\langle k^2 \rangle$. Actually, when the expression (3.3.1) is computed using the information collected empirically, a better estimation of q_p can be obtained. This is because the value of $\langle k^2 \rangle$ is overestimated, being easier for an agent to explore highly connected nodes than peripheral ones during early trips. With all the previous remarks, the adaptive algorithm can be formulated as follows (see Fig. 3.6):

- 1) Set $q \sim 0.1$ and let the agent perform its first round trip.
- 2) Calculate $E(t_1) = V(t_1)/t_1$ and let the agent perform another trip.
- 3) Calculate the new value of the efficiency and check if $\Delta E(t_2) = E(t_1) - E(t_2) < 0$. If it is not the case, let the agent explore again, until the condition $\Delta E(t_r) < 0$ is satisfied.
- 4) Calculate

$$\langle k_2 \rangle(t_r) = \sqrt{\sum_{i \in V(t_r)} k_i^2 / V(t_r)},$$

and then

$$q_{UB} = 1/\sqrt{\langle k_2 \rangle(t_r) - 1}.$$

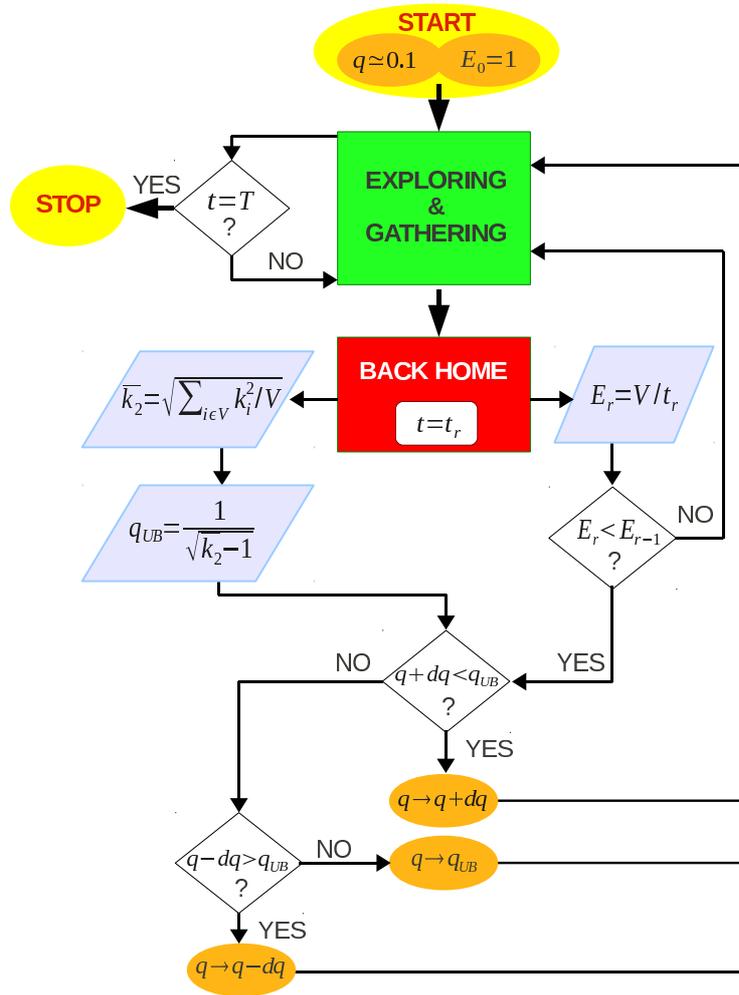


Figure 3.6: The figure represents, in a *flowchart* a possible implementation of the algorithm described in the main text for the adaptive strategy. Here the efficiency E takes the initial value $E_0 = 1$. The notation is simplified in order to make the diagram easier to read: $E(t_r)$ is indicated as E_r and $V(t_r)$ just as V .

- 5) Check if $q + dq < q_{UB}$, where dq is a small positive quantity (in general $dq = 0.01$ is a good choice).
- 6) If the condition (5) is satisfied, update the value of q adding dq : $q \rightarrow q + dq$.
- 7) If the condition (5) is not satisfied, but $q < q_{UB} < q + dq$, update the value of q so that $q \rightarrow q_{UB}$.
- 8) If $q > q_{UB}$ then:
 - 8a) if $q - dq < q_{UB}$ then $q \rightarrow q_{UB}$,
 - 8b) if $q - dq > q_{UB}$ then $q \rightarrow q - dq$.

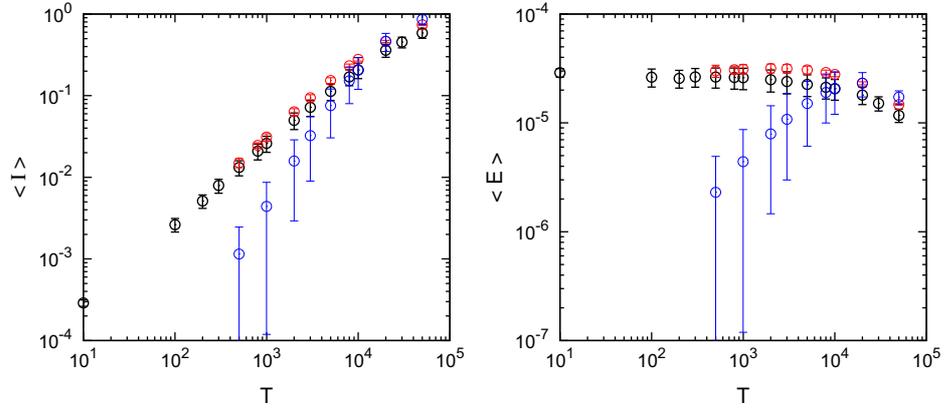


Figure 3.7: Left panel: the total efficiency $\langle \bar{E} \rangle = \langle \bar{V} \rangle / T$ as a function of the searching time T in the case of three searching strategies (averaged over all the nodes and 100 realizations for each of them): $q \gg 1$ (blue), $q = q_p$ (red) and the adaptive strategy (black). Right panel: mean information $\langle \bar{I} \rangle$ as a function of T , again for the three best searching strategies (represented with the same colors as above). Error bars represent the dispersion (standard deviation) among the values obtained for different home-nodes.

Figure 3.7 shows results for the final efficiency $E(T)$ and the information gathered $I(T)$ for the three best strategies: the adaptive one, $q \rightarrow \infty$ and $q = q_p$ (although this is not really a *strategy* since we need to know the precise value of q_p). Both quantities confirm that, unless T is more than twice the network size $N = 10000$, the best performance is obtained for $q = q_p$. Nevertheless, our adaptive strategy gives results that are very close to those obtained for $q = q_p$ and always better than those obtained for $q \gg 1$ (at least for $T \leq 2N$) both in terms of efficiency and in terms of the total amount of information recovered.

All these results are coherent with the description of the walkers' behavior commented on in the previous section. In particular, it is reasonable that when $q \gg 1$ the efficiency initially increases with T since in this case the shorter the searching duration, the larger the probability that an agent gets lost. On the contrary, for q_p and the adaptive strategy, which precisely aims at capturing the behavior of the agents at q_p , the information is mainly collected by means of quite short round trips. Consequently, increasing the searching time reduces the efficiency because it increases the chance to visit many times the same nodes. In any case, when $T \gg N$ and $I \sim 1$, the problem of visiting already visited nodes becomes relevant also for the strategy $q \gg 1$. Finally, it is worth stressing that while for $q \gg 1$, the dispersion among the values $E^{(i)}$ and $I^{(i)}$ for different home-nodes is very high, in the case of the other two strategies, the same does not happen. This is a clear indication of the fact that the adaptive strategy recovers one of the most interesting features of the agents' behavior at q_p , namely, the homogeneity of the performance starting from different home-nodes. We next discuss one potential application of the searching strategies previously discussed. This would also allow for a better distinction of what strategy is the best.

3.5 Recovering the degree distribution.

An important global descriptor of every network is its degree distribution $P(k)$. However, this information is not always at hand. For instance, suppose you belong to a network of which you only know your local neighborhood (like an online social network or a city map). The problem is then to know what is your position in the network as far as the degree is concerned or to make an exploration that allow you to gather information about the entire map. In other words, we want to study if the sample of nodes visited by an agent is more or less representative of the global system, at least with regard to its $P(k)$.

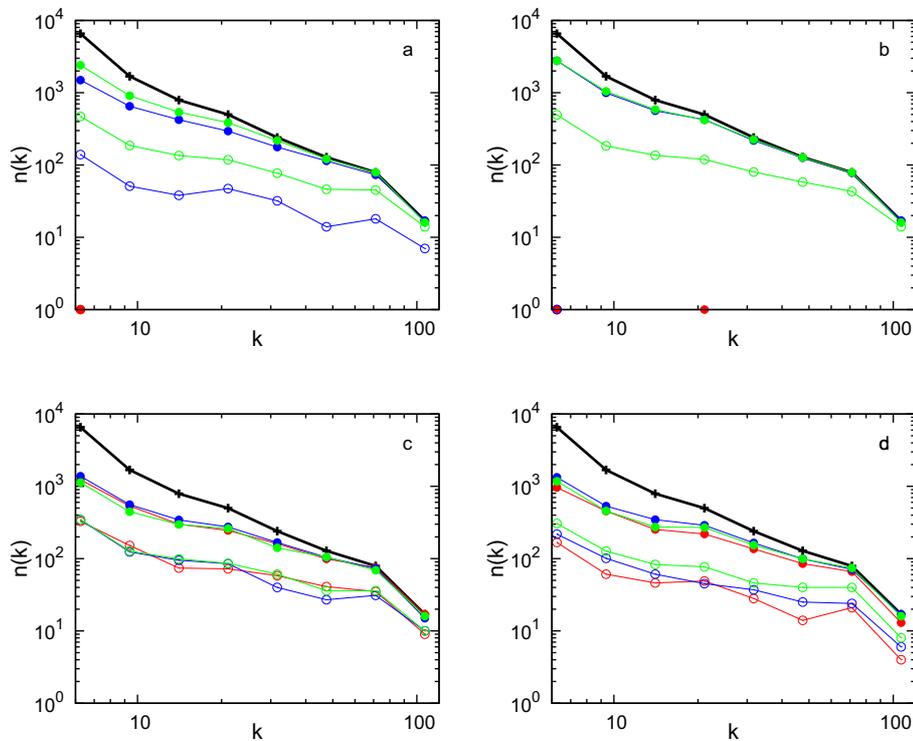


Figure 3.8: Number of nodes of degree k as measured by a walker searching during $T = 2000$ (void circles) or $T = 10000$ (filled circles) time steps, in a single realization of the process. Red color refers to a home-node of degree $k = 5$, blue to a home-node with degree $k = 22$ and green to the node with the largest degree in the network ($k = 100$). The searching strategies are, respectively: $q = 1$ in panel (a), $q \gg 1$ in (b), $q = q_p$ in (c) and the adaptive strategy in (d). The black line is the real degree distribution.

In Fig. 3.8 we plot the number of nodes of degree k , $N(k)$, found in a typical realization of the different strategies, for two different values of T and for different choices of home-nodes. As we expected, the usual random walker and the agent with $q \gg 1$ are very bad when the home-node has a small degree (red curves). On the contrary, the performance of the adaptive protocol and that of the walker when $q = q_p$ are almost not affected by the walkers' lifetimes

(at least for the considered values) and by the degree of the home-node. Note that panels (a) and (b) represent the common situation in which a random walker starting at a lowly connected home node gets lost. Indeed, for such cases, the only information brought back is the degree of the node from which the walker started the exploration of the network. However, as it is also appreciated in the figure, when the home node has a relative high degree, setting $q \gg 1$ constitutes the best strategy for an accurate estimation of $N(k)$. Nevertheless, as the walker "does not know" what is the connectivity of its home node in relation to the rest of the network, the latter strategy seems to be, as a rule of thumb, a bad choice. Additionally, the figure also shows that in general what is difficult for an agent to recover are the most peripheral nodes of the network. Consequently, the nodes with a small degree are usually under represented while the heavy tail of the degree distribution is reconstructed with high accuracy.

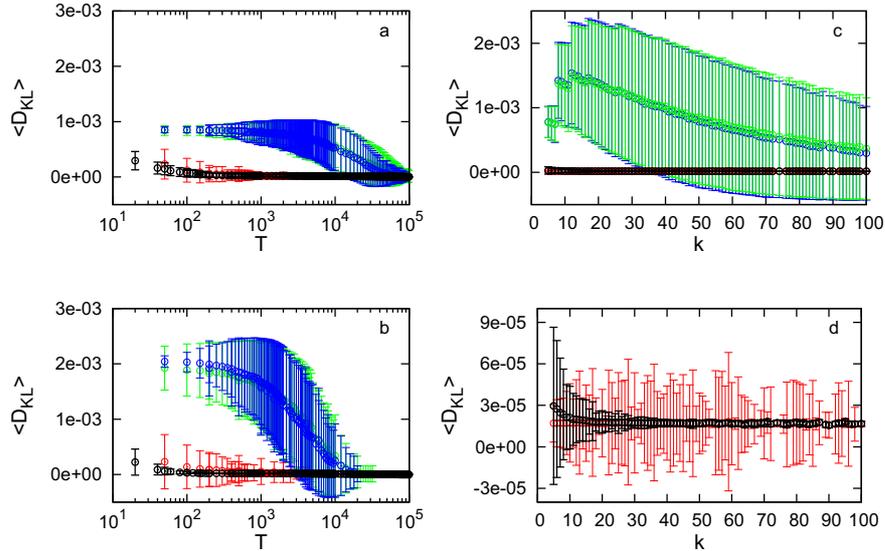


Figure 3.9: Kullback Divergence (D_{KL}) of the measured (normalized) degree distribution with respect to the real one. Different colors refer to different searching strategies: blue for $q \gg 1$, green for $q = 1$ (pure random walk), red for $q = q_p$ and black for the adaptive strategy. In panels (a)-(b) D_{KL} is plotted as a function of the searching duration T , averaged over 10^4 realizations. The considered home-node are, in panel (a), the node with the largest degree ($k = 100$) and, in panel (b), a node with degree $k = 5$. In panels (c)-(d), D_{KL} is plotted as a function of the degree of the home-node, averaged over $[1/p(k_H)]$ realizations for each starting node of degree k_H . The searching time is $T = 2000$. In panel (d) the adaptive strategy is plotted on a smaller scale together with the strategy $q = q_p$.

In order to quantify the accuracy of the reconstructed networks, we calculate the Kullback-Leibler Divergence or relative entropy [66], a non-symmetric measure of the difference between two probability distributions. This is a standard method to evaluate how different an experimentally estimated distribution is from the real one. For the probability distributions P

and Q of a discrete random variable their KL divergence is defined to be

$$D_{\text{KL}}(P\|Q) = \sum_k P(k) \log \frac{P(k)}{Q(k)}, \quad (3.5.1)$$

where $P(k)$ is the real distribution and $Q(k)$ the estimated one. Using this measure, we explore how the accuracy of the reconstruction depends on the searching strategy, the walkers' lifetimes and the degree of the home node. In what follows, we report results for the mean values, averaged over many realizations, and for the deviations around the means.

In Fig. 3.9, panels (a)-(b), we plot D_{KL} as a function of T for four different strategies (including the standard random walk) and two different starting nodes (corresponding to, respectively, maximum and minimum degree,). As expected, $D_{\text{KL}} \rightarrow 0$ when $T \rightarrow \infty$, in all the considered cases. The $q = q_p$ strategy and the adaptive protocol perform much better than the other two settings, with less dispersion and a very much weaker dependence on the degree of the home node. Hence, these last two strategies are more suited if we aim at recovering $P(k)$, especially when T is not too long. Moreover, even if they both are good, the adaptive strategy is better than $q = q_p$, with a very small dispersion. Thus, although it is not possible to perform better than $q = q_p$ in terms of nodes visited, the adaptive strategy does better in terms of the accuracy of $Q(k)$, that is, when it comes to reconstruct $P(k)$.

We have also analyzed the dependency of D_{KL} on the degree of the home nodes for fixed values of T . Figure 3.9, panels (c)-(d), displays D_{KL} as a function of k for all the strategies, in the case of a short lifetime ($T = 2000$). Differences among strategies are really noteworthy, while for the larger lifetime ($T = 10000$) we verified that they persist just in the case of quite small degrees of the home nodes. Finally, the adaptive strategy is in general the best option, being $q = q_p$ slightly better only in the case of home nodes with degree $k < \langle k \rangle$.

3.6 Conclusions

In this chapter, we have presented a model for network search and exploration in which walkers evaluate at each time step whether to go farther from a home node or get back with the information retrieved up to that moment. These probabilities depend on a single parameter q , that has been shown to exhibit an optimal value, $q = q_p < 1$ ($q = 1$ corresponds to the markovian random walk limit) for exploration times comparable to the system size. When the walkers are allowed to explore the network indefinitely or during long times, the optimal value turns out to be $q = \infty$. However, although the amount of information recovered for the latter choice could be maximal, the results are highly dependent on the degree of the home node: the smaller the degree of the node assigned to the walker, the less information the walker can get back home. As a matter of fact, for most of the nodes (recall that in a scale-free network most of the nodes are poorly connected), $q = \infty$ is not the best strategy.

Capitalizing on the behavior of the walkers as a function of q , we have also proposed an alternative algorithm in which the agents are allowed to tune the value of the parameter q to optimize the information retrieved. Through numerical simulations, we have shown that this mechanism allows an exploration as efficient as that performed setting $q = q_p$. Nevertheless, the adaptive scheme has the advantage that the value of q is changed dynamically, and therefore it overcomes the problem of fixing an *a priori* unknown optimal value q_p . We believe that this adaptive search protocol could be a valuable addition to the current literature as it performs optimally with a minimum (local) information about the network structure.

As a demonstration of the potentialities of the algorithms explored in this work, we have made use of the different searching strategies to address the problem of network discovery. As expected, the adaptive mechanism is the one whose performance, in terms of the quality and quantity of the information retrieved, is the best. Whether or not these kinds of strategies can be further developed and applied to the exploration of real networks is out of the scope of the present work, but we identify at least two scenarios in which they can be useful: the discovery of new connections in communication networks and the exploration of planar networks (i.e., city networks) using minimal local information. Notice that the network we used as a benchmark has a very strong cut-off and the largest hub is connected with just 1% of the nodes, something that is possible also in planar networks, which cannot be scale-free, strictly speaking. The results obtained in this chapter are quite general, holding whenever the maximum degree k_{max} is (at least) one order of magnitude larger than $\langle k \rangle$.

We hope that our work guide future research along these lines.

Chapter 4

Synchronization of moving integrate and fire oscillators

Among the many emerging phenomena we observe in nature, synchronization is one of the most paradigmatic examples. It can be roughly understood as the collective dynamics of units whose internal state evolves periodically in time and when they interact tend to synchronize their internal variables [31]. The achievement of the final synchronized state (if any) strongly depends on the interaction pattern of the system [16, 37]. Up to now, synchronization has been mainly analyzed in fixed topologies but we are witnessing the first evidences that links between agents can evolve in time [67, 68, 69]. The case in which this evolution of the network topology is an effect of the agents mobility is a particularly interesting case [27, 28, 29]. The effect of this changing patterns of interaction on synchronization features has been analyzed in different settings, for instance in chemotaxis [70], mobile ad hoc networks [71], wireless sensor networks [72], and the expression of segmentation clock genes [73].

In the recent literature, studies on synchronization in dynamically evolving complex networks have been mainly concentrated in the case when the topology changes very fast. This is the so-called fast-switching approximation (FSA) [74, 75, 76, 77], which replaces the real interaction between agents by the “mean field assumption” that all agents interact with an effective strength that corresponds to the probability that any pair of agents are connected.

Recently, it has been proposed a general framework of mobile oscillator networks where agents perform random walks in a two-dimensional (2D) plane [78]. It has been shown that FSA fails when the time scale of local synchronization is shorter than the time scale of the topology change due to the agent motion. New behaviors arise due to the interplay between instantaneous network topology, agent motion, and interaction rules. This framework, that reduces to FSA when velocity is high enough, is valid for models whose evolution can be well approximated by linear dynamics. This actually holds for models such as populations of Kuramoto oscillators [45, 39], whose evolution, after a short transient time, is very well described by a set of linear equations that can be solved in terms of spectral properties of the

Laplacian matrix [79].

Here, we focus on a dynamical system, a population of integrate and fire oscillators (IFO), where linearization is not a good approximation, since the evolution takes place in two different time scales. One for the slow evolution of the internal state variables (the phase and the orientation) and the other for the instantaneous interaction between the units (pulse coupling). During the last years it has been shown that the interaction structure plays a fundamental role in the dynamics of IFO networks. Zillmer *et al.* observed different dynamical regimes due to network connectivity in a system formed by inhibitory integrate-and-fire neurons that were randomly connected. Also, the underlying network structure can affect the speed with which the system reaches the synchronized state, as studied by Grabow *et al.* Usually, IFO have been used to model neural systems but we can also find some examples of applications in other fields, as for example in economy [80]. Models where the oscillators do not remain fixed, and the network of interactions changes with time can find a direct application in biological systems such as flashing fireflies, anurans and bush-crickets, that interchange light or sound signals while searching for potential mates [81, 82].

In the present case we show that the interplay between agents motion and phase evolution towards a synchronized state presents different asymptotic behaviors, reminiscent of the observation in Kuramoto oscillators [78] and agents using communication protocols [83].

In the next section we introduce a general model in which agents interact with neighbors within a certain distance, as proposed in [78] and [83]. We characterize the behavior of the system for different regions of the parameter space (velocity of the agents and range of interaction) and later we identify the different microscopic mechanisms that lead the system to a globally synchronized state.

Then, we focus on a limiting case. In Sec. 4.2, we introduce a minimal model where each unit interacts with its nearest neighbor only. We analyze in details this extreme situation that, once fixed the interaction rule, allows to explore other dimensions of the parameter space by varying the coupling strength and the size of the system (number of oscillators), apart from with the velocity. First, in Sec. 4.3, we briefly describe how the strict rule we adopt produces new kind of interaction patterns that affect the synchronization time dependence on the velocity. We identify three different regimes, according to the synchronization properties of the system. Namely, we distinguish a slow regime, a fast switching limit, and, between them, an anomalous intermediate region where the synchronization time diverges. A qualitative characterization of the first two cases is provided in Sec. 4.4. Then, we determine the corresponding regions of the parameters space for the fast limit (Sec. 4.5) and for the slow regime (Sec. 4.6), also providing an estimator of the synchronization time. In Sec. 4.7, we take stock of which conditions help or prevent the achievement of the coherent state, putting forward some empirical explanation about what makes the system unable to synchronize in the intermediate region. In Sec. 4.8, we sketch some possible improvements of the interaction rule, evaluating their costs and their

effects on the efficiency of the synchronization mechanisms. Finally, Sec. 4.9 is devoted to summarize the obtained results.

4.1 The geometric model

We propose a setting in which a population of N IFOS [81] move at a constant velocity V in a bidimensional plane of size L with periodic boundary conditions. Each agent has two degrees of freedom corresponding to an internal phase $\phi_i \in [0, 1]$ and orientation $\theta_i \in [0, 2\pi]$, both randomly set in a uniform manner at the initial configuration.

The phases of the agents increase uniformly with period τ ,

$$\frac{d\phi_i}{dt} = \frac{1}{\tau}, \quad (4.1.1)$$

until they reach a maximum value of 1, when a firing event occurs. Then the phase is reset. Upon this event at time t , the firing oscillator influences its *nearest neighbors* (oscillators at minimal distance) producing an update in their phases by a factor ϵ :

$$\phi_i(t^-) = 1 \Rightarrow \begin{cases} \phi_i(t^+) = 0 \\ \phi_{nn}(t^+) = (1 + \epsilon)\phi_{nn}(t^-) \\ \theta_i(t^+) \in [0, 2\pi] \end{cases} . \quad (4.1.2)$$

Notice that we have kept the dynamical evolution of the units at its maximal simplicity since we are mostly interested in the interplay between motion (and hence construction of a dynamical network) and internal dynamics and how synchronization emerges as a collective property of the system.

The fire event, its propagation, and the phases and orientation updates it implies, have to be considered as instantaneous events. No *ordinary* phase evolution or motion takes place in the meanwhile.

The system is synchronized when we encounter in the system a succession of consecutive firing events (avalanche) equal to the system size N , since after this fact all oscillators will remain synchronized forever because all of them will have the same period τ , regardless of the interactions. For the sake of clarity we define the (discrete) time T , as the number of times a given oscillator (that we will identify with oscillator 1 in our computer simulations) has fired. This allows us to define T_{sync} as the number of cycles this reference oscillator takes to enter the synchronized state (i.e., the number of updates needed for an avalanche of size N to occur).

Neighbor selection upon a firing event is performed according to a geometric condition, as shown in Fig. 4.1: every agent scans a circular area of radius R around it and shots the neighbors therein.

We introduce a parameter $r \in [0, 1]$ that indicates the fraction of the system available for

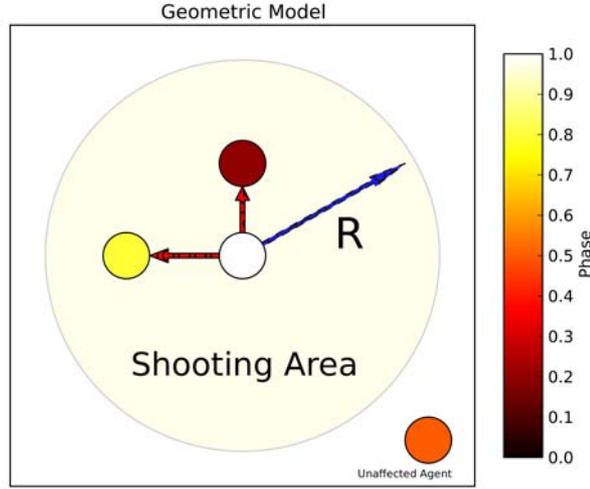


Figure 4.1: The model of interaction between oscillators, based on geometrical constraints. Only those within a distance R are affected by the firing of the central one.

interaction and relates both R , L variables and the average degree of the nodes of our evolving network

$$r = \frac{\pi R^2}{L^2} \quad \langle k^{out} \rangle = (N - 1)r. \quad (4.1.3)$$

Hereafter, we use fixed parameters $L = 100$, $\tau = 1$, $N = 50$ and $\epsilon = 0.02 \sim \mathcal{O}(1/N)$, while analyzing the explicit dependence on the mobility parameters, r and V .

We present in Fig. 4.2 the results of our simulations. A preliminary observation points out that the roles and importances of V and r change throughout our map.

Notice that the type of proposed interaction range in this system has been reported to show statistical properties similar to a continuous percolation [78], that in the case of static oscillators occurs for approximately $r_c \approx 4.51/(N - 1) = 0.09$ [84, 85]. In our range of study, we hope to observe some traces of this percolation as well as saturation properties observed in other moving oscillator systems at high speeds [78]. Actually, for high enough values of r ($r \gtrsim r_c$), the synchronization time, T_{sync} , is almost unaffected by the values of the speed V . Although this time is dependent on r , its range of possible values is much narrower (by orders of magnitude) than below r_c . When r is smaller than this critical value, the velocity plays the crucial role since for a fixed value of r the synchronization time changes by a factor of 10^4 .

To bring further insight to the map, we show a profile of the final “energy cost” rT_{sync} of the synchronization dependence on r for various velocities (right panel). Considering that each interaction between a pair of oscillators has a unit cost (proportional to ϵ), this final cost variable balances the range of interactions (the number of shots to different neighbors at each firing event) and the number of firing events that the system needs to synchronize (equal on

average to $N T_{sync}$).

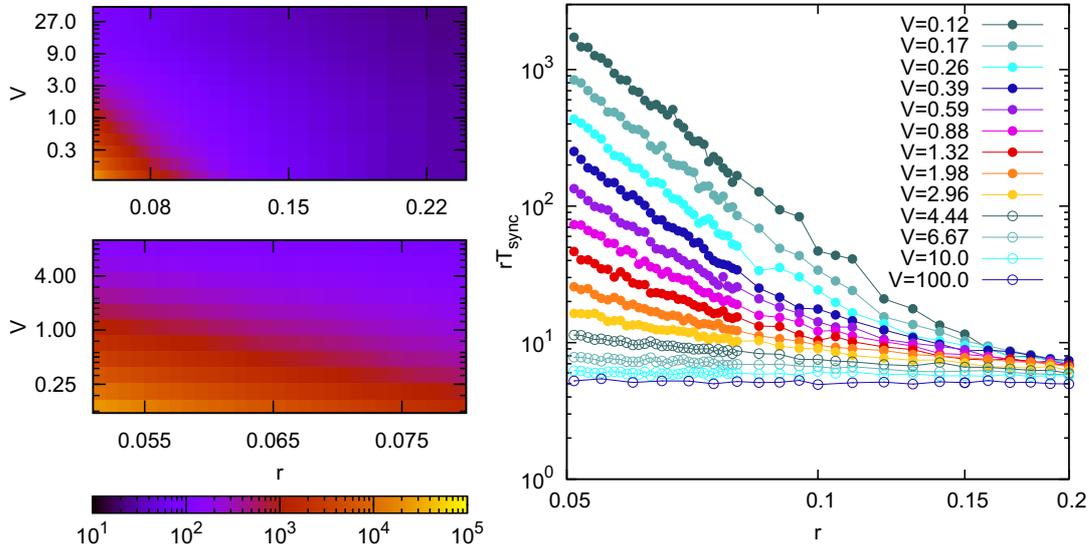


Figure 4.2: On the left: Heat map of the synchronization time as a function of r and V ; in the top picture a large region of the parameters space has been considered, while the bottom figure is a zoom on the most critical region ($r < 0.08$, $V < 0.25$). On the right: Profiles of the efficiency of the system rT_{sync} against r , for several values of V .

4.1.1 Synchronization Mechanisms

It seems clear that mobility of agents helps to minimize energy consumption but as we approach the critical percolating value of r_c , we observe that a range of behaviors emerge. On one hand, we observe for high velocities that the efficiency of the process remains roughly constant independently of r , since the extreme mobility of the agents compensate the reduced range of its interactions and successfully diffuses the synchronization process around the system, a process that is equivalent to the observation in other settings [75, 78, 83]. On the other hand, if the mobility of the agents is reduced, then the path to synchronization is more lengthy as well as more energy consuming. Synchronization is still possible (below the percolation static limit r_c) but the time to achieve it grows very fast, even resulting in an effectively infinite time¹.

These results lead us to identify different regimes and the consequent transitions between them. On the one hand, we find a “topological” transition at the critical static value r_c , since above it there is basically no velocity dependence whereas below r_c the influence of V is determinant. On the other hand, below r_c , where synchronization is made possible by the mobility of the agents, we can identify a transition (depending on both r and V) that separates

¹In our simulations some realizations of the experiment did not reach synchronization, even working with a reduced number of oscillators and the described boundary conditions, fact that induces us to think in this direction.

two dynamical regimes.

In order to study the distinct mechanisms that the system uses in its path to synchronization, we introduce an order parameter $\eta(T) = \langle \cos(2\pi\phi(T)) \rangle$ that is an increasing function that measures the overall synchrony of our system, ranging from a uniform phase distribution of our oscillators ($\eta(T) = 0$) to complete synchronization² ($\eta(T) = 1$).

To couple our V and r control parameters we also introduce the number of distinct interactions per oscillator N_c^i (accumulated encounters with different agents by a oscillator i). This value is bounded (on average) between a minimal starting value of $\langle N_c \rangle_0 = (r - 1)N$ and a maximal value of $\langle N_c \rangle_{\max} = N - 1$ and it provides information about the evolution of our system's synchronizing mechanism. Since the bounding of this value depends on r , we introduce a normalized magnitude χ ,

$$\chi(T) = \frac{\langle N_c \rangle(T) - \langle N_c \rangle_0}{\langle N_c \rangle_{\max} - \langle N_c \rangle_0} \quad \langle N_c \rangle = \frac{1}{N} \sum_{i=1}^N N_c^i, \quad (4.1.4)$$

that is a quantification of the *mixing* of our system. When the mixing is minimal ($\langle N_c \rangle = \langle N_c \rangle_0$) χ is 0. As the system mixes, i.e. the oscillators increase its average number of contacted neighbors, it can grow up to its maximum value $\chi = 1$, i.e. $\langle N_c \rangle = \langle N_c \rangle_{\max}$.

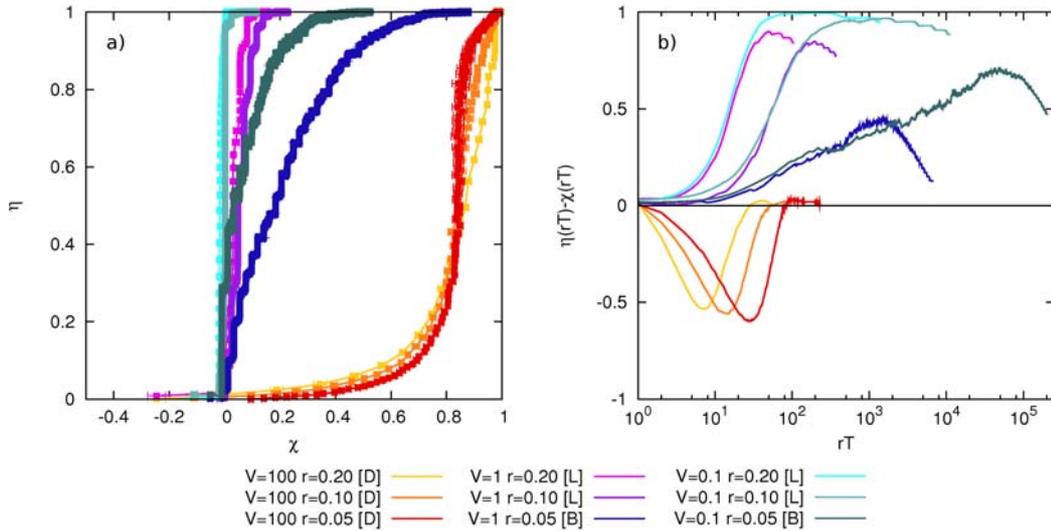


Figure 4.3: Panel a): η against χ for several values of r and V . Panel b): the difference between the two order parameters ($\eta - \chi$) as a function of rT . Letters [D], [L] and [B] stand respectively for “diffusive”, “local” and “bounded” regimes. The values of η and χ at each time instant have been calculated averaging over 1000 realizations.

We have calculated the time evolution of both η and χ for different values of $V \in \{0.1, 1, 100\}$

²Note that the average is calculated upon a firing event by the reference oscillator, hence our order parameter is an average of the phase difference of the other oscillators with respect to this one.

and $r \in \{0.05, 0.1, 0.2\}$. In Fig. 4.3 one can see the evolution of both parameters measured at the same time instants (and sufficiently averaged over enough realizations) together with the difference between them over time $\eta(T) - \chi(T)$ that give insight about the topological evolution of our system as it synchronizes. The figures show three clearly distinct patterns.

For high velocities we observe in fig. 4.3a) a gradual increase of the order parameter and a minor influence on r at fixed V , indicating that the synchronization emerges evenly on the system in a global fashion, due to the quickly changing topology of the network (neighbors of a given oscillator change rapidly). This regime (which we call *diffusive*) requires the inter-contact of the majority of the system, but this circumstance is rapidly achieved due to the strong mobility of the agents. In fact, this regime is optimal as far as the synchronization time is concerned, since the interactions are more effective. These conclusions were obtained for populations of Kuramoto oscillators [78], for which this regime corresponds to the region of validity of the FSA.

In the opposing case, when velocities are small enough, this behavior is completely lost and a step function appears, indicating that the slow mobility of the agents allows them to synchronize locally with their neighbors creating *islands* of synchrony. The sudden increase of the order parameter occurs at a regular pace, fact that points out that whenever the islands are disbanded (change of neighbors), they still transmit the local synchrony to the neighboring groups, mechanism that allows for system synchrony while keeping χ in small values. The initial height of the steps is dependent on r and decreasing as $\chi(T)$ grows due to the limited range of $\eta(T)$ available states.

Finally, as we decrease r approaching the critical value r_c , we observe a transition from a *local* to a *bounded* regime, where the synchronizing time is so long that again allows for the interaction of the majority of the agents among themselves upon synchronization time (due to the bounded nature of the system). In this regime, the range of interaction is very reduced, and so is the size of the clusters, so an agreement between the multiple clusters created (if any) comes after almost all the system has interacted. Consequently, the increasing of η with χ is slower (many small steps, see fig. 4.3a)) while the final value χ becomes larger (fig. 4.3).

In fig. 4.3b) we provide an explicit time evolution of the difference $\eta(T) - \chi(T)$ in order to make the three regimes and the influence of r better identified.

It is interesting to study the final mixing of our system upon synchronization as shown in Fig. 4.4. This value $\chi(T = T_{sync}) \equiv \chi_{sync}$ together with T_{sync} characterize the evolution of the system towards the synchronized final state. These features depend on both r and V . At fixed velocity the final mixing of the system decreases as the interaction range grows. This is caused by the fact that although an increased r induces more mixing (as oscillators find new neighbors more easily) it also drastically reduces (below the critical values r_c) the synchronizing time T_{sync} thus reducing the chances of encounters between different oscillators. Above r_c we observe a saturation of the values as the dependence of T_{sync} in r and V is practically lost.

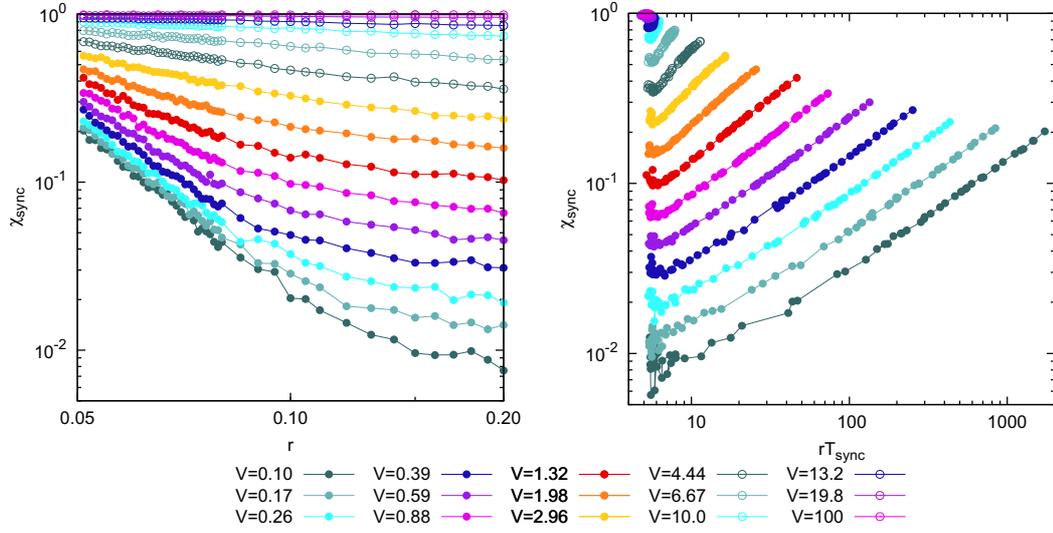


Figure 4.4: On the left: the average value of χ at the synchronization time (χ_{sync}) as a function of r for several values of V . On the right: χ_{sync} as a function of rT_{sync} for several values of V . All the means have been performed over 200 realizations.

Figure 4.4 b) shows a change between the relation pattern of χ_{sync} and cost rT_{sync} ranging from the *diffusive* regime (mixing independent of rT_{sync}) to the curves where both the *local* and *bounded* regimes are shown. It is important to note the similarity of the observed shapes (for a wide range of V values) where we find the local phase concentrated around the minimum value of χ_{sync} that gradually grows in a power law fashion as the performance of the system decreases (it consumes more energy to synchronize).

The introduction of the new parameter χ_{sync} allows us to present a phase diagram of our system relating the overall performance (in terms of energy cost) with the synchronizing mechanism used (Fig. 4.5). We identify the *diffusive* regime in the zone of high velocities $V \sim \mathcal{O}(10)$ where both values of χ and rT_{sync} are almost independent of r . This zone falters into the *bounded* one as V and specially r decrease, where both the mixing and the efficiency of the system is reduced (energy cost increased). Finally for small enough velocities the local zone is clearly visible with low values of system mixing. From the map one clearly observes that the most beneficial synchronizing mechanisms (in terms of energy consumption) are the diffusive and local ones.

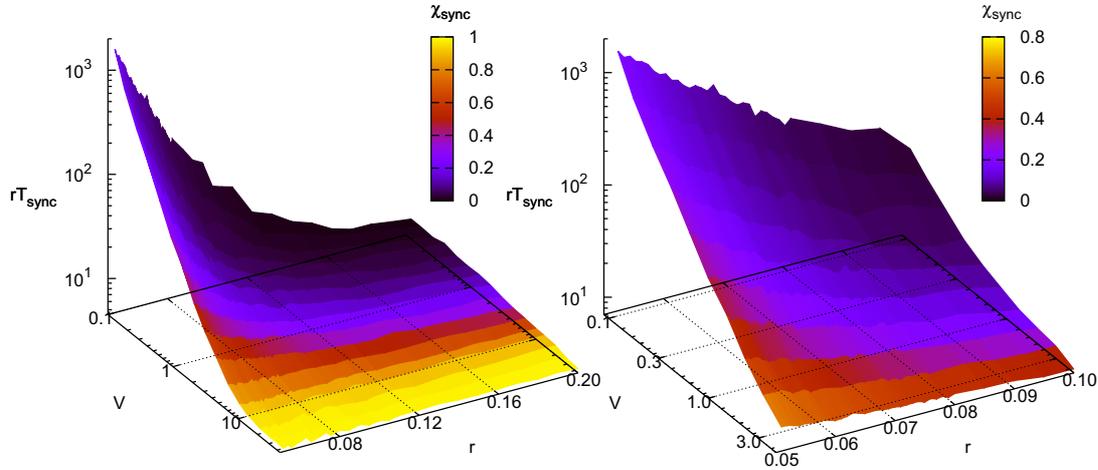


Figure 4.5: On the left: The cost rT_{sync} as a function of r and V . The heat map of χ_{sync} has been superimposed to the surface $rT_{sync}(r, V)$. On the right: the same quantities, zoomed in the most sensitive region of the parameter space.

4.2 Minimal topological model

In the previous section, we have analyzed the system behavior keeping both the coupling strength ϵ and the system size N constant, by changing solely the velocity of the agents and the interaction range. We have seen that the motion of the agents is what ensures that the system will be able to synchronize. If the agents are static, a minimal fixed topology will be required to connect them (what is called a giant component in network terminology). In contrast, a population of moving agents (even when the interaction range is small) will eventually synchronize.

Now, we are interested in studying the dependence of the synchronization time on the coupling strength and on the size of the system. Therefore, in order to limit the number of free parameters, we need to fix the interaction range. Since we are interested in observing the maximum possible diversity of behavior when varying the velocity, we want the system to lie below the percolation threshold. Hence, we consider that the simplest rule we can adopt is a minimal distance rule, so that each oscillator interacts with its nearest neighbor only. An additional advantage of this choice is that the total energy cost can now be directly identified with the synchronization time, thus simplifying the evaluation of the system's performance. Finally, having reduced the multiplicity of the *firing event* to a single shot at a single oscillator, this minimal model is suitable to represent phenomena where the *signal* is transferred from one agent to another instead of spreading around simultaneously towards many agents. This can be the case of communication protocols with special restrictions for security/privacy reasons.

Additionally, we decided to change the orientation updating rule, relating it to the signals

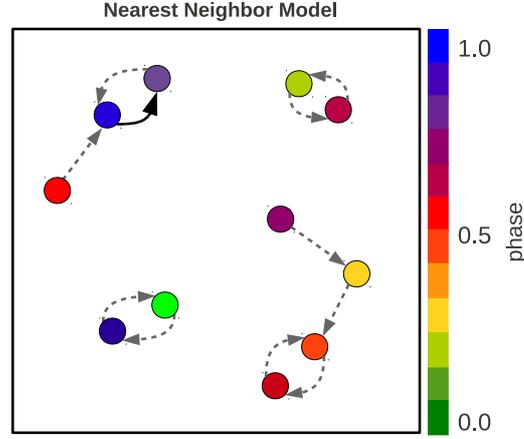


Figure 4.6: Snapshot of the system in the incoherent state. Each disk is an oscillator, gray dashed arrows stand for first-neighbor relationships, while the continuous black one represents a firing event that is taking place at that precise instant.

an oscillator receives, instead of the signal it sends. Notice that, for small values of the velocity, the displacement of a given oscillator increases linearly with time only if it does not change its orientation at each time step. On the contrary, if it is permanently changing the direction of its motion, the displacement will be proportional to the square root of the time, similarly to a brownian motion but with finite time step. According to our new interaction rule, it is assured that each agent, every time it *fires*, reaches some another unit, while the same is not true for the incoming shots, whose time separation is a random variable. Oscillators can in principle remain a long time without any incoming neighbor, thus delaying the achievement of the synchronization. In order to improve the efficiency of the system, which is critical at low velocities³, we chose to allow the oscillators that are not receiving any signal to keep going straight on, preventing at the same time those with one or more *in* neighbors to do the same.

Thus, upon a firing event at time T , the firing oscillator influences its *nearest neighbor* (oscillator at minimal distance) producing a random reorientation of its motion and an update in its phase by a factor ϵ :

$$\phi_i(t^-) = 1 \Rightarrow \begin{cases} \phi_i(t^+) = 0 \\ \phi_{nn}(t^+) = (1 + \epsilon)\phi_{nn}(t^-) \\ \theta_{nn}(t^+) \in [0, 2\pi] \end{cases} . \quad (4.2.1)$$

Notice that the chosen interaction rule has two important features that differentiate it from the previous one, deeply influencing the dynamical behavior of the system. First of all, the resulting interaction pattern is directed, since there is no necessity for the first-neighbor relation to be symmetric. Second, the outgoing degree of each oscillator (number of reached unit when

³We have seen that the energy cost dramatically increases when both V and r are small (see Fig. reffig5g)

it fires) at each time T is fix and equal to 1, while the *in* degree (number of units that fire to it) can vary between zero and five⁴.

In the following sections, we are going to study the behavior of the synchronization time as a function of the velocity V . We also investigate the parametric dependence of $T_{sync}(V)$ on N and ϵ , while keeping L fix. The dependence on L is trivial since, in the same way as τ sets the time scale, similarly, the only role of side of the box is that of fixing the space scale. The nature of the adopted interaction rule is such that the density of the oscillators does not play any role. Changing the linear dimension of the box, we are just rescaling the distribution of the distance between an oscillator and its first neighbor. Consequently, we are rescaling the time that any agent needs to cover the characteristic length of the system. Modifying the velocity in the opposite sense, it is possible to exactly compensate the variation of L . Indeed, it can be shown that any quantity that depends on V is a function of $v = V/L$ and the relation $f(V; aL) = f(V/a; L)$ always holds (see Fig. 4.7, inset). Hereafter we take $L = 400$.

4.3 Time to synchronize

The chosen minimal interaction rule is such that the system lies far below the static percolation transition. Therefore, without motion, global synchronization is not achievable. Since it is the non-null velocity of the oscillators that enables the system to reach the coherent state, we could expect T_{sync} (the average time the system needs to synchronize) to be a decreasing function of V , such that $T_{sync} \rightarrow \infty$ when $V \rightarrow 0$ and $T_{sync} \rightarrow T_f > 0$ (a constant value) when V is high enough⁵

Although both limits are correct, T_{sync} is not a monotonous function of V . On the contrary, there is an intermediate region of values of the velocity where the synchronization time diverges.

Fig. 4.7 shows that, for $V < V_s$, the synchronization time decreases as a power of V when the latter increases. Then, the decreasing slows down and T_{sync} has a minimum at $V = V_m > V_s$. Beyond this value, the synchronization time gets larger and larger, until the system enters a region where it is unable to reach the coherent state in a finite time (gray area in Fig. 4.7). For even larger values of the velocity, $V > V^*$, it decreases abruptly, finally reaching its asymptotic value when $V = V_f$.

⁴Consider a point A in a planar euclidean space. Then add another point P_1 , plus a second one P_2 and locate it closer to A than to P_1 . Then, add a third point closer to A than to P_1 and P_2 . While for P_2 there is no problem, for P_3 it depends on the positions of the other points whether this is possible or not. It is easy to show that the maximum number of points one can locate is six, while the conditions on the positions of the rest of the points become stricter and stricter. Finally, the sixth point can be added only if the previous five point are located at the vertices of a hexagon, a condition that has null probability if the positions are randomly extracted. Therefore we say that the maximum number of oscillators that can have the same nearest neighbor, and so the maximum number of possible in-neighbor of the same agent, is five [86].

⁵When V is such that the agents, at each time step, cover a distance of the same order of magnitude as of the linear dimensions of the box, then the interactions are completely randomized (fast switching approximation). A further increasing of V does not produce any effect.

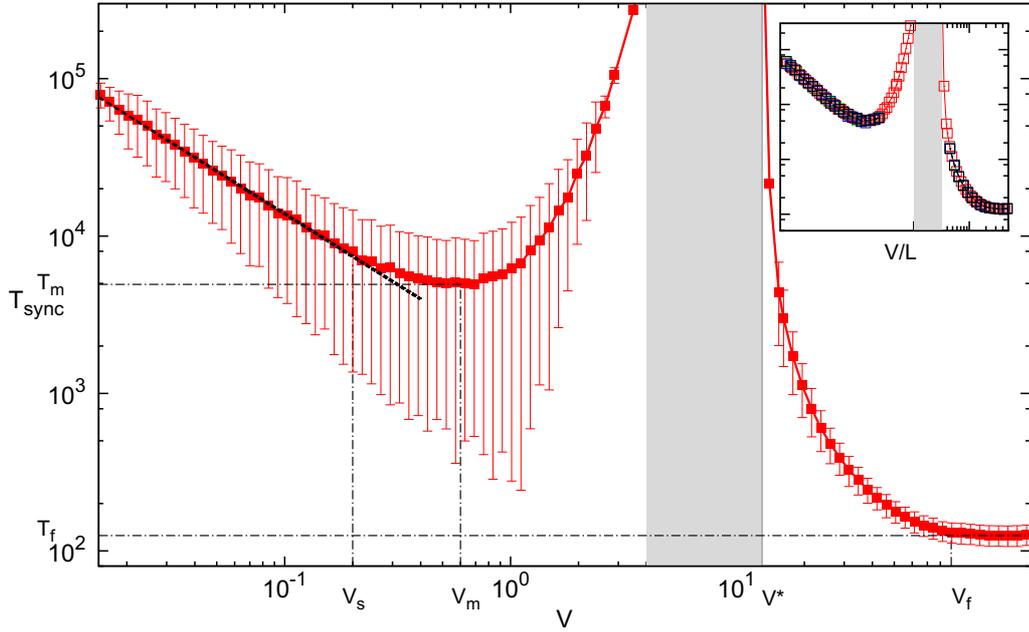


Figure 4.7: The average synchronization time T_{sync} as a function of V , for $L = 400$, $N = 20$, $\epsilon = 0.1$. In the following, when not otherwise states, the values of the parameters are those used in this figure. In the inset: T_{sync} against V/L , for $L = 1200$ (black), 800 (blue), 400 (red) and 200 (green). Averages are performed over 2000 realizations.

Hence, we distinguish three main regions. A “no-synchronization zone” in the middle, a “left region” at smaller values of V , and a “right region” at larger ones. In the left region, we can separate two sub-regions: on the left, at $V < V_s$, there is what we call the *slow regime*, and, on the right, a transition zone. The same happens for the right region. On the left, we find a transition region, while on the right, at $V > V_f$, the system enters the *fast limit*, where T_{sync} no longer depends on V .

4.4 Characterizing the system behavior

Before entering into a more specific discussion on the synchronization time, a general characterization of the system behavior can be useful.

In Fig. 4.8 we represent the cumulative individual interaction network (CIN) of an oscillator (labeled “0”), for two realizations of the synchronization process at two different velocities. Such a network has been constructed in the following way: whenever an oscillator i fires at oscillator 0 and oscillator 0 fires at oscillator j , a link between i and j is added. If the link already exists, its weight is increased. In the case 0 does not receive any shot, we put a link between 0 and j . A reciprocal shot is represented as a self-link. We repeat this process until the system reaches the synchronized state. By drawing its CIN, it is possible to visualize the

role played by a given unit in the *signal spreading*. In panel A, the velocity is $V = 2V_f$, while in panel B it is $V = V_m$ (see Fig. 4.7). Notice that, in panel A, all the nodes are represented in the picture while, in panel B, there are just a few of them. This means that, when the oscillators are moving fast, each one of them plays a global role, sending and receiving signals throughout the whole system. On the contrary, when the velocity is slow, each oscillator plays a local role mediating the interactions among a small number of units that are the same all the time, no matter how long the synchronization process could be.

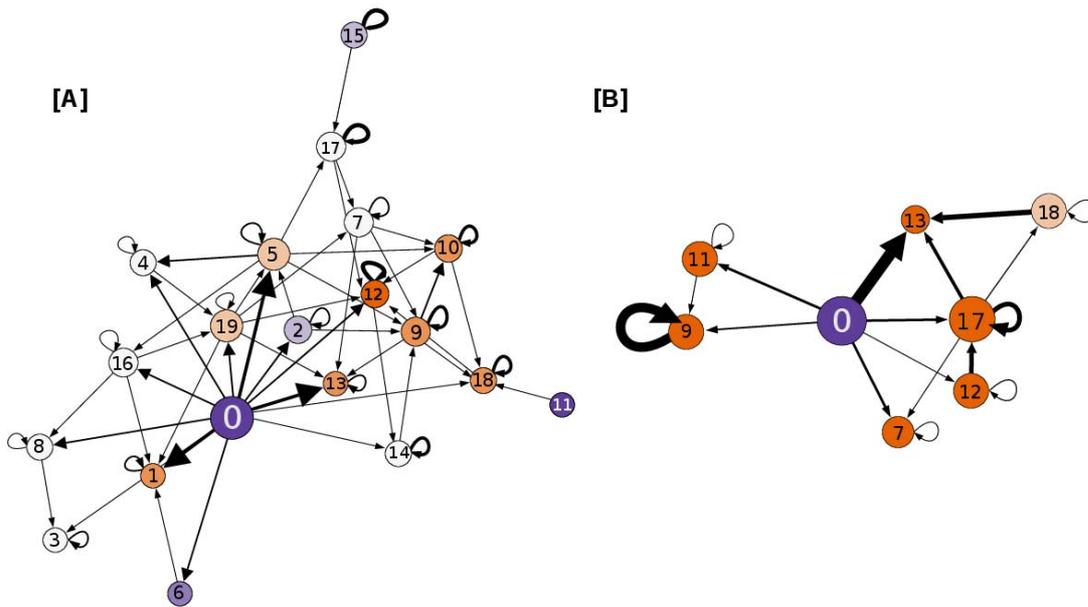


Figure 4.8: Final ($T = T_{sync}$) network of the interactions mediated by a single oscillator (labeled "0"), in the fast limit, at $V = 2V_f$ (panel A) and at $V = V_m$ (panel B) respectively. Node color changes from purple to orange increasing the in-degree. Size increases with increasing out-degree. The weights of the links are proportional to occurrence of the interactions.

In Fig. 4.9 we represent the cumulative total interaction network (CTN), where each link between i to j stands for a shot between i and j , occurred at any time $T < T_{sync}$. Any repetition implies an increasing in the link's weight. Drawing the CTN allows to visualize the interactions that took place in the system while it was evolving toward the coherent state. In the fast limit (panel A), the resulting topology is an all-to-all network, with almost homogeneous weights. In the other case ($V = V_m$, panel B), the network has a single connected component of size N , that is a necessary condition for synchronization, but the out-degrees and in-degrees are heterogeneously distributed and their mean value is quite small.

Summarizing, this preliminary observations point out that the fast regime can be understood as an *homogeneous* regime, while slow velocity enhance heterogeneity among units. Hence, the mechanisms that allow the system to synchronize have to be different in the two cases. Our

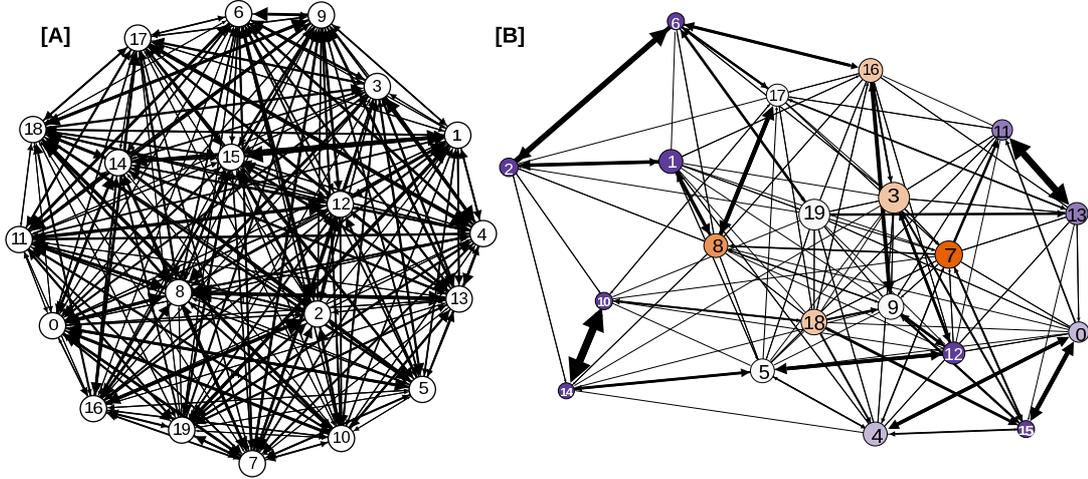


Figure 4.9: Final ($T = T_{sync}$) total interaction networks, in the fast limit (panel A) and at $V = V_m$ (panel B) respectively. Node color changes from purple to orange increasing the in-degree. Size increases with increasing out-degree. The weights of the links are proportional to occurrence of the interactions.

hypothesis is that the system has different *strategies* to reach the coherent state in the left and in the right region, while neither of them work in the intermediate region.

In order to clarify this point, in Fig. 4.10 we have plotted $\eta(T)$ against T (blue line) for the same values of the velocity used in Fig. 4.8 and 4.9 and a single realization for each one. Then, to have an insight of what happens at the local scale, we also plotted the quantity $\lambda(T) = \cos(2\pi\phi_{nn}(T))$ (red line), where $\phi_{nn}(T)$ is the phase of the oscillator to which the reference oscillator fired at time T . The black line represents $m(T)$, the total number of oscillators that have been out-neighbors of the oscillator of reference until that moment.

Again, two deeply different behaviors correspond to the different regimes. In the fast case (panel A), since the velocity is high, interactions are completely rewired at each time T . Therefore $m(T)$ increases very rapidly, and therefore ϕ_{nn} is just a random variable extracted among $N - 1$ possible ones. This means that λ is exactly the same as η , but with less statistics. Both quantities increase together (more or less noisily) because, by means of the firing events, the whole phase distribution becomes narrower. Then, η and λ reach a value that is very close to 1 and they do not fluctuate any more. However, full synchronization has not been achieved yet. The system lies in an *almost-synchronized* state, all the oscillator having almost the same phase, just with a very little dispersion of their values. Remarkably, the system spends around one half of the total synchronization time in this state before being able to produce an avalanche of size N .

In the slow case (panel B), things go completely different way. The behaviors of η and λ appear to be uncorrelated, being $\lambda = 1$ almost all the time. The considered oscillator (as anyone else) spends a quite long time with each one of its neighbors, usually being able to synchronize

with it before leaving. At the beginning ($T < 500$), whenever it starts firing toward a new oscillator (black vertical lines in Fig. 4.10), λ experiences an abrupt decreasing while $m(T)$ increases, that means that it is the first time the reference oscillator meets that neighbor. Later ($T > 500$), the chances to change a neighbor for another one it has already known, and that has almost the same phase, increase. The phase distribution has become narrower and, especially at local scale, among the units the oscillator of reference can meet, the dispersion is small. Consequently, neighbor changes do not affect λ anymore. It suggests that, unlike in the fast regime, global coherence can be achieved through local synchronization.

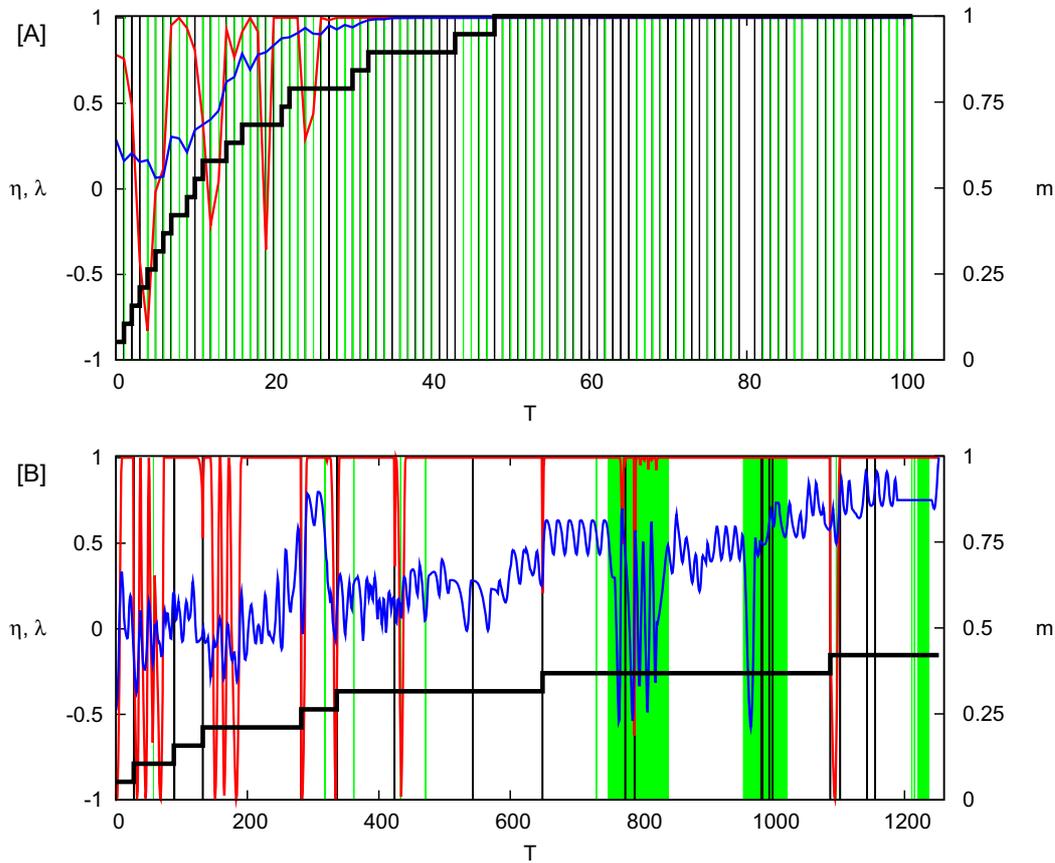


Figure 4.10: Global (η , blue line, left axis) and Individual (λ , red line, left axis) order parameters are plotted together with the Individual Mixing (m , black line, right axis) as a function of time, from $T = 0$ to the synchronization time. Green vertical lines mark a change of the in-neighbor, while the black lines stand for a change of the out-neighbors. In panel A, the velocity is $V = 2V_f$. In panel B, it is $V = V_m$. Where green lines are so dense that they form a green band, it means that the oscillator has more than one in-neighbor simultaneously.

4.5 Fast limit

Here we want to study the asymptotic behavior of the synchronization time at $V > V_f$ (see Fig. 4.7), where it no longer depends on V . The value V_f , at which the velocity-dependence of dynamical quantities ceases to exist, can be empirically estimated. A heuristic argument leads us to set⁶ $V_f = L/\sqrt{N-1}$ (see Appendix B.1).

Since the first preliminary observations (Fig. 4.7-4.9), it was clear that the fast limit was the most efficient. We described it as the most homogeneous regime of a model that is intrinsically deeply heterogeneous, having lost even a very general feature of this class of models, that is, an undirected connectivity pattern. Indeed, the fast motion of the units randomizes the interaction pattern enabling a relevant reduction of the differences among the oscillators, and erasing in very few steps every memory of the initial positions and orientations. In this sense, it appears to be similar to the static all-to-all model since, at the end of the synchronization process, every unit will have interact almost equally with any other. Therefore, as usually happens in all-to-all IFO systems, only values of $\epsilon < 1/N$ make sense. Indeed, when $\epsilon > 1/N$, the quantity $T_f(N, \epsilon)$ is no longer monotonous in ϵ (see Fig4.11, panel A). Moreover, keeping ϵ fixed ($\epsilon = 0.1$), and varying N from $N_{min} = 10$ to $N_{max} = 80$, we observe that for $N > 20$ the behavior of $T_{sync}(V; N, \epsilon)$ at V slightly less than V_f starts to change and the dependence on ϵ seems to become less and less important (see Fig4.11, panel B).

Here, we will restrict ourselves to the region $\epsilon \leq 1/N$ and we will empirically derive the form of the functional dependence of $T_{sync}(V > V_f; N, \epsilon) = T_f(N, \epsilon)$ on N and ϵ .

Fixing $N = 20$ and varying $\epsilon \in [10^{-3}, 10^{-1}]$, we find that values $T_f(20, \epsilon)$ can be fitted by a function of the type $f(\epsilon) = C\epsilon^{-\beta}$, with $\beta = (0.95 \pm 0.01)$ (Fig4.12, panel A).

Then, by varying N , we find that a function of the type $f(N) = cN^a$ fits the values of $C(N)$, with $c = (1.4 \pm 0.1)$ and $a = (0.37 \pm 0.01)$. Finally, we can write the general expression for $T_f(N, \epsilon)$:

$$T_f(N, \epsilon) = cN^\alpha \epsilon^{-\beta}, \quad (4.5.1)$$

with $\alpha = (0.37 \pm 0.01)$ and $\beta = (0.95 \pm 0.01)$ (see Fig. 4.12, panel B). This relation has been verified for $\epsilon \sim 1/N$ up to $N = 150$.

4.5.1 Assessment of the energy cost

In Sec. 4.1.1, we showed that, whenever the velocity is high enough, the energy cost appears to be minimal, even if the interaction range is very small. However, a quantitative demonstration of this qualitative observation was not provided. The present model, that represents the extremal case $r = 1/(N-1)$ (without fluctuation), can be useful in order to check the hypothesis of a complete independence of the efficiency on r in the fast regime.

⁶Since τ has been fixed equal to 1 once for all, for the seek of notation simplicity, hereafter it will not be included in any expression.

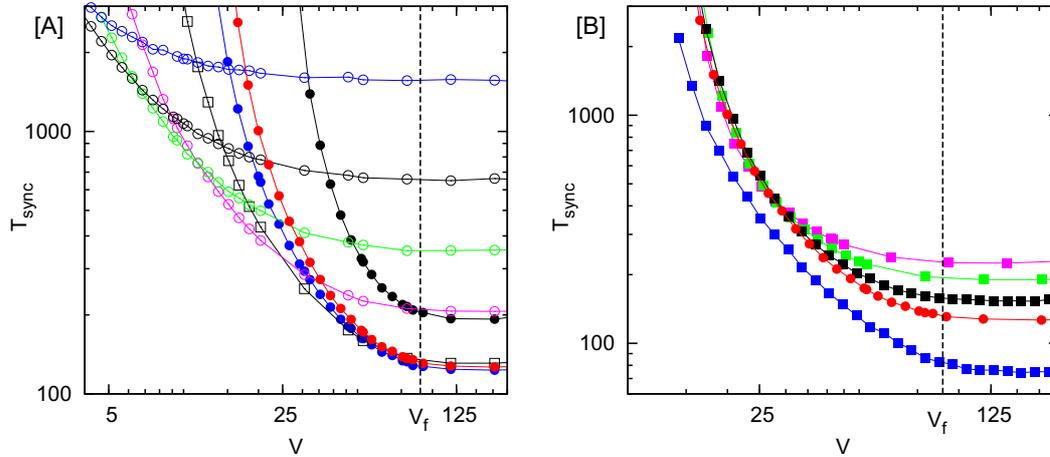


Figure 4.11: Synchronization time dependence on the velocity in the right region. Panel A: each curve corresponds to a different value of ϵ , while the size of the system is kept fixed at $N = 20$. Filled circles correspond to values of $\epsilon > 1/N$ (blue: $\epsilon = 0.08$, red: $\epsilon = 0.1$, black: $\epsilon = 0.2$), squares to $\epsilon = 1/N = 0.05$ and empty circles to $\epsilon < 1/N$ (blue: $\epsilon = 0.002$, black: $\epsilon = 0.005$, green: $\epsilon = 0.01$, pink: $\epsilon = 0.02$). Panel B: each curve corresponds to a different value of N (blue: $N = 10$, red: $N = 20$, black: $N = 30$, green: $N = 50$, pink: $N = 80$), while the coupling strength is $\epsilon = 0.1$.

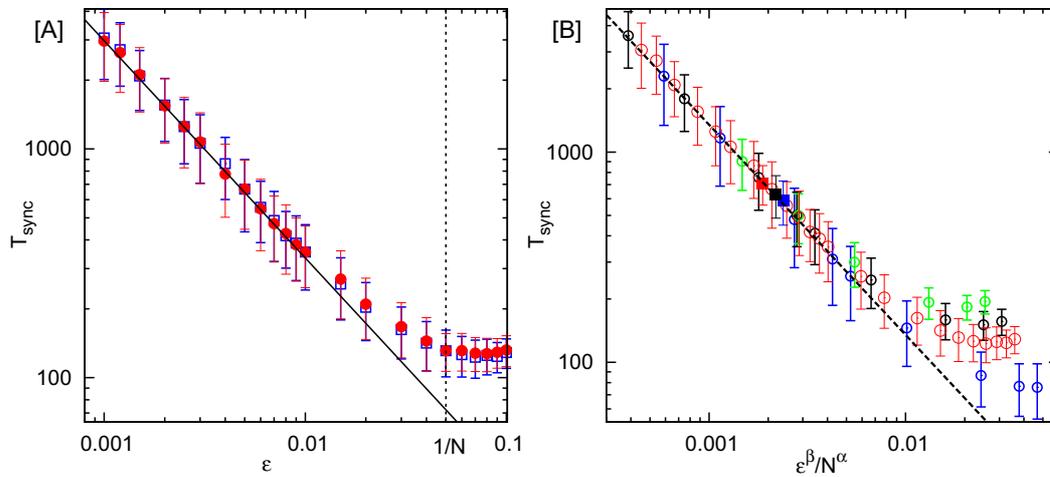


Figure 4.12: Panel A: synchronization time at $V = V_f$ (red filled circles) and $V = L/T$ (blue empty squares) as a function of ϵ , for $N = 20$. Black line stands for a fit function of the type $f(\epsilon) = C/\epsilon^\beta$, where C is a constant depending on N and $\beta = (0.95 \pm 0.01)$. Panel B: synchronization time at $V = L/T$ against $x = \epsilon^\beta/N^a$, with $a = (0.38 \pm 0.01)$, for several value of N . Circles are for $N \leq 50$ (blue: $N = 10$, red: $N = 20$, black: $N = 30$, green: $N = 50$), and single filled squares are respectively for $N = 80$ (blue), $N = 100$ (black) and $N = 150$ (green), for whom only the point at $\epsilon = 0.01$ has been calculated. Black dashed line stands for the fit function $f(x) = c/x$, where $c = (1.4 \pm 0.1)$.

Here we want to compare the energy cost of synchronization in two extremal conditions:

the fast limit of our minimal model and the all-to-all case⁷, which represents the opposite situation of a maximum value of r ($r = 1$). In this way we can verify if the efficiency is r -independent and, at the same time, we check out whether the fast limit of our minimal model is as efficient as a fully connected system.

Using the definition of the cost of the synchronization adopted in Sec. 4.1, we have that it is equal to $\Gamma_1^f = T_{sync}/(N - 1) \simeq T_f/N$ for our minimal model, while it is $\Gamma_{ALL} = \frac{N}{N-1} T_{sync}^{ALL} \simeq T_{sync}^{ALL}$ in the all-to-all case. Hence, in order to have the same efficiency, the synchronization time of the minimal model has to be N times larger than that of the fully connected case. This would imply that our oscillators, firing N times, achieve the same effect the oscillators in the all-to-all model achieve with a single shot. Obviously, the number of pair interactions is the same in both cases, and this makes the hypothesis plausible. However, between the present model and an all-to-all system of oscillators with an N times larger internal period there are some other important differences. In our model, firing events are delayed (not simultaneous) since a given oscillator fires at different targets at different times. Moreover, while interactions in the all-to-all case occur at constant time intervals, it is not the case for the fast limit of the minimal model since the time separation between two shots $i \rightarrow j$ is a random variable. Therefore, the energy cost can be the same only if these differences do not affect the synchronization time.

In order to perform this comparison, we have measured, by means of numerical simulations, the synchronization time in the all-to-all case (see Appendix B.2), obtaining

$$T_{sync}^{ALL} = \frac{2}{N\gamma\epsilon} \simeq \Gamma_{ALL}, \quad (4.5.2)$$

where $\gamma = (0.80 \pm 0.02)$. From Eq. (4.5.1), we have $\Gamma_1^f = c/(N^{1-\alpha}\epsilon^\beta)$, where $1 - \alpha = (0.63 \pm 0.01)$. Hence, the all-to-all model can be considered slightly more efficient. It follows that, even if the performances are very similar, an oscillator with only one neighbor need to fire something more than N times in order to achieve the same effect an oscillator in the all-to-all model achieves with a single shot.

We can deduce that the little efficiency loss we observe is due the previously described secondary differences. This means that heterogeneous interactions make the achievement of the coherent state slightly more lengthy. The take home message is that an isotropic interaction pattern without delays reduces the number of interactions the system needs in order to synchronize⁸. In conclusions, our analysis points out that, in order to improve the efficiency of the synchronization, it is strongly recommended to minimize all the possible sources of heterogeneity.

⁷Since in the all-to-all model the motion of the agents does not play any role, there is no problem comparing it with any other model with a different value of r , at any value of V .

⁸We have also checked that, among the patterns with fixed out-degree $k_{out} = 1$, the ring is always the best option (see Appendix B.3). Unfortunately, this configuration is not consistent with the minimal distance rule.

4.6 Slow regime

When the velocity is very small ($V \leq V_s$), we observe that T_{sync} decreases as a (negative) power of V (see Fig. 4.12). Then we can make the following ansatz:

$$T_{sync}(v; N, \epsilon) = C_s(N, \epsilon) V^{-\alpha(N, \epsilon)}. \quad (4.6.1)$$

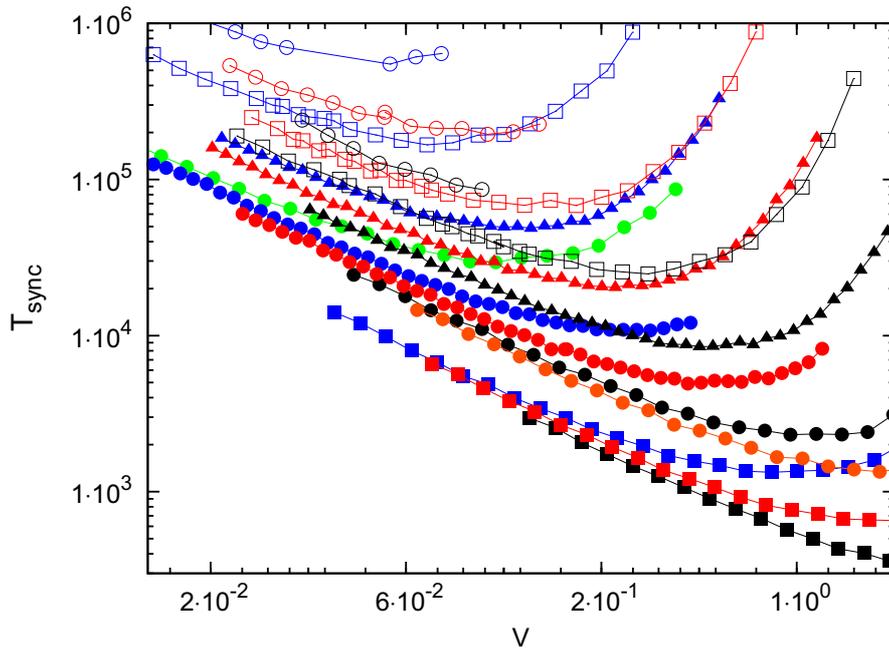


Figure 4.13: T_{sync} dependence on V in the slow regime for several values of N (filled squares: $N = 10$, filled circles: $N = 20$, triangles: $N = 30$, empty squares: $N = 40$, empty circles: $N = 50$) and ϵ (orange: $\epsilon = 0.3$, black: $\epsilon = 0.2$, red: $\epsilon = 0.1$, blue: $\epsilon = 0.05$, green: $\epsilon = 0.02$). All the figures of this section are obtained averaging over 2000 (for $N \leq 20$) or 1000 ($N > 20$) realizations.

Given the minimal interaction rule of our model, the system is not able to synchronize in the static case, that is $T_{sync} \rightarrow \infty$ when $V \rightarrow 0$. Indeed, as we have previously noticed, it is the motion of the units that allows the synchronization.

We suppose that when the velocity is small enough, the mechanics of synchronization can be understood in terms of a coalescence toward a globally coherent state through (i) the achievement of local synchrony and (ii) the modification of the interaction pattern. Consequently, we consider that the dynamics of the system is characterized by two time scales: the one of local synchronization (phases dynamics) and that of neighbor changing (motion). The slow regime can be defined as that region where the second time scale is larger than the first one and therefore local synchrony is usually achieved before the interaction topology undergoes a change [78, 83]. If such a hypothesis is correct, then

- i) T_{sync} must be proportional to this second characteristic time, a purely kinetic quantity that does not depend on the coupling constant ϵ ;
- ii) it has to scale with the ratio of the two time scales.

In order to verify our hypothesis we have to estimate both time scales.

The rate at which a given oscillator changes its neighbor is a function of N and V that can be calculated under some approximation (see Appendix B.4). Here we can use the estimator (B.4.12), that can be rewritten as a function of $v = V/L$:

$$T_{out}(v; N) = \frac{\pi g'}{4v\sqrt{N-1}}, \quad (4.6.2)$$

where $g' = 0.83$ is a correction factor empirically determined. We notice that, if T_{out} is the mean time an oscillator spends with his neighbor, $T_{out}^{(N)} = T_{out}/N$ is the average time separation between two sequential neighbor change events in the system, and hence can be regarded as a good estimator of the topological change time scale.

The other quantity we have to calculate is the mean synchronization time of an average connected clusters in the static case. We observe that the nearest neighbor rule is such that the average size of connected components in the system does not depend on density the of agents, taking the “universal value” $S_{cc} = 3.2$ (see Appendix B.1). Therefore, we can assume that the typical time needed for the achievement of local synchrony can be estimated as the average (over initial phases) synchronization time of three oscillators $T_3(\epsilon)$ arranged in the only configuration allowed by our topological rule (two symmetric shots plus a single one, being $p = 0$ the probability of a triangle, see the inset of Fig. 4.14 and Appendix B.3 for more details). The function $T_3(\epsilon)$, for $\epsilon \leq 0.3$, can be approximated as $T_3(\epsilon) = \kappa/\epsilon$, $\kappa = (2.0 \pm 0.1)$ (see Fig. 4.14). We notice that, at least in the range of parameters we are taking into account ($\epsilon \in [0.02, 0.3]$, $N \in [10, 100]$), the value V_s (see Fig. 4.7) is such that in general when $V < V_s$ it is also true that $T_3(\epsilon) \leq T_{out}^{(N)}$. This fact corroborates our previous statements.

We can now write a preliminary expression for T_{sync} in the slow regime:

$$T_{sync}(v; N, \epsilon) = T_{sync}\left(\frac{V}{L}; N, \epsilon\right) = N_{out}(N)T_{out}^{(N)}\left(\frac{V}{L}, N, \epsilon\right) = \frac{\pi g' N_{out}(N)}{4N^{3/2}v}, \quad (4.6.3)$$

where $N_{out}(N)$ is the average number of neighbor changes that are necessary in order to reach the synchronized state. Our hypothesis implies that $N_{out}(N)$ only depends on N , since it is the size of the system that determines how many times the interaction pattern has to be modified if order to enhance the coherence from local to global scale.

Comparing Eq. (4.6.1) and Eq. (4.6.3) we have

- a) $C_s(\epsilon, N) = N_{out}(N)N^{-3/2}$;

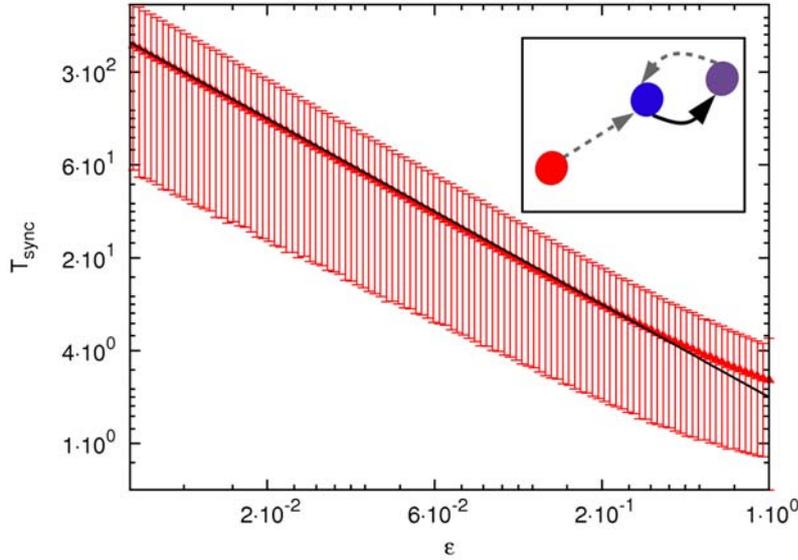


Figure 4.14: Average synchronization time for a group of $N = 3$ static oscillators dependence on the velocity. The black line stands for a fit function $f(\epsilon) = \kappa/\epsilon$. Averages have been performed over a regular sample of the initial phase values in the square $[0, 1] \times [0, 1]$, keeping fixed at 1 that of the oscillator of reference. We have considered different initial conditions. In the inset: a sketch of the connectivity pattern.

b) $\alpha(\epsilon, N) = 1$.

To write the scaling relation for T_{sync} in terms of the variable

$$r(v; N, \epsilon) = T_3/T_{out}^{(N)} = \rho v N^{3/2}/\epsilon, \quad (4.6.4)$$

where $\rho = 4\kappa g'/\pi$, we have to substitute v with r in Eq.(4.6.3). Thus we obtain

$$T_{sync}(v; N, \epsilon) = T_{sync}(r; N)N^{3/2}/\epsilon = \frac{r}{v}T_{sync}(r; N). \quad (4.6.5)$$

4.6.1 Checking the hypothesis.

In order to verify our hypothesis, we have to check expression *a*) and *b*), and the scaling relation (4.6.5). We expect to observe, plotting T_{sync} dependence on r , that all the curves with different ϵ that correspond to the same value of N collapse, thus obtaining a single curve for each value of N . Moreover, we expect the shape of such curves to be the same, while they should differ for a multiplicative factor $N_{out}(N)$, that is an increasing function of N to be empirically determined. Notice that in this new framework, determining $v_m(N, \epsilon)$ and $v_s(N, \epsilon)$ means to check out the corresponding values of r , r_m and r_s , that are the same for all the considered N and ϵ . Finally, we should verify that $T_{sync} \sim r^{-1}$ at $r < r_s$ (or, that is the same, $T_{sync} \sim V^{-1}$ at $V < V_s$).

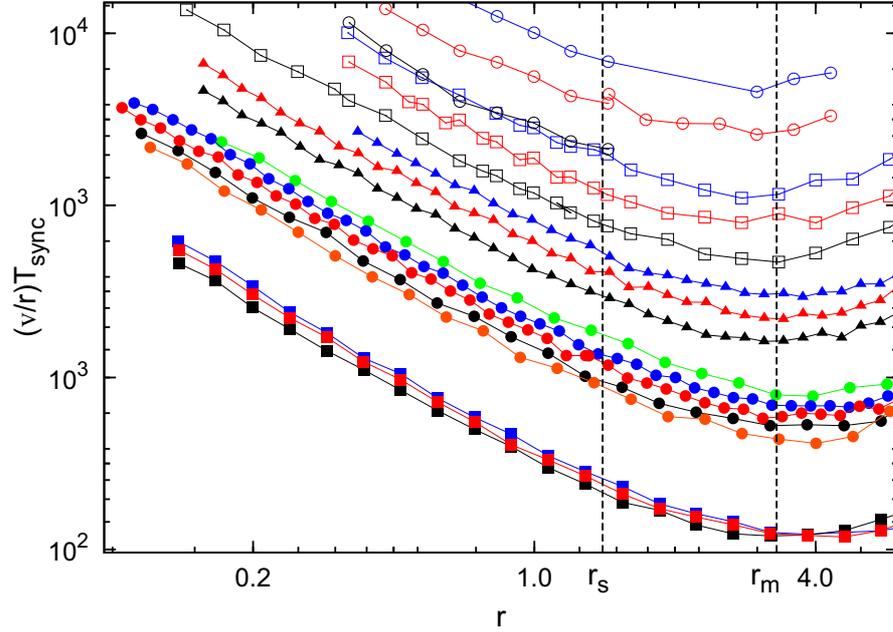


Figure 4.15: Rescaled synchronization time $T_{sync}(r) = (v/r)T_{sync}(v)$ against r in the slow regime for several values of N and ϵ . Colors and symbols are the same as in Fig. 4.13.

First of all, we plot $T_{sync}(r) = \frac{v}{r}T_{sync}(v)$ against r . As we expect, all the minima are arranged along a vertical line and the curves look like vertical translations of the same function (Fig. 4.15), at least at $r \leq r_m$. Hence, the first part of our hypothesis has been verified. We find out $r_m = (3.3 \pm 0.3)$ and $r_s = (1.4 \pm 0.2)$. Anyway, points corresponding to the same value of N and different values of ϵ do not collapse onto the same curve. For each one of the considered values of N , we can observe more or less separated group of close but not superposed curves. This dispersion can be regarded as a minor deviance from the predicted behavior, since actually the curves belonging to the same group are almost collapsed. Consider that, while the spread of values of ϵ is around one order of magnitude, the largest value of N is just five times the smallest one. However, the separation among curves of the same group seems to increase with N (see Fig. 4.15).

In order to quantify the discrepancy between our theoretical prediction and the real behavior of the system, and to check the relation $T_{sync} \propto V^{-1}$, we try to fit raw data (T_{sync} vs V) at $V < V_s$ with a function of the type $f(V) = cV^{-a}$.

The obtained values of a are quite close to 1, but they show a weak decreasing when N increases. In particular, for $N = 10$ the parameter takes value $c = (0.95 \pm 0.02)$, while for $N = 40$ it takes value $c = (0.85 \pm 0.02)$. At the same time, apparently, it starts decreasing when $\epsilon < 0.05$, a too small value that does not allow to formulate any hypothesis about the form of this dependence.

The values of c , instead, are clearly grouped into separated set, each one corresponding to a different value of N , but inside each sets the variability associated to ϵ is significant. Actually, c increases with N (strong dependence) and decreases with ϵ (weak dependence). The empirical form we find for this weak dependence is that of a negative power with exponent $\beta(N)$, ranging from $\beta(10) = (0.20 \pm 0.01)$ to $\beta(40) = (0.50 \pm 0.02)$ (see Fig. 4.16).

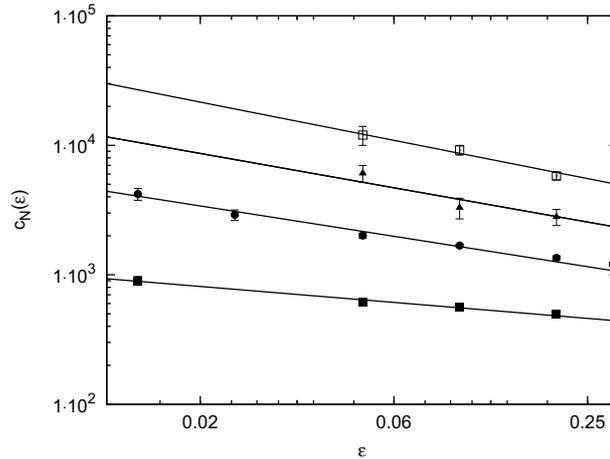


Figure 4.16: Coefficient c dependence on ϵ for several values of N (filled squares: $N = 10$, filled circles: $N = 20$, triangles: $N = 30$, empty squares: $N = 40$). Black lines stand for fit function of the type $f_N(\epsilon) = c_0 \epsilon^{\beta(N)}$, with $\beta(N) = \{0.20; 0.33; 0.42; 0.50\}$ respectively for $N = \{10; 20; 30; 40\}$.

Consequently, we can conclude that our heuristic argument is mostly correct, although there are some limitations. On the one hand, even in this slow regime, the coupling constant has considerable influence on the synchronization time, whose relevance increases at larger N . On the other hand, the dependence on the velocity is not exactly of the expected kind, even if quite close to it. However, r results to be the appropriate scale variable.

4.6.2 Empirical proposal of a scaling relation

Through an empirical analysis, we have been able to derive the following relation:

$$T_{sync}(r; N, \epsilon) \epsilon^{\Lambda(N)} N \log N = K r^{-B}, \quad (4.6.6)$$

where $\Lambda(N) = N/\Lambda_0$, with $\Lambda_0 = (100 \pm 5)$, is the scaling factor that allows the perfect collapse of curves corresponding to the same N (see Fig. 4.17). Other parameters $K = (500 \pm 5)$ and $B = (0.90 \pm 0.05)$. This equation holds whenever $N \geq 20$ (4.18), but it is also possible that for larger N new terms are going to appear in Λ as well as in K . Unfortunately, for $N > 50$, simulations of the slow regime become extremely lengthy⁹. In conclusion, in the

⁹Notice that the computational time scale roughly as the total number of avalanches ($\sim N T_{sync}$) times the number of pair distances we need to calculate ($\sim N^2$), so that it is more or less proportional to $N^3 T_{sync}$

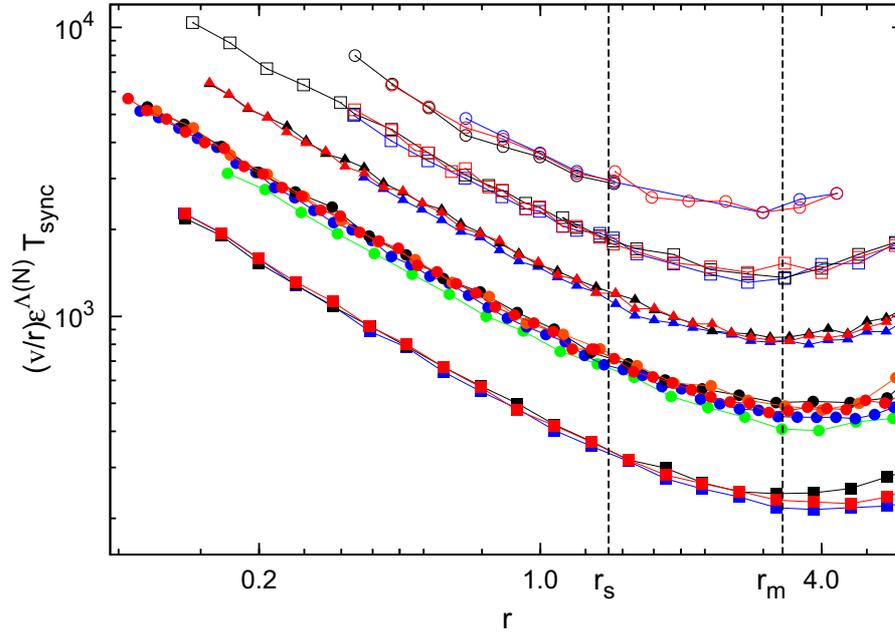


Figure 4.17: Rescaled synchronization time $T_{sync}(r) = (v/r)T_{sync}(v)$ times the empirical scaling function $\epsilon^{\Lambda(N)}$ against r for several values of N and ϵ . Colors and symbols are the same as in Fig. 4.13.

slow regime, to separate the dependence of the synchronization time on the coupling strength on the one side, and on the size of the system on the other side, is almost impossible. The interplay of both quantities is much more complicated than in the fast limit case. Despite this intrinsic complexity, there exists a universal value r_m at which T_{sync} has a local minimum and a universal value r_s below which $T_{sync} \sim r^{-B}$.

4.6.3 From local coherence to global synchrony

In order to better clarify the underlying mechanisms, let us consider the case of a very small system of just four oscillators composed by two pairs of units already synchronized. The first one includes the reference oscillator and so its phase is $\phi_1 = 0$. The second one has a different phase ϕ_2 . Interchanging two units between the pairs, we obtain two identical objects, that is two pairs each one composed by an oscillator with phase ϕ_1 and the other one with phase ϕ_2 . At the end, when local synchronization will be reached again, all the units will share the same phase. Global synchrony will have been achieved.

Now consider the case $N = 6$. If there are pairs only, by means of interchanging units among them, we always obtain two pairs that share the same phase plus a third one with a different phase value. Obviously, by means of subsequent recombination, their values can approach each other, but the process may be arbitrary long. On the contrary, if at a given

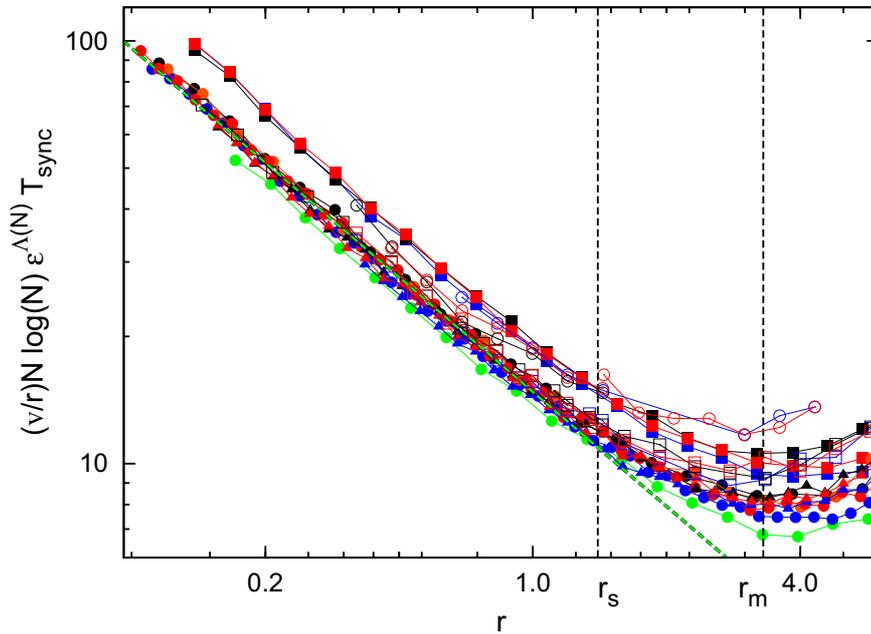


Figure 4.18: Rescaled synchronization time $T_{sync}(r) = (v/r)T_{sync}(v)$ times the empirical scaling function $N \log(N) \epsilon^{\Lambda(N)}$ against r for several values of N and ϵ . Green dashed line stands for a fit function of the type $f(r) = Kr^{-B}$, with $K = (500 \pm 5)$ and $B = (0.90 \pm 0.05)$. Colors and symbols are the same as in Fig. 4.13.

moment the pair with a different phase value breaks and its two oscillators start firing each one at an oscillator belonging to a different pair, global synchrony is immediately achieved. Indeed, in this particular case, these two oscillators proceeding from the broken pair, are able to impose their common phase to the whole population. They have no in-neighbor, so they keep their phase unchanged, while they are able to set the pace of the other oscillators. Consequently, when local synchrony is achieved, all the units will share their same phase.

In more complicated scenarios, where there are several groups of two, three, four and sometimes five or more oscillators, the complexity of the process increases awfully. There are, however, some general features that can be worth to outline:

- 1) in any connected component of size larger than 2, there is always at least one oscillator that has null in-degree and hence its phase is not changed by the shots of the others. Then, the only way towards the local synchrony is through the progressive adjustment of the phases of the other units to that of this one;
- 2) whenever inside the same connected component there are two or more units with null in-degree that do not share the same phase, complete local synchrony is not possible;
- 3) when an oscillator with null in-degree leaves the group it belongs to a new one, if (and only if) in the new group it still has null in-degree and there is no other unit without

in-neighbors, then it will be able to synchronize its original group with the new one;

- 4) continuous units interchanges among pairs of oscillator progressively narrows the phase distribution.

More in general, we can say that

- 5) every change in the interaction pattern causes a temporary damage in terms of local coherence;
- 6) the large majority of the modifications an interaction pattern can undergo contribute little or nothing to the enhancing of global synchrony.

In conclusion, we can state that this coalescence from local to global synchrony needs so many steps (topological changes) in order to be achieved because many of them are useless or even damaging. We can describe it as a random trial and error process, without any learning.

At the end, the system is able to synchronize thanks to a reduced number of really useful topological changes, like those at point 3), plus a huge number of small contributions, like those at point 4).

4.7 Wasting time. The no-synchronization region.

In the previous sections, we have seen that the fast regime is the most homogeneous and therefore the most efficient. It is similar to the static all-to-all case with an N times larger period, even if not exactly the same because of a certain degree of residual heterogeneity. The phase dispersion is almost monotonously reduced during the in time evolution of the system and synchrony is achieved directly at a global scale, similarly to what occurs in other system of interacting moving agents [78, 83].

In the slow regime instead the system reaches the synchronized state through a very tortuous pattern, achieving and breaking the local synchrony a huge number of times, and finally converging to a globally coherent state. This mechanism reminds a walk through a rough landscape.

In the intermediate region, to find a way towards the achievement of a globally coherent state appears to be even more difficult.

If starting from $V = V_f$ (fast limit) we reduce the velocity a bit, the oscillators will start to fire at the same target consecutively ($T_{out} > 1$). These repetitions can be regarded as pure waste of time, since their are too few in order to enhance the local coherence. They are just increasing the effective period in the all-to-all analogy from $T = N\tau$ to $T = N T_{out}^{(i)}\tau$. Moreover, since $T_{out}^{(i)}$ is a random variable that varies in time and that is different for every oscillator (even if they all have the same mean value), the amount of heterogeneity introduced by these repeated interactions is enormous (out of control). And so, the synchronization time diverges. However,

to understand the details of the mechanism that makes the synchronization time increase so abruptly is still an open problem.

When the velocity is further reduced, then repeated interactions among small groups of units start to allow the system to reach a certain degree of local synchrony. At a first moment, it will be the case for a fraction of *lucky* pairs only, those having the shortest synchronization time. Then, also some groups of three units, plus some other pairs, will be allowed to synchronize, and so forth. Consequently, the system can again reach the global coherent state. At even slower velocity, the improvement of the local synchrony allows the system to synchronize faster and faster, until it reaches a local minimum (where possibly almost all the pairs are able to synchronize) as explained in Sec. 4.6.

4.8 Synchronizing faster, synchronizing cheaper

In this section we want to discuss, on the basis of what we have learned up to now, how to improve the efficiency of the system. Let us start noting that synchronization mechanisms in the slow regime and in the fast limit are deeply different. Consequently we can expect that a modification able to reduce the synchronization time in one case, could damage in the other. For instance, in the fast case, it would help if the oscillators could look for the first and the second neighbor, eventually firing at that with whom they have been interacting less. This is a quite simple compensation mechanism able to homogenize the interaction pattern. However, the result is doubtful in the case of the slow regime since this would probably prevent the achievement of local coherence, damaging the ability of the system to synchronize. Similarly, it could be useful, in the slow regime, to make the agents stay at rest for a while when they meet a new neighbor. They could wait until they synchronize with their neighbor or move anyway after a maximum waiting time¹⁰. In this way, we hope to optimize the ability of the system to combine the achievement of local coherence and topological change. Obviously, in the fast limit, such a rule would be absolutely damaging.

It is also worth to stress that all this additional rules, even if they are neutral in terms of the efficiency defined in Sec. 4.1 since the number of shots per agent is kept unchanged, actually imply other kind of cost. In order to be able to recognize with whom they have been interacting more or less, or to distinguish whether the current neighbor is the same they had at the previous time step or not, our agents need to store information. Usually, in real world, this can not be done for free. Therefore, it could be appropriate, in order to compare the efficiency of different models, to define a more general cost function.

A complementary step towards a more realistic evaluation of the cost of the synchronization can be taken by including the cost of the agents motion, that is, their kinetic energy. Whenever

¹⁰Since there exist local configurations that are *frustrated* and cannot reach a complete coherent state, it is important to prevent the oscillators from falling into frozen configuration by letting the agents move after a certain time, even if they have not synchronized with their neighbor.

the agents have a mass, this is a non-null quantity that, in principle, should be added to the interaction term in the energy cost expression. Notice that, by adopting this new framework, previous conclusions about which is the most efficient regime could change. In Fig. 4.19 we show the total energy cost as a function of the velocity for $N = 20$ and $\epsilon = 0.1$ (the same values used in Fig. 4.7), using the expression

$$\Gamma = T_{sync} + \lambda V^2,$$

with $\lambda = 1$ and $\lambda = 10$.

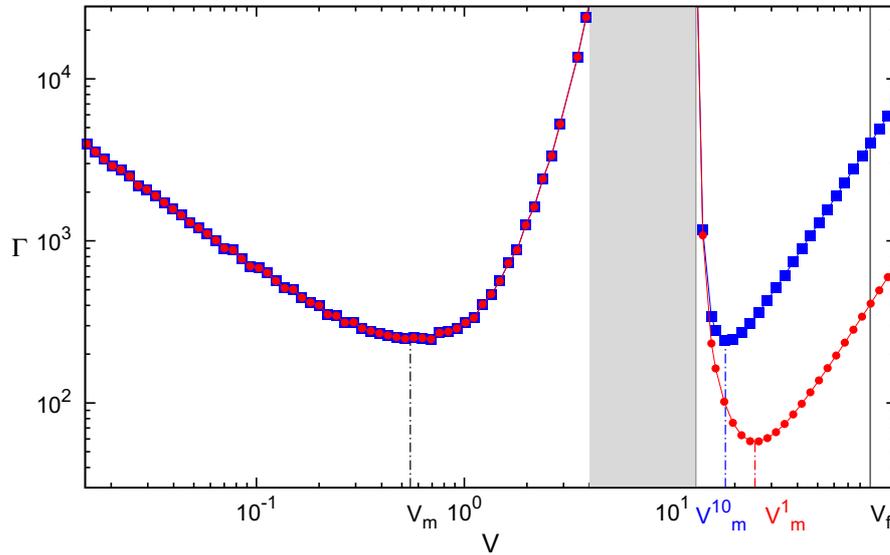


Figure 4.19: Efficiency dependence on the velocity for $N = 20$, $\epsilon = 0.1$, and $\lambda = 1$ (red circles) and $\lambda = 10$ (blue squares).

In this case, there are two minima, one at $V = V_m$, the other in the right region, at a value of $V \leq V_f$ that depends on λ . Indeed, while in the left region the velocity is so low that the kinetic contribution to the cost is unimportant, in the right one it is much more significant. Hence, there exists a precise value of V , depending on λ , such that a further increasing of the velocity implies a waste of kinetic energy that the decreasing of the synchronization time, if any¹¹, cannot compensate. Therefore, if when defining the efficiency we take into account the energy required to make the agents move at the desired velocity, it is no longer assured that the fast limit represents the best option.

¹¹Notice that for $V > V_f$ the synchronization time does not decrease any more.

4.9 Conclusions

Following the recent literature on complex systems, one of the hottest topics is the relation between dynamics and topology of interactions. In particular, there are many evidences that the patterns of interaction change rapidly with time, thereby completely altering and conditioning the dynamical properties of the system.

Here we have proposed a framework in which agents move on a plane and are allowed to interact in a pulsing way. Each agent, representing a phase oscillator, moves at a common velocity and changes its internal phase with a common period. When this internal phase reaches a threshold value, the oscillator “fires”, thus resetting its own phase. We considered two different models. In the first one, each oscillator, when firing, resets the orientation of its own motion and interacts with the neighbors within a certain distance by changing their phases. In the second one, the oscillator who fires modifies the phase and at the same time resets the orientation of its nearest neighbor, while itself continuing to move in the same direction.

For the first model, keeping all geometrical parameters constant, we analyze the system behavior by changing solely the velocity of the agents and the interaction range. We measure the time needed for the system to synchronize as a function of the velocity and the fraction of population they interact with. We have introduced a new order parameter that stands for the fraction of different units each oscillator has interacted with. This order parameter allows to introduce a novel phase diagram that enables us to identify three different mechanisms leading to synchronization: i) the diffusive mechanism, where the system very quickly reaches synchronization by means of extremely effective interactions of each oscillator with a large fraction of the population; ii) the local mechanism, where agents move very slowly, but the interaction range is large enough so that local synchronization is sufficient to lead the system to the globally synchronized state; and finally iii) the bounded mechanism, characterized by very slow motion and short range of interaction, which allow a high degree of accumulated mixing during a very long synchronizing time. As expected, the minimal energy cost, which stands for the total number of signals emitted by the population to reach the synchronized state, is achieved when agents move very fast, irrespective of the interaction range; however, synchronization in systems with low mobilities dramatically depend on the interaction range.

For the second model, we study the dependence of the synchronization time on the velocity, also varying the size of the system and the coupling strength. Since the interaction rule is so restrictive that the system lies far below the static percolation threshold, it is the non-null velocity of the oscillators that enables the system to reach the coherent state. Hence, we could expect the synchronization time to be a decreasing function of V , but we find that there is an intermediate region where the synchronization time diverges. We have been able to characterize the mechanisms that allow the system to synchronize both in the slow and in the fast regime, also calculating the value of the velocity at which the system enters these two regions as a function of the system’s size N and the coupling strength ϵ . The mechanism of the fast limit of

this model does not differ from the diffusive mechanism (i) of the previous one, while the slow regime resembles the bounded regime (iii) but with a very small velocity since, at the end of the synchronization process, each oscillator has interacted only with a finite fraction of the other oscillators. It has been also possible to provide an empirical estimator of the synchronization time as a function of N and ϵ in the fast limit, and as a function of V , N and ϵ in the slow regime. Then, we qualitatively discuss the reasons why the system cannot synchronize within a finite time in the intermediate region. Finally, we consider that, measuring the cost as the number of shots needed for the system to synchronize, the fast limit is always the most efficient regime. Whereas, if we take into account the energy required to make the agents move at the desired velocity, i.e., a kinetic term proportional to the square of the velocity, this is no longer the case. The new energy cost displays two minima, one corresponding to the local minimum the synchronization time has in the left region, the other one located in the right region, below the fast limit.

The identification of the mechanisms that relate mobility and interaction to synchronization will undoubtedly be crucial in similar models of populations of moving agents which can be applied, for instance, to the field of wire-less communications. Indeed, understanding these phenomena will help to design optimal protocols to dissect more realistic settings by defining appropriate interaction rules.

Chapter 5

Conclusions and perspectives

5.1 Conclusions

In this thesis we have neither developed a new theory about some aspect of complex networks, nor have we studied an applied problem of interest. What we have done, hopefully, has been to enhance the “theoretical toolbox” that researchers can use to build new knowledge about applied as well as general issues. The models we investigate are as simple in their definition as rich in the behaviors they display. They are suitable to be modified to model real situations, but from each one them we can also extract useful insights about general features of the class of systems they belong to.

- In Chapter 2, we address a problem common to many experimental situations, where the *a priori* unknown connectivity of a particular network is inferred from purely dynamical measurements. We consider a system of Kuramoto oscillators where we perturb a setting of identical units by allowing one of them, the pacemaker, to have a different natural frequency. If the natural frequency of the pacemaker is not very different from the value of the rest of the population, the system enters in a frequency-locked state and all the units evolve with the same effective frequency. On the contrary, if the frequency difference is too large, this is no longer possible and the system undergoes transition between a coherent and a incoherent state. We show how to capitalize the dependence of the synchronization properties on the topology of the underlying complex network to gather information about the latter.

We find that, measuring the critical frequency, it is possible to estimate the degree of each pacemaker. Moreover, slightly above the transition point, we can obtain the hierarchical structure of the whole network by measuring correlations between effective frequencies. Finally, far above the critical point, the oscillators that are directly connected the pacemaker can be easily recognized, so that we can recover the connectivity pattern.

This problem has been addressed by different works proposing different approaches. Our

method is based on the change of the frequency of a single unit and its efficiency may change depending on the practical purposes it is applied to.

- In Chapter 3, we propose and study a model of information retrieval in a complex network based on a biased random walk, whose bias to walk forward or backward is tuned by a parameter. The key point of the algorithm is that the information retrieved by the walker is only available when it first arrives to its home departing node. Our main finding is that, unless the walker is allowed to explore the network during a very long time, there is a finite value of the parameter for which the information retrieved is maximal. Moreover, only at this precise value the performance of the explorer is independent on the degree of the home-node. For other choice of the parameter, to recover a satisfactory amount of information when starting from a poorly connected nodes is very unlikely, even if the walker can perform a very long searching.

We have also proposed an adaptive scheme in order to overcome the problem of fixing an *a priori* unknown optimal value. This adaptive search protocol performs optimally, both in terms of the quality and quantity of the information retrieved, with a minimum (local) information about the network structure whenever. Moreover it works equally good with all kinds of topology whenever there is minimal amount of heterogeneity in the degree distribution and hence could be suitable to be applied to the exploration of real artificial networks.

- Finally, in Chapter 4 we further explore the relation between dynamics and topology of interactions focusing on a case in which the patterns of interaction change rapidly with time. We propose a framework in which a population of integrate-and-fire oscillators moves on a plane, considering two possible interaction rules. According to the first one, each oscillator, when firing, resets the orientation of its own motion and interacts with the neighbors within a certain distance. According to the second one, each oscillator interacts with its nearest neighbor only, changing its phase and also resetting its orientation.

For both models, we analyze the system behavior by changing the velocity of the agents and other parameters that characterize the system, such as the interaction range, the coupling strength or the size of the population. We measure the time needed for the system to synchronize as a function of the velocity and study its parametric dependence on the rest of parameters.

We have been able to characterize the mechanisms that allow the system to synchronize under different dynamical conditions, and in some cases we also calculate the value of the velocity at which the system enters each one of these regions of the parameter space.

For the second model, it has been possible to provide an empirical estimator of the synchronization time when the motion of the agents is very slow or very fast. At intermediate

values of the velocity we find that the synchronization time diverge and we qualitative discuss what, at the microscopical level, prevents the system to synchronize. Finally, we consider that, measuring the cost as the number of shots needed for the system to synchronize, the fast motion is always the optimal choice. On the contrary, if we take into account the kinetic energy required to make the agents move, to identify the most efficient regime is not so trivial since the total cost displays two minima whose relative height depends on the system's parameters.

Understanding the mechanisms that relate mobility and interaction to synchronization, identifying those general features that do not depend on the details of the interaction rule, can help, for instance, to design optimized communication protocols.

5.2 Perspectives

During the development of the work presented in this thesis, a number of possible applications, additional questions, and related issues have been raised and could be further investigated in the future.

- It can be worth to check the robustness of the method for the detection of functional communities in networks of Kuramoto oscillators presented in Sec. 2.4. If we could prove that, as we expect, the results are not sensitive to the details of the functional form of the interactions, then the method could be applied to a lot of situations of interest, regardless of the specific nature of the considered network.
- The walkers introduced in Chapter 3 change their behavior in a nontrivial way when changing the parameter that tune the probability they go forward or backward. In particular, there exists a precise value of the parameter for which the agents, in general, visit many nodes and then come back to their *home*. Beyond the issue of network exploring, we are interested in understanding more general consequences of this peculiar behavior. For instance, we could analyze the new phenomenology that may arise substituting usual random walks with our walkers in a metapopulation model for epidemic spreading on complex networks. It could be also interesting to consider each walker as a phase oscillator allowed to interact with other oscillators in the same node only, and each node population as an initially synchronized community. We may ask whether the value of the parameter that is optimal for the network exploration is also able to enhance the achievement of global coherence throughout the whole network.
- Finally, further investigation is required to better understand some features of the minimal model introduced in Sec. 4.2. It can be worth to identify the necessary and sufficient conditions for the dependence of the synchronization time on the velocity to become non-monotonous. Currently, we are investigating in this direction and it is already proved

that a pulse interaction is required. Indeed, when replacing the integrate-and-fire oscillators with Kuramoto oscillators, the synchronization time behaves as a monotonously decreasing function of the velocity, for any values of the coupling strength.

Appendix A

Pacemakers in a Cayley tree of Kuramoto oscillators

In this appendix we analyze in details, as a simple, analytically solvable application of what we studied in the first chapter, a system of Kuramoto oscillators with identical frequencies in a Cayley tree. Heterogeneity in the frequency distribution is introduced in the root of the tree, allowing for analytical calculations of the phases evolution. This simple case can be regarded as a starting point in order to understand how to extract topological features of the connectivity pattern from the dynamic state of the system, and viceversa, for the general situation of a set of phase oscillators located on a tree-like network.

A.1 The model

We consider as the network structure of our system a Cayley tree of variable radii and coordination numbers (Fig. A.1).

The nodes follow the dynamics described by the Kuramoto model of phase oscillators [45, 39]. All of them have the same natural frequency (taken to be zero without loss of generality) except for the one located at the root of the tree, that has a different natural frequency ω . We will refer to this node as the pacemaker. The equations are

$$\dot{\phi}_p = \omega + \sum_j a_{pj} \sin(\phi_j - \phi_p); \quad (\text{A.1.1})$$

$$\dot{\phi}_i = \sum_j a_{ij} \sin(\phi_j - \phi_i) \quad \forall i \neq p, \quad (\text{A.1.2})$$

where the first one is for the pacemaker, the second for all the rest. The matrix a_{ij} is the corresponding adjacency matrix for the network.

Notice that for our natural frequencies distribution it is never possible to reach synchronization in the meaning of an equal phases state, that is a state where all the interaction terms

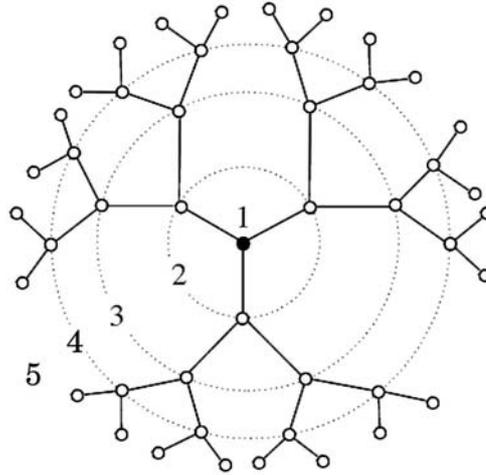


Figure A.1: Cayley tree with coordination number $q = 3$ and radius $R = 5$. The pacemaker is located at the root of the tree (filled circle).

vanish. It is also important to stress it does not depend on the choice we made of a null natural frequency for all the $N - 1$ oscillator except the pacemaker. Indeed, we can change the mean value of the natural frequencies distribution just rotating the frame, without any modification of dynamical properties of the system. For instance we could choose a zero mean distribution, that can be obtained applying the following transformation to the phases: $\varphi_i \rightarrow \varphi_i + \Omega t$, in which $\Omega = \omega/N$ is the first moment of the distribution.

A.2 Characterizing the global behavior

The transition from an incoherent to a synchronized state is usually described by means of an order parameter introduced by Kuramoto:

$$z(t) = \frac{1}{N} \sum_j \exp(i\phi_j). \quad (\text{A.2.1})$$

The amplitude of this quantity is proportional to $1/N$ in the incoherent regime whereas it grows with decreasing ω in the phase-locked state and asymptotically it tends to 1 as ω tends to 0 (phase-locking state). Anyway, since our system can undergo through a phases synchronization only at local level, this parameter is not an appropriate one in order to describe its global dynamical behavior.

Then, following [87], we analyze an order parameter already introduced and used in [41]

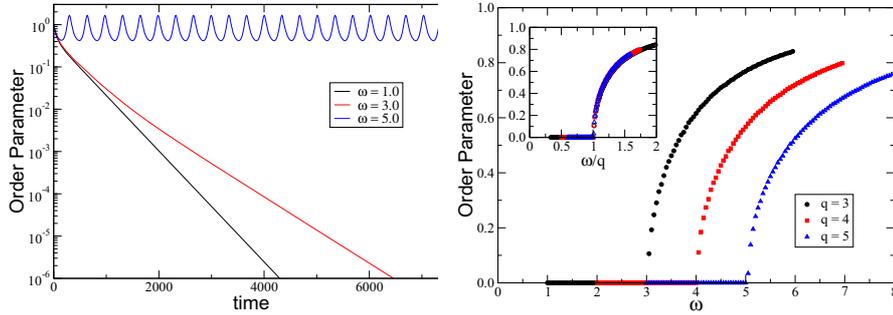


Figure A.2: (Left) Effective Frequency Dispersion Order Parameter Δ_ω (A.2.2) vs. time. (Right) Normalized Frequency Dispersion Order Parameter r_ω (A.2.3) vs. natural frequency of the pacemaker ω .

and [46] that measures the effective frequency dispersion:

$$\Delta_\omega = \sqrt{\frac{1}{N} \sum_{i \in N} [\dot{\phi}_i - \langle \omega \rangle]^2} \quad (\text{A.2.2})$$

In Fig. A.2 (left) we present the time evolution of this order parameter for a Cayley tree with a coordination number $q = 3$, and radius $R = 5$, with a total of $N = 94$ nodes. The results correspond to initial conditions with all the phases equal to zero. For frequencies below a critical value ω_c the order parameter decays exponentially to zero, revealing that for long times there is no dispersion in the effective frequencies and the system becomes synchronized. Above ω_c the system does not reach equilibrium, and the order parameter presents an oscillating behavior.

Note also that the frequency dispersion presented in Eq.(A.2.2) is not normalized. We divide the frequency dispersion by its maximum allowed value $\frac{\omega}{N} \sqrt{N-1}$, and thus obtain a normalized *order parameter*:

$$r_\omega = \sqrt{\frac{1}{N-1} \sum_{i \in N} \left[\frac{\dot{\phi}_i}{\langle \omega \rangle} - 1 \right]^2} \quad (\text{A.2.3})$$

In Fig. A.2 (right) we present the stationary value of the normalized order parameter as a function of the pacemaker frequency ω . As noted, for values above ω_c the system is not in equilibrium, however it is possible to define a mean value around which the order parameter oscillates. The figure shows the behavior observed for three different values of the coordination number, $q = 3, 4$ and 5 . As q grows, the critical value also grows accordingly. The inset in the figure shows the data collapse obtained by scaling the natural frequency of the pacemaker with the coordination number, ω/q .

A.3 Results

In order to analyze the relation between the topology of the connectivity pattern and the ability of the system to reach a synchronized state, we can notice that when we increase the regulator frequency ω then there will be some oscillators cannot keep the phase difference, breaking the synchronized state. The left-hand side of eq.(A.1.2) is indeed bounded because of the sine terms, whereas the right term increase as the regulator frequency is increased. The same happens in eq.(A.1.1). Consequently, there will be a transition from a synchronized to an incoherent state, at a precise value of ω . Thus we can define the natural frequency of the pacemaker critical value as follows:

$$\omega_p^c/\omega > \omega_p^c \# \text{ frequency locking state}$$

i.e. ω_c is that value of the frequency of the regulator above which no synchronized state exists. In the case of a Cayley tree topology, it is possible to analytically calculate this quantity. We start noticing that, since the existence of a steady state is a necessary condition for a frequency-locked state, we can compute the constant value all the effective frequencies are going to take in a stationary condition. First of all we notice, following [87], that summing up all the Eq.s (A.1.1)-(A.1.2) we have:

$$\sum_{i=1}^N \Omega_i = \omega \quad (\text{A.3.1})$$

where Ω_i are the stationary effective frequencies. Consequently, if they all take the same value, this will be ω/N . In this case the system is in a synchronized state where all the frequencies are equal and constant. Now we can calculate, as a function of the radius R and of the connectivity number q , which is the maximum value ω such that all the oscillators can reach an effective frequency equal to ω/N . Since the left side of phase evolution equation for the pacemaker (A.1.1) is bounded, it is clear that a bound exists. Indeed, it is just when all the sine terms in that expression take values -1 that we have the maximum ω able to satisfy this equation. So we found the following condition for the critical frequency:

$$\omega_c \leq q \frac{N}{N-1} = q + \left[\sum_{r=0}^{R-1} (q-1) \right]^{-1}. \quad (\text{A.3.2})$$

where in the last equality we used the expression $N = 1 + q \sum_{r=0}^{R-1} (q-1)$.

Locating the pacemaker at the root is a special situation of the more general case discussed in [87]. There it was shown how this particular configuration, where all the neighbors of the pacemakers keep a phase difference from it equal to $-\pi/2$, is not always allowed. Indeed, the interplay of other equations, different from that of the pacemaker, may cause a breaking of the stationary condition earlier.

In order to take into account all these contributions, according with [87], for an arbitrary topology of the interactions, we can generalize inequality (A.3.2) in the following form:

$$\omega_c \leq N \left[\frac{K_{out}}{N_{out}} \right]_{min} \quad (\text{A.3.3})$$

where $\left[\frac{K_{out}}{N_{out}} \right]_{min}$ is the minimum ratio we can find for a connected component of the connectivity pattern including the pacemaker, being N_{out} the number of nodes outside the considered group and K_{out} the number of "external links" connecting the component with the rest of the network. This expression reduces to (A.3.2) if we take into account just the pacemaker itself, since in this case we have $N_{out} = N - 1$ and K_{out} coincides with the degree of the pacemaker, q in our case. Doing so, we are not minimizing the ratio in Eq. (A.3.3), we are instead maximizing its denominator but often it is a very good approximation. Moreover, if the connectivity pattern has a tree-like topology, that is there is no cycle on it, Eq. (A.3.4) is not just a bound, but the exact expression of the critical frequency.

The precise case of a Cayley tree topology of the connectivity pattern has been analyzed in ([87]), where it is provided an explicit expression of the inequality (A.3.3) for a tree of radius R and coordination number q , where the pacemaker is located at distance r from the center:

$$\omega_c^{(r)} = \frac{N}{N - \sum_{i=0}^{R-r} (q-1)^i}, \quad (\text{A.3.4})$$

It is easy to show that the new bound simply coincides with (A.3.2) when the distance between the pacemaker and the center of the Cayley tree is $r = 0$. Consequently, we can rewrite Eq. (A.3.2) in the form of an equality:

$$\omega_c = q + \left[\sum_{r=0}^{R-1} (q-1)^r \right]^{-1}, \quad (\text{A.3.5})$$

it is the critical value for the natural frequency of a pacemaker located at the root of a Cayley tree.

In order to better describe the dynamical behavior of the system above the critical point it is useful to compute the order parameter defined by Kuramoto. Indeed, even if it is not useful in order to identify the transition point, this quantity provides interesting qualitative information about the global time evolution of the phases when the coherent state is broken.

First, we can ask ourselves which is the behavior of this quantity exactly at the critical point. In order to do this, we can re-write the order parameter defined in (A.2.1) as

$$z(t) = \frac{1}{N} [\exp(i\phi_p) + \sum_{j=1}^q \exp(i\phi_j) + \sum_{l=1}^{N-q-1} \exp(i\phi_l)], \quad (\text{A.3.6})$$

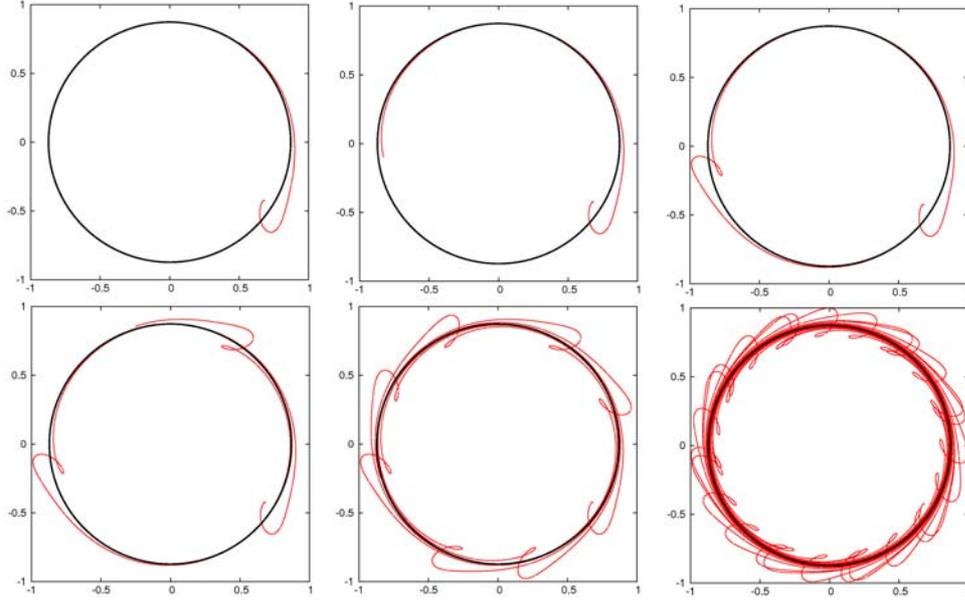


Figure A.3: Values taken in the complex plane by parameter (A.3.9) for a Cayley tree for radius $R = 4$ and coordination number $q = 3$, during time windows of increasing duration. The pacemaker natural frequency is $\omega = 1.01\omega_c$.

where the units labeled j are the first neighbors of the pacemaker, and those labeled l all the rest. Now, using the critical condition for the synchronization $\phi_j - \phi_p = -\frac{\pi}{2}$, we have:

$$z(t) = \frac{1}{N} [\exp(i\phi_p) + \sum_{j=1}^q \exp(i(\phi_p - \frac{\pi}{2})) + \sum_{l=1}^{N-q-1} \exp(i\phi_l)] \quad (\text{A.3.7})$$

$$z(t) = \frac{1}{N} [\exp(i\phi_p)(1 + q \exp(-i\frac{\pi}{2})) + \sum_{l=1}^{N-q-1} \exp(i\phi_l)] \quad (\text{A.3.8})$$

$$z(t) = \frac{1}{N} [\exp(i\phi_p)(1 - iq) + \sum_{l=1}^{N-q-1} \exp(i\phi_l)] \quad (\text{A.3.9})$$

In general, this quantity is not computable analytically since the phases are mutually coupled by transcendental equations. However, if we consider a star shaped network, that is a one level Cayley tree ($R = 1$), then this last equation becomes

$$z(t) = \frac{1}{N} [\exp(i\phi_p)(1 - iq)] \quad (\text{A.3.10})$$

and we have an analytical expression for the behavior of the Kuramoto order parameter. For this simple case it is also possible to achieve an analytical expression for the time evolution of

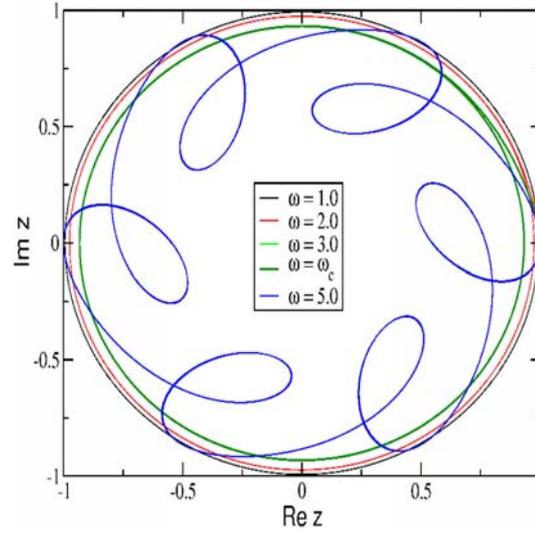


Figure A.4: Curves described in the complex plane by the order parameter (A.3.10) for a star of degree $q = 3$ and different value of the pacemaker natural frequency. That corresponding to the critical value is analytically computed while the other ones are obtained by numerical simulations.

effective frequencies:

$$\dot{\phi}_{i \neq p} = -\sin \left[2 \arctan \left(\left[A \tan \left(-\frac{A}{2}t + B \right) - N \right] \frac{1}{\omega} \right) \right]; \quad (\text{A.3.11})$$

and

$$\dot{\phi}_p = \omega + (1 - N)\dot{\phi}_{i \neq p}, \quad (\text{A.3.12})$$

where $A = \sqrt{\omega^2 - N^2}$ and $B = \arctan(\frac{N}{A})$. Plotting the values the order parameter (A.3.9) and (A.3.10) in the complex plane, for a sufficiently large time window, we can see what happens below and above the critical point. In the first case we have a circle whose radius decreases as we increase the natural frequency of the pacemaker. At the critical point this circle takes its minimal size. If we further increase the value of ω then we get something that is completely different, although the figure is still a closed curve with central symmetry. In Fig. A.3 we plot a sequences of images of the values taken by parameter (A.3.9) during time windows of increasing duration.

In Fig. A.4 we plot the final frame of some sequences analogous to the previous one, for different values of ω , in the case of a star. All these plots are obtained by means of numerical simulation, except for the critical order parameter for the star in Fig. A.4. Both figures show how, above the critical point, these values approach initially a circle smaller than the critical one, then they depart from it rapidly. It implies that the effective frequencies of the all oscillators approach each other during a certain time lag, but then they can not reach a stable constant value. This situation is also confirmed in Fig. A.5, where we plot expressions (A.3.11) and

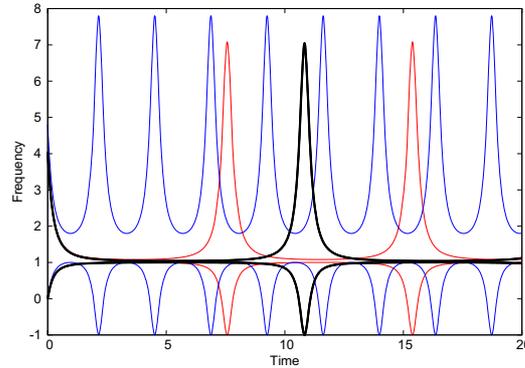


Figure A.5: Time evolution of the effective frequencies of the pacemaker (up) and its neighbors (down) in star of coordination number $q = 3$ for different value of the pacemaker natural frequency $\omega > \omega_c$: $\omega = 1.01\omega_c$ (black line), $\omega = 1.05\omega_c$ (red line), $\omega = 1.20\omega_c$ (blue line).

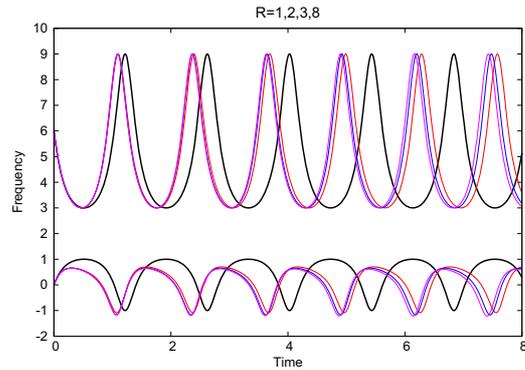


Figure A.6: Comparison of the time evolution of the effective frequencies of the pacemaker (up) and its neighbors (down) in Cayley trees of coordination number $q = 3$ and different radius. The radii are respectively $R = 1$ (black line, analytically computed), $R = 2$ (red line), $R = 3$ (blue line) and $R = 8$ (pink line).

(A.3.12) for several values of $\omega \geq \omega_c$. It is easy to notice how all the frequencies reach the same meta-stable value if ω is not too far from ω_c , keeping it during a time lag that decreases increasing ω , until it disappears. Then the frequency of the pacemaker and the ones of the other oscillators get completely apart. In all cases, the frequencies show a periodical trend with a common period.

Finally we want to compare this analytically solved simple case (the star) with the general case of a Cayley tree of arbitrary radius R (see Fig. A.6). There are two main consideration we may do about the general case. First, we notice that the time evolution of the effective frequency of the pacemaker and its neighbors, changing the radius R (and consequently the size N), given the same ω , differ from the corresponding quantities for a star of equal q , just for the oscillation period. Actually, they share the same maximum and minimum value, and it is possible to transform one curve into another one rescaling time $t \rightarrow t' = \alpha t$. However, there is

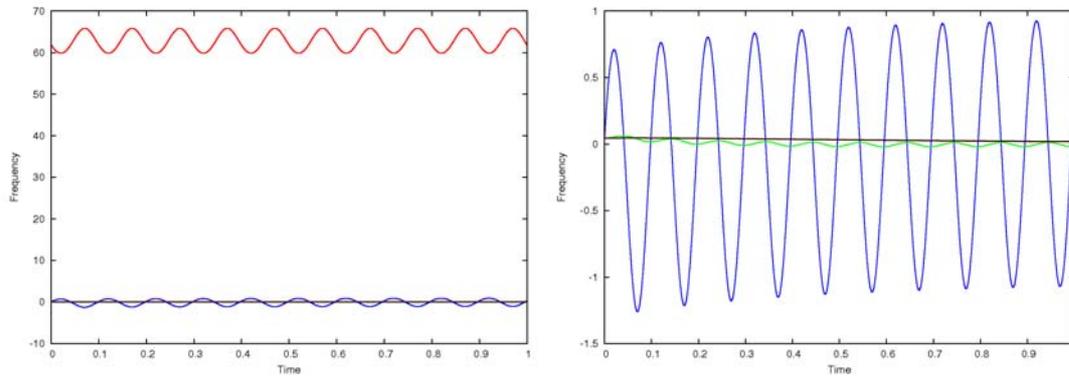


Figure A.7: Comparison of the effective frequencies time evolution in a Cayley tree of radius $R = 4$ and coordination number $q = 3$. The red line (left panel) corresponds to the pacemaker, the blue one to the nodes of the first shell, the other to the second one (green) and the third one (black). In the right panel we have removed the pacemaker effective frequency zooming on the other ones in order to show the hierarchical organization of the amplitudes of the frequencies oscillation. The pacemaker natural frequency was 20 times larger than its critical value.

no trivial expression for the rescaling factor α . Then we observe that, at least for large enough ω , the perturbation introduced by the pacemaker on the dynamics in the more external shells can be neglected (see Fig. A.7). The frequencies of these units oscillates with a very small amplitude, when compared with those of the pacemaker and the most internal shell, around a mean value that is the same for all the oscillator that are not the leading one.

In terms of the order parameter it means that if we consider a Cayley tree with increasing radius, the last term in Eq.(A.2.1) becomes clearly dominant, and eventually, the perturbation introduced by the pacemaker will not be detected (see Fig. A.7, right panel).

A.4 Conclusions

Analytical results are quite scarce when dealing with dynamical properties of nonlinear systems embedded in complex topologies. We have presented here a paradigmatic model with some assumptions that relax topological complexity, but nevertheless maintain the inherent dynamical complexity. Cayley trees or Bethe lattices have been subject of many analysis in the literature because its recurrent topology facilitates the achievement of exact results. In our case, taking a Cayley tree as a reference structure, we have computed analytically the effect of a single pacemaker located at the root of the tree for a system of Kuramoto oscillators. We have then found exact values for the critical frequency of the pacemaker above which the system is not able to synchronize and also exact values for the order parameter introduced originally by Kuramoto. These results concern the relation between topology and dynamics in a system of interacting units in a tree structure. They provide important clues on this relation in a number of situations where the connectivity patterns are tree-like or have an almost tree-like topology,

such as technological systems, like computer networks [21], or biological systems, such as prokaryotic gene regulatory networks [88].

Appendix B

Estimations and micro-characterization for the minimal IFOS model

This appendix is devoted to the derivation of some results we used throughout the chapter 4. In Sec. B.1, we derive the average size S_{cc} of connected components in the model introduced in Sec. 4.2, then discussing the stability of these local structures when dealing with moving agents. In Sec. B.2, we analyze the synchronization time T_{sync}^{ALL} in a model of integrate-and-fire oscillators whose phases evolve according to the rule introduced in the chapter 4 and whose interaction pattern is a fully connected graph. Sec. B.3 is devoted to discuss the synchronization properties of those small (weakly) connected clusters which characterize the interaction patterns of the model presented in Sec. 4.2. We empirically derive the dependence of the synchronization time (when existing) on the coupling strength. Finally, in Sec. B.4, we provide an analytical estimator for the average time an oscillator spends with its nearest neighbor (T_{out}) as a function of V and N .

B.1 Average size of connected components

Here we want to calculate the average size of the connected components in the Nearest Neighbor Graph (NNG) for a set of N point $\{P_i\}$, $i = 1, \dots, N$, randomly distributed in a box periodic with boundary condition. The *NNG* is a directed graph with P being its vertex set and with a directed edge from P_i to P_j whenever P_j is a nearest neighbor of P_i ¹.

Let us starting noting that in the NNG, as well as in any directed graph where every vertex has out-degree $k_{out} = 1$, only cycles of length 2 are possible and each *weakly connected*

¹Franco P. Preparata and Michael Ian Shamos (1985). Computational Geometry - An Introduction. Springer-Verlag. 1st edition: ISBN 0-387-96131-3; 2nd printing, corrected and expanded, 1988: ISBN 3-540-96131-3; Russian translation, 1989: ISBN 5-03-001041-6.

*component*² has exactly one 2-cycle[86]. Thus, knowing the probability that a given point is in a 2-cycle of the NNG, we can predict the number of connected components and therefore their average size.

Consider a particle located at the center of a box of side L ³. The probability of finding its nearest neighbor P_j at a distance between xL and $(x + dx)L$ is given by the product of the probability of finding anyone of the possible $(N - 1)$ particles between Lx and $L(x + dx)$ times the probability of having none of the remaining ones at a smaller distance, that is, the probability that $(N - 2)$ units lie outside this disk $D_i(Lx)$ of radius Lx centered in P_i :

$$P(x)dx = (N - 1)(1 - \pi x^2)^{N-2}2\pi x dx. \quad (\text{B.1.1})$$

The necessary condition for P_i to belong to a 2-cycle is that it is also the nearest neighbor of its nearest neighbor P_j . This condition is satisfied if in the disk $D_j(Lx)$, centered in P_j , there is no other particle. Hence, the probability distribution of the distance between two reciprocal nearest neighbors can be obtained by replacing, in expression (B.1.1), the area of $D_i(Lx)$ with that of $A_{i,j}(Lx) = D_i(Lx) \cup D_j(Lx)$ [86]. Finally we have

$$P(x)dx = (N - 1) \left[1 - x^2 \left(\frac{4\pi}{3} + \frac{\sqrt{3}}{2} \right)^2 \right]^{N-2} 2\pi x dx. \quad (\text{B.1.2})$$

Finally, we have to integrate Eq.(B.1.2) over all the possible values of x . Since we have a radial distribution, while our system is in a box, we need to approximate it with a disk by setting the upper limit of integration equal to $x^* = 1/\sqrt{\pi}$, so that $\pi(x^*L)^2 = L^2$. Then, using standard approximations that hold when $N \gg 1$ while N/L^2 remains a finite quantity, we obtain

$$P_{2c} = \frac{6\pi}{8\pi + 3\sqrt{3}} \simeq 0,6215. \quad (\text{B.1.3})$$

Therefore, the expected number of 2-cycles is $N_{2c} = N P_{2c}/2$. Since, as previously stated, each pair of reciprocal neighbors belongs to a different connected component, the average size of a connected cluster in the system is given by

$$S_{cc} = 2/P_{sl} \simeq 3,218. \quad (\text{B.1.4})$$

B.1.1 Significance in the case of moving agents

This value can be regarded as an universal constant of this kind of graphs since it neither depends on N , nor on L or N/L^2 . However, its significance is limited when the particles are

²A directed graph is weakly connected if replacing all of its directed edges with undirected edges produces a connected (undirected) graph. In the NNG there weakly connected components only, except for the case of pairs of reciprocal nearest neighbors.

³Thanks to the periodic boundary conditions, every particle is virtually at the center of the box.

allowed to move. Here we consider the model introduced in Sec. 4.2. Our aim is to determine the maximum value of V at which this quantity makes sense. We want to measure the fraction $\langle s \rangle$ of shots that, during the evolution of the system, can be regarded as a reply to a previous interaction between the same pair of units in the opposite direction, thus closing a cycle. Therefore, we need a good, general, definition of “reply-shot”. The main problem is that we have limited time resolution. We are not able to order firing events that occur between two shots of the oscillator of reference. Hence, we have to establish a formal ordering. The simplest way to do this is by ordering the oscillators according with their labels. We consider that a firing event by unit i , occurred at time T , comes before (after) an other firing event by oscillator j , occurred at the same time T , if $i < j$ ($i > j$). Thus, the conditions for a given shot $i \rightarrow j$, occurred at time T , to be the reply to a previous one $j \rightarrow i$ are:

- 1) j fired at i at time $T - 1$ if $i < j$;
- 2) j fired at i at time T if $i > j$.

According to this definition, if a symmetric link between two oscillators, that is, a 2-cycle is established at time T and broken at time $T + k$, all the shots from both the units have to be considered as reply shot, except the first one, that is that of the oscillator with the smallest ordering number at time T . Thus, when $V \ll 1$, the average fraction of replies is almost equal to P_{2c} . On the contrary, when the oscillators are moving fast, all the possible pairs of reciprocal neighbor are rewired at each time step. In this case, only the second one of the previously exposed conditions can be satisfied due to a favorable geometrical configuration. The first condition can be satisfied only by chance, with mean probability $p_1 = \frac{1}{2N}$. Consequently, the average fraction of replies will approach $P_{2c}/2$ (plus a correction of order N^{-1}). In this situation, the meaning of the connected components in the interaction pattern is lost. They no longer represent a set of units whose phases evolve in a interdependent way, because the *mean life* of the 2-cycle is simply too short. By measuring the average fraction of reply-shots for increasing values of V , we are able to localize the transition from an almost static situation, where the connected components play an important role in the dynamical evolution of the system toward the coherent state, to a condition of permanent random rewiring.

Fig. B.1 shows that in the left region, and in particular for $V \leq V_m$, there is absolutely no problem using P_{2c} and therefore S_{cc} . In the same figure, we also plotted the average final mixing $\chi = \frac{1}{N} \sum_i n_i(T_{sync})$, where n_i is the number of different neighbors with whom oscillator i has interacted until T_{sync} . Both $\langle s \rangle$ and χ undergo a transition that starts close to the no-synchronization region, displaying an opposite behavior. The first quantity is constant in the slow regime, then starts decreasing and finally, when the system enters the fast regime, it reaches its minimal value. The second one is minimal in the slow regime, slowly increasing until the system enters the no-synchronization region, and finally it reach the maximum value $\chi = 1$ at the entrance of the right region.

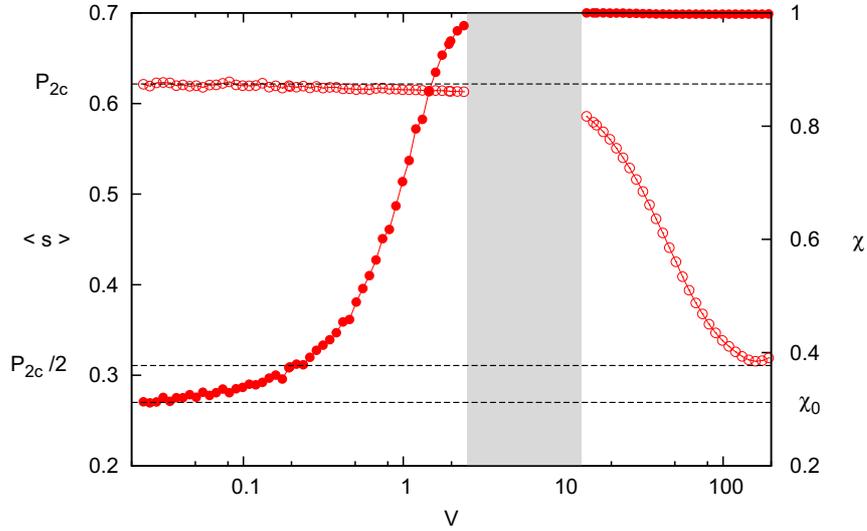


Figure B.1: The average symmetry $\langle s \rangle$ (empty circles, left axis) and the global final mixing χ (filled circles, right axis) are plotted against the velocity V . The grey rectangle marks the no-synchronization region where no empirical data is available.

B.2 Synchronization time in a fully connected IFOS system

We consider a system of N integrate-and-fire of the type introduced in Sec.4.1 on an all-to-all interaction pattern. When any of the oscillators fires, it reaches simultaneously all the remaining $N - 1$ modifying their phase according to the rule $\phi_j \rightarrow (1 + \epsilon)\phi_j$. In this case, the mobility of the agents does not play any role.

We have numerically computed the average synchronization time for this model with $N \in \{10, 20, 30, 50, 80, 100, 150\}$, averaging over 2×10^4 realizations, varying value of the coupling strength between $\epsilon_{min} = 2/N$ and $\epsilon_{max} = 1$. In Fig. B.2 we plotted the synchronization time dependence on ϵ . The standard deviation of the sample is irreducible increasing the number of repetitions since it is a direct consequence of different initial conditions giving rise to different paths to synchronization in the phases landscape. In the inset of Fig. B.2.1 we can see that all the points collapse into a single curve when plotted against $N^\gamma \epsilon$, with $\gamma = (0.80 \pm 0.02)$. Finally, we found the empirical expression

$$T_{sync}^{ALL} = \frac{2}{N^\gamma \epsilon}. \quad (\text{B.2.1})$$

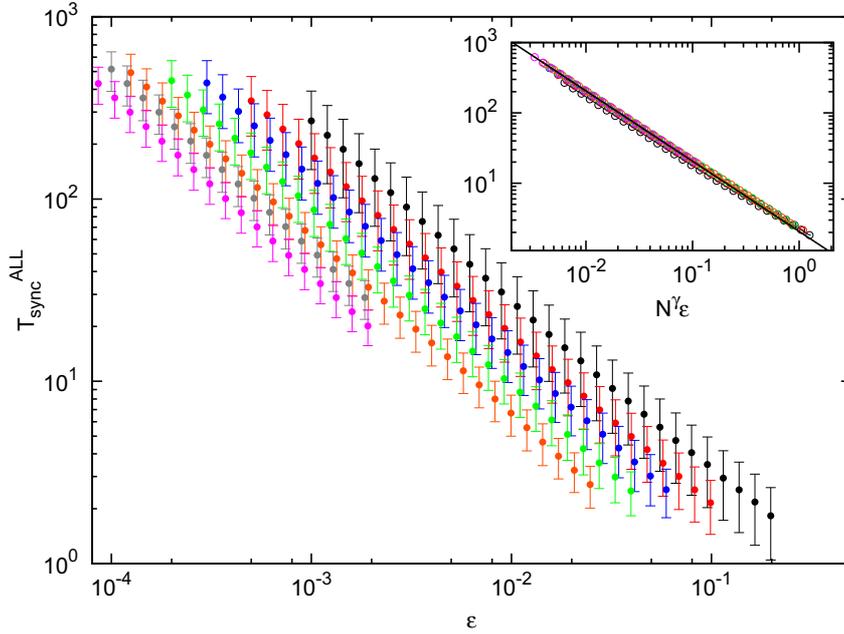


Figure B.2: Synchronization time dependence on ϵ for the same value of N (black: $N = 10$, red: $N = 20$, blue: $N = 30$, green: $N = 50$, orange: $N = 80$, grey: $N = 100$, pink: $N = 150$). Error bars represent the dispersion (standard deviation) among the values obtained for different realizations (different initial conditions). In the inset: Synchronization time T_{sync}^{ALL} as a function of $N\gamma\epsilon$ for various values of N . Black solid line stands for Eq. (B.2.1).

B.3 Local cluster synchronization time

In this section we want to explore the typical time spent by the stable clusters formed in our minimal model. We have already seen that in slow regime $V < V_m$ such clusters emerge with a typical size of $S_{cc} = 3.218$ and hold the key to the eventual synchronization of the system.

Here we consider groups of size $N = 2, 3, 4$. We take into account only those configurations that may attain synchronization. For $N = 2$ and $N = 3$ all configurations are able to synchronize, while for $N = 4$ the oscillators with $k_{in} = 0$ have to share the same initial phase.

For the pair case $N = 2$, the synchronization time T_2 as a function of the initial phase difference $\Delta\phi$ can be computed in a straightforward way, as well as its mean value,

$$T_2(\epsilon) = \frac{1}{\epsilon(\epsilon + 2)}. \quad (\text{B.3.1})$$

For the case $N = 3$ we can consider two given configurations, a triangle (that has null probability) and a pair plus a third oscillator with $k_{in} = 0$, that is the configuration we considered in Sec. 4.6. Hereafter, we will refer to this configuration as 3-chain (see figure B.3).

Finally, for the case $N = 4$, we consider two different cases, both with non-null probability,

while we neglect the square configuration (see figure B.3). The first one, is a configuration with only one agent having a null in-degree (4-chain). In second one, there are two agents with $k_{in} = 0$ (and with the same initial phase) which fire at two different oscillators which are reciprocal neighbors (star). There is also a third configuration that can be present in the system. In this case the oscillators with null in-degree share the same nearest neighbor (semi-star). Differently from the first two ones, this last connected component only synchronizes for a subset of the possible initial conditions, even if the oscillators that are not receiving any input are already synchronized. Notice that for any configuration may exist a set of initial conditions that prevents the system to synchronize, but in all the cases except this one this set has null-measure. For instance, for a pair of reciprocal neighbors there exists a unique fix-point at $\Delta\phi = [1 - 1/(2 + \epsilon)]$ such that the phase difference does not change during the time evolution of the system. In Fig. B.4 we show the average synchronization time as a function of ϵ for each

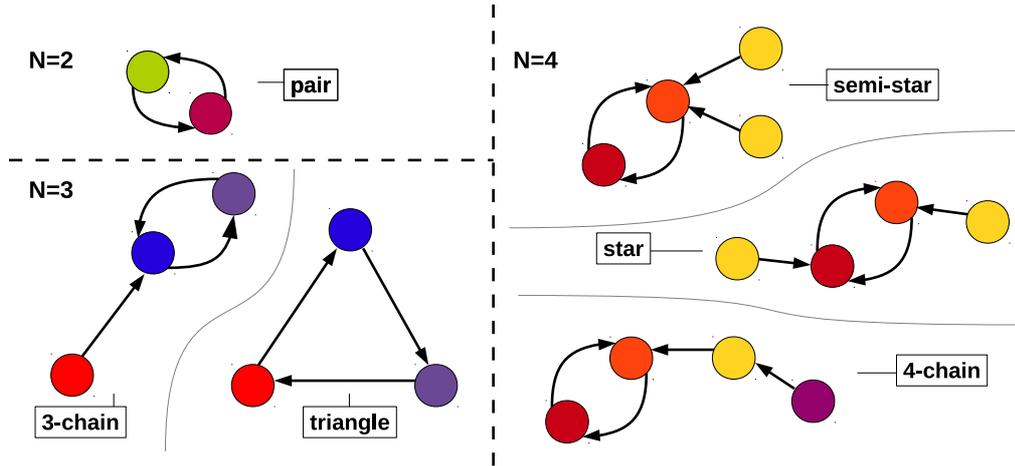


Figure B.3: Connected cluster of size $N = 2, 3, 4$.

one of the considered motifs.

B.3.1 Local synchrony time scale

Fig. B.4 shows that, for the cases of a star, a 3-chain and a 4-chain, the synchronization time is almost always the same, with a constant relative error (standard deviation). The triangle is quite faster to synchronize. As we discussed in Sec.4.5, an isotropic connection pattern help the system to achieve a coherent state in a shorter time. The case $N = 2$ is characterized by a larger dispersion (see the bottom inset), being the standard deviation of the values of T_{sync} obtained for different initial conditions of the same order of magnitude as the mean value.

Consequently, to identify the typical time scale of the local synchrony is not an easy task. On the one hand, there are configurations that cannot synchronize, or that can synchronize only if some of the units had already the same phase; on the other hand, there is a relevant

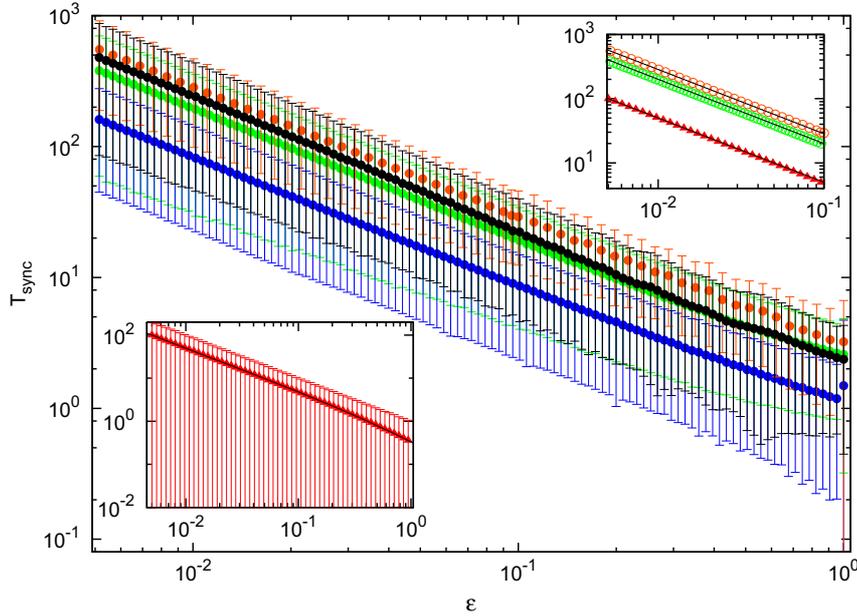


Figure B.4: Synchronization time as a function of ϵ for a 3-chain (green), a triangle (blue), a 4-chain (orange) and a 4-star (black). In the inset below, the synchronization time dependence on ϵ for a pair of reciprocal neighbors. Error bars stand for the standard deviation of the values obtained for different initial conditions. In the inset above: data for a pair (red triangles), a 3-chain and a 4-chain, in the interval $\epsilon \in [0.005; 0.1]$ fitted by a function (black lines) of the type $f(\epsilon) = a_N/\epsilon$, with $a_2 = (0.50 \pm 0.01)$, $a_3 = (2.00 \pm 0.01)$ and $a_4 = (2.85 \pm 0.01)$. Each point has been calculated averaging over a sample of 100 equispaced values of the phase of each oscillator (except that of reference, whose phase we set equal to 1).

fraction of pairs that synchronize very quickly. Actually, we should count out all the frustrated configuration since they do not play any role in the increasing of local coherence. Even if the mean size of the connected components is $S_{cc} \simeq 3.2$ (see Appendix B.1), the mean size of a *good* cluster able to synchronize is probably smaller than 3. Therefore, the typical time scale of local synchrony would likely be smaller than T_3 , the synchronization time of the 3-chain, as assumed in Sec. 4.6. However, the only effect of this possible correction is a multiplicative constant factor in the definition of the scaling variable r introduced in Sec. 4.6, Eq. 4.6.4. Indeed, the dependence on ϵ of the synchronization time for all the considered configuration is of the type $T_{sync} \sim \epsilon^{-1}$, at least when the coupling strength is small enough (see the top inset). The scaling argument provided in Sec. 4.6 holds even if we cannot set a local synchrony time scale with a good precision since what really matter is the dependence of this time scale on ϵ that has been univocally identified.

B.4 Estimation of the (out-)neighbor change time

Here we want to calculate the average time an oscillator spends with its neighbor $T_{out}(V; N, \epsilon)$ or, in other words, the average time an oscillator needs to change its neighbor.

Let us start considering a unidimensional version of our system: N oscillators move on a ring of length L and fire at their first neighbor when their phase is equal to 1. To perform our calculation we suppose that all the particles are fixed and equispaced, being $l = L/(N - 1)$ the distance between two consecutive ones. Only a single oscillator, labeled i , is allowed to move.

The ring can therefore be considered as divided into $(N - 1)$ line segments $\{l_j\}_{j \neq i}$, each one of whom is centered in a different unit j . If i lies in l_j , this implies that j is its first neighbor.

Now we want to write the probability that i changes its neighbor within a time window $\Delta T = 1$, that is, the frequency of the neighbor change events. For sake of simplicity, let us suppose that the system is already synchronized.

According to the rules of the model, i is supposed to move straightway during a time lag $\tau = 1$, covering a distance $\Delta s = V$. At each time T , it would change its orientation ($V \rightarrow -V$) with probability $p = 1/2$ because of the shot it receives from its neighbor. Notice that the existence of an in-neighbor that fires at i is assured by construction.

In order to change its neighbor, oscillator i has to cross one of the end points of l_j .

Therefore, it is going to change its neighbor between time $T - 1$ and time T if (i) it lies at a distance $d \leq \Delta s$ from one of the two end points at time $T - 1$ and (ii) it is moving toward that same point. The joint probability to satisfy condition (i) and (ii) is given by:

$$P_{out}^{1D} = 2 \frac{\Delta s}{l} \frac{1}{2} = \frac{V(N - 1)}{L}, \quad (\text{B.4.1})$$

where the factor 2 is due to the two possible exiting directions.

Eq.(B.4.1), being a probability per unit of time, is the inverse of the average time separation between two neighbor changes, that is the quantity we are looking for. Thus we can write

$$T_{out}^{1D} = \frac{L}{V(N - 1)}. \quad (\text{B.4.2})$$

In order to generalize the calculation to the 2-dimensional case, instead of a ring divided in line segments, we have to consider a box of side L where $(N - 1)$ particles are arranged in a lattice. Thus we have $(N - 1)$ squares of side $l = L/\sqrt{N - 1}$, and a fixed oscillator located at the center of each one of them. Oscillator i changes its neighbor when it exits a square to enter one of the four adjacent ones. It will cross a given side between time $T - 1$ and T if, at time $T - 1$, the component of V perpendicular to that side is larger than the distance s separating i from that same boundary. Adopting an appropriate axis system, the coordinates of the four vertices of the considered square are, respectively, $A \equiv (0; 0)$, $B \equiv (l; 0)$, $C \equiv (l; l)$ and $D \equiv (0; l)$. Let us focus on one of the four possible sides, namely \overline{BC} (see Fig. B.5, left panel).

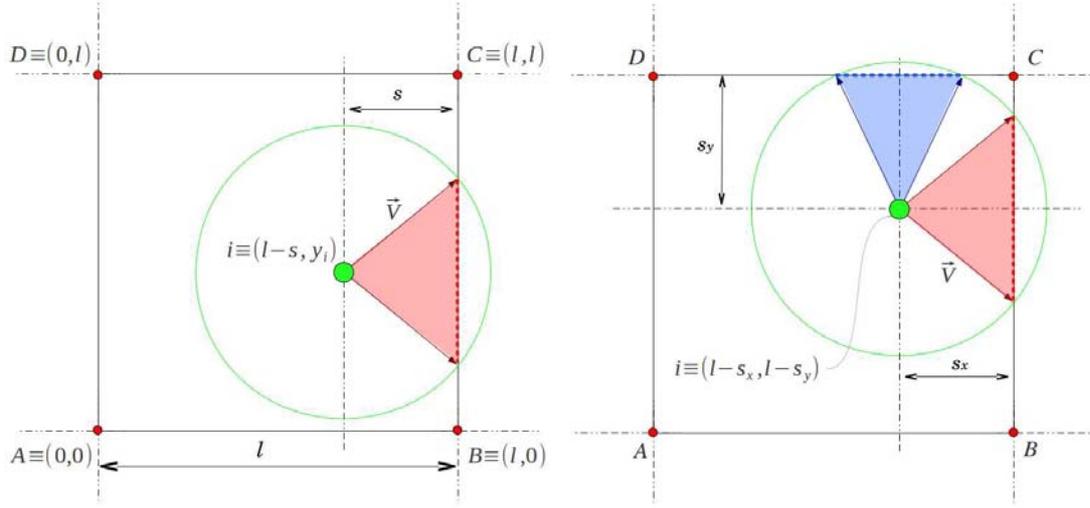


Figure B.5: Oscillator i exiting from a square of side l . The green circles of radius V and center in i marks the possible positions that the oscillator will take at time $T + 1$. In the left panel, the position and velocity (module) are such that, if the velocity vector (orientation) lies in the red area, i will exit the square through the border \overline{BC} at the next time step, otherwise it will remain inside the square. In the right panel, there is also the chance for i to cross the border \overline{CD} if the velocity vector is in the blue area. Depending on the orientation of its motion, i could cross \overline{BC} or \overline{CD} or none of them.

For this side, previous requirement can be expressed as $(V \cos \theta) > s$ and the probability of satisfying this condition is given by

$$P_{BC}(out|s) = \frac{1}{\pi} \arccos\left(\frac{s}{V}\right), \quad \text{if } s \leq V \quad (\text{B.4.3})$$

$$P_{BC}(out|s) = 0 \quad \text{otherwise.}$$

Finally, we can write:

$$P_{out} = \int 4P_1(out|s)p(s)ds = \frac{4}{\pi} \int_0^V \arccos\left(\frac{s}{V}\right) \frac{\sqrt{N-1}}{L} ds = \frac{4}{\pi} \frac{V\sqrt{N-1}}{L}, \quad (\text{B.4.4})$$

where the factor 4 is due to the four possible exiting directions. From the inverse of Eq.(B.4.4) we obtain

$$T_{out} = \frac{\pi}{4} \frac{L}{V\sqrt{N-1}}. \quad (\text{B.4.5})$$

Despite that in deriving estimator (B.4.5) we used some strong approximations, it turns out to be remarkably good (dashed line in Fig. B.6).

However, we have completely neglected the fact that, under certain conditions, oscillator i is able to exit the square through more than one side during a time interval $\Delta T = 1$. In other words, it can be the case that both $(V \cos \theta)$ and $(V \sin \theta)$ are larger than the distance

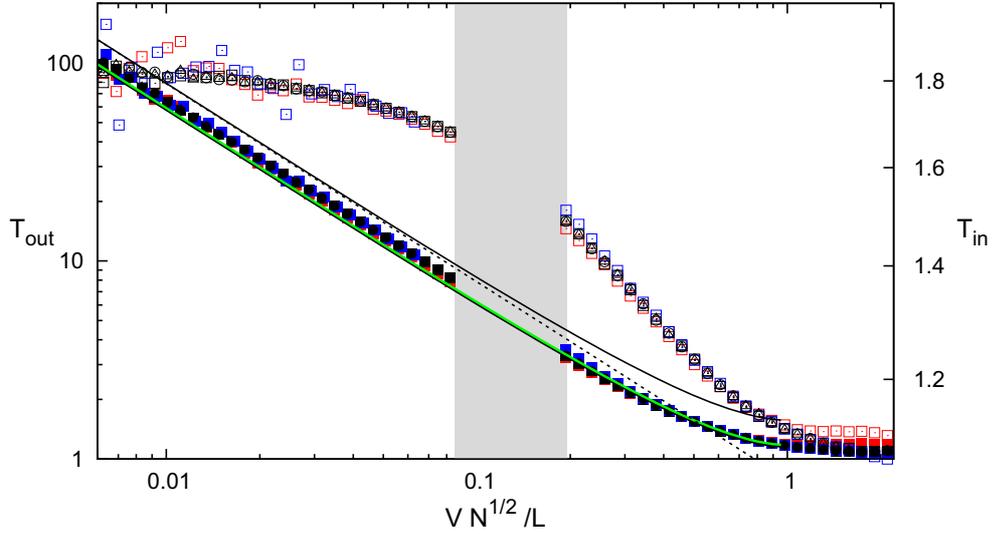


Figure B.6: The average out-neighbor time T_{out} (filled symbols, left axis) and the average in-neighbor time T_{in} (empty symbols, right axis) are plotted against the quantity $V\sqrt{N-1}/L$, for some values of N (red: $N = 10$, black: $N = 20$, blue: $N = 30$) and L (circles: $L = 200$, squares: $L = 400$, triangles: $L = 800$). The black dashed line stands for the theoretical estimator (B.4.5), while the solid line is the estimator (B.4.11). The green solid line correspond to Eq.(B.4.13).

from the respective orthogonal sides. This situation, that is quite unlikely when the velocity is small, becomes very frequent when V is of the same order of magnitude as l . This implies that we are not allowed to sum the probabilities corresponding to each exit direction to get the total exiting frequency, as we do in (B.4.4). On the contrary, we have to take into account all the different possible situations, in order to properly express the probability to cross a certain boundary. Considering again the side \overline{BC} (see Fig. B.5, right panel), the conditions to exit the square by this side can be written as

$$\begin{cases} V \cos \theta > l - x; \\ V \sin \theta < l - y; \\ -V \sin \theta < y; \end{cases} \quad (\text{B.4.6})$$

where $P \equiv (x, y)$ is the position of oscillator i at time $T - 1$. The conditional probability to exit the square by the boundary \overline{BC} , given the position of i , can then be expressed as

$$\begin{aligned} P_{BC}(out|x, y) &= \\ &= \frac{1}{2\pi} \left[\min \left\{ \arccos \left(\frac{l-x}{V} \right); \arcsin \left(\frac{l-y}{V} \right) \right\} + \min \left\{ \arccos \left(\frac{l-x}{V} \right); \arcsin \left(\frac{y}{V} \right) \right\} \right]. \end{aligned}$$

Since it is

$$\begin{aligned} \arccos\left(\frac{l-x}{V}\right) &= \arcsin\left(\frac{l-y_M}{V}\right) \Rightarrow y_M = l - \sqrt{V^2 - (l-x)^2}; \\ \arccos\left(\frac{l-x}{V}\right) &= \arcsin\left(\frac{y_m}{V}\right) \Rightarrow y_m = \arccos\left(\frac{l-x}{V}\right); \arcsin\left(\frac{y}{V}\right); \end{aligned} \quad (\text{B.4.7})$$

we obtain

$$P_{BC}(out|x, y) = \begin{cases} \frac{1}{2\pi} \left[\arccos\left(\frac{l-x}{V}\right) + \arcsin\left(\frac{y}{V}\right) \right], & y < y_m; \\ \frac{1}{\pi} \arccos\left(\frac{l-x}{V}\right), & y_m < y < y_M; \\ \frac{1}{2\pi} \left[\arcsin\left(\frac{l-y}{V}\right) + \arccos\left(\frac{l-x}{V}\right) \right], & y > y_M. \end{cases} \quad (\text{B.4.8})$$

Integrating over the region of the square where the first condition holds ($y < y_m$), we can write the corresponding exit probability as

$$\begin{aligned} P_{BC,1}(out) &= \frac{1}{2\pi l^2} \int_{l-V}^l \int_0^{y_m} P_{BC}(out|x, y) dy dx \\ &= \frac{1}{2\pi l^2} \int_{l-V}^l \left[2 \arccos\left(\frac{l-x}{V}\right) \sqrt{V^2 - (l-x)^2} + l - x - V \right] dx \\ &= \frac{V^2}{2\pi l^2} \left(\gamma - \frac{1}{2} \right), \end{aligned} \quad (\text{B.4.9})$$

where $\gamma = (4 + \pi^2)/8$.

For the second region ($y_m < y < y_M$) we have

$$\begin{aligned} P_{BC,2}(out) &= \frac{1}{\pi l^2} \int_{l-V}^l \int_{y_m}^{y_M} P_{BC}(out|x, y) dy dx \\ &= \frac{1}{2\pi l^2} \int_{l-V}^l \arccos\left(\frac{l-x}{V}\right) \left[l - \sqrt{V^2 - (l-x)^2} \right] dx \\ &= \frac{V}{\pi l} - \gamma \frac{V^2}{\pi l^2}. \end{aligned} \quad (\text{B.4.10})$$

For the third region ($y > y_M$) we notice that, since the system is symmetric with respect to the axis $x = l/2$, the probability $P_{BC,3}$ has to be the same as the probability $P_{BC,1}$.

Finally, we can write

$$T_{out} = \frac{1}{4} \left[\frac{V\sqrt{N-1}}{\pi L} - \frac{V^2(N-1)}{2\pi L^2} \right]^{-1}. \quad (\text{B.4.11})$$

Estimator (B.4.11) behaves the same as the empirically measured quantity it should predict (continuous black line in Fig. B.6), differing from it just for a constant factor. This discrepancy is due to the underestimated average first-neighbor distance induced by the used simplified

configuration. Indeed, the first moment of the distribution (B.1.1) is $\langle x_{nn} \rangle \simeq l/2$, where the approximation holds for $N \gg 1$ (it is already exact for any practical purpose when $N = 10$). On the contrary, considering a uniformly distributed particle inside a square of side $l = L/\sqrt{N-1}$, the average distance from the center is given by $\langle s \rangle = l[(\sqrt{2} + \log(\sqrt{2} + 1))/6] \simeq 0.299l$. As a consequence, we are overestimating the distance the oscillator i has to cover in order to change its neighbor, and so the corresponding time T_{out} .

Hence, we decide to try a fit function of the form (B.4.11) with a multiplicative free parameter g . We find that with $g = 0.75$ (green line in Fig. B.6) the function fits perfectly data points corresponding to velocity values of the right region, while for those of the left one a factor $g' = 0.83$ is better.

This is because our empirical data refer to the entire duration of the dynamical evolution of the system, from a completely random initial state, until the achievement of the synchronization. In the incoherent state, the time separation between two consecutive fires of the same oscillator is in general smaller than 1 (being equal to one only if it is not receiving any shot). This implies that the distance covered by any unit during the time lasting between two shots of the oscillator of reference is a bit smaller than $s = V$, and consequently T_{out} is slightly larger. Anyway, the effect of this correction is not the same in all the regimes, being much more relevant in the slow case. Indeed, in the right (fast) region, the system spends half of the time it needs to synchronize in an almost-coherent state (see Sec.4.4) and the measured T_{out} is very close to that we could have measured in the synchronized state.

In conclusions, we can write the following predictors T_{out}^{left} and T_{out}^{right} respectively for the region on the left (including the slow regime region $V < V_s$) and that on the right of the no-synchronization interval:

$$T_{out}^{left} = \frac{g'}{4} \left[\frac{V\sqrt{N-1}}{\pi L} - \frac{V^2(N-1)}{2\pi L^2} \right]^{-1} \simeq 4g' \frac{\pi L}{V\sqrt{N-1}} \quad (\text{B.4.12})$$

$$T_{out}^{right} = \frac{g}{4} \left[\frac{V\sqrt{N-1}}{\pi L} - \frac{V^2(N-1)}{2\pi L^2} \right]^{-1}, \quad (\text{B.4.13})$$

where $g = 0.75$ and $g' = 0.83$.

B.4.1 The fast limit case. Heuristic derivation of V_f

About the fast limit ($V > V_f$), we have to notice that in our derivation we have considered only values of the velocity $V \leq l$. We were implicitly assuming that there is no necessity for a predictor of T_{out} beyond that value of V . Indeed, in the real system, all the particles move at the same time, but with independent orientations. If, at each time step, the displacement of every oscillator is such that the distance between the previous and new position is l for all of them, this should ensure a complete re-arrangement of the units in the different squares. This means a complete rewiring of the connectivity pattern. Then, the probability of taking a new

neighbor is given simply by $P_{out} \simeq (1 - 1/N)$, that is the same as saying that if someone of the oscillators maintains its previous neighbor, it is just for chance. Hence, for the fast regime, we can write:

$$T_{out}^f = \frac{N}{N-1} \quad (\text{B.4.14})$$

Since beyond this value of the velocity, every dynamical quantities we can measure, including T_{sync} , no longer depends on V , we expect the system to enter the fast regime at this precise value of V , that is $V_f = L/\sqrt{N-1}$.

B.4.2 The in-neighbor change time

Together with T_{out} , in Fig. B.6, we also plotted T_{in} . Differently from T_{out} , T_{in} is bounded and does not diverge when $V \rightarrow 0$. Notice that, if an oscillator i has a number of in-neighbor $k_{in}^{(i)} > 1$, the it will be changing $k_{in}^{(i)}$ in-neighbors per time unit. Assuming that all the connected components in the system are groups of three units, during any time windows ΔT , there is one third of the population with $k_{in} = 2$. So, during this time lag, the total number of neighbor changes is given by

$$N_{in}(\Delta T) = \frac{2N\Delta T}{3},$$

where we have neglected any additional changes due to the dynamics (static limit). Then, the average in-neighbor change time is

$$T_{in} = \frac{1}{P_{in}} = \lim_{\Delta T \rightarrow \infty} \frac{N\Delta T}{N_{in}(\Delta T)} = \frac{3}{2}.$$

This simplistic calculation accounts very well for the order of magnitude of T_{in} in the slow regime (see Fig. B.6).

Resumen

Introducción

Durante siglos, se ha considerado que el conocimiento científico se identifica con la posibilidad de predecir la evolución futura de los fenómenos a partir del conocimiento de las leyes que los rigen, y que el objetivo de todo esfuerzo científico es el descubrimiento de esas leyes. Este programa está profundamente influenciado por el gran éxito que la Física tuvo a partir del siglo XVI. La Física, una disciplina experimental en la que las predicciones teóricas se comparan con los experimentos, se diferencia de otras ciencias naturales por el papel fundamental de las matemáticas. Los físicos describen los fenómenos mediante el uso de modelos matemáticos y sus predicciones se obtienen mediante el razonamiento matemático [7]. La trayectoria de una bala, la velocidad de un cuerpo en caída libre, el tiempo necesario para cubrir una órbita determinada, fueron las primeras cantidades que se pudieron calcular con precisión arbitraria, mediante la aplicación de las leyes de la dinámica y de la gravitación universal formuladas por Newton. El éxito empírico de este método ha ido reforzando la creencia de que todos los fenómenos físicos, a pesar de sus apariencias heterogéneas, puedan ser explicados en términos de leyes *simples* y *universales* que permitan predecir, por medio de cálculos, la secuencia de eventos que se darán en un sistema dado, siempre y cuando las condiciones iniciales sean conocidas [8].

A principios del siglo XIX, Laplace escribía que un matemático infinitamente inteligente sería capaz de predecir con certeza el futuro mediante la observación del estado actual del universo utilizando su conocimiento de las leyes del movimiento. Según este punto de vista, para entender un sistema, es necesario descomponerlo en sus componentes mínimos. Por el contrario, una representación detallada es inútil y engañosa, pues los accidentes no cambian la *naturaleza* del fenómeno y, por tanto, su evolución. Una vez realizado este proceso de *limpieza*, todo se vuelve simple y predecible. Esta visión, generalmente etiquetada como *reduccionista*, gobernó el pensamiento científico durante la mayor parte de la historia de la ciencia y alcanzó su apogeo durante el siglo XX con el nacimiento de la biología molecular y de la física de partículas. Su papel en el desarrollo de la ciencia moderna es innegable. Gracias a este enfoque, la física moderna ha alcanzado el nivel más alto de precisión en la descripción matemática de la realidad, basando su descripción en una representación adecuada de la materia en términos de

estructura atómica y subatómica. Además, este mismo método permitió llegar a la estructura de las macromoléculas biológicas y a la lógica de sus funciones. En aquellos años, gran parte de la comunidad científica compartía la opinión de Francis Crick, quien pensaba que conocer “en todos sus detalles todos los pasos químicos que tienen lugar en la célula durante el ciclo celular” sólo era cuestión de tiempo, y después “no habrá nada más que saber acerca de la propia célula, y el mecanismo de su vida será decodificado por completo.”

Sin embargo, en los años 60 las cosas empezaron a cambiar. Sin duda, la gran mayoría de los científicos activos admitía que “el funcionamiento de nuestras mentes, de nuestros cuerpos y de toda la materia animada e inanimada de la cual tenemos un conocimiento detallado, está [...] controlado por el mismo conjunto de leyes fundamentales” [9]. Se empezaba también a poner de manifiesto que la comprensión de fenómenos a mayor escala a partir de las escalas inferiores es a menudo imposible. Como señaló P. W. Anderson en 1972 [9],

“la hipótesis reduccionista no implica en modo alguno una hipótesis constructorista: la capacidad de reducir todo a leyes simples y fundamentales no implica la posibilidad de reconstruir el universo a partir de estas leyes. [...] La hipótesis constructorista se quiebra cuando es confrontada con las dificultades gemelas de escala y complejidad. Resulta ser que el comportamiento de grandes y complejos agregados de partículas elementales no puede ser entendido en términos de una simple extrapolación de las propiedades de unas pocas partículas. Por contrario, en cada nivel de complejidad aparecen propiedades completamente nuevas, y el entendimiento de los nuevos comportamientos requiere de investigación que considero tan fundamental en su naturaleza como cualquier otra [...] y que mostrará cómo el todo no sólo es más que la suma de las partes, sino que es muy diferente.”

Sistemas que muestran tales características están formados por muchos elementos en interacción y se conocen generalmente como sistemas complejos. Tradicionalmente, la ciencia fundamental explora lo que está más allá de la percepción cotidiana de los humanos, es decir, lo muy pequeño y lo muy grande. La peculiaridad de los sistemas complejos es que tienen que ver con una clase de fenómenos de importancia fundamental en los cuales el sistema y el observador pueden evolucionar en escalas de tiempo y espacio comparables [10]. Existen muchos ejemplos de sistemas complejos para los cuales se han desarrollado modelos, entre ellos encontramos las colonias de hormigas, las economías y las estructuras sociales humanas, el clima, el sistema nervioso, las células y los seres vivos, incluidos los seres humanos, y también la modernas infraestructuras de energía o de telecomunicaciones. Se entiende que tales sistemas muestran propiedades dinámicas emergentes inherentemente relacionadas con la topología del subyacente patrón de conexiones entre las partes constituyentes. Las propiedades más importantes de los sistemas complejos se deben más a la forma de las relaciones entre las partes y en la organización dinámica global que en la naturaleza de sus constituyentes. Por eso, para entenderlos, es preciso estudiar tanto el todo como sus partes [12].

As que, por un lado, no es posible explicar un nivel de organización sólo en términos de los niveles inferiores. Por otro lado, queda claro que, más que los detalles inherentes a los componentes, lo que necesitamos es un mapa de las interacciones. Los sistemas complejos en muchos casos se pueden describir adecuadamente a través de sus redes de contactos, es decir, en términos de nodos (que representan los componentes del sistema) y de enlaces (sus interacciones). De esta manera es posible capturar sus características esenciales en una representación simple y general.

En 1998, Watts y Strogatz [13] presentaron la primera evidencia de las que llamaron redes complejas. Hasta entonces, ya se habían utilizado redes de sistemas dinámicos acoplados para modelar osciladores biológicos, uniones de Josephson, redes neuronales, redes de control genético y muchos otros sistemas auto-organizados. Pero, mientras que normalmente la topología de las conexiones se suponía totalmente regular, o completamente aleatoria, Watts y Strogatz se dieron cuenta de que “muchas redes biológicas, tecnológicas y sociales se encuentran en algún lugar entre estos dos extremos.” Este es el caso, por ejemplo, de la red neuronal del gusano *Caenorhabditis elegans*, de la red eléctrica del oeste de Estados Unidos, y del grafo de las colaboraciones de los actores de cine. Todos estos patrones de conexión muestran rasgos comunes que no pueden explicarse mediante modelos matemáticos simples. En particular, la presencia de atajos hace estas redes extremadamente pequeñas en términos de distancias medias entre nodos, mientras que al mismo tiempo guarda la huella de su origen ordenado a través de un coeficiente de agrupamiento (*clustering*, relacionado con la existencia de triángulos en la red) mayor de lo esperado [11]. Estos dos fenómenos no se podían explicar simultáneamente mediante los modelos utilizados hasta la época (totalmente ordenados o totalmente desordenados). Un poco más tarde, Barabási y Albert [14], mostraron que la distribución del número de conexiones que salen de un nodo dado (por ejemplo, ordenadores en Internet, las páginas de la *World Wide Web*, o el número de artículos escritos por los científicos) está sesgada, lo que significa que, mientras que la gran mayoría de los elementos tienen muy pocas conexiones, existen algunos actores en las redes que están inesperadamente muy conectados [11]. Desde entonces, las redes complejas [15, 16, 17, 18] se han convertido en un marco importante y ampliamente utilizado para la comprensión de aspectos tanto dinámicos como topológicos de sistemas tales como el cerebro [19], las redes de interacción proteína-proteína [20], o Internet y la WWW [21].

Al principio, cuando las redes complejas eran un campo científico recién nacido, casi todo el esfuerzo investigador estaba dirigido a describir sistemas mediante una red y analizar las características topológicas de los grafos resultantes. Fue así posible identificar características importantes compartidas por sistemas pertenecientes a una misma clase, lo que en física se conoce como clases de universalidad. En años más recientes, el creciente interés en este enfoque, gracias también a un progreso tecnológico favorable, ha llevado a la acumulación de una cantidad cada vez mayor de datos. Bases de datos de redes habituales codifican la in-

formación acerca de cuales pares de nodos están conectados y cuales no, a veces con alguna especificación adicional sobre de la fuerza de las conexiones o la *importancia* de las unidades. Actualmente, existe una enorme cantidad de estos datos disponibles. Esta situación ha permitido el surgimiento de nuevas preguntas y, por lo tanto, la diversificación de la actividad científica. Entre ellas podemos destacar tres cuestiones generales que han estado recibiendo mucho interés: (i) la información disponible es siempre fiable y completa? (ii) cómo un patrón de interacción complejo puede afectar el surgimiento de comportamientos colectivos? Y (iii) cual es el papel de la movilidad en el marco de las redes complejas?

- i) Por lo que se refiere a la pregunta (i), hoy en día se sabe que muchas de las redes analizadas sólo se conocen parcialmente. Pensemos, por ejemplo, en las redes artificiales, como las redes sociales on-line o la misma Internet, que se componen de millones de nodos heterogéneos y no idénticos. En este tipo de redes de gran tamaño, un mapa completo es muy difícil de conseguir [22]. En consecuencia, proporcionar herramientas eficientes para su exploración se ha convertido en un desafío crucial.

Por otro lado, en muchas situaciones, las conexiones entre los elementos no son accesibles directamente, en el sentido que no son objetos físicos, como un cable, un hiperenlace o un axón. Muchas de las redes naturales, como las redes genéticas, cerebrales o ecológicas, tienen que ser entendidas como representaciones de las relaciones existentes entre las diferentes partes del sistema considerado. En todos estos casos, para poder trazar un patrón de conectividad, los enlaces tienen que ser inferidos midiendo correlaciones de cantidades relacionadas con el comportamiento de las unidades. Obviamente, diferentes métodos y experimentos pueden dar lugar a resultados diferentes. Es este el caso, por ejemplo, de las redes de regulación génica. Al fin de construir patrones de conexión cada vez más completos, a menudo se necesita combinar datos obtenidos por medio de diferentes procedimientos experimentales, no siempre coherentes entre ellos. Por lo tanto, en los últimos años, el problema de la inferencia de topologías de red a partir de medidas dinámicas ha sido objeto de intensa investigación, tanto experimental como teórica. Recientemente, se ha presentado un marco matemático y computacional para hacer frente al problema de la fiabilidad de los datos en redes complejas [23]. En particular, este enfoque pretende identificar de forma fiable las interacciones que faltan y las que sobran en observaciones de redes ruidosas.

- ii) Acerca de la pregunta (ii) tenemos que decir que el hecho de que la topología del patrón de conexión afecte a la evolución dinámica de los elementos es precisamente lo que permite la reconstrucción de muchas redes. Por lo tanto, es evidente que los comportamientos colectivos, es decir, los comportamientos que caracterizan el sistema como un todo y que existen gracias a las interacciones entre las partes, pueden cambiar dependiendo las características de la topología de conexión. Hoy en día, una parte importante del esfuerzo científico está dirigido a comprender cómo las propiedades dinámicas globales

se relacionan con la dinámica de las unidades y las interacciones entre ellas. Uno de los primeros ejemplos de esta relación a ser investigados fue la transición entre orden y caos que las redes booleanas [25] muestran al aumentar el número de entradas (conexiones) de cada nodo. Más recientemente, el impacto de la topología de la red ha sido objeto de gran interés en relación, solo por citar uno entre muchos otros fenómenos, a la propagación de enfermedades [26]. Sin embargo, se podrían proporcionar muchos otros ejemplos importantes.

- iii) Hay dos tipos de movilidad que se pueden considerar en el marco de las redes complejas. Por un lado, puede haber diferentes tipos de caminantes moviéndose a lo largo de la red para realizar un gran número de posibles tareas. Modelos basados en agentes se han utilizado para estudiar fenómenos como la difusión de epidemias, la formación de opiniones o el tráfico en Internet. Por otro lado, en algún otro caso de interés, es la propia red a ser generada por las interacciones entre agentes móviles que constituyen los nodos del patrón de conectividad, así que la topología de la red evoluciona por efecto de la movilidad. Esta situación es común a muchos sistemas, tanto artificiales como naturales, y las aplicaciones van desde la robótica [27, 28] hasta la ecología [29].

Esta tesis se ha desarrollado siguiendo estas tres líneas, que están estrictamente relacionadas entre sí. Hemos profundizado tres casos de estudio, cada uno de los cuales se ocupa de dos de los macro-temas mencionados. Consideramos la cuestión del carácter incompleto de la información disponible tanto en el caso de redes naturales (Capítulo 2) como de redes artificiales (Capítulo 3). Nos centramos en la sincronización de los osciladores de fase acoplados (Capítulos 2 y 4) en cuanto comportamiento emergente paradigmático, investigando en profundidad cómo los diferentes patrones de conexión puedan afectar la consecución de un estado coherente a nivel global. Por último, incluimos agentes móviles en dos marcos diferentes, utilizándolos como exploradores de redes desconocidas (Capítulo 3) y considerándolos como unidades que interactúan y son capaces de establecer conexiones con sus vecinos (Capítulo 4).

Si la emergencia se refiere “al surgimiento de estructuras, patrones y propiedades nuevas y coherentes”, y “los fenómenos emergentes se conceptualizan como algo que tiene lugar en el nivel macro, en contraste con el nivel micro de los componentes y procesos de los cuales surgen” [30], la sincronización se puede considerar como el comportamiento emergente más interesante, abundante y bien definido. Reconocidos por primera vez en 1665 por Christiaan Huygens, los fenómenos de sincronización se encuentran en la naturaleza, en la ingeniería y en la vida social. “La variación sincrónica de los núcleos celulares, de las emisiones sincrónicas de las neuronas, el ajuste de la frecuencia cardíaca con los ritmos respiratorio y/o locomotor, las diferentes formas de comportamiento cooperativo de los insectos, animales e incluso humanos”, son fenómenos universales y pueden ser entendidos dentro de un mismo marco común basado en la moderna dinámica no lineal [31]. Nuestro entorno está lleno de objetos

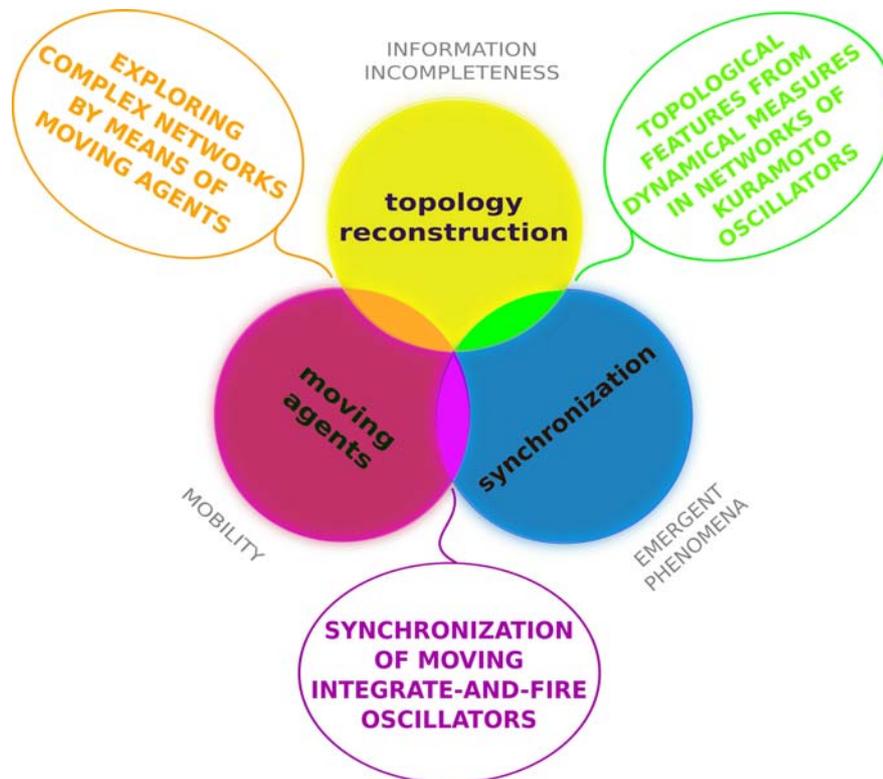


Figure i: Mi tesis en un diagrama.

oscilantes. Sistemas tan diversos como un público aplaudiendo y un equipo de radiocomunicación comparten la característica común de producir ritmos. Por lo general, estos sistemas no están cerrados e interactúan con otros objetos. “Esta interacción puede ser muy débil [...] pero sin embargo a menudo causa una transición cualitativa: un objeto ajusta su ritmo en conformidad con los ritmos de otros objetos. [...] Este ajuste de los ritmos debido a una interacción es la esencia de la sincronización” [31].

En esta tesis se han considerado osciladores acoplados con amplitudes fijas que interactúan a través de la adaptación mutua de sus fases. En el capítulo 2, se estudia el problema de la reconstrucción de una red de interacción desconocida, cuyos nodos son osciladores de Kuramoto. El modelo de Kuramoto, propuesto por primera vez por Yoshiki Kuramoto [32], ofrece un ejemplo paradigmático de las transiciones de no-equilibrio entre un estado incoherente y uno sincronizado. Su formulación fue motivada por el comportamiento de sistemas de osciladores químicos y biológicos, y ha encontrado aplicaciones diversas, por ejemplo en neurociencias. En el modelo se asumen varias hipótesis, entre las cuales que el acoplamiento sea débil, que los osciladores sean idénticos o casi idénticos, y que las interacciones dependan sinusoidalmente de la diferencia de fase entre cada par de objetos. En esta tesis se analizan poblaciones de osciladores casi idénticos en redes de interacción arbitrarias. Nuestro objetivo es extraer

características topológicas del patrón de conectividad mediante medidas puramente dinámicas, basándonos en el hecho de que en una red heterogénea la dinámica global no sólo está afectada por la distribución de las frecuencias naturales, sino también por la ubicación de los diferentes valores. Con el fin de realizar un estudio cuantitativo, nos centramos en una distribución de frecuencias muy simple, asumiendo que todas las frecuencias sean iguales excepto una, la del nodo que llamamos marcapasos. Luego, se analiza el comportamiento dinámico del sistema justo en el punto de transición y ligeramente por encima de él, así como muy lejos del punto crítico, cuando se encuentra en un estado altamente incoherente. La información topológica recogida varía desde características locales, como la conectividad individual de los nodos, a la estructura jerárquica de los grupos funcionales, e incluso a la matriz de adyacencia completa.

En el capítulo 4, en cambio, se presenta un modelo de osciladores *integrate-and-fire* (integra y dispara) que son agentes móviles, desplazándose libremente en un plano. La fase de los osciladores evoluciona linealmente en el tiempo y, cuando alcanza un valor umbral, estos disparan a sus vecinos eligiéndolos de acuerdo a un cierto rango de interacción (Sec. 4.1) o simplemente seleccionando lo que se encuentre más cerca (Sec. 4.2 y siguientes). De este modo, la red de interacción es un objeto dinámico por sí mismo, ya que se recrea en cada paso de tiempo en consecuencia del movimiento de las unidades. Dependiendo de la velocidad del movimiento, del número medio de vecinos, de la fuerza de acoplamiento y el tamaño de la población de agentes, identificamos diferentes regímenes. Caracterizamos estos regímenes en términos del tiempo que el sistema necesita para alcanzar el estado coherente, que en algunos casos somos capaces de predecir, y también proporcionamos una detallada aunque cualitativa descripción de los mecanismos que permiten aumentar el grado de sincronización del sistema.

Los agentes móviles se emplean también en el capítulo 3, donde juegan el papel de exploradores de redes artificiales desconocidas, cuya misión es recuperar información acerca de la estructura topológica. El problema de la reconstrucción de redes se suele abordar por medio de algoritmos basados en el enrutamiento de la búsqueda y del tráfico [33, 34, 35] realizados, en muchos casos, por agentes móviles que exploran el espacio topológico. Aquí proponemos un modelo en el que *random walkers* (caminantes aleatorios), con nodos de origen previamente asignados, navegan a través de la red durante un período de tiempo fijo. Consideramos que la exploración es exitosa si el caminante devuelve la información obtenida a su nodo de partida (*nodo-casa*), de lo contrario, no se recuperan los datos. En cada paso de tiempo, los caminantes tiene la opción, con cierta probabilidad, de ir hacia atrás acercándose a su casa o bien ir más lejos. Se demuestra que hay una solución óptima para este problema en términos de la información recuperada promedio y del grado de los nodos de origen y se diseña una estrategia adaptativa basada en el comportamiento del *random walker*. Por último, se comparan diferentes estrategias que surgen del modelo en el contexto de la reconstrucción de la red.

Resultados y conclusiones

Extrayendo características topológicas de medidas dinámicas en redes de osciladores de Kuramoto

Se ha demostrado recientemente que sistemas de osciladores de Kuramoto no idénticos pueden alcanzar un grado de sincronización que depende fuertemente de la topología de la red compleja subyacente. Aquí, estas propiedades dinámicas se han utilizado para recopilar información sobre los patrones de conectividad, en particular mediante el establecimiento de diferentes tipos de correlaciones entre la evolución dinámica de los osciladores. Es importante subrayar que este es el caso de mayoría de las situaciones experimentales, donde la conectividad *a priori* desconocida de una red concreta se infiere a partir de mediciones puramente dinámicas.

Cuando los osciladores son idénticos (todos ellos tienen la misma frecuencia natural) cualquier configuración topológica tiene un atractor único, que es el estado completamente sincronizado, lo que significa que los osciladores terminan teniendo todos la misma frecuencia eficaz y fases idénticas. Este estado no ofrece ninguna información sobre la topología. Nosotros perturbamos esta configuración, permitiendo que uno de los osciladores tenga una frecuencia natural diferente que el resto. Esta unidad se llama el marcapasos de la red. Tal perturbación provoca que el estado final deje de ser de la sincronización de fases. Pero si la frecuencia natural del marcapasos no es muy diferente del valor del resto de la población, el sistema todavía mantendrá un cierto grado de coherencia, es decir, el sistema en su conjunto puede evolucionar con la misma frecuencia efectiva. Sin embargo, si la diferencia se hace más grande, el sistema será incapaz de encontrar cualquier tipo de sincronización. El umbral entre el primer caso y el último es un valor bien definido, que es estrictamente dependiente de la ubicación del marcapasos en la red. En este contexto, podemos utilizar las correlaciones entre las frecuencias efectivas de los osciladores en el estado incoherente para reproducir la conectividad de la red. Además, se demuestra que las correlaciones dinámicas en diferentes situaciones, cerca o lejos del punto crítico, aportan información complementaria de la red.

1. Trabajando alrededor del punto crítico se puede estimar el grado de cada marcapasos utilizando sólo su frecuencia crítica.
2. Ligeramente por encima del punto de transición, la estructura jerárquica de la red (en relación a los módulos funcionales) puede obtenerse a partir de las correlaciones entre las frecuencias eficaces. Un refinamiento adicional nos permite recuperar toda la red de conectividad con buen grado de exactitud.
3. Muy por encima del punto crítico es posible reconocer cuáles de los osciladores están directamente conectados al marcapasos a partir de medidas muy cortas de la evolución temporal de las frecuencias eficaces. De esta manera podemos recuperar el patrón de conectividad, y este método resulta ser mucho más preciso y más eficiente que el anterior.

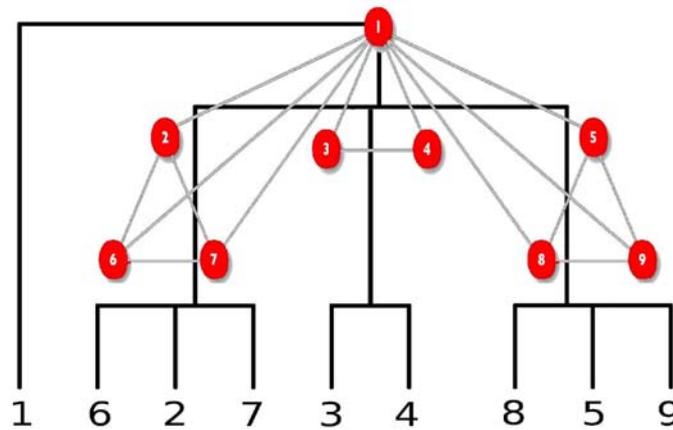


Figure ii: Ejemplo de un dendograma representante la estructura jerárquica de una red, obtenido mediante medidas de correlaciones dinámica justo por encima del punto crítico.

En resumen, hemos analizado diferentes maneras de poner en relación las propiedades dinámicas de los nodos individuales y la topología de la red. Las propiedades topológicas inferidas de la dinámica pueden ser locales (la existencia de un enlaces entre dos nodos), así como globales (la organización jerárquica de los nodos de la red funcional). En particular, para una red libre de escala, y si los grados de los nodos se conocen (o han sido estimados a partir de las frecuencias críticas), considerando el 30 % de los posibles marcapasos, si se seleccionan los nodos más conectados, será suficiente información como para reconstruir aproximadamente el 90 % de los enlaces.

Otros trabajos han considerado el problema de la reconstrucción de la red a partir de información dinámica. Al igual que en nuestra propuesta, pero con objetivos específicos, Tegner *et al.* [56] han analizado la respuesta dinámica de una red de regulación génica al variar de los niveles de expresión de determinados genes. Por otro lado, di Bernardo *et al.* [57] han considerado el efecto global de diferentes tipos de perturbaciones para inferir la topología de la red. Más recientemente, este enfoque ha sido seguido también por Gorur Shandilya y Timme [58]. Estos autores asumen que se disponga de cierta información acerca de la evolución dinámica de las unidades aisladas y sobre el acoplamiento. Nuestro método, basado en la modificación de la frecuencia de sola unidad mediante la cual se confiere contenido informativos a las correlaciones entre los nodos, puede ser más eficaz en los sistemas oscilatorios. En cualquier caso, para fines prácticos, el método elegido dependerá los detalles específicos de la configuración experimental e incluso una combinación de diferentes enfoques puede ser la opción más adecuada.

Explorando redes complejas mediante caminantes adaptativos

Hemos presentado un modelo para la búsqueda y exploración de redes en el que agentes móviles evalúan en cada paso de tiempo si deben ir más allá, alejándose su nodo-casa, o volver con la información obtenida hasta ese momento. Estas probabilidades dependen de un único parámetro q el cual, para tiempos de exploración comparables al tamaño del sistema, presenta un valor óptimo, $q = q_p < 1$ ($q = 1$ corresponde al límite de camino aleatorio markoviano). Por contrario, cuando a los caminantes se les permite explorar la red de forma indefinida o durante tiempos más largos, el valor óptimo resulta ser $q = \infty$. Sin embargo, aunque la cantidad de información recuperada adoptando la última opción podría ser máxima, los resultados son altamente dependientes del grado del nodo-casa: cuanto menor sea el grado del nodo asignado al caminante, menor será la información que ese podrá llevar a casa. De hecho, para la mayoría de los nodos (recordemos que en una red libre de escala la mayor parte de los nodos están muy mal conectados) $q = \infty$ no es la mejor estrategia.

Aprovechando el comportamiento de los caminantes en función de q , hemos propuesto un algoritmo alternativo en el que los agentes pueden ajustar el valor del parámetro q para optimizar la información recuperada. A través de simulaciones numéricas, hemos demostrado que este mecanismo permite una exploración tan eficiente como la que se obtiene con la configuración $q = q_p$. Sin embargo, el esquema adaptativo tiene la ventaja de que el valor de q se modifica de forma dinámica, por lo que supera el problema de tener que fijar un valor óptimo q_p *a priori* desconocido. Creemos que este protocolo adaptativo de búsqueda podría ser una adición valiosa a la literatura actual, ya que su rendimiento es óptimo con un mínimo (local) información sobre la estructura de la red. Como demostración de las potencialidades de los algoritmos analizados en este trabajo, hemos comparado diferentes estrategias de búsqueda y exploración de redes. Como se esperaba, el mecanismo adaptativo tiene el mejor rendimiento en términos de calidad y cantidad de la información recuperada. Establecer si este tipo de estrategias pueden ser ulteriormente desarrolladas y aplicadas a la exploración de redes reales sería ir más allá del objetivo del presente trabajo. De toda forma, podemos identificar por lo menos dos escenarios en los que nuestro algoritmo puede ser de utilidad: el descubrimiento de nuevas conexiones en las redes de comunicación y la exploración de redes planares (es decir, redes de ciudades), utilizando solo información local. Nótese que en la red que se ha utilizado como referencia, el nodo con el grado máximo está conectado con tan sólo el 1% del resto de los nodos, algo que se puede dar también en las redes planas, aunque estas no puedan ser libre de escala. Los resultados obtenidos son muy generales, siendo válidos siempre cuando el grado máximo k_{max} sea (al menos) un orden de magnitud mayor que $\langle k \rangle$.

Sincronización de osciladores móviles

Siguiendo la literatura reciente sobre los sistemas complejos, uno de los temas más candentes es la relación entre la dinámica y la topología de las interacciones. En particular, hay muchas

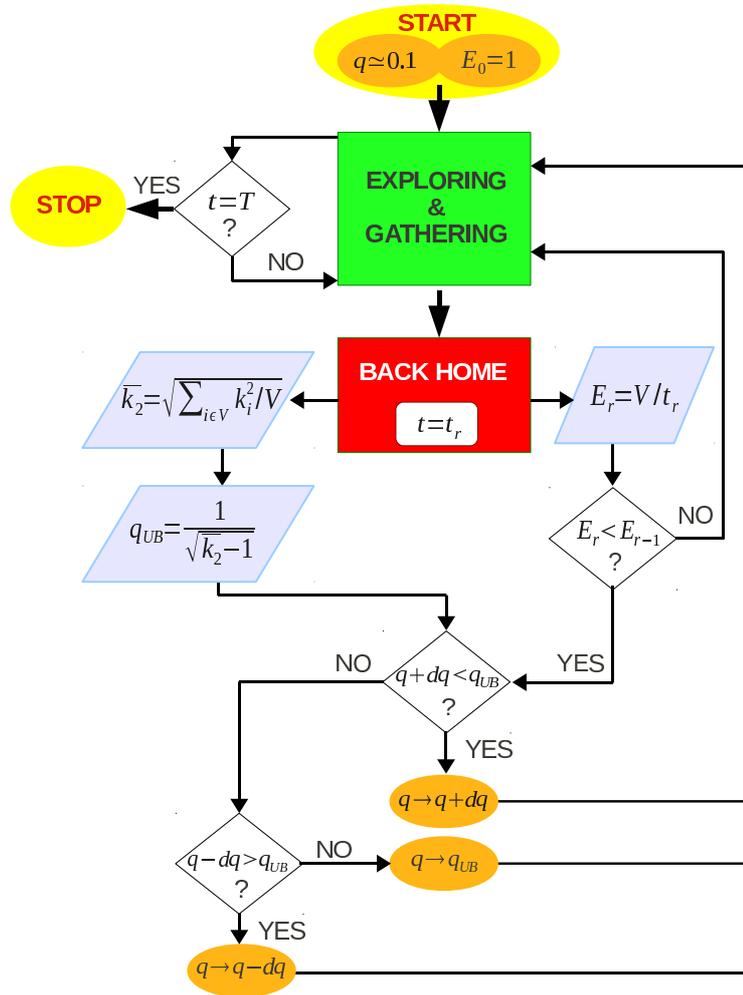


Figure iii: Diagrama de flujo del algoritmo adaptativo.

evidencias de que los patrones de interacción cambian rápidamente con el tiempo, alterando y condicionando por completo las propiedades dinámicas del sistema. Aquí hemos propuesto un marco en el que los agentes se mueven en un plano y se les permite interactuar de una manera pulsátil. Cada agente, que representa un oscilador de fase, se mueve a una velocidad común (en módulo) y cambia su fase interna con período común. Cuando esta fase interna alcanza valor umbral, el oscilador “hace fuego”, así reiniciando a cero su propia fase. Se han considerado dos modelos diferentes. En el primero, cada oscilador, al disparar, reinicia al azar la orientación de su movimiento e interactúa con los vecinos dentro de una cierta distancia, modificando sus fases. En el segundo, el oscilador que dispara modifica la fase y al mismo tiempo reinicia la orientación de su vecino más próximo, mientras que él continúa moviéndose en la misma dirección.

Para el primer modelo, manteniendo todos los parámetros geométricos constantes, se ha

analizado el comportamiento del sistema cambiando únicamente la velocidad de los agentes y el rango de interacción. Hemos medido el tiempo necesario para que el sistema sincronice en función de la velocidad y de la fracción de la población con la que los agentes interactúan. Hemos introducido un parámetro de orden nuevo, a lo que llamamos *mixing*, que representa la fracción de las diferentes unidades con las cuales cada oscilador ha interactuado. Este parámetro de orden permite introducir un diagrama de fase nuevo que nos permite identificar tres mecanismos de sincronización diferentes. i) El mecanismo difusivo: a altas velocidades y rango arbitrario, el sistema alcanza muy rápidamente la sincronización por medio de interacciones extremadamente efectivas de cada un oscilador con una gran fracción de la población.

ii) El mecanismo local: los agentes se mueven muy lentamente, pero el rango de interacción es bastante grande, así que la sincronización local es suficiente para llevar el sistema al estado globalmente sincronizado manteniendo limitado el valor del *mixing*.

iii) El mecanismo limitado que, siendo caracterizado por velocidades muy lentas y interacciones de corto alcance, implica tiempos de sincronización muy largos y, de consecuencia, un *mixing* mayor.

Como era de esperar, el costo mínimo de energía, que representa el número total de señales emitidas por la población para alcanzar el estado sincronizado, se consigue cuando los agentes se mueven muy rápidamente, independientemente del rango de interacción. Sin embargo, la sincronización en los sistemas con movilidades bajas si que depende, y dramáticamente, del rango de interacción.

Para el segundo modelo, se ha estudiado la dependencia del tiempo de sincronización de la velocidad, variando también el tamaño del sistema y la intensidad del acoplamiento. Dado que la regla de interacción es tan restrictiva que el sistema se encuentra muy por debajo del umbral de percolación estática, es la velocidad no nula de los osciladores la que permite que el sistema alcance el estado coherente. Por lo tanto, podríamos esperar que el tiempo de sincronización fuera una función decreciente de V , pero nos encontramos con que hay una región intermedia en la que el tiempo de sincronización diverge. Hemos podido caracterizar los mecanismos que permiten al sistema de sincronizar, tanto en el régimen lento como en el rápido. También hemos calculado, en función del tamaño del sistema N y la fuerza de acoplamiento ϵ , los valores de velocidad a los cuales el sistema entra en estas regiones. El mecanismo del límite de movimiento rápido de este modelo no difiere del mecanismo difusivo (i) del anterior, mientras que el régimen lento es similar al régimen limitado (iii) cuando la velocidad es extremadamente pequeña ya que, en el régimen lento de este modelo, el *mixing* final no es muy grande.

Ha sido también posible proporcionar un estimador empírico del tiempo de sincronización en función de N y ϵ en el límite rápido, y en función de V , N y ϵ en el régimen lento. Luego, hemos analizado cualitativamente las razones por las cuales el sistema no pueden sincronizar dentro de un tiempo finito en la región intermedia.

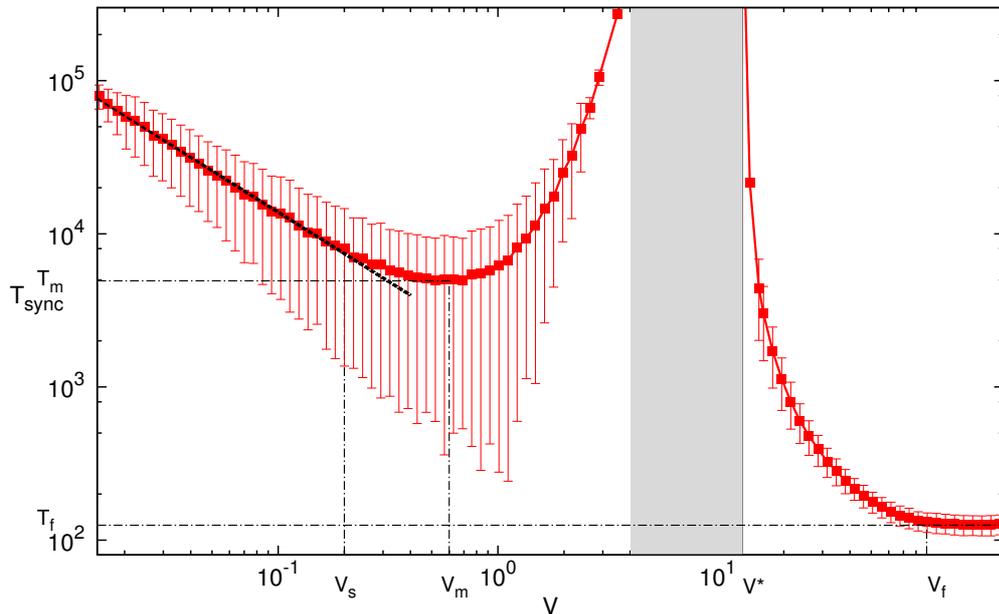


Figure iv: Tiempo de sincronización en función de la velocidad para el segundo modelo considerado (interacción a primeros vecinos).

Finalmente, se ha considerado que, midiendo el coste en el número de disparos necesarios para que el sistema sincronice, el límite rápido es siempre el régimen más eficiente. Por contrario, si se tiene en cuenta también la energía necesaria para hacer que los agentes se muevan a la velocidad deseada, es decir, el término cinético proporcional al cuadrado de la velocidad, esto ya no es cierto. El nuevo coste energético total muestra dos mínimos, uno de los cuales correspondiente al mínimo local del tiempo de sincronización en la región izquierda, el otro localizado en la región derecha, por debajo del límite rápido.

La identificación de los mecanismos que relacionan movilidad y interacción con la sincronización, será sin duda crucial en modelos similares de poblaciones de agentes móviles que pueden ser aplicados, por ejemplo, al campo de la comunicación inalámbrica. De hecho, la comprensión de estos fenómenos ayudará a diseñar protocolos óptimos para configuraciones más realistas mediante la definición de normas de interacción adecuadas.

Perspectivas

Durante el desarrollo del trabajo presentado en esta tesis, se han planteado una serie de posibles aplicaciones, preguntas adicionales, y otros temas relacionados que podrían ser investigado en el futuro.

- Puede valer la pena comprobar la robustez del método para la detección de las comunidades funcionales en redes de osciladores de Kuramoto presentado en la sección 2.4. Si se pudiera demostrar que, como es de esperar, los resultados no son sensibles a los

detalles de la forma funcional de las interacciones, entonces el método podría aplicarse a una gran cantidad de situaciones de interés, independientemente de la naturaleza específica de red considerada.

- Los caminantes introducidos en el capítulo 3 cambian su comportamiento de manera no trivial variando el parámetro que ajusta la probabilidad de ir hacia adelante o hacia atrás. En particular, existe un valor preciso del parámetro por el cual los agentes, en general, visitan muchos nodos y luego vuelven a su *casa*. Más allá del problema de la exploración de redes, estamos interesados en comprender las consecuencias más generales de este comportamiento peculiar. Por ejemplo, podríamos analizar la nueva fenomenología que pueda surgir sustituyendo los *random walks* habituales con nuestros caminantes en un modelo de meta-población para la propagación de epidemias en redes complejas. Podría ser también interesante considerar los caminantes como un osciladores de fase a los que se le permite interactuar solamente con otros osciladores en el mismo nodo, asumiendo que la población de cada nodo sea una comunidad inicialmente sincronizada. Podemos preguntarnos si el valor del parámetro que es óptimo para la exploración es también capaz de ayudar a la consecución de la coherencia global en la red.
- Por último, se necesita más investigación para entender mejor algunas de las características del modelo mínimo introducido en la sección 4.2. Puede merecer la pena proseguir la investigación al fin de identificar las condiciones necesarias y suficientes para que la dependencia del tiempo de sincronización de la velocidad deje de ser monótona. Actualmente, estamos investigando en esta dirección y ya está demostrado que la interacción tiene que ser pulsátil. En efecto, al sustituir los osciladores *integrate-and-fire* con osciladores de Kuramoto, el tiempo de sincronización se comporta como una función monótonamente decreciente de la velocidad, para cualquier valor de la fuerza de acoplamiento.

Bibliography

- [1] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente. Synchronization processes in complex networks. *Physica D*, 224:27–34, 2006.
- [2] S. Fortunato and M. Barthelemy. Resolution limit in community detection. *Proc. Natl. Acad. Sci. USA*, 104(1):36–41, 2007.
- [3] E. Ravasz and A.-L. Barabási. Hierarchical organization in complex networks. *Phys. Rev. E*, 67(2):026112, Feb 2003.
- [4] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA*, 99:7821, 2002.
- [5] J. W. Scannell, G. A. P. C. Burns, C. C. Hilgetag, M. A. O’Neil, and M. P. Young. The connectional organization of the cortico-thalamic system of the cat. *Cereb. Cortex*, 9:277–299, 1999.
- [6] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente. Synchronization Reveals Topological Scales in Complex Networks. *Phys. Rev. Lett.*, 96:114102, 2006.
- [7] G. Parisi. Complex systems: a physicist’s viewpoint. *Physica A*, 263:557–64, 1999.
- [8] M. Cini. *Un paradiso perduto. Dall’universo delle leggi naturali al mondo dei processi evolutivi*. Feltrinelli Editore, Milan, Italy, 1994.
- [9] P.W. Anderson. More is different. *Science*, 177:393–396, 1972.
- [10] G. Nicolis and C. Rouvas-Nicolis. Complex systems. *Scholarpedia*, 11:1473, 2007.
- [11] M. San Miguel, J.H. Johnson, J. Kertesz, K. Kaski, A. Díaz-Guilera, R.S. MacKay, P. Loreto, V. an Erdi, and D. Helbing. Challenges in complex systems science. *ArXiv Nonlinear Sciences e-prints*, April 2012.
- [12] B. Goodwin. *How the Leopard changed Its Spots: The Evolution of Complexity*. Princeton University Press, Princeton, New Jersey, 2001.

-
- [13] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature (London)*, 393:440, 1998.
- [14] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74(1):47–97, Jan 2002.
- [15] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev.*, 45:167–256, 2003.
- [16] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Phys. Rep.*, 424:175–308, 2006.
- [17] G. Caldarelli. *Scale-Free Networks*. Oxford University Press, Oxford, UK, 2007.
- [18] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Critical phenomena in complex networks. *Rev. Mod. Phys.*, 80:1275–1336, 2008.
- [19] G. Sporns. *Networks of the Brain*. MIT Press, USA, 2010.
- [20] S. Boccaletti, V. Latora, and Y. Moreno. *Handbook on Biological Networks*. World Scientific Lectures Notes in Complex Systems, Singapore, 10.
- [21] R. Pastor-Satorras and A. Vespignani. *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge University Press, Cambridge, UK, 2004.
- [22] M. Gjoka, M. Kurant, C.T. Butts, and A. Markopoulou. A walk in facebook: Uniform sampling of users in online social networks. *Proceedings of IEEE INFOCOM 2010*, 1, 2010.
- [23] R. Guimerà and M. Sales-Pardo. Missing and spurious interactions and the reconstruction of complex networks. *Proc. Nat. Acad. Sci. USA*, 106:22073–22078, 2009.
- [24] A. Barrat, M. Barthelemy, and A. Vespignani. *Dynamical processes on complex networks*. Cambridge University Press, 2008.
- [25] S.A. Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.*, 22:437–467, 1969.
- [26] M. D.F. Shirley and S. P. Rushton. The impacts of network topology on disease spread. *Ecological Complexity*, 2(3):287–299, 2005.
- [27] A. Buscarino, L. Fortuna, M. Frasca, and A. Rizzo. Dynamical network interactions in distributed control of robots. *Chaos*, 16:015116, 2006.
- [28] H. G. Tanner, A. Jadbabaie, and G. J. Pappas. Stable flocking of mobile agents part i: dynamic topology. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 2, pages 2016–2021 Vol.2, 2003.

- [29] J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson. From Disorder to Order in Marching Locusts. *Science*, 312:1402–1406, 2006.
- [30] J. Goldstein. Emergence as a construct: History and issue. *Emergence: Complexity and Organization*, 1:49–72, 1999.
- [31] A. Pikovsky, M. Rosenblum, and J. Kurths. *Synchronization*. Cambridge University Press, Cambridge, UK, 2001.
- [32] Y. Kuramoto. Self-entrainment of a population of coupled nonlinear oscillators. In H. Araki, editor, *International Symposium on Mathematical Problems in Theoretical Physics, Lecture Notes in Physics, Vol. 39*,, pages 420–422, New York, NY, USA, 1975. Springer.
- [33] R. Guimera, A. Diaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas. Optimal network topologies for local search with congestion. *Phys. Rev. Lett.*, 89:248701, 2002.
- [34] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno. Dynamics of jamming transitions in complex networks. *Europhys. Lett.*, 71:325, 2007.
- [35] S. Sreenivasan, R. Cohen, E. Lopez, Z. Toroczkai, and H. E. Stanley. Graph partitioning induced phase transitions. *Phys. Rev. E*, 75:036105, 2007.
- [36] G. V. Osipov, J. Kurths, and C. Zhou. *Synchronization in oscillatory networks*. Springer, Berlin, Germany, 2007.
- [37] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. Synchronization in complex networks. *Phys. Rep.*, 469:93–153, 2008.
- [38] Arthur T. Winfree. *The Geometry of Biological Time*. Springer-Verlag, Berlin, Germany, 1980.
- [39] J. A. Acebrón, L. L. Bonilla, C. J. Pérez-Vicente, F. Ritort, and R. Spigler. The Kuramoto model: A simple paradigm for synchronization phenomena. *Rev. Mod. Phys.*, 77:137–185, 2005.
- [40] A. Díaz-Guilera and A. Arenas. Phase patterns of coupled oscillators with application to wireless communication. *Lect. Not. Comp. Sci.*, pages 184–191, 2008.
- [41] L. Buzna, S. Lozano, and A. Díaz-Guilera. Synchronization in symmetric bipolar population networks. *Phys. Rev. E*, 80:066120, 2009.
- [42] H. Kori and A. S. Mikhailov. Entrainment of Randomly Coupled Oscillator Networks by a Pacemaker. *Phys. Rev. Lett.*, 93:254101, 2004.

- [43] F. Radicchi and H. Meyer-Ortmanns. Entrainment of coupled oscillators on regular networks by pacemakers. *Phys. Rev. E*, 73:036218, 2006.
- [44] F. Mori. Necessary condition for frequency synchronization in network structures. *PRL*, 104:108701, 2004.
- [45] Y. Kuramoto. *Chemical oscillations, waves, and turbulence*. Springer-Verlag, New York, NY, USA, 1984.
- [46] I. Sendiña-Nadal, J.M. Buldu, I. Leyva, and S. Boccaletti. Phase locking induces scale-free topologies in networks of coupled oscillators. *PLoS ONE*, 3(7):e2644, 2008.
- [47] P. M. Gleiser and D. H. Zanette. Synchronization and structure in an adaptive oscillator network. *Europ. Phys. J. B*, 53:233–238, 2006.
- [48] A. Arenas, A. Fernandez, and S. Gomez. An optimization approach to the structure of the neuronal layout of *C. elegans*. *HandBOOK on Biological Networks*, 10:243–257, 2009.
- [49] P. H. A. Sneath and R. R. Sokal. The principles and practice of numerical classification. *Numerical Taxonomy*, 1973.
- [50] L. Zemanová, C. Zhou, and J. Kurths. Structural and functional clusters of complex brain networks. *Physica D*, 224:202–212, 2006.
- [51] C. Zhou, L. Zemanová, G. Zamora, C. C. Hilgetag, and J. Kurths. Hierarchical organization unveiled by functional connectivity in complex brain networks. *Phys. Rev. Lett.*, 97:238103, 2006.
- [52] J. L. Rodgers and W. A. Nicewander. Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42(1):59–66, Feb 1988.
- [53] R. Sharan, I. Ulitsky, and R. Shamir. Network-based prediction of protein function. *Molecular Systems Biology*, 3, 2007.
- [54] S. C. Janga, Javier J. Díaz-Mejía, and G. Moreno-Hagelsieb. Network-based function prediction and interactomics: The case for metabolic enzymes. *Metabolic Engineering*, 13:1–10, 2011.
- [55] M. Timme. Revealing network connectivity for response dynamics. *Phys Rev Lett.*, 98, 2007.
- [56] J. Tegner, M. K. S. Yeung, J. Hasty, and J.J. Collins. Reverse engineering gene networks: Integrating genetic perturbations with dynamical modeling. *Proc. Natl. Acad. Sci. USA*, 100:5944–5949, 2003.

- [57] D. di Bernardo, M.J. Thompson, T.S. Gardner, S.E. Chobot, E.L. Eastwood, A.P. Wojtovich, S.J. Elliott, S.E. Schaus, and J.J. Collins. Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks. *Nat Biotech*, 23(3):377–383, 2005.
- [58] S. Gorur Shandilya and M. Timme. Inferring network topology from complex dynamics. *New Journal of Physics*, 13(1):013004, 2011.
- [59] S.-J. Yang. Exploring complex networks by walking on them. *Phys. Rev. E*, 71:016107, Jan 2005.
- [60] L. da Fontoura Costa and G. Travieso. Exploring complex networks through random walks. *Phys. Rev. E*, 75:016102, Jan 2007.
- [61] J. I. Perotti and O. V. Billoni. Smart random walk: the cost of knowing the path. *ArXiv Condensed Matter e-prints*, February 2012.
- [62] K. Avrachenkov, B. Ribeiro, and D. Towsley. Improving random walk estimation accuracy with uniform restarts. Rapport de recherche RR-7394, INRIA, September 2010.
- [63] M.P. Viana, J.L.B. Batista, and L.D.F. Costa. How many nodes are effectively accessed in complex networks? *CoRR*, abs/1101.5379, 2011.
- [64] M. Catanzaro, M. Boguñá, and R. Pastor-Satorras. Generation of uncorrelated random scale-free networks. *Phys. Rev. E*, 71:027103, Feb 2005.
- [65] J. J. Binney, A. J. Fisher, N. J. Dowrick, and M. E. J. Newman. *The Theory of Critical Phenomena*. Oxford University Press, Oxford, UK, 1992.
- [66] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22(1):79–86, 1951.
- [67] R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *P. IEEE*, 95(1):215–233, 2007.
- [68] J.-P. Onnela, J. Saramaki, J. Hyvonen, G. Szabo, D. Lazer, K. Kaski, J. Kertesz, and A. L. Barabasi. Structure and tie strengths in mobile communication networks. *Proc. Natl. Acad. Sci. USA*, 104:7332–7336, 2007.
- [69] M. Valencia, J. Martinerie, S. Dupont, and M. Chavez. Dynamic small-world behavior in functional brain networks unveiled by an event-related networks approach. *Phys. Rev. E*, 77:050905(R), 2008.
- [70] D. Tanaka. General chemotactic model of oscillators. *Phys. Rev. Lett.*, 99:134103, 2007.

- [71] K. Römer. Time synchronization in ad hoc networks. *Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing*, pages 173–182, 2001.
- [72] F. Sivrikaya and B. Yener. Time synchronization in sensor networks: a survey. *Network, IEEE*, 18:45–50, 2004.
- [73] K. Uriu, Y. Morishita, and I. Yoh. Random cell movement promotes synchronization of the segmentation clock. *Proc. Natl. Acad. Sci. USA*, 107(11):4979–4984, March 2010.
- [74] I. V. Belykh, V. N. Belykh, and M. Hasler. Blinking model and synchronization in small-world networks with a time-varying coupling. *Physica D*, 195:188–206, 2004.
- [75] M. Frasca, A. Buscarino, A. Rizzo, L. Fortuna, and S. Boccaletti. Synchronization of moving chaotic agents. *Phys. Rev. Lett.*, 100:044102, 2008.
- [76] M. Porfiri, D. J. Stilwell, E. M. Bollt, and J. D. Skufca. Random talk: Random walk and synchronizability in a moving neighborhood network. *Physica D: Nonlinear Phenomena*, 224:102 – 113, 2006.
- [77] D. J. Stilwell, E. M. Bollt, and D. Gray Roberson. Sufficient conditions for fast switching synchronization in time-varying network topologies. *SIAM J. App. Dyn. Syst.*, 5:140–156, 2006.
- [78] N. Fujiwara, K. Kurths, and A. Díaz-Guilera. Spectral analysis of synchronization in mobile networks. *AIP Conference Proceedings*, 1389(1):1015–1018, 2011.
- [79] N. Fujiwara, J. Kurths, and A. Díaz-Guilera. Spectral analysis of synchronization in mobile networks. *AIP Conference Proceedings*, 1389(1):1015–1018, 2011.
- [80] P. Erola, A. Díaz-Guilera, S. Gomez, and A. Arenas. Trade synchronization in the world trade web. *International Journal of Complex Systems in Science*, 1:1–4, 2011.
- [81] R. Mirollo, E. Strogatz, and H. Steven. Synchronization of Pulse-Coupled biological oscillators. *SIAM J. Appl. Math.*, 50:1645–1662, 1990.
- [82] G. M. Ramirez-Avila, J.-L. Deneubourg, J.-L. Guisset, N. Wessel, and J. Kurths. Firefly courtship as the basis of the synchronization-response principle. *Euro. Phys. J. B*, 94(6):60007, June 2011.
- [83] A. Baronchelli and A. Díaz-Guilera. Consensus in networks of mobile communicating agents. *Phys. Rev. E*, 85:016113, 2012.
- [84] J. Dall and M. Christensen. Random geometric graphs. *Phys. Rev. E*, 66:016121, 2002.

-
- [85] P. Balister, B. Bollobás, and M. Walters. Continuum percolation with steps in the square or the disc. *Random Struct. Algor.*, 26:392–403, 2005.
- [86] D. Eppstein, M. S. Paterson, and F. Yao. On nearest-neighbor graphs. *Discrete and Computational Geometry*, 17:263–282, 1997.
- [87] L. Prignano and A. Díaz-Guilera. Extracting topological features from dynamical measures in networks of kuramoto oscillators. *Phys. Rev. E*, 85:036112, 2012.
- [88] G. Balazsi, A.-L. Barabasi, and Z. N. Oltvai. Topological units of environmental signal processing in the transcriptional regulatory network of escherichia coli. *Proc. Nat. Acad. Sci. USA*, 102:7841–7846, 2005.

Publications list

Publications within the scope of this thesis

- L. P. and A. Díaz-Guilera,
Extracting topological features from dynamical measures in networks of Kuramoto oscillators.
Physical Review B, 85, (2012) 3: 036112.
- P. M. Gleiser, L. P., C. J. Pérez-Vicente and A. Díaz-Guilera,
Pacemakers in a Cayley tree of Kuramoto oscillators,
Int. J. of Bifurcation and Chaos, (2011) *in press*.
- L. P., O. Sagarra, P. M. Gleiser and A. Díaz-Guilera,
Synchronization of moving integrate and fire oscillators,
Int. J. of Bifurcation and Chaos, (2011) *in press*.
- L. P., Y. Moreno and A. Díaz-Guilera,
Exploring complex networks by means of adaptive walkers.
arXiv : 1203.1439v1 (2012) *submitted*.
- L. P., O. Sagarra, and A. Díaz-Guilera,
Synchronization of moving oscillators with a minimal interaction rule.
(2012) *in preparation*.

Other publications

- L. Angelani, C. Conti, L. P., G. Ruocco and F. Zamponi,
Mode-locking transitions in nanostructured weakly disordered lasers,
Physical Review B, 76, (2007) 6: 064202.
- L. P. and M. Serva,
Genealogical trees from genetic distances.
European Physical Journal, B 69, (2009), 455463.

-
- F. Petroni, L. P. and M. Serva,
Family trees: languages and genetics. In "Inhomogeneous Random Systems - Systèmes Aleatoires Inhomogenes", (Paris 2008).
Markov Processes and Related Fields 15, (2009), 417-440.
 - L. P., C. J. Pérez-Vicente and A. Díaz-Guilera,
Synchronization time of an all-to-all system of integrate and fire oscillators.
(2012) *in preparation*.