CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL
Unidad Mérida

DEPARTAMENTO DE FÍSICA APLICADA

“Estructura y dinámica de redes complejas: aplicación al crecimiento de redes y extinción de enfermedades”

TESIS

Que presenta

Ezequiel Arceo May

Para obtener el grado de

Doctor en Ciencias

en la especialidad de

Física Aplicada

Director(es) de Tesis:

Dr. Rodrigo Huerta Quintanilla
Dr. Cristian Fernando Moukarzel Rodríguez

Mérida, Yucatán, México Agosto de 2017

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL
Unidad Mérida

DEPARTAMENTO DE FÍSICA APLICADA

“Structure and Dynamics of complex networks: application to network growth and disease extinction”

TESIS

Que presenta

Ezequiel Arceo May

Para obtener el grado de

Doctor en Ciencias

en la especialidad de

Física Aplicada

Director(es) de Tesis:

Dr. Rodrigo Huerta Quintanilla
Dr. Cristian Fernando Moukarzel Rodríguez

Mérida, Yucatán, México Agosto de 2017

Structure and Dynamics of complex networks: application to network growth and disease extinction

M.C. Ezequiel Arceo May

Thesis advisors
Dr. Rodrigo Huerta-Quintanilla
Dr. Cristian F. Moukarzel R.

August 10, 2017

Agradecimientos

Las últimas líneas de este proyecto son las más importantes, un reconocimiento a todas las personas sin las cuales alcanzar esta meta habría sido imposible. Extiendo mi agradecimiento:

- A mi gran familia, por todo su apoyo y comprensión.
- Al CINVESTAV IPN Unidad Mérida, como institución, y a todo su personal humano.
- Al CONACYT, por la beca otorgada a lo largo del programa de doctorado.
- A Zhirnay Rodríguez Pinelo, por toda la ayuda brindada.

A mis asesores, Dr. Rodrigo Huerta-Quintanilla y Dr. Cristian Moukarzel gracias por depositar en mí su confianza, tener paciencia y darme la oportunidad de desarrollar este trabajo.

A todos ustedes gracias por su confianza, amistad, cariño, crítica, apoyo y consejo.

Contents

Agradecimientos 3

Contents 5

List of Figures 9

List of Tables 17

Resumen 19

Abstract 21

Introduction 23

Organization 25

I Network Structure 27

1 Network Theory basics 29
1.1 Network measures . . . 29
1.1.1 Degree distribution . . . 30
1.1.2 Mixing patterns and degree correlations . . . 31
1.1.3 Geodesic distances and paths . . . 32
1.1.4 Clustering coefficient . . . 33
1.1.5 Components . . . 34
1.2 Types of networks . . . 35
1.2.1 Random Networks . . . 35
1.2.2 Generalized random networks . . . 39
1.2.3 Small-world networks . . . 39
1.2.4 Scale-Free networks . . . 41

2 Growing networks 45
2.1 Problem definition . . . 45
2.2 Network growth modelation . . . 47
2.2.1 Degree distribution tail behavior . . . 47

2.2.2 Erdös-Rényi Networks . . . 48
2.2.3 Exponential Networks . . . 49
2.2.4 Scale-Free Networks . . . 52
2.2.5 Scale-Free – Exponential Roll-Off . . . 54
2.3 Results . . . 56
2.3.1 Simulation algorithm . . . 56
2.4 Conclusions and Discussion . . . 58

II Network Dynamics 61

3 Dynamical aspects 63
3.1 Master equation . . . 63
3.2 Network embedding . . . 64

4 Network Resilience 67
4.1 Percolation . . . 67
4.1.1 Percolation in arbitrary topologies . . . 68
4.1.2 Average component size . . . 69
4.2 Percolation in Complex Networks . . . 70
4.2.1 Failure Tolerance . . . 71
4.2.2 Targeted removal tolerance . . . 72

5 Synchronization 75
5.1 General Framework . . . 76
5.2 Identical and linearly coupled oscillators . . . 77
5.2.1 Structurally enhanced synchronization . . . 78
5.3 Integrate-and-fire oscillators . . . 80
5.4 The Kuramoto Model . . . 81
5.5 Kuramoto Model with long range interactions . . . 83
5.6 The Kuramoto model on complex networks . . . 84
5.6.1 Small-world effects . . . 84
5.6.2 Degree distribution effects . . . 85
5.6.3 Modularity effects . . . 86

6 Epidemiological models . . . 89
6.1 Deterministic homogeneous models . . . 89
6.1.1 SI model . . . 91
6.1.2 SIS model . . . 92
6.1.3 SIR model . . . 93
6.1.4 SIRS model . . . 94
6.1.5 SIRS model with fixed refractory time . . . 95
6.2 Epidemics in homogeneous networks . . . 96
6.2.1 Net infection rate . . . 96
6.2.2 Epidemic outbreak size . . . 97
6.3 Epidemics on networks and percolation . . . 97

6.4 Epidemics in heterogeneous uncorrelated ... 98
6.4.1 SI Model 98
6.4.2 SIS Model 100
6.4.3 SIR Model 103
6.4.4 SIRS model in Small-World networks 105
6.5 Epidemics in heterogeneous correlated ... 109
6.5.1 Pair (State) correlations 109

7 Excitable media 115
7.1 Characteristics of excitable media 115
7.2 Examples of excitable media 116
7.2.1 Calcium waves in the cell 116
7.2.2 Waves in the Belousov–Zhabotinsky reaction 117
7.2.3 Waves of electrical activity in the heart 118
7.2.4 Waves of spreading depression and persistent activity in the brain 118
7.2.5 Waves of infection in a population 119
7.2.6 Waves in Lotka-Volerra model 120
7.3 The Hodgkin-Huxley model 121
7.3.1 Integrate-And-Fire neurons in small-world networks 123
7.4 The FitzHugh-Nagumo model 125
7.5 The Chialvo Model 127
7.5.1 Chialvo model in small-world networks 129
7.6 Effects of complex topologies 130

8 SIRS model with long-range interactions 133
8.1 Model definition 135
8.1.1 Building networks with power law length distribution 135
8.1.2 SIRS model specification 135
8.1.3 Definitions and notation 137
8.1.4 Acceleration of the dynamics 138
8.1.5 Normal distribution over a circle 138
8.2 Dynamic Analogies 139
8.2.1 SIRS nodes as oscillators 139
8.2.2 SIRS and SIR model with removal and renewal as models for recurrent epidemics 140
8.3 Analysis of simulation results 140
8.3.1 Types of extinctions 140
8.3.2 Phase dispersion distribution at extinction time 143
8.3.3 Dynamic behavior before extinction 147
8.3.3.1 Densities in phase space 147
8.3.3.2 Coherence and phase variance 149
8.3.4 Characteristics of Stage 2 152
8.3.4.1 Coherence and phase variance 153
8.3.5 Modeling the distribution of extinction times 154

8.3.6 Mean lifetime scaling with system size $N = L^d$ 158
8.3.7 Mean lifetime collapse 162
8.4 Conclusions and Discussion 164

Appendices **167**

A Generating power-law distributed radius **169**

B Average link-length **171**

C Directories and datafiles **173**

Bibliography **175**

List of Figures

1.1 Graph with 8 nodes and 10 links. The set of nodes is $V = \{$ 1, 2, 3, 4, 5, 6, 7, 8 $\}$. The set of links is $E = \{$ (1,1), (1,2), (1,3), (2,3), (3,4), (4,5), (4,6), (6,8), (5,8), (7,6) $\}$. Among E $(7,6)$ is a directed link, $(1,1)$ is a reflexive link or loop, and all other links are non-directed. 30

1.2 Mixing patterns in networks. An assortative network presents a increasing behavior of $k_{nn}(k)$. An disassortative network presents a decreasing behavior of $k_{nn}(k)$. Uncorrelated networks presents a constant $k_{nn}(k) = \kappa$. 32

1.3 Graphic definition of clustering coefficient. In the image $v(i) = \{1, 2, 3\}$. The possible links among the nodes is $v(i)$ are $\{$ (1, 2), (2, 3), (1, 3) $\}$, just $(1, 2)$ and $(2, 3)$ (marked with dashed lines). In this example $C_i = 2/3$. 33

1.4 Examples of network components; (Upper row, from left to right) cycles of order 3, 5 and 7. (Central row, from left to right) cliques of order 5, 6 and 7. (Bottom row, from left to right) trees of order 10, 30 and 50. 34

1.5 Degree distribution of a Random network $G_{1000,0.05}$. 36

1.6 On the left, a Random network $G_{100,0.005}$. On the right a Random network $G_{100,0.015}$. The presence of a giant component is evident for the $G_{100,0.015}$, because $p = 1.5N^{-1} > N^{-1}$. 37

1.7 Phase transition in the size of the giant component χ_0 in the model $G_{N,p}$ with $N = 1000$ and varying p. For convenience the horizontal axis is the average degree. The transition point at $\langle k \rangle_c = 1$ is rounded due to finite size effects. 38

1.8 Evolution of the average path length l in the model $G_{N,p}$ with $N = 1000$ and varying p. Observe the discontinuity at the critical average degree $\langle k \rangle_c = 1$. 38

1.9 Evolution of a regular network with 20 nodes and average degree 4 to a small world network with $p = 0.7$. 40

1.10 Relative clustering coefficient and relative average path length as a function of p for a small world network with 1000 nodes and average degree 10. Values relative to that at $p = 0$. 40

1.11 Degree distribution of a Scale-Free network from the Barabasi-Albert model with $N = 1000$ and $m = 1$. 42

1.12 Construction of a deterministic scale-free network. The number n indicates the iteration step, N is the number of nodes of the network. In this iterative process N increases with n. Taken from [1, chapter 4] 43

2.1 **Comparative plot of degree distribution parameters in Exponential Networks.** A) Open circles: calculated vs. fitted values for the parameter k_o in simulated Exponential Networks. The line corresponds to the identity $k_o = \hat{k}_o$; B) Open circles: numerical values of average initial degree $m\,a$ vs. $k_{min} - \Delta$, the line corresponds to the identity $m\,a = k_{min} - \Delta$ line. 50

2.2 **Parameter relation in Scale-Free Networks.** A) Open circles: $\frac{(\hat{\gamma}-2)\langle k\rangle - \hat{\kappa}}{\hat{\gamma}-1}$ vs $m\,a$ for simulated Scale-Free Networks. Line: $\frac{(\gamma-2)\langle k\rangle - \kappa}{\gamma-1} = m\,a - \frac{1}{2}$ line; B) individual behavior of parameters $\hat{\gamma}$ and $\hat{\kappa}$ with average initial degree. 53

2.3 $k_{min} - \left(\frac{\gamma-2}{\gamma-1}\right)\Delta$ **for Exponential and Scale-Free** ($b \neq 0.00$) **Networks.** We used equations (2.10), (2.25), (2.34) and degree distribution data from simulated networks with different parameters $(a, b, m = 5)$ to appreciate the influence of k_{min}. This expression is almost constant over a wide range of k_{min} while shows a maximum as $a \to 0$ and $b \to 1$. This maximum is emphasized in Scale-Free Networks ($b > 0$) with $a \simeq 0$. $k_{min} - \left(\frac{\gamma-2}{\gamma-1}\right)\Delta$ fluctuates about $\langle k_{ini}\rangle = ma$ in Exponential Networks ($b = 0.00$) as well as in Scale-Free Networks when $a \simeq 1$. 55

2.4 **Parameters relation in the exponential approximation for the degree distribution of Scale-Free networks with** $b \simeq 0$**.** According to equation (2.37) as $b \to 0$ the ratio κ/γ converges to k_o for every value of a, indicating a Scale-Free – Exponential roll-off. 56

4.1 Graphical representation of a percolation process. Open circles are absent nodes, and gray links are removed links associated to absent nodes. 68

4.2 Networks under error and targeted removal of nodes. The parameter registered are the size of the giant component S (open symbols) and the average cluster size (filled symbols) as a function of the removed fraction f. In the left side we observe the case of a random network with $N = 10,000$ and $\langle k\rangle = 4$. In the right side the response of a scale-free network with $N = 10000$ and $\langle k\rangle = 4$. The value of γ is not available. Taken from [2]. 71

4.3 Percolation threshold vs. γ for scale-free networks in the continuous approximation. The inset show the same data, from an exact discrete approach. Taken from [3, section 6.5]. 73

5.1 Sketch of a typical master stability equation. The synchronized state is stable if for all λ_i, $\sigma\lambda_i \in [\alpha_1, \alpha_2]$. Taken from [3, section 7.2]. 78

5.2 Evolution of the eigenratio λ_{max}/λ_2 as a function of the connectance D in small-world networks with $N = 100$ and several values of m. Filled circles correspond to deterministic rings, the dashed line corresponds to random networks. Taken from [3, section 7.2]. . . 79

5.3 Phase dynamics of an integrate and fire oscillator. The situation corresponds to the response of an oscillator j when another oscillator i fires. In the left panel the coupling is excitatory $\epsilon_{ij} > 0$. In the right panel the coupling is inhibitory $\epsilon_{ij} < 0$. Taken from [3, section 7.3]. 80

5.4 Mean-field parameters in the Kuramoto model: Geometric representation of oscillators dynamics, the arrow is $re^{i\Theta}$ (Left). Qualitative evolution of coherence $z(t)$ as a function of the coupling strength K (Center). Maximum coherence as a function of the coupling strength K for a Cauchy distribution of oscillator frequencies (Right). 82

5.5 Critical coupling strength K_c (left panel) and synchronization time T (right panel) as a function of the rewiring probability p. Adapted from [4]. 85

5.6 Kuramoto oscillators in a modular network G. Network G has $N = 256$ nodes, grouped in 16 compartments for the first level of community. The second level of community groups the first level in four compartments. The internal degree of nodes at the first level is k_{in_1}. The internal degree of nodes at the second level is k_{in_2}. The number of links with any community of the network is k_{out}. For the present data, total degree $k = k_{in_1} + k_{in_2} + k_{out}$ is fixed to 18. Networks are labeled as $k_{in_1} - k_{in_2}$. Time evolution of the number of disconnected synchronized components of the dynamic connectivity matrix (top). Rank index i versus the inverse of the correponding – ascending ordered – eigenvalue in the Laplacian spectrum of the dynamic connectivity matrix (bottom). Taken from [5]. 87

6.1 Schematic representation of epidemiological models. SI model (top left), SIS model (bottom left), SIR model (top right) and SIRS model (bottom right). These models vary in the number of allowed states from $\{S, I, R\}$, and the number of (constant rate) transitions. β is a binary transition rate, while γ and δ are unary transition rates. 91

6.2 Time evolution of a SIR model with parameters $\beta = 1, \gamma = 0.4, s_0 = 0.99, i_0 = 0.01$ and $r_0 = 0$. Taken from [6, section 17.3]. 94

6.3 Average degree of the newly infected nodes for SI model as a function of re-scaled time by t/τ. Reference lines are drawn at κ for $t/\tau \ll 1$ and m for $t/\tau \gg 1$. Simulation where performed on scale-free networks in the Barábasi-Albert model with $N = 10^4$ and $m = 4(\text{bottom}), 14(\text{top})$. Taken from [7]. 100
6.4 Time evolution of the infection fraction $n_{inf}(t)$ for different values of disorder parameter (shown in legends). Simulations were performed with parameters $\tau_I = 4, \tau_R = 9, K = 3, N = 10^4$ and $n_{inf}(t) = 0.1$. Taken from [8]. 107
6.5 Order parameter as a function of the disorder for different network sizes (shown in legends). Each point is a time average over 2000 time steps. Simulations were performed with parameters $\tau_I = 4, \tau_R = 9, K = 3$ and $n_{inf}(t) = 0.1$. Taken from [8]. 108
6.6 Order parameter as a function of the disorder parameter for different average degrees (shown in legends). Simulations were performed with parameters $\tau_I = 4, \tau_R = 9, N = 10^4$ and $n_{inf}(t) = 0.1$. Taken from [8]. 108
6.7 Graphical representation of pair correlations. Note that i and j are first order neighbors, also j and k. i and k are second order neighbors, and we assume they are not correlated. 111
6.8 First- and second-order approaches in comparison with direct simulations for the SI model. The first- and second-order approaches give similar results in tree-like networks with low clustering coefficient, or *transitivity*. For networks with high transitivity the first-order approach shows a bad agreement with simulation data (dots). Taken from [6, section 17.10]. 113

7.1 Spiral wave patterns in the Belousov–Zhabotinsky reaction. Taken from [9, section 2.5]. 117
7.2 Spreading pattern of malaria in northern Bengal, India. Lighter regions mark places or earlier disease activity reported from January 206 to February 2009. Gray scale encodes the mean relative phase angle, i.e. the average phase angle of a given location relative to the spatial average over all regions. Taken from [10, section 4.7]. 120
7.3 Schematic representation of the cell membrane potential as an electrical circuit. The lipid bilayer is represented by a capacitor C_m that maintains the potential $V_m = V$ across the cell membrane. Ion-channels act as voltage dependent rheostats for controlling the inward Na^+ and outward K^+ currents. Taken from [10, section 2.2]. 121
7.4 Dynamics of the Hodgkin-Huxley model variables. Time evolution of voltage V (top) and gate probabilities (bottom) as a response to $I_{ext} = 4\mu$A applied for 5 ms. Taken from [10, section 2.2]. 123
7.5 Failure (extinction of the dynamics) rates as a function of p, normalized by $p_c(N)$. The inset shows the failure rate as a function of p. Several system sizes are shown. Taken from [11]. 124

7.6 Failure (extinction of the dynamics) rates as a function of p for system of size $N = 1000$ and $\tau_D = 0 : 6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8$ (from left to right). Taken from [11]. 124

7.7 Time evolution of variables u and v in the FitzHugh-Nagumo model (left) and the corresponding trajectory in the (u, v) plane (right). Parameter values were $u_c = 0.25, k = 1, \epsilon = 0.1$, and the initial condition were $(0, -0.2)$. Taken from [10, section 1.2]. 125

7.8 Dynamic behaviors of the FitzHugh-Nagumo model. Excitable dynamics (top) and oscillatory dynamics (bottom). Taken from [10, section 2.3]. 126

7.9 Nullclines of the Chialvo model. In dependence of the parameter values, one (top) or three (bottom) intersections are possible. . . . 128

7.10 Regime $0 < p < p_c^l$ of the Chialvo model in small-world networks. a) circular waves in the transient. b) self-sustained spiral waves. c) Time behavior of activity. Taken from [10, section 8.3]. 130

7.11 Fraction of system with spiral waves, F_S as a function of a) N for $p = 0.05$ and b) p with $N = 300 \times 300$. Taken from [10, section 8.3]. 130

7.12 Fraction of system with spiral waves, F_S as a function of $(p - p_c^l)N$. The inset shows the dependence of p_c^l with the diffusion coefficient D. Taken from [10, section 8.3]. 131

7.13 Dependence of p_c^u with the system size N. Te inset shows the traction of ceased systems after a time $t = 20,000$, F_C as a function of $(p - p_c^u)$. Taken from [10, section 8.3]. 131

8.1 Graphical definition of infectible nodes. Consider two neighboring nodes i and j, such that i is infectious. j may be refractory or susceptible. A $S \to I$ transition is possible only if the gray time band if non-zero, i.e. $t_i + \tau_i - (t_j + \tau_i + \tau_r) = t_i - t_j - \tau_r > 0$, this leads to the relation $t_i - t_j > \tau_r$, as given by equation 8.2. In such case we say that node j is infectible. Note that node j will become susceptible before node i heals. 136

8.2 Normal distribution $N(\Theta, \sigma^2)$ folded over a circle. 138

8.3 Active period circle. The black dot marks the actual location of $t - t_i$ in the active period. Gray dots mark possible locations for $t - t_i$. The red-and-blue dot marks the position where the oscillator enters (red) and leaves (blue) the active period. t_i is the last infection time of oscillator i. The red region represents the Infectious state, while the blue region represents the Refractory state. 139

8.4 Extinction states in phase space for 1000 networks in dimension $d = 2$ with $L = 20$ and $p_{ann} = 0$. Two values of α/d are shown; $\alpha/d = 0$ (A) and $\alpha/d = 1.25$ (B). For non-zero annealing Case c) extinctions are absent, and Case a) extinctions ore more uniformly distributed over the $\mathbf{S} \leftrightarrow \mathbf{R}$ border. 141

8.5 Extinction states in phase space for 1000 networks in dimension $d = 2$ with $L = 40$ and $p_{ann} = 0$. Two values of α/d are shown; $\alpha/d = 0$ (A) and $\alpha/d = 1.25$ (B). For non-zero annealing Case c) extinctions are absent, and Case a) extinctions ore more uniformly distributed over the $\mathbf{S} \leftrightarrow \mathbf{R}$ border. 142

8.6 z^2 vs. σ^2 corresponding to 1000 extinctions for networks in dimension $d = 2$ with $L = 20$ an $p_{ann} = 0$. Two values of α/d are shown; $\alpha/d = 0$ (A) and $\alpha/d = 1.25$ (B). Vertical lines indicates the location of σ_c^2 (solid line) and $\sigma_c^2/4$ (dashed line). Most extinctions happens to the left of σ_c^2. Some outliers extinctions appears to the right of σ_c^2 only for small system sizes and large values of α. . . . 144

8.7 z^2 vs. σ^2 corresponding to 1000 extinctions for networks in dimension $d = 2$ with $L = 40$ and $p_{ann} = 0$. Two values of α/d are shown; $\alpha/d = 0$ (A) and $\alpha/d = 1.25$ (B). Vertical lines indicates the location of σ_c^2 (solid line) and $\sigma_c^2/4$ (dashed line). In this series of simulations all extinctions happens to the left of σ_c^2. 145

8.8 Susceptible fraction s vs. z^2 and σ^2 corresponding to 1000 extinctions for networks in dimension $d = 2$ with $L = 40$ and $p_{ann} = 0$. Two values of α/d are shown; $\alpha/d = 0$ (A) and $\alpha/d = 1.25$ (B). The lines at the bottom plane indicates the location of σ_c^2 (solid line) and $\sigma_c^2/4$ (dashed line). 146

8.9 Densities (i, s, r) in phase space corresponding to 1000 networks in dimension $d = 2$ with $L = 40$, $p_{ann} = 0$ and $\alpha/d = 0$. Forward time averages (A) and time-to-extinction averages (B). 147

8.10 Densities (i, s, r) in phase space corresponding to 1000 networks in dimension $d = 2$ with $L = 40$, $p_{ann} = 0$ and $\alpha/d = 1.25$. Forward time averages (A) and time-to-extinction averages (B). 148

8.11 Time behavior of z^2, σ^2 and persistence for networks in dimension $d = 2$, with $L = 40$, $\alpha = 0.00$, and $p_{ann} = 0$. In forward time (A) and in time-to-extinction (B). In (B) the y axis is logarithmic to better appreciate variations with respect to t'. Horizontal doted lines indicate the location of σ_c^2 (upper) and $\exp(-\sigma_c^2)$ (lower) on each plot. 150

8.12 Time behavior of z^2, σ^2 and persistence for networks in dimension $d = 2$, with $L = 40$, $\alpha = 2.50$, and $p_{ann} = 0$. In forward time (A) and in time-to-extinction (B). In (B) the y axis is logarithmic to better appreciate variations with respect to t'. Horizontal doted lines indicate the location of σ_c^2 (upper) and $\exp(-\sigma_c^2)$ (lower) on each plot. 151

8.13 Time averages of active fraction (A), susceptible fraction (B) and infectible fraction (C) vs. α for networks in dimension $d = 2$, with $L = 40$ and $p_{ann} = 0$. These quantities are almost constant for $\alpha \leq 2$, while change slowly for $\alpha > 2$. 152

8.14 z^2 vs. σ^2 corresponding to self-sustained oscillations for networks in dimension $d = 2$ with $p_{ann} = 0$. Values of α are shown for every data point; $L = 20$ (A) and $L = 40$ (B). Vertical lines indicate the location of σ_c^2 (solid line) and $\sigma_c^2/4$ (dashed line). 153

8.15 Density distribution of nodes over the active period circle. States are color coded; S = yellow, I = red, R = blue. The yellow circle and band indicates the extremes of the phase range that exists in an equivalent linear distribution. 154

8.16 Time evolution of the infectible fraction for 1000 networks in dimension $d = 2$ with $L = 40$, $p_{ann} = 0$, and $\alpha = 0$ in forward time t (A), and in time-to-extinction t' (B). The three stages discussed in the previous section are distinguishable here. It worth noting that, in Stage 3, the mean oscillation minima departs from zero exponentially. Increasing the infectible fraction is equivalent to increasing the dynamics surface. Larger surfaces provides more "reactive" nodes to the dynamics. Disease spread is enhanced with large infection fluctuations. This condition leads to faster extinctions. 155

8.17 (A) Temporal behavior of the minima x_{min} of infectible fraction for 1000 networks in dimension $d = 2$ with $L = 40$ and $p_{ann} = 0$. Two values of α are shown: $\alpha = 0$ and $\alpha = 2.5$. Infectible fraction minima behave as $x_{min} = x_{min,\infty}\left(1 - e^{-at'}\right)$. In (B) we plot $x_{min,\infty}/(x_{min,\infty} - x_{min})$ to make clearer the exponential behavior of minima. Parameters $x_{min,\infty}$ and a vary with L and α. 156

8.18 Persistence fits for data of networks with $L = 20$ in dimension $d = 2$. Three values of α/d are shown in each sub-figure; $\alpha/d = 0.00, 1.00, 1.25$. Static case with $p_{ann} = 0.00$ (A), and annealed networks with $p_{ann} = 0.05$ (B). **Symbols represent simulation data, while dotted lines of the same color are their respective best fits of equation** (8.20). The correspondence is striking. Data fits in $d = 1$ yield similar results. 158

8.19 $\ln\left(t_{1/2}N^{1/2}\right)$ as a function of the system size N for networks in $d = 2$ (A) and $d = 1$ (B). Several values of α/d are shown in each sub-figure, from left to right; $\alpha/d = 2.0, 1.5, 1.0, 05, 0.0$. Different levels of annealing were considered. $\ln\left(t_{1/2}N^{1/2}\right)$ is asymptotically linear with N. Annealing reduces the asymptotic slope of $\ln\left(t_{1/2}N^{1/2}\right)$. 159

8.20 $\ln\left(t_{1/2}N^{1/2}\right)$ as a function of the system size α for networks in $d = 2$. Two levels of annealing were considered; $p_{ann} = 0.00$ (A), and $p_{ann} = 0.05$ (B). $\ln\left(t_{1/2}N^{1/2}\right)$ increases with α, and this is enhanced by the system size N. Annealing reduces growth of $\ln\left(t_{1/2}N^{1/2}\right)$. Simulations in $d = 1$ yield similar results. 160

8.21 $\log(y)$ vs. Λ for networks with $p_{ann} = 0.00$ (A) and $p_{ann} = 0.05$ (B). Values of $\log(y)$ for different values of L collapse into a common behavior when plotted against Λ. The common behavior resembles a power-law $\log(y) = A\Lambda^{\beta}$ with exponent $\beta \simeq 3.5$. The power-law scaling is clearer for annealed dynamics data (B). Annealing appears to only affect the factor A, reducing it. 163

B.1 λ vs. α/d for different values of L in $d = 2$. As stated by equation B.3, for $\alpha \leq d$ we observe that λ varies little. In the range $d < \alpha \leq d + 1$ it is clear that λ decays with α. For $\alpha > d$ we can observe a "saturation" of λ, i.e it becomes independent of α. 171

List of Tables

2.1 **Real and simulated parameters for the collaboration graph of movie actors (MAN).** Parameter values with superscript r were used as input for our algorithm to simulate the collaboration graph of movie actors. Parameter values with superscript s were the result of our simulation over 20 runs. 58

2.2 **Real and simulated parameters for real world Exponential Networks.** Parameter values with superscript r were used as input for our algorithm to simulate Exponential Networks. Parameter values with superscript s were the average results of our simulations over 100 runs. Values in column 5 were calculated using the values in columns 3 and 4 in equation (2.15). *Data from [12]. 59

8.1 Completed simulations in $d = 1$ (top) and $d = 2$ (bottom), with cpu time usage per series. 161

Resumen

Esta tesis presenta trabajo original concerniente a dos ramas de la teoría de redes:

El Capítulo 2 de la Parte I, fue motivado por la necesidad de modelar la distribución de grado exponencial reportada en tres escuelas secundarias en Yucatán México por Huerta-Quintanilla et al. [13]. La redes con distribución de grado exponencial no son comunes. Como resultado, desarrollamos un modelo de crecimiento de redes que permite el ajuste fino de los parámetros de distribución en redes exponenciales y libres de escala. Los parámetros ajustables en el proceso de crecimiento son la regla de conexión y el grado inicial de los nodos nuevos.

El Capítulo 8 de la Parte II fue motivado por las limitaciones en nuestro entendimiento sobre el proceso de extinción en enfermedades recurrentes. Empleamos un modelo SIRS espacial con periodo refractario fijo para estudiar el efecto del embebido espacial en d dimensiones y la distribución de longitudes de enlace en forma de ley de potencia con exponente de decaimiento α. Como resultado

1) Identificamos tres etapas dinámicas con respecto a la coherencia de fases. La primera etapa es Stage 1, un corto periodo transitorio de coherencia creciente. La segunda etapa es Stage 2, caracterizada por valores sostenidos de coherencia promedio. Una tercera etapa, Stage 3, es visible cuando el sistema es observado desde el tiempo de extinción. Todas las extinciones muestran una distribución unimodal de fluctuaciones de fase.

2) Proveemos un relación cerrada para la distribución de tiempos de extinción, basado en un mapeo de la extinción como un problema de primer paso (First-Passage Time problem), donde la frontera dinámica , la fracción "infectible", es considerada como un caminante aleatorio con una trampa en el origen. La relación resultante es notablemente exacta en el campo medio, de otro modo solo es una aproximación. La calidad de la aproximación se reduce con el tamaño del sistema.

3) Encontramos una relación de escala para el tiempo de vida medio que nos permite colapsar múltiples series de datos. Esta relación depende del tamaño del sistema N y el exponente de decaimiento α.

Abstract

This thesis present original work touching two branches of network theory.

Part I, Chapter 2 was motivated by the need to model the exponential degree distributions of Enmity Networks documented in three elementary schools in Yucatan Mexico by Huerta-Quintanilla et al. [13]. Exponential degree distributions are not common in real world networks. This results in the depelopment of a parametric model of network growth that allows to fine tune parameters of the final degree distribution. The adjustable parameters of the growth process are the attachment rule and the initial degree of newly added nodes.

Part II, Chapter 8 was motivated by the gap in our understanding of extinction times in recurrent epidemics. We analyzed the effects of i) spatial embedding and a ii) power-law link-length distribution (with decay exponent α) on the SIRS model with fixed refractory period. Dynamics is developed over small-world networks with nodes on a d-dimensional lattice. We were able to

1) Identify three dynamic stages with respect to phase coherence. The first stage is Stage 1, a short transient of increasing coherence from a very low initial value. The second stage is Stage 2 (the longest stage) characterized by sustained coherence. A third stage is only visible on averages measured from extinction time, this is Stage 3, a short transient of increasing synchronization. This hyper-synchronization leads to extinction. All extinctions display an unimodal distribution of phase fluctuations.

2) Provide a closed relation for the distribution of extinction times, based on the mapping of extinction to a First-Passage Time problem, where the size of the dynamical frontier, the "infectible" fraction, is considered as a random walker with a trap at zero. The approach is strikingly accurate in the mean field, an approximation otherwise. The quality of the approximation decreases with the system size.

3) Find a scaling relation for the mean lifetime to collapse multiple data sets. This relation is dependent on the system size N and the decay exponent α.

Introduction

"Many objects of interest in the physical, biological, and social sciences can be thought of as networks and, thinking of them in this way can often lead to new and useful insights." from Mark Newman, one of the most renowned authors in the field of complex networks, may be conceived as a paraphrase to the cliché "the whole is more than the sum of its parts".

A quick look around will reveal a plethora of interacting entities. Outside, we see people building communities at different scales, power grids conveying the electricity that moves our daily lives, supply chains feeding productive actors of the economic system, information networks keeping us up-to-date with the news, trophic networks maintaining our ecosystem, etc.

Inside, we are aware of the physiological orchestration that maintains life: cell sending and receiving an innumerable quantity of chemical messages, organs depending on precise chemical concentrations to be able to function, neurons firing electrical and chemical signal to convey information.

Deep inside the previous examples is the fact that what we are from a physical stand point, and what we are capable of doing, depend somehow on an infrastructure. An scheme to organize interactions, this is what we call a *network*. A network is tagged as *complex* when the scheme of interaction is non trivially described by a simple rule, such as 'all elements interact with one another' or 'all elements interact through a regular lattice'. The adjective complex also refers to networks where the number of entities is large and the coordination numbers uneven.

Basic network theory deals with network as pure infrastructure that is either static or evolves. However, the consensus that the outcome of a networked system depends both on its constituents and the way they interact sets a greater concern: relate structure to function, opening the possibility to make predictions and excerpt control.

This thesis present original work touching two branches of network theory.

Part I, Chapter 2 was motivated by the need to model the exponential degree distributions of Enmity Networks documented in three elementary schools in Yucatan Mexico by Huerta-Quintanilla et al. [13]. Exponential degree distributions are not common in real world networks. Most real networks present fat tail in the form of power-law or exponentially truncated power-laws [14, 15, 16].

The result of this chapter is a parametric model of network growth that allows to fine tune parameters of the final degree distribution. The adjustable parameters

of the growth process are the attachment rule and the initial degree of newly added nodes.

Part II, Chapter 8 was motivated by the gap in our understanding of extinction times in recurrent epidemics, so as the effect of spatial embedding and long-range interactions over them.

We analyzed the effects of i) spatial embedding and a ii) power-law link-length distribution (with decay exponent α) on the SIRS model with fixed refractory period. Dynamics is developed over small-world networks with nodes on a d-dimensional lattice.

The results of this chapter are:

- The identification of three dynamic stages with respect to phase coherence. The first stage is Stage 1, a short transient of increasing coherence from a very low initial value. The second stage is Stage 2 (the longest stage) characterized by sustained coherence. A third stage is only visible on averages measured from extinction time, this is Stage 3, a short transient of increasing synchronization. This hyper-synchronization leads to extinction. All extinctions display an unimodal distribution of phase fluctuations.

- A closed relation for the distribution of extinction times, based on the mapping of extinction to a First-Passage Time problem, where the size of the dynamical frontier, the "infectible" fraction, is considered as a random walker with a trap at zero. The approach is strikingly accurate in the mean field, an approximation otherwise. The quality of the approximation decreases with the system size.

- A scaling relation for the mean lifetime to collapse multiple data sets. This relation is dependent on the system size N and the decay exponent α.

Organization

This work is divided into two parts

Part I focuses on the structural analysis of complex networks:

- Chapter 1 is aimed at introducing the reader to the basics of *Network Theory*.
- In Chapter 2 we develop a *parametric model of network growth*, which allows us to obtain networks with fine tuned degree distributions.

Part II focuses on dynamical processes over complex networks:

- Chapter 4 deals with resilience, to describe what happens when some parts of a network are effectively removed. It also addresses the effects of different removing strategies.
- Chapter 5 is aimed to introduce reader to *synchronization* as a mean to study the behavior of interacting oscillatory units.
- Chapter 6 is aimed to introduce reader to paradigmatic epidemiological models.
- Chapter 7 is aimed at introducing the reader to a special class of dynamical systems called *excitable media*.
- In Chapter 8 we take into account the spatial dimension where a network is embedded, then analyze the extinctions of a *SIRS spatial model with long-range interactions following a power-law decay*.

The last chapter in each part includes a "Conclusions and Discussion" section because these chapters constitute the publishable results of the thesis work. Chapter 2 in Part I was published under the supervision of Dr. Rodrigo Huerta-Quintanilla in the Journal of Complex Networks. Chapter 8 in Part II is in process of manuscript writing before submission to a publisher, under the supervision of Dr. Cristian F. Moukarzel.

Part I

Network Structure

Chapter 1

Network Theory basics

"Network" is today a pervasive word in our jargon, perhaps we automatically think in social networks or even computer networks when we heard it. In few words a network is a diagrammatic representation of the underlying structure of relations among the elements of a system.

For small networks a graph drawn on paper is an effective way for representation and visual analysis, however, for complex networks, which maps the interactions of a large number of individuals, we use the mathematical framework of Network Theory. In the big picture Network Theory could be divided into to areas, *network structure* which deals with the morphological measures of networks (size, order, paths, etc.) and *network dynamics* which deals with network evolution (i.e. morphological changes of a network over time) and dynamical processes happening on top of a network (i.e. processes on the system encoded by a network).

These areas are not two pieces apart, the structure and dynamics of and on networks are interrelated: network evolution is driven by a set of dynamical rules for attachment and growth which in the end influence morphological characteristics such as degree distribution, assortativity, clustering coefficient, path length distribution, average path length, etc.; morphological characteristics of a network play a key role over dynamical processes over it, such as search, information diffusion, epidemic spreading, etc. [17, 18, 19, 20, 6, 3].

1.1 Network measures

Mathematically, a *network* is represented by a *graph*. A graph G is a pair of sets, $G = \{V, E\}$, where V is a set of N elements called *vertices/nodes* (labeled from 1 to N); E is a set of M ordered pairs of the form $\{(i, j)|i, j \in V\}$ called *edges/links*. There could be links of different kind in a graph:

- Directed link: a link is a *directed link* if the existence of (i, j) does not implies the existence of (j, i). Graphs which contain directed links are called *digraphs*.

- Non-directed link: a link is a *non-directed link* if the existence of (i, j) implies the existence of (j, i). Graphs which contain only non directed links are called *simple graphs*.

- Reflexive link: a *reflexible link* is a link of the form (i, i), connecting a node with itself. This kind of link is also called a *loop* or *self-link*.

- Multiple link: a link (i, j) is a *multiple link* if it is present more than once in a graph.

Graphs with loops and/or multiple links are collectively called *pseudographs*. In this work we deal with *simple graphs* only, or *graphs* for short. The terms couples *network-graph*, *vertice-node*, *edge-link* will be used interchangeably.

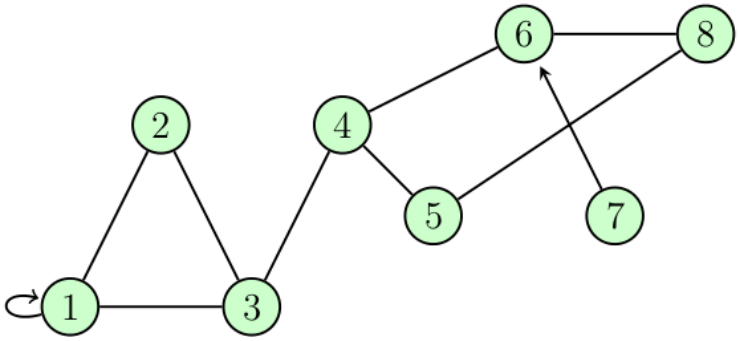

Figure 1.1: Graph with 8 nodes and 10 links. The set of nodes is $V =$ { 1, 2, 3, 4, 5, 6, 7, 8 }. The set of links is $E =$ { (1,1), (1,2), (1,3), (2,3), (3,4), (4,5), (4,6), (6,8), (5,8), (7,6) }. Among E $(7, 6)$ is a directed link, $(1, 1)$ is a reflexive link or loop, and all other links are non-directed.

1.1.1 Degree distribution

The set of nodes connected to node i is, $v(i)$, denoted as the *neighborhood* of node i. The cardinality of this set, k_i, is the *degree* of node i. Unlike lattices, nodes in a complex network display a variety of degree values. Indeed, node degree is distributed accordingly to a *degree distribution* p_k

$$p_k = \frac{n_k}{N}, \tag{1.1}$$

where n_k is the number of nodes having degree k in a network of size N. p_k encodes the probability that a randomly chosen node has degree k. Two well know examples of degree distributions are the Poisson distribution of *Random networks*, and the power-law (fat-tail) distribution of *Scale-Free newtorks*. [1]

[1] Detailed discussion about the construction of such networks will be given in Chapter 2.

Given p_k we are able to calculate some statistical measures of a network, such as the average degree $\langle k \rangle$

$$\langle k \rangle = \sum_k k p_k, \tag{1.2}$$

although it only depends on N and M [2]

$$\langle k \rangle = \frac{2M}{N}, \tag{1.3}$$

i.e. this global average give us information about the links to nodes ratio in a network. Equation 1.3 is valid for any network, independently of its degree distribution.

1.1.2 Mixing patterns and degree correlations

For a given network, not all the information about node organization is encoded by its degree distribution; empirical evidence from ecology, epidemiology, and social science suggests that a node does not connect to other node irrespective of their properties. There exists mixing patterns [21]

1. *Assortative mixing*: if nodes tend to connect to other nodes with similar properties, and
2. *Dissortative mixing*: if nodes tend to connect to other nodes with dissimilar properties.

When the property of interest is node degree, the conditional probability $p_{k'|k}$, i.e. the probability that a node of degree k is connected to a node of degree k', encodes degree correlations. For practical reasons $p_{k'|k}$ is generally estimated in terms of the *nearest neighbors degree*. For node i it is

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in v(i)} k_j, \tag{1.4}$$

where $v(i)$ stands for the neighborhood of node i. Thus, the *average nearest neighbors degree* of nodes with degree k [22, 23] is

$$k_{nn}(k) = \frac{1}{N_k} \sum_{i/k_i=k} k_{nn,i} = \sum_{k'} k' p_{k'|k}. \tag{1.5}$$

In the absence of degree correlations $p_{k'|k}$ does not depends on k

$$P_{k'|k} = \frac{k' p_{k'}}{\langle k \rangle},$$

then

[2] This fact is know as the "Hand-shaking Lemma"

$$k_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle} = \kappa \tag{1.6}$$

is a constant determined by the two first moments of the degree distribution. This constant is called *heterogeneity parameter*, a measure of the degree heterogeneity on uncorrelated networks.

When degree correlations are present networks can be classified into one of two classes:

- If $k_{nn}(k)$ is an increasing function of k, then nodes with high degree have a larger probability of being connected with other large degree nodes. This is an *assortative network*, i.e, with assortative mixing.
- If $k_{nn}(k)$ is a decreasing function of k, then high degree nodes have a majority of neighbors with low degree, while the opposite holds for low degree nodes. This is an *disassortative network* ,i.e.,with diassortative mixing.

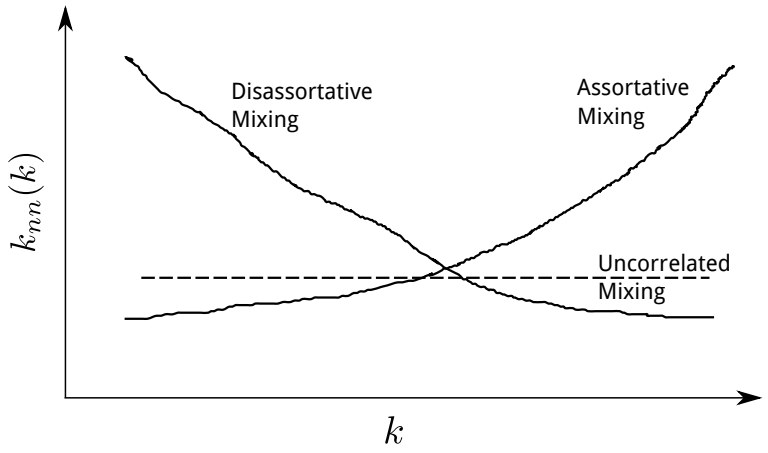

Figure 1.2: Mixing patterns in networks. An assortative network presents a increasing behavior of $k_{nn}(k)$. An disassortative network presents a decreasing behavior of $k_{nn}(k)$. Uncorrelated networks presents a constant $k_{nn}(k) = \kappa$.

In general, networks with different mixing patterns presents qualitative different dynamical behaviors in processes such as synchronization and epidemic spreading. Mixing patterns even affects network resilience under attack (targeted node deleting) [3].

1.1.3 Geodesic distances and paths

In complex networks, it is possible to define some analogous to distance in terms of links, which exists in between nodes and connect them. This distance is denoted as *geodesic distance*.

The *geodesic distance* $d(i, j)$ between nodes i and j is the length of the shortest path between i and j if this path exists, else it is infinite. There exists a *path* of length k between nodes i and j if node j could be reached from node i by following a sequence of k links.

In term of geodesic distances we can define some topological metrics of the network such as the *diameter* D, i.e. the longest geodesic distance in a network, and the *average path length* l, i.e. the average geodesic distance between all node pairs in a network.

There is one reason why paths are of invaluable importance: in order to reach node i from node j a path **must** exists between them. This is relevant in any search or diffusion process within a network. Without a path among two nodes, they cannot interact neither directly nor indirectly.

1.1.4 Clustering coefficient

If node i has k_i neighbors, these neighbors could share among them at most $k_i(k_i - 1)/2$ edges, but only a fraction C_i of those links exists. The (average) clustering coefficient of the network, C, is the average over these fractions

$$C = \frac{1}{N} \sum_i C_i, \tag{1.7}$$

with $0 \leq C \leq 1$. C is the probability that two neighbors of a node are also neighbors of each other.

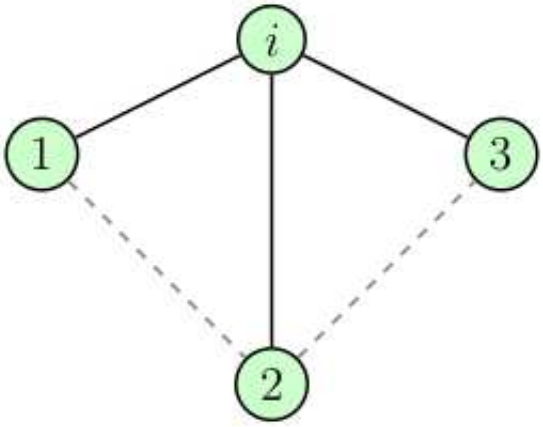

Figure 1.3: Graphic definition of clustering coefficient. In the image $v(i) = \{1, 2, 3\}$. The possible links among the nodes is $v(i)$ are { (1, 2), (2, 3), (1, 3) }, just $(1, 2)$ and $(2, 3)$ (marked with dashed lines). In this example $C_i = 2/3$.

This quantity is a *global* measure of *local* link density, that refers to the tendency to form triangles in the network. A more informative quantities is the explicit dependence of the clustering coefficient on the degree, $C(k)$, defined as

$$C(k) = \frac{1}{n_k} \sum_{i/k_i=k} C_i. \tag{1.8}$$

A functional $C(k)$ is indicative for the presence of a complex structure in the three-vertex correlation pattern. In some cases $C(k)$ encodes the hierarchical structure of a network [23, 24].

As will see later, many real world network display high values of clustering, and this fact has motivated the development of specific network models [25, 26].

1.1.5 Components

A network could be made of a set of sub-networks or components, with no paths of any order among the nodes in different components.

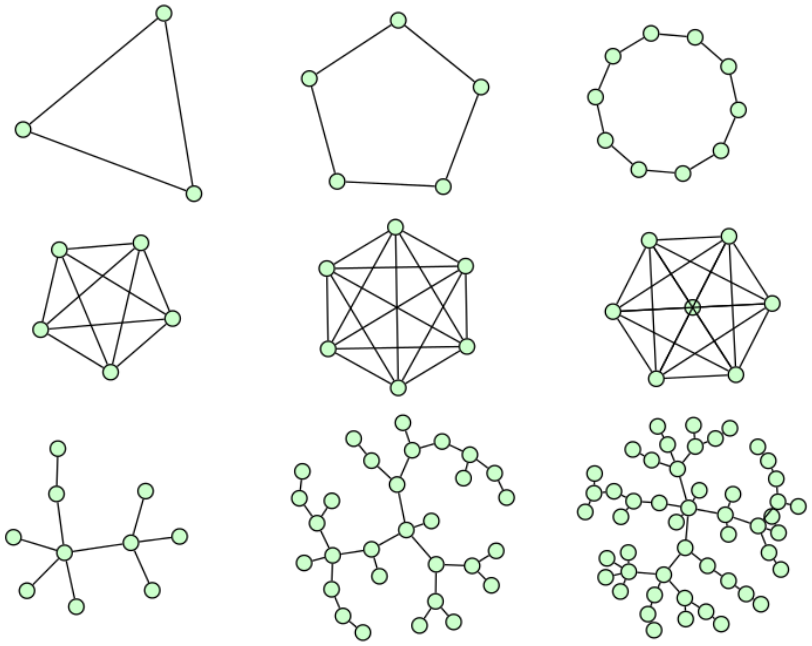

Figure 1.4: Examples of network components; (Upper row, from left to right) cycles of order 3, 5 and 7. (Central row, from left to right) cliques of order 5, 6 and 7. (Bottom row, from left to right) trees of order 10, 30 and 50.

The most simple kinds of components are

- *cycles*: a k-th order cycle is a network with k nodes and k edges forming a single closed path.
- *cliques*: a k-th order clique is a network with k nodes and $k(k-1)/2$ edges. A clique is also called a "complete network" because every node i is connected to the rest of nodes.

- *trees*: a k-th order tree is a network with k nodes and $k-1$ edges, without closed paths.

The components of a network can be as complex as the network where they belong. A network with only one component (i.e. the network itself) is said to be a *connected network*.

Not only separated components are important for the structural characterization of a network, other subsets of a network are relevant too. At the local level the identification of *communities* has promoted a large amount of work in the field [27, 28, 29]. Basically, communities are subsets of nodes that are highly connected among themselves and poorly connected to nodes from the outside, this clearly indicates the need for criteria to specify what "highly" and "poorly" connected are.

The existence of communities raises another characteristic of a network, *modularity*. In terms of modularity, networks are classified as

- *non-hierarchical* if its communities are disjoint, and
- *hierarchical* if their communities are nested ones inside one another.

Clearly, these are sharp cases, because another possibility is to get a network with an overlapping community structure.

1.2 Types of networks

1.2.1 Random Networks

Random networks were perhaps the first attempt for describe real world networks [30, 31]. Random network models assume an absolute lack of knowledge about the principles that guide the creation of links between nodes.

Start with a network of N isolated nodes, if any link out of the $N(N-1)/2$ possible links exists with *wiring probability* p we get a random network denoted as $G_{N,p}$. Alternatively, we can randomly distribute M links among node pairs to obtain a random network denoted as $G_{N,M}$. The number of graphs with N nodes and exactly M links equals the number of ways of picking the positions of the links from the $\binom{N}{2}$ possible pairs of nodes. In the model $G_{N,p}$, the probability of obtaining a single network with M links is

$$P(G) = p^M (1-p)^{\binom{N}{2}-M}. \tag{1.9}$$

Then the total probability to obtain any $G_{N,M}$ is

$$P(G_{N,M}) = \binom{\binom{N}{2}}{M} p^M (1-p)^{\binom{N}{2}-M}, \tag{1.10}$$

which is a Binomial distribution, in which the mean number of links is

$$\langle M \rangle = \binom{N}{2} p. \tag{1.11}$$

Similarly, the probability $P(k)$ of creating a node of degree k is equal to the probability that it is connected to k other nodes and not connected to the remaining $(N-1)-k$ nodes

$$P(k) = \binom{N-1}{k} p^k (1-p)^{(N-1)-k}, \tag{1.12}$$

which is a Binomial distribution, with mean value k

$$\langle k \rangle = (N-1)\, p. \tag{1.13}$$

i.e. for a given node i the average degree is just the number of possible neighbors, the other $N-1$ nodes, multiplied by the probability p that a link exists between i and any other node j.

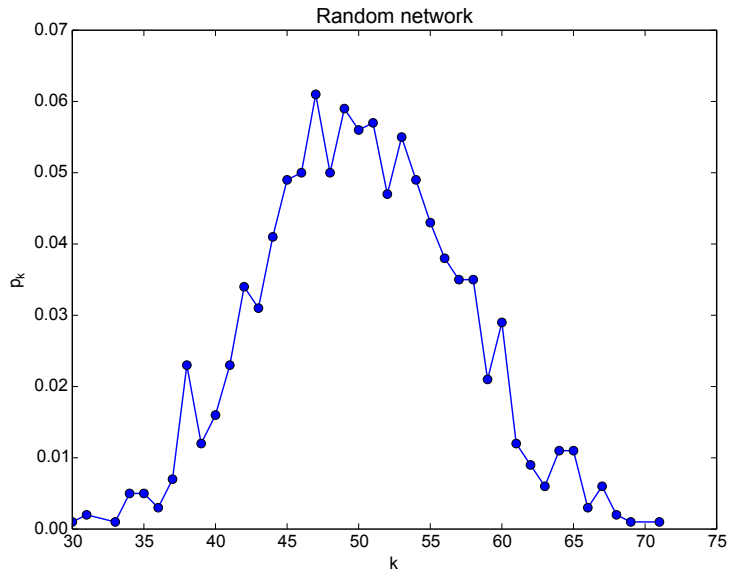

Figure 1.5: Degree distribution of a Random network $G_{1000,0.05}$.

In the limit $N \to \infty$ both models yield networks with Poisson degree distribution (figure 1.5)

$$p_k = e^{-z} \frac{z^k}{k!}, \tag{1.14}$$

where $z = \langle k \rangle$.

In a random network the probability that any two nodes are neighbors is

$$p = \frac{\langle k \rangle}{N-1}, \tag{1.15}$$

this means that the probability of two neighbors of a node to be connected is p, thus, for a random network the clustering coefficient is

$$C = p = \frac{\langle k \rangle}{N-1}, \tag{1.16}$$

it decreases as N^{-1} for fixed average degree and approaches zero as $N \to \infty$. This implies a lack of local cohesiveness in the model.

The average path length of a random network can be estimated as follows: when the presence of cycles is negligible (i.e. $C \sim 0$, with the conditions described above) the average number of nodes at distance $d = 1$ from a node (its neighbors) is $\langle k \rangle$; the average numbers of neighbors at distance $d = 2$ is $\langle k \rangle^2$, and so on. If at distance $d = r_G$ we get $\langle k \rangle^{r_G} \sim N$ then

$$l \sim r_G = \frac{\ln N}{\ln \langle k \rangle} \tag{1.17}$$

is an approximation to the average path length. The scale of l with $\ln N$ is the signature of the *small-world* effect: for example, a connected network with as many as $N = 10^6$ nodes could have average path length of order 6, or less.

While a fixed *wiring probability* p controls the average degree, clustering coefficient and average path length of a network, a varying p induces structural changes on it:

- When $p < 1/N$ the network is composed of many component of negligible size. See left side of figure 1.6.
- When $p > 1/N$ the network has a *giant component* connecting a non-negligible fraction of the nodes, χ_0. See right side of figure 1.6.

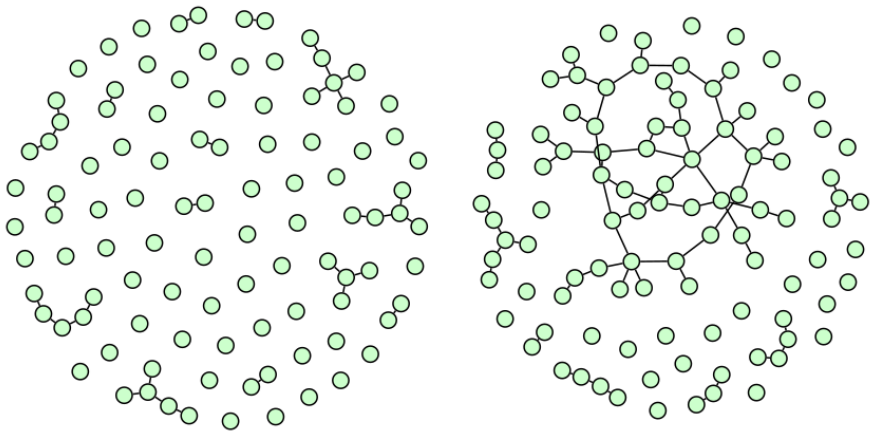

Figure 1.6: On the left, a Random network $G_{100,0.005}$. On the right a Random network $G_{100,0.015}$. The presence of a giant component is evident for the $G_{100,0.015}$, because $p = 1.5N^{-1} > N^{-1}$.

The fraction χ_0 presents a phase transition equivalent to a percolation transition in infinite dimension [3] at $p = 1/N$, figure 1.7.

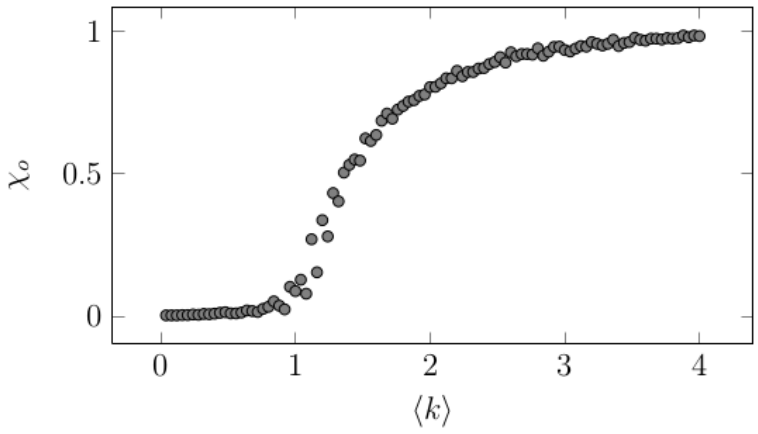

Figure 1.7: Phase transition in the size of the giant component χ_0 in the model $G_{N,p}$ with $N = 1000$ and varying p. For convenience the horizontal axis is the average degree. The transition point at $\langle k \rangle_c = 1$ is rounded due to finite size effects.

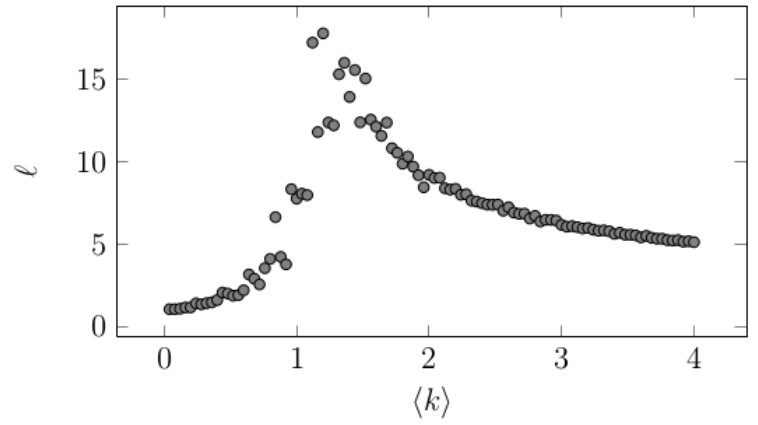

Figure 1.8: Evolution of the average path length l in the model $G_{N,p}$ with $N = 1000$ and varying p. Observe the discontinuity at the critical average degree $\langle k \rangle_c = 1$.

The average path length behavior presents a discontinuously at $\langle k \rangle_c = 1$: as $\langle k \rangle$ increases from 0 to 1 l also increases due to the nucleation and growth of even larger clusters, mainly trees-like clusters; at $\langle k \rangle = \langle k \rangle_c$ the giant component appears; as $\langle k \rangle$ increases from 1 l decreases because each new node added to the giant component contributes with an average of $\langle k \rangle > 1$ links, this generates more new paths as the link density increases. See figure 1.8.

[3] More details about this will be given in Chapter 4

Most real world networks shows structural and dynamical properties which differ from that of a random network with the same number of nodes and links. Then, new and more elaborate network models are needed in order to capture the properties of real networks.

1.2.2 Generalized random networks

The random network model of the previous subsection can be extended to allow the construction of random graph of predefined degree distribution, not just Poisson distributions. The process consist in giving a fixed degree sequence $\{k_i\}_{i=1}^{N}$ to assign a degree k_i to node i. Then randomly distribute M links among nodes so that every node ends the process with its predefined degree [32, 33]. Clearly the degree sequence sum must be an even number

$$\sum_i k_i = 2M, \tag{1.18}$$

because each link is shared by exactly two nodes. The resulting network has a degree distribution p_k predefined by the degree sequence. The average cluster coefficient is given by [34, chapter 2]

$$C = \frac{1}{N}\frac{\left(\langle k^2\rangle - \langle k\rangle\right)^2}{\langle k\rangle^3}, \tag{1.19}$$

which approaches zero as $N \to \infty$, such as in a Poisson Random network. On the other hand, the average path length is

$$l \approx 1 + \frac{\ln\left(N/\langle k\rangle\right)}{\ln\left(\left(\langle k^2\rangle - \langle k\rangle\right)/\langle k\rangle\right)}. \tag{1.20}$$

Equations 1.19 and 1.20 confirm that a generalized Random network with degree p_k displays small-world properties, and such properties are consequence of the randomness of the construction process.

1.2.3 Small-world networks

In many real-world complex networks it has been observed that the average path length is relatively small in comparison with the size of the network, while the clustering coefficient is large in comparison to a random network of the same size and order; the difference in clustering coefficient values is often several orders of magnitude [35, Table I]. The neural network of the worm Caenorhabditis elegans, the power grid of the western United States, the collaboration graph of film actors [25], friendship networks [26] and the network of world airports [15] are just a few examples of such kind of networks, generally referred to as *small-world networks*. Small-world is a term coined by Watts and Strogatz in the 90's [25], whom were the first to developed a network model yielding networks with small average path length and high clustering coefficient.

To construct a small-world network in accordance to the model of Watts and Strogatz [25] we start with a regular network where each node is connected to its $2r$ nearest neighbors. With probability p one end from each link is freed and attached to a randomly chosen node in the network. The process permit the establishment of long range link called *shortcut* connecting distant nodes. In a small world network both C and l shows a characteristic decreasing behavior (figure 1.10) while interpolates between a regular network at $p = 0$ and a random network at $p = 1$ (see figure 1.9).

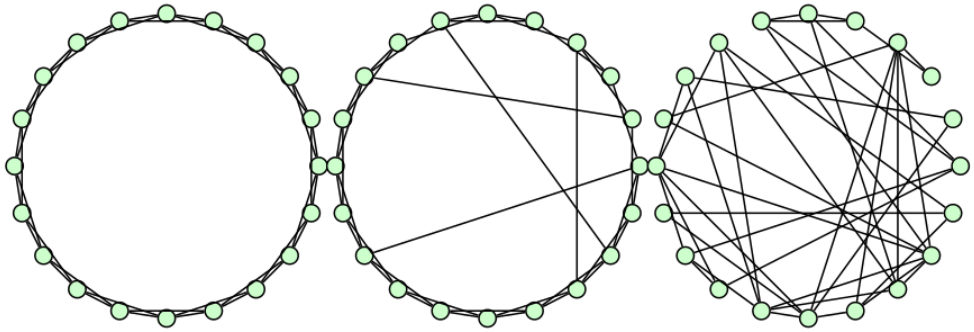

Figure 1.9: Evolution of a regular network with 20 nodes and average degree 4 to a small world network with $p = 0.7$.

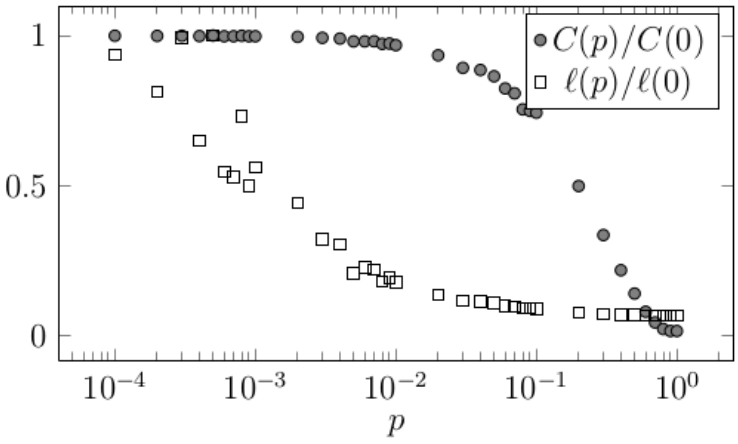

Figure 1.10: Relative clustering coefficient and relative average path length as a function of p for a small world network with 1000 nodes and average degree 10. Values relative to that at $p = 0$.

The dependence of the clustering coefficient with p is [36]

$$C(p) = \frac{3(r-1)}{2(2m-1)}(1-p)^3. \tag{1.21}$$

Alternatively to the model of Watts and Strogatz, a small-world network could be constructed by adding *shortcuts* between randomly chosen pairs of nodes. These shortcuts are added with probability ϕ per link in the underlying network, from a total of $N\langle k\rangle/2$ connections, this is the so called Newman-Watts model [26]. In this model the dependence the average path length is

$$l = \frac{N}{r} f(Nr\phi), \tag{1.22}$$

where the scaling function is

$$f(x) = \frac{1}{2\sqrt{x^2+2x}} \tanh^{-1} \sqrt{\frac{x}{x+2}}. \tag{1.23}$$

In general, it is accepted that a network displays the small-world effect if

$$l \sim \ln(N). \tag{1.24}$$

For an arbitrary network, the average path length is bounded by

$$1 \leq l \leq \frac{N+1}{3},$$

where $l = 1$ is obtained for a complete graph or clique, and $l = \frac{N+1}{3}$ is obtained by a path of order N [17].

1.2.4 Scale-Free networks

The development of random networks and small world network models were aimed to build networks which replicate the properties of real world networks, particularly small average path lengths and high clustering coefficients. However, these model does not account for networks with fat tail degree distributions such as power laws, shifted power laws, exponentially truncated power laws, or even exponential distributions. Power law degree distributions has been observed in the World Wide Web (WWW), the collaboration graph of movie actors, the electrical power grid of the western United States [14], the interaction network of proteins in the metabolism of the yeast S. cerevisiae [37] among many other social, biological, technological and model systems, see refs [38, Table 1] and [35, Table II]. The scientific collaboration networks [39], the WWW in-degree network [16] and the network of movie actors [15] present degree distributions in the form of exponentially truncated power laws. The electric power grid of Southern California displays an exponential degree distribution [15].

The fat tail in the degree distribution of a network implies the existence of nodes with abnormally high degree, these nodes are called *hubs*. As we will see in Chapter 4 they are particularly important in maintaining connectivity in scale-free networks.

Barabási and Albert [14] proposed the first network model that allows to build networks with degree distribution tails scaling as a power law. Their model includes two key facts

- networks grow in time, i.e. the number of nodes and links increases.
- nodes are connected to each other according to an attachment rule $\Pi(k_i, t)$.

In the Barabási and Albert dynamics the initial condition is a network with m isolated nodes. A new node is added at each time step. The new node attaches to m previously existing nodes. Each new link from the new node is set with probability $\Pi(k_i, t)$.

They show that, if the attachment rule is egalitarian, such that

$$\Pi(k, t) = \frac{1}{N(t)} \tag{1.25}$$

then, the resulting network has an exponential degree distribution

$$p_k \propto e^{-k/k_0} \tag{1.26}$$

and we call this network an *Exponential network*.

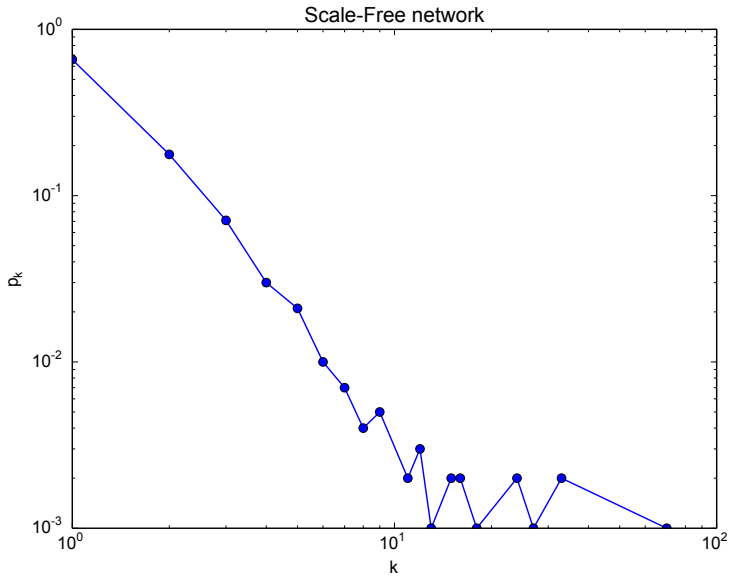

Figure 1.11: Degree distribution of a Scale-Free network from the Barabasi-Albert model with $N = 1000$ and $m = 1$.

On the other hand, if the attachment rule gives preference to highly connected nodes, such that

$$\Pi(k, t) = \frac{k}{M(t)} \tag{1.27}$$

then, the resulting network has a power law degree distribution

$$p_k \propto k^{-\gamma} \tag{1.28}$$

and we call this network a *Scale-Free network*, see figure 1.11.

The (average) clustering coefficient in this case is [40]

$$C(N,m) = \frac{m}{8}\frac{(\ln(N))^2}{N}, \tag{1.29}$$

which decays as N^{-1}, as in the case of random networks. Barabási et al. [41] developed a deterministic model to obtain scale-free networks with non-negligible clustering coefficient at large size. The iterative construction procedure, easily get from figure 1.12, resembles the construction of a fractal and generates networks with hierarchical structure.

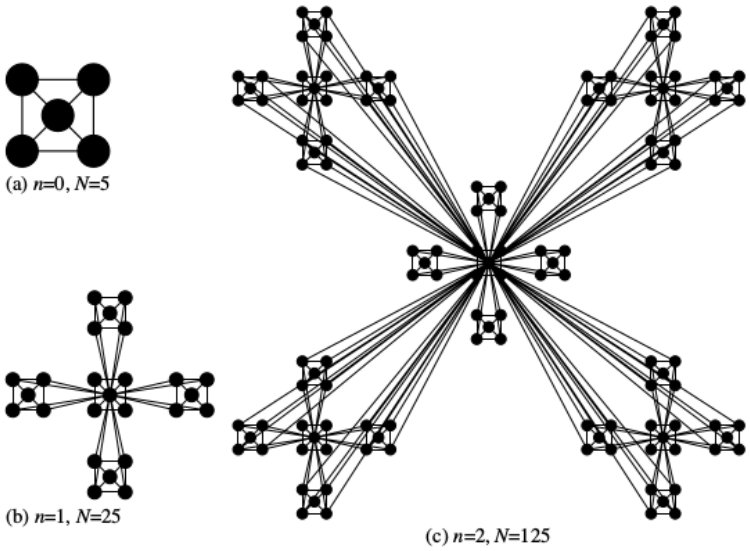

Figure 1.12: Construction of a deterministic scale-free network. The number n indicates the iteration step, N is the number of nodes of the network. In this iterative process N increases with n. Taken from [1, chapter 4]

Hierarchical structures appears in metabolic network [24] [42, chapter 7] and the internet [1, chapter 4] [23]. Remarkably, in both cases the clustering coefficient as a function of degree has the form

$$C(k) \sim k^{-\beta}, \tag{1.30}$$

with $\beta \approx 1$. This fact has motivated the use of the functional form of $C(k)$ as a prove to detect hierarchical structure in networks.

The average path length scales as [34, chapter 1]

$$l \sim \frac{\ln(N)}{\ln\ln(N)}. \tag{1.31}$$

Scale-Free networks in the real world have a variety of γ values. Several models have been devoted to fine tune γ in order to reproduce such values, see the review of Albert and Barabási for a resume of them [35, Table III].

Chapter 2

Growing networks

A real world network can change its topology over time by birth and death of nodes and links as a consequence of any processes allowed to has this effect over the network. The dynamical rules which drive these changes have influences over structural characteristics of a network, allowing us to stear changes in the persuit of specific structural features.

A wide variety of network growth models are now available that mimic structural and dynamical features of real world networks. These have been built on complex network research done over the past half century, beginning with the early work of Erdös and Rényi [30, 31]. Degree distribution is a particularly relevant structural feature since it is the first basis for networks classification: Most networks are right-skewed, and are known as *Scale-Free* Networks when their degree distribution follows a power-law [14]. This power-law can be truncated by an exponential factor [15, 16]. Far less common are *Exponential* networks (with an exponential distribution) and *Poisson* networks (with a Poisson distribution).

Several mechanisms have been proposed to explain network formation and evolution: preferential attachment [14]; relinking [43]; node aging and linking cost [15]; information filtering [16]; and popularity vs. similarity competition [44], among many others (see [35] and [45] for a literature review of complex networks).

2.1 Problem definition

We have detected a general limitation on available network growth models: most of them assume that, in a given time step, all new edges are either attached to the new node (i.e., all new nodes enter the network with degree $k = m$, for an integer $m > 0$) [14, 46, 43, 47, 48, 16, 49, 12], or attached to non-specified nodes (i.e., all new nodes enter the network with degree $k = 0$) [50, 51]. In other words, the initial degree of new nodes is set to a constant value without considering the possible diversity of its values or its influence on the long-run network degree distribution.

We illustrate the effects of the initial degree of new nodes using a series of examples. For the sake of simplicity, we consider exponential network growth:

starting with n_0 nodes and m_0 edges at $t = 0$. An exponential network grows by adding one node and $m \geq 1$ edges in a single time step for $t \geq 1$. Edges are set between node pairs with no biases or preferences, meaning any node is equally likely to be connected to any other node in the network. There will be $N(t) = n_o + t$ nodes and $M(t) = m_o + mt$ edges at time t. In the long-run ($t \to \infty$) the average degree $\langle k \rangle(t) = 2M(t)/N(t)$ approaches the value $\langle k \rangle = 2m$.

Let E_0 and E_1 be two exponential networks with the same set of parameter values ($n_0 = 1, m_0 = 0, m = 1$), connected in slightly different ways. E_0 will grow by adding a new node of degree $k_{ini} = 0$ and a new edge between a randomly selected node pair at each time step, while E_1 will grow by adding a new node of degree $k_{ini} = 1$ which attaches to a randomly-selected pre-existing node at each time step. In the long-run both networks have a degree distribution that takes the form

$$p_k \propto e^{-k/k_o} \qquad k \geq k_{min}, \tag{2.1}$$

with a different set of parameters k_o and k_{min} despite having grown under almost equal conditions to the same number of nodes ($N = 10^4$) and edges. We estimate k_o using the least square fitting of equation (2.1) for each degree distribution with the appropriate k_{min}: $k_{min} = 0$ and $\hat{k}_o = 2.46$ for E_0; $k_{min} = 1$ and $\hat{k}_o = 1.44$ for E_1. From this point on we denote fitted values with a hat. In all of the parameter estimations we used a least square fitting in gnuplot.

The exact degree distribution for networks E_0 and E_1 has been solved analytically: for E_0 Callaway et al. [51] calculated (for $\delta = 1$ on their model) $k_o = [\ln(3/2)]^{-1} \simeq 2.46$; for E_1 Dorogovtsev & Mendes [20, section 8] calculated $k_o = [\ln 2]^{-1} \simeq 1.44$ (see also [18, section 2.5]). Fitted values agree with these calculations.

The only difference in the growth process for E_0 and E_1 is the initial degree of new nodes, which clearly affects the final degree distribution. In solving for p_k, and hence k_o, both Callaway et al. [51] and Dorogovtsev & Mendes [20, section 8] rely on master equations, assuming a constant initial degree, $k_{ini} = 0$ and $k_{ini} = 1$, respectively. However, the master equation approach is unclear, if applicable at all, for a general case of exponential network growth with nodes of random initial degree that preserves the relative edges-to-nodes growth rate m/n. For example if a third exponential network $E_{1/2}$ is grown, using the same growth parameters as in E_0 and E_1, any new node would have the initial degree $k_{ini} = 1$ with probability $1/2$, or initial degree $k_{ini} = 0$ with probability $1 - 1/2 = 1/2$; nodes with initial degree $k_{ini} = 0$ and $k_{ini} = 1$ are randomly distributed over time. The master equation approach does not account for this kind of grown network.

The examples above demonstrate that new node initial degree is relevant to the long-run degree distribution of a growing network. This is a subtle topic within network growth research and has not received close attention in previous models. The present study aim was to fill this gap in understanding the effects of initial degree on network growth by

- setting the general relation satisfied by the parameters set in a network's

degree distribution, and

- identifying the role of new node initial degree on the degree distribution of Exponential and Scale-Free growing networks

2.2 Network growth modelation

An alternative to Master Equations is needed if a random initial degree is to be included in a network growth model. We opt for a general treatment which deals with a posteriori information, and then relates degree distribution to initial degree in a network growth process.

2.2.1 Degree distribution tail behavior

A network $R(N, M)$ is defined as a set of N nodes connected in pairs by a set of M edges. The number of edges that node i shares with other nodes is called the degree of node i, denoted as k_i. There are n_k nodes with degree $k = 0, 1, \ldots, N-1$ and these numbers satisfy

$$\sum_k n_k = N, \tag{2.2}$$

$$\sum_k k\, n_k = 2M. \tag{2.3}$$

For large networks (i.e. those with thousands, millions or even billions of nodes and edges) the degree distribution is a reliable large-scale statistical property [45]. The degree distribution p_k is the probability that a randomly selected node has degree k, and it relates to n_k as

$$p_k = \frac{n_k}{N}, \tag{2.4}$$

therefore equations (2.2) and (2.3) become

$$\sum_k p_k = 1, \tag{2.5}$$

$$\sum_k k\, p_k = \langle k \rangle. \tag{2.6}$$

Complex networks of real world interest are called Poisson, Exponential or Scale-Free Networks in response to their distributions' tail behavior, i.e. the behavior of their degree distributions for degrees equal to or greater than some minimum degree, k_{min}. We approximate general tail behavior by

$$p_k = C f(k; \mathbf{P}) + \epsilon_k \qquad k \geq k_{min}, \tag{2.7}$$

where $f(k;\mathbf{P})$ is a known function with parameters set $\mathbf{P}$; C is an appropriate constant; and ϵ_k is the approximation error. Aiming for simplicity in our analysis we will not include any possible functional or irregular behavior of p_k for $k < k_{min}$ but their numerical values.

Assuming a set $\mathbf{P}^*$ such that

$$\sum_{k \geq k_{min}} \epsilon_k \to 0,$$

equations (2.5) and (2.6) can be rewritten in a single expression as

$$g(k_{min};\mathbf{P}^*) = \langle k \rangle + \Delta, \tag{2.8}$$

where

$$g(k_{min};\mathbf{P}^*) = \frac{\sum_{k \geq k_{min}} k f(k;\mathbf{P}^*)}{\sum_{k \geq k_{min}} f(k;\mathbf{P}^*)}, \tag{2.9}$$

is a function of k_{min} and the parameters set $\mathbf{P}^*$, and

$$\Delta = \frac{\sum_{k<k_{min}} (\langle k \rangle - k)\, p_k}{1 - \sum_{k<k_{min}} p_k}, \tag{2.10}$$

is a numerical factor. The parameters relation expressed in equation (2.8) holds for any network, independent of how it has grown. Its use requires knowing i) the numerical values of the degree distribution for $k < k_{min}$, and ii) the functional form $f(k;\mathbf{P})$.

Many authors have stressed the utility of using parameter relations to express the general behavior of network growth models [43, 49] or to compare numerical results with empirical data [16]. We will use equation (2.8) to infer relations between a network's degree distribution parameters set and the initial degree of new nodes.

2.2.2 Erdös-Rényi Networks

An Erdös-Rényi Network has a Poisson degree distribution for $k \geq 0$, so

$$p_k = e^{-z}\frac{z^k}{k!} \qquad k \geq 0. \tag{2.11}$$

Using equation (2.8) with $f(k;z) = e^{-z}\frac{z^k}{k!}$ and $k_{min} = 0$ directly produces the classic result

$$z = \langle k \rangle. \tag{2.12}$$

i.e. the degree distribution's parameter z of an Erdös-Rényi Network is equal to the average degree [35, III D].

2.2.3 Exponential Networks

According to equation (2.1), the degree distribution tail for an Exponential Network behaves as $p_k \propto e^{-k/k_o}$. Using equation (2.8) with $f(k;k_o) = e^{-k/k_o}$ then

$$g(k;\mathbf{P}) = \frac{1}{e^{1/k_o}-1} + k_{min} = \langle k\rangle + \Delta,$$

allowing us to solve for k_o

$$k_o = \left[\ln\left(\frac{\langle k\rangle + \Delta - k_{min} + 1}{\langle k\rangle + \Delta - k_{min}}\right)\right]^{-1}. \tag{2.13}$$

We can now estimate k_o for networks E_0 and E_1 using equation (2.13) and simulation data; as k_{min} is not specified in equation (2.13) we decided, as an initial creterion, assign to k_{min} the value m.

For E_0 : $\langle k\rangle \approx 2$, $k_{min} = 0$, $\Delta = 0$ and

$$k_o \simeq \left[\ln\left(\frac{2+0-0+1}{2+0-0}\right)\right]^{-1} = \left[\ln\left(\frac{3}{2}\right)\right]^{-1}.$$

For E_1 : $\langle k\rangle \approx 2$, $k_{min} = 1$, $\Delta = 0$ and

$$k_o \simeq \left[\ln\left(\frac{2+0-1+1}{2+0-1}\right)\right]^{-1} = [\ln 2]^{-1}.$$

These results agree with those of Callaway et al. [51], and Dorogovtsev & Mendes [20, section 8] for networks E_0 and E_1 respectively.

For $E_{1/2}$: $\langle k\rangle \approx 2$, $k_{min} = 1$. In this case $\Delta \neq 0$ and can be calculated with equation (2.10)

$$\Delta = \frac{\langle k\rangle p_0}{1-p_0} \simeq \frac{2(0.2)}{1-0.2} = 0.5,$$

so

$$k_o \simeq \left[\ln\left(\frac{2+0.5-1+1}{2+0.5-1}\right)\right]^{-1} = \left[\ln\left(\frac{5}{3}\right)\right]^{-1} \simeq 1.95(76),$$

which agrees with $\hat{k}_o = 1.95(55)$ using $k_{min} \geq 1$ for network $E_{1/2}$. As we can see in Figure 2.3, for exponential networks k_{min} is not uniquely determined but it can be chosen within a wide range of values above m without changing the value of (2.13).

To determine the dependence of k_o on initial degree, we simulate eleven Exponential Networks, $E_{i/10}$ $(i = 0, 1, \ldots, 10)$ using the parameter values $n_0 = 5, m_0 = 0, m = 5$. With probability $a = i/10$ a new edge is attached to the new node. Any new edge-end not attached to the new node attaches to a randomly selected node, including the new node. The final number of nodes was $N = 10^4$ and self loops were avoided.

The factor $\Delta - k_{min}$, a numerical value which depends on the non-tail portion of the degree distribution, is the only difference between networks $E_{i/10}$ when k_o is calculated from equation (2.13), while the average initial degree $\langle k_{ini} \rangle = m\,a$ is the only difference in the growth process. We used Figure 2.1 as a comparative plot to search for a trend between these terms.

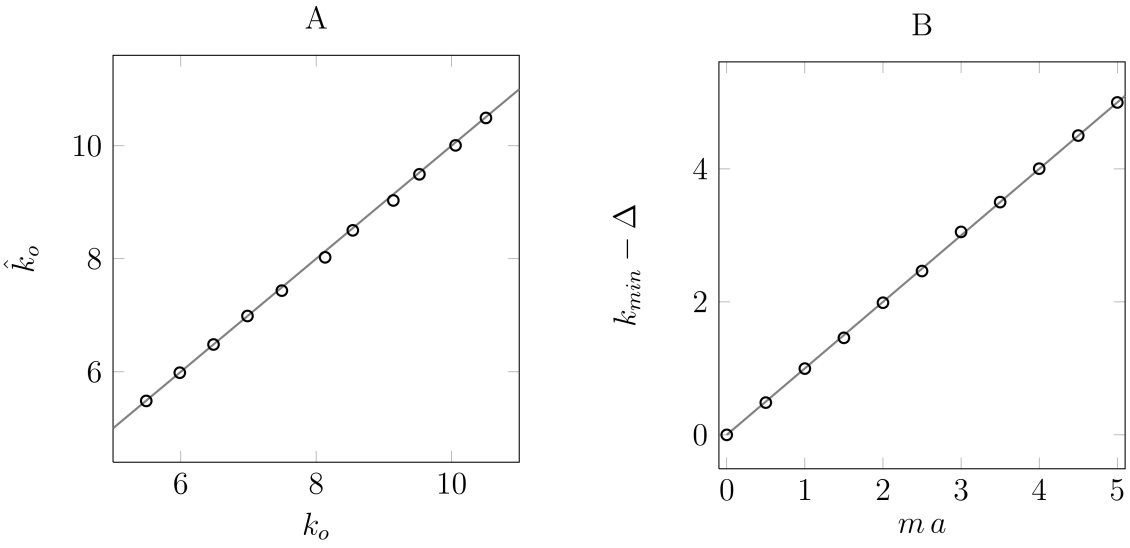

Figure 2.1: Comparative plot of degree distribution parameters in Exponential Networks. A) Open circles: calculated vs. fitted values for the parameter k_o in simulated Exponential Networks. The line corresponds to the identity $k_o = \hat{k}_o$; B) Open circles: numerical values of average initial degree $m\,a$ vs. $k_{min} - \Delta$, the line corresponds to the identity $m\,a = k_{min} - \Delta$ line.

Figure 2.1 A shows the relation between k_o and $\hat{k}_o$ (open circles) very close to the identity $\hat{k}_o = k_o$. Figure 2.1 B shows the relation between $m\,a$ and $k_{min} - \Delta$ (open circles) very close to the identity $k_{min} - \Delta = m\,a$. With these results we can set the *observational* relation

$$m\,a \simeq k_{min} - \Delta, \tag{2.14}$$

and use it to write an approximation to equation (2.13) as

$$k_o \simeq \left[\ln \left(\frac{\langle k \rangle - \langle k_{ini} \rangle + 1}{\langle k \rangle - \langle k_{ini} \rangle} \right) \right]^{-1}. \tag{2.15}$$

Equation (2.15) states a direct relation between average initial degree $\langle k_{ini} \rangle$ and the degree distribution parameters set of Exponential Networks. In the specific case $\langle k_{ini} \rangle = 0$ equation (2.15) reduces to the general result of Callaway et al. for $0 \leq \langle k \rangle \leq 2$ [51], although our analysis assumed a binomial distribution of initial degrees instead of a Kronecker delta distribution; it seems that if a initial degree distribution preserves the average then equation (2.15) remains the same.

To better understand these results we will draw a simple analytical model. Consider a network starting with n_0 nodes and m_0 edges at $t = 0$. It grows by adding a new node and m new edges in a single time step; as $k_{ini} \leq m$ edge-ends are initially attached to the new node then there are $2m - k_{ini}$ edge-ends to be allocated in accordance to a given set of attachment probabilities. We are interested in the long-run degree distribution p_k.

For exponential networks the attachment probability is

$$\Pi = \frac{1}{N(t)}. \tag{2.16}$$

The number of nodes with degree k_{ini} at time t, $n_{k_{ini}}(t)$, will increase by one at each time step, and decrease if one of them is attached to an edge-end. Thus

$$n_{k_{ini}}(t+1) = n_{k_{ini}}(t) + 1 - (2m - k_{ini}) n_{k_{ini}}(t)\,\Pi. \tag{2.17}$$

In contrast, for each $k > k_{ini}$ the number of nodes with degree k at time t, $n_k(t)$, increases if a node of degree $k-1$ is attached to an edge-end or decreases if a node of degree k is attached to an edge-end.

$$n_k(t+1) = n_k(t) + (2m - k_{ini}) n_{k-1}(t)\,\Pi - (2m - k_{ini}) n_k(t)\,\Pi. \tag{2.18}$$

In the long-run, these equations have the solution

$$n_k(t) = t p_k \qquad k \geq k_{ini}. \tag{2.19}$$

We substitute eqs. (2.16) and (2.19) into eqs. (2.17) and (2.18) and solve recursively

$$p_k = \frac{1}{2m - k_{ini} + 1} \exp\left(-k\left[\ln\left(\frac{2m - k_{ini} + 1}{2m - k_{ini}}\right)\right]\right) \qquad k \geq k_{ini}, \tag{2.20}$$

then

$$k_o = \left[\ln\left(\frac{\langle k\rangle - k_{ini} + 1}{\langle k\rangle - k_{ini}}\right)\right]^{-1}. \tag{2.21}$$

Notice that equation (2.15) has the same form as equation (2.21) but the former takes into account an averaged initial degree of new nodes rather than discrete values. Based on this result and those displayed in Figure 2.1, we consider that equation (2.15) is a generalization of equation (2.21): when initial degree is a constant value, i.e. a Kronecker delta, we recover equation (2.21) from (2.15).

2.2.4 Scale-Free Networks

A Scale-Free Network has a *shifted power-law* degree distribution tail represented by [43]

$$p_k \propto (k+\kappa)^{-\gamma} \qquad k \geq k_{min}. \tag{2.22}$$

Using equation (2.8) with $f(k;\kappa,\gamma) = (k+\kappa)^{-\gamma}$ the sum $g(k_{min};\mathbf{P})$ can not be expanded in a straightforward way. We can write [52, equation B.12]

$$\int_{k_{min}-1/2}^{\infty} (x+\kappa)^{-\gamma} dx = \sum_{k=k_{min}}^{\infty} (k+\kappa)^{-\gamma} + \frac{1}{24} \sum_{k=k_{min}}^{\infty} \frac{d^2}{dk^2}(k+\kappa)^{-\gamma} + \ldots,$$

$$\int_{k_{min}-1/2}^{\infty} x(x+\kappa)^{-\gamma} dx = \sum_{k=k_{min}}^{\infty} k(k+\kappa)^{-\gamma} + \frac{1}{24} \sum_{k=k_{min}}^{\infty} \frac{d^2}{dk^2} k(k+\kappa)^{-\gamma} + \ldots.$$

Assuming a negligible contribution of derivatives and setting aside terms that depend explicitly on k_{min} and Δ. Then equation (2.8) yields

$$\frac{(\gamma-2)\langle k\rangle - \kappa}{\gamma-1} \simeq k_{min} - \frac{1}{2} - \left(\frac{\gamma-2}{\gamma-1}\right)\Delta. \tag{2.23}$$

We simulate eleven Scale-Free networks, $S_{i/10}$ $(i = 0, 1, \ldots, 10)$ using the values $n_0 = 5, m_0 = 0, m = 5$. The growth procedure is the same as for exponential networks, with the addition of preferential attachment: the second end of an edge is attached to node i with probability

$$\Pi_i = \frac{k_i + 1}{\sum_j (k_j + 1)},$$

where k_i is the degree of node i. The term $+1$ is called initial attractiveness, $A \geq 0$ [50], and allows young/isolated nodes to acquire new links. The final number of nodes was $N = 10^4$ and self loops were avoided.

Equation (2.22) was fitted to the degree distributions of network $S_{i/10}$ for $k \geq 0$ and the values $(\hat{\gamma}, \hat{\kappa})$ substituted on the left side of equation (2.23); for all these cases we set $k_{min} = 0$ in order to get $\Delta = 0$ and then simplify our analysis. Figure 2.2 A shows the relation between parameters $(\hat{\gamma}, \hat{\kappa})$ and $m\,a$ (circles) to be very close to line

$$\frac{(\gamma-2)\langle k\rangle - \kappa}{\gamma-1} \simeq m\,a - \frac{1}{2}. \tag{2.24}$$

This is our first approximation for a relation between the degree distribution parameters set and initial degree in Scale-Free Networks. In conjunction with equation (2.23), equation (2.24) allow us to write the *observational* relation

$$m\,a \simeq k_{min} - \left(\frac{\gamma-2}{\gamma-1}\right)\Delta. \tag{2.25}$$

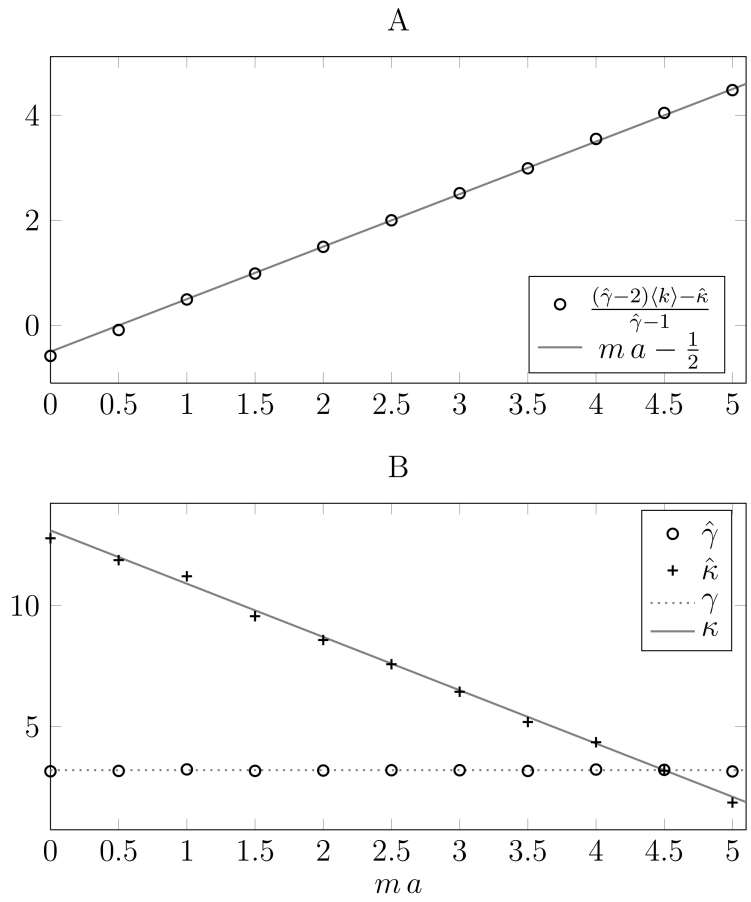

Figure 2.2: Parameter relation in Scale-Free Networks. A) Open circles: $\frac{(\hat{\gamma}-2)\langle k\rangle-\hat{\kappa}}{\hat{\gamma}-1}$ vs $m\,a$ for simulated Scale-Free Networks. Line: $\frac{(\gamma-2)\langle k\rangle-\kappa}{\gamma-1} = m\,a-\frac{1}{2}$ line; B) individual behavior of parameters $\hat{\gamma}$ and $\hat{\kappa}$ with average initial degree.

Figure 2.2 B shows the parameters set behavior of network $S_{i/10}$; $\hat{\gamma}$ is constant while $\hat{\kappa}$ decreases linearly with $m\,a$ for a constant $\langle k\rangle$. As can be read in reference [52] fitting power-law data is not trivial, and a distribution like that in equation (2.22) does not has a well-defined k_{min}. We also estimated γ and κ for $0 \leq k_{min} \leq 20$: the numeric value of left side in equation (2.23) converges to $ma-\frac{1}{2}$ as soon as $k_{min} \geq 5$. This is the same behavior as in figure 2 A), independently of k_{min}.

To better understand these results we will draw a simple analytical model (similar to that developed for Exponential networks) which includes a combination of preferential attachment (PA) and random attachment (RA). This combination has been addressed by a number of authors [53, 54, 55] and [56, section 5.3]. We will use a slightly different approach for our analysis. We begin with a network starting with n_0 nodes and m_0 edges at $t = 0$. This grows by adding a new node and m new edges in a single time step; k_{ini} edge-ends are initially attached to the new node. From the $2m - k_{ini}$ available edge-ends, K_{RA} will be allocated with an attachment probability given by equation (2.16), and the others K_{PA} will be

allocated with the attachment probability

$$\Pi_i = \frac{k_i + A}{\sum_j k_j + A}. \tag{2.26}$$

Using a continuous mean-field approach like that proposed by Barabási et al. [46], we can solve the equation of temporal change in node i's degree

$$\frac{\partial k_i}{\partial t} = \langle K_{PA}\rangle \Pi_i + \langle K_{RA}\rangle \Pi, \tag{2.27}$$

to obtain the degree distribution p_k for $t \gg m_0$. Clearly

$$\langle k_{ini}\rangle + \langle K_{PA}\rangle + \langle K_{RA}\rangle = \langle k\rangle = 2m. \tag{2.28}$$

The solution for p_k has the form given by equation (2.22) with

$$\gamma = 1 + \frac{\langle k\rangle + A}{\langle K_{PA}\rangle}, \tag{2.29}$$

$$\kappa = A + \frac{\langle K_{RA}\rangle}{\langle K_{PA}\rangle}(\langle k\rangle + A). \tag{2.30}$$

Substituting equations (2.29) and (2.30) into equation (2.28) and solving for $\langle k_{ini}\rangle$ produces

$$\frac{(\gamma-2)\langle k\rangle - \kappa}{\gamma - 1} = \langle k_{ini}\rangle, \tag{2.31}$$

which equals equation (2.24) if

$$\langle k_{ini}\rangle = m\,a - \frac{1}{2}. \tag{2.32}$$

i.e. as a consequence of discarding fluctuations in k_{ini} within the mean-field approach the average initial degree has a correction term $-1/2$.

2.2.5 Scale-Free – Exponential Roll-Off

Consider the previous mean-field results, if we define $\langle k_{ini}\rangle = ma - \frac{1}{2}$, $\langle K_{PA}\rangle = b(2m - \langle k_{ini}\rangle)$ and $\langle K_{RA}\rangle = (1-b)(2m - \langle k_{ini}\rangle)$ with $a, b \in [0,1]$ then equation (2.29) yields

$$\gamma = 1 + \frac{2m + A}{b(2m - ma + \frac{1}{2})} = 1 + \frac{2 + \frac{A}{m}}{b(2 - a + \frac{1}{2m})}, \tag{2.33}$$

diverging as $b \to 0$. In that case $\langle K_{PA}\rangle = 0$ and there is no preferential attachment. Then we obtain a network with exponential degree distribution as given by equation (2.1).

According to equation (2.33), as $b \to 0$

$$\frac{\gamma-2}{\gamma-1} = 1 - \frac{b(2-a+\frac{1}{2m})}{2+\frac{A}{m}} \to 1, \tag{2.34}$$

the *observational* relations (2.14) and (2.25) become the same (see figure 2.3), indicating an Scale-Free to Exponential roll-off.

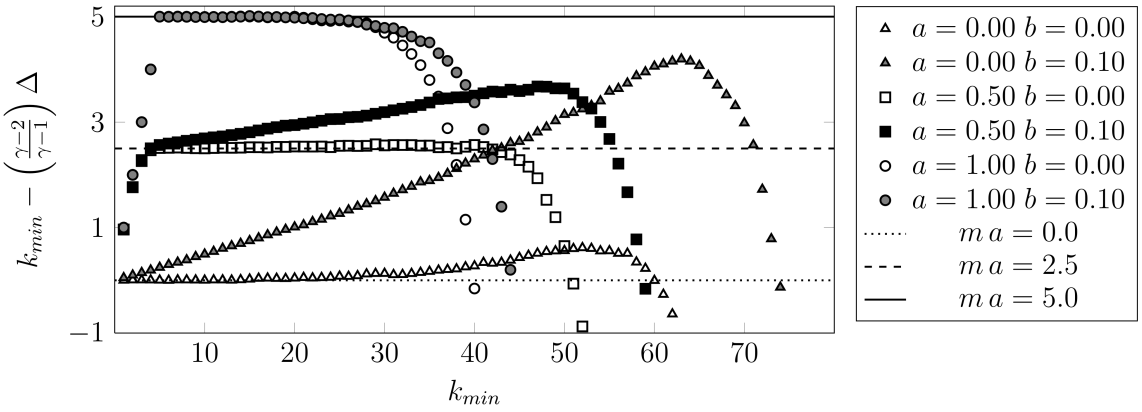

Figure 2.3: $k_{min} - \left(\frac{\gamma-2}{\gamma-1}\right)\Delta$ **for Exponential and Scale-Free** ($b \neq 0.00$) **Networks.** We used equations (2.10), (2.25), (2.34) and degree distribution data from simulated networks with different parameters $(a, b, m = 5)$ to appreciate the influence of k_{min}. This expression is almost constant over a wide range of k_{min} while shows a maximum as $a \to 0$ and $b \to 1$. This maximum is emphasized in Scale-Free Networks ($b > 0$) with $a \simeq 0$. $k_{min} - \left(\frac{\gamma-2}{\gamma-1}\right)\Delta$ fluctuates about $\langle k_{ini}\rangle = ma$ in Exponential Networks ($b = 0.00$) as well as in Scale-Free Networks when $a \simeq 1$.

This roll-off takes place when γ become large, the degree distribution decays fast, so a first order Taylor Series expansion of equation (2.22) about $k = 0$ is representative for (the most probable) low k values:

$$(k+\kappa)^{-\gamma} \simeq \kappa^{-\gamma} \mathrm{e}^{-\gamma k/\kappa} \qquad k \sim 0, \tag{2.35}$$

as $\gamma \to \infty$ the degree distribution at small k is exponential with parameter

$$k_o \simeq \frac{\kappa}{\gamma}. \tag{2.36}$$

Using equations (2.29) and (2.30) we can write

$$\frac{\kappa}{\gamma} = \frac{Ab(2-a+\frac{1}{2m}) + (1-b)(2-a+\frac{1}{2m})(2m+A)}{b(2-a+\frac{1}{2m}) + 2 + \frac{A}{m}} = k^*. \tag{2.37}$$

Figure 2.4 shows that equation (2.36) is valid for $b \sim 0$; as $b \to 0$ equation (2.37) converges to (2.15).

In the limit case of $a \to 1$, $b \to 0$ and $A \to 0$

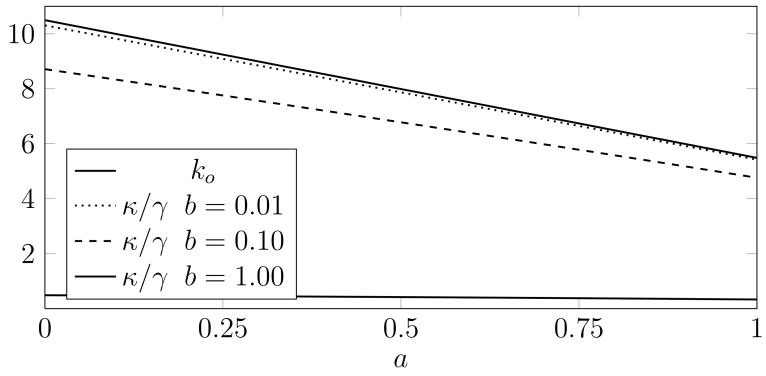

Figure 2.4: Parameters relation in the exponential approximation for the degree distribution of Scale-Free networks with $b \simeq 0$. According to equation (2.37) as $b \to 0$ the ratio κ/γ converges to k_o for every value of a, indicating a Scale-Free – Exponential roll-off.

$$k^* = m + \frac{1}{2}, \tag{2.38}$$

this value is similar to the mean-field solution of the Model A proposed by Barabási, Albert and Jeong [46]; from equation (2.15) we expected

$$k^* = k_o(a = 1.00) = \left[\ln\left(\frac{m+1}{m}\right)\right]^{-1}, \tag{2.39}$$

both solutions converge for $m \gg 1$.

2.3 Results

2.3.1 Simulation algorithm

We used Exponential and Scale-Free degree distributions in equation (2.8), as well as a series of network growth simulations, to obtain equations (2.15) and (2.24). These parameters set relations are consistent with those obtained by means of master equations and the mean-field approach (equations (2.21) and (2.31), respectively).

Equations (2.15) and (2.24) allow us to define an *effective* minimum degree in terms of average degree and the parameters in the degree distribution tail of a network, independent of how it has grown. We can use this effective minimum degree as an average initial degree to simulate a growing network with the same degree distribution of a real world network with average degree $\langle k \rangle$ whose degree distribution matches equations (2.1) or (2.22).

The algorithm requires a set of input parameters: number of nodes N, average degree $\langle k \rangle$, initial attractiveness A , Intitial attachment probability a, and Pref-

erential attachment probability b. Parameters $(N, \langle k \rangle, A, a, b)$ are required for simulating Scale-Free Networks; setting $b = 0$ we simulate Exponential Networks

Let

$$n_0 = \lceil \frac{\langle k \rangle}{2} \rceil,$$

$$m = \frac{\langle k \rangle}{2},$$

and r, s, u be random numbers borrowed from a uniform distribution $U(0, 1)$.

1. Start with n_0 isolated nodes.

2. Add one new node and L new edges, where

$$L = \begin{cases} n_0 - 1 & \text{if } r \leq n_0 - m, \\ n_0 & \text{if } r > n_0 - m, \end{cases}$$

   This ensures that in the long-run the average degree is $\langle k \rangle \simeq 2m$, which could be a non-integer quantity.

3. If $s < a$ attach one edge-end of a new edge to the new node. Do this with every new edge. This allows new nodes to have an initial degree in the range $[0, L]$ with average initial degree $\langle k_{ini} \rangle = ma$.

4. With probability

$$\Pi_i = \begin{cases} \frac{k_i + A}{\sum_j (k_j + A)} & \text{if } u \leq b, \\ \frac{1}{N(t)} & \text{if } u > b, \end{cases}$$

   attach every available edge-end to a randomly selected node i. Self loops are avoided.

   Repeat steps 2 to 4 $N - n_0$ times.

Parameters a and b can be estimated as follows by first fitting equations (2.1) or (2.22) to the real network degree distribution and then:

- using $\hat{k_o}$ as well as $\langle k \rangle$ in equation (2.15), or ($\hat{\gamma}$,$\hat{\kappa}$) as well as $\langle k \rangle$ in equation (2.24), and solving for a.

- using $\langle k \rangle$ and $\hat{\kappa}$ as well as an arbitrary $A \leq \hat{\kappa}$ in the relation

$$b = \frac{\langle k \rangle + A}{\langle k \rangle + \hat{\kappa}}. \tag{2.40}$$

Our model does not proposes any new linking mechanism but a change in the initial degree (with the parameter a) and a mixture of random and preferential attachment (with the parameter b). It offers an alternative to others models. Our model, such as that in reference [43], allow us to simulate exponential and scale-free networks, but in our case the roll-off is explicitly tracked by equations (2.36) and (2.37) when b goes to zero (see figure 4). For specific (a, b) values our model is equivalent to previous models:
- The model in [14] considers preferential attachment with constant initial degree equal to m. In our model this is equivalent to set $(a, b) = (1, 1)$.
- The model A in [46] considers random attachment with constant initial degree equal to m. In our model this is equivalent to set $(a, b) = (1, 0)$

We have demonstrated the usability of our model's algorithm by simulating Exponential and Scale-Free networks which reproduce the degree distribution of real world networks.

In the first example, we simulate the collaboration graph of movie actors (MAN), this is a Scale-Free network with $N^r = 212250$ nodes (actors), $M = 3045787$ edges, average degree $\langle k\rangle^r = 28.69$, $\gamma^r = 3.07$ and $\kappa^r = 31.68$ [43]. Using these values in equations (2.24) and (2.40) results in the parameter values $a = 0.002$ and $b = 0.49$ ($A = 1$, arbitrarily chosen): new nodes will have initial degree $k_{ini} \approx 0$ and about a half of the total edge-ends will be allocated using preferential attachment. The parameters set resulting from this simulation closely agrees with that of the real collaboration graph of movie actors (Table 2.1).

Table 2.1: Real and simulated parameters for the collaboration graph of movie actors (MAN). Parameter values with superscript r were used as input for our algorithm to simulate the collaboration graph of movie actors. Parameter values with superscript s were the result of our simulation over 20 runs.

N^r	$\langle k\rangle^r$	γ^r	κ^r	$\langle k\rangle^s$	$\hat{\gamma}^s$	$\hat{\kappa}^s$
212250	28.69	3.07	31.68	28.69	3.23	35.06

Although a majority of real world networks often exhibit fat-tail degree distributions there are real world networks with exponential degree distribution; for example, the North America Power Grid Network, $NAPGN$, [57]; the Email Network of the University Rovira i Virgili in Spain, $ENURV$, [58]; the Worldwide Marine Transportation Network, $WMTN$, [59]; and the Enmity Networks (*Enmity2* and *Enmity3*) documented in three elementary schools in Yucatan Mexico [13]. We simulated these Exponential Networks, the resulting parameters set is summarized in Table 2.2, and these resulting values closely agree with their real counterparts.

2.4 Conclusions and Discussion

From the perspective of complex networks, use of parameters for relaxing the restriction of constant initial degree to allow randomness, as well as the combina-

Table 2.2: Real and simulated parameters for real world Exponential Networks. Parameter values with superscript r were used as input for our algorithm to simulate Exponential Networks. Parameter values with superscript s were the average results of our simulations over 100 runs. Values in column 5 were calculated using the values in columns 3 and 4 in equation (2.15). *Data from [12].

Network	N^r	$\langle k \rangle^r$	k_o^r	a	$\langle k \rangle^s$	$\hat{k}_o^s$
Enmity2	226	2.24	2.10	0.53	2.22	2.15
Enmity3	419	2.86	2.55	0.54	2.85	2.58
NAPGN	14099	2.79	2.00	0.89	2.77	2.00
ENURV	1133	9.62*	9.20	0.19	9.57	9.34
WMTN	676	7.60*	7.20*	0.23	7.55	7.35

tion of preferential and random attachment, make our network growth algorithm flexible and applicable to a broad range of growth conditions. Parameter a identifies average initial degree and average degree as key factors in specifying the parameters set in the degree distribution of a growing network: most previous models implicitly rely on changing the average degree while keeping a constant initial degree.

The meaning of parameters a and b in our model could provide insight into network growth: parameter b represents the probability of setting an edge-end with preferential attachment; and parameter a arises from the need to allow a random initial degree within the network growth simulation. The latter parameter is susceptible to interpretation since it could be

- a measure of a new node efficiency to attach itself to the pre-existing network; no real process is 100 % efficient. This is a kind of new node fitness.

  For example, if a person is exposed to a new social network of friendship relationships, s/he could meet many people within the network but not necessarily establish a friendship relationship with all of them. In the most general case s/he only establishes a fraction a of all the possible links.

- a measure of network efficiency for the attachment of new nodes.

The characteristic represented by parameter a and the analysis of degree distribution behavior yields a fundamental innovation: the use of the full degree distribution. We demonstrate the relation between the tail's functional behavior $f(k, \mathbf{P})$ and the information from the non-tail portion, as represented by factors Δ and k_{min} in equation (2.8).

From a dynamical point of view, equations (2.15) and (2.24) state that *the tail behavior in the degree distribution of a growing network depends both on average degree and average initial degree.* In conjunction with equation (2.12), equations (2.15) and (2.24) offer valuable information about how these two averages directly influence the parameters set of a network in a growth process. Formation of a

Poisson network is caused by growth in the number of edges with a constant number of nodes over time, there is no average initial degree to take into account and the average degree is the only way to change parameter z. Equation (2.12) illustrates this dependence relation. An average initial degree exists as Exponential and Scale-Free Networks grow in number of nodes and edges, and it is needed to specify the degree distribution's parameters set. Equation (2.15) shows that average degree and average initial degree have an equal effect in changing the parameter k_o in the Exponential Network degree distribution. In contrast, average degree and average initial degree play different roles in changing parameters γ and κ in the Scale-Free Network.

Our results as expressed in equations (2.15) and (2.24) can be used to drive a growing non-static Exponential or Scale-Free network to a degree distribution with a particular parameters set. The resulting network's properties can be compared to those of real world networks of interest. This contrasts with the default procedure of using the Molloy-Reed algorithm [32] to produce a static random network with a given degree distribution. Our approach acknowledges the growing character of a network.

Part II

Network Dynamics

Chapter 3

Dynamical aspects

We study networks in order to better understand the systems they represent. The structural characterization of such networks is the first step toward this goal. However, the knowledge about how structure affects or influences the dynamics happening on networked systems is the next step. This knowledge would allow us to make predictions about the future behavior of a system, or even drive its dynamics.

Unfortunately, progress in this area has been slow. The whole idea of using networks is to bring back more realism to the over-simplified models we used in the attempt to describe systems in our world. This might mean adding an extra amount of difficulty to the problem at hand.

In Chapter 1, we deal with networks composed of nodes and links in an abstract sense. It is time to remember that nodes represent the *elements* or entities of a system, and the links represent the *interactions* among such elements.

3.1 Master equation

For the study of network dynamics we often use basic techniques, such as the Master Equation approach. However, a complete analytic solution using a Master Equation is not always achievable, even for very simple dynamical processes. Other techniques such as mean-field and continuous deterministic approximations often provide viable approaches to understand at least the basic features of the process under study [3].

The *Master Equation* consists of the evolution equation for the probability $P(\sigma, t)$ of finding the system at time t in a given configuration σ. This probability has to be normalized

$$\sum_{\sigma} P(\sigma, t) = 1.$$

In the continuous time approximation the master equation is

$$\frac{\partial P(\sigma, t)}{\partial t} = \sum_{\sigma'} \left[P(\sigma', t) W(\sigma' \to \sigma) - P(\sigma, t) W(\sigma \to \sigma') \right]. \tag{3.1}$$

The two terms in the right side represent, respectively, the gain and loss contributions for the probability distribution of the system in a given configuration σ. $P(\sigma, t)$ in principle allows us to calculate average values at time t for any function of the system's state $A(\sigma)$

$$\langle A(t) \rangle = \sum_{\sigma} A(\sigma) P(\sigma, t).$$

As the complete solution of the Master Equation is rarely achievable, we rely on approximation schemes such as projections over specific quantities of interest. For example

$$N_x(t) = \sum_{\sigma} \sum_{i} \delta_{\sigma_i, x} P(\sigma, t), \tag{3.2}$$

represents the average number of nodes in state x at time t ($\delta_{i,j}$ is the Kroneker delta function). This equation is deterministic and does not include fluctuations.

In conjunction with the mean field (MF) approach we can write a set of equations

$$\frac{\partial N_x(t)}{\partial t} = F_x(N_1, N_2, \ldots, N_x), \tag{3.3}$$

where $x = 1, 2, \ldots, e$ indexes the available states of a node, and the functional form of F_x depends on the specific interactions among nodes.

This approach was used by us in two specific applications:

- We used recurrence relations over projections for the calculation of the general degree distribution p_k of grown networks with random attachment and variable initial degree in section 2.2.3, and
- we used mean field for the description of known relations for the epidemiological SI, SIS, SIR and SIRS models in Chapter 6.

3.2 Network embedding

Networked systems are either tangible such as the internet, or intangible such as the Word Wide Web. Evidently, tangible networks are embedded into a physical space. This embedding imposes physical constrains to either structure and dynamics of a network. Thus, spatial dimension plays an important role in both network evolution and network dynamics.

Let us take an example from the study of urban street networks, where links are streets and nodes are street intersections. In this case, it is known that street networks are planar structures, i.e. its graph can be drawn on a plane without any link crossing. Other planar networks are: power grid networks, the US Natural Gas Pipeline Compressor Stations, and the neighboring countries networks from a map [17].

Biological systems also show consequences of the physical embedding. The neural network of the worm C. elegans, where nodes are neurons and links are synaptic connections between neurons, is constrained by the small volume of its anatomy. When thinking of this network as a signal processing circuit, to avoid signal interference due to link crossing, a multilayer structure composed of planar sub-networks could be a suitable evolving strategy. This non-planar multilayer structure is present in the C. elegans neural network, and in the neocortex of mammals brains, with neurons distributed among 6 layers labelled from I to VI [60, Chapter 3 and references there in].

Another kind of physical restrictions exists in the brain: neurons and brain regions physically close have higher probability to be connected while connections between remote components are less likely. This because long axonal projections are expensive in terms on building materials and energy cost, so it is common for neuronal networks to optimize wiring via the small-world topology, or to show complex networks features such as highly connected hubs and modularity [61]. Anatomical studies of cortical connectivity, performed by Staiger on 58 pyramidal cells of the mouse cortex, report exponential distributions of the lengths of axon collaterals and projections [62, chapter 17], however, a re-display of his data in log-log scale shows power-law distribution of distances, $p_l \propto l^{-\alpha}$, with a decay exponent $\alpha = 2.2$ [63, chapter 3].

Chapter 4

Network Resilience

Networked systems are networked for a purpose. The interactions of their elements allow the realization of some process: the network is functional. One of the fundamental issues is to know the structural conditions of a functional network. In other words, we can ask which structural conditions break the functionality of a network. In the first place, the occurrence of some processes over a network requires the existence of a connected structure. It is evident that these processes cannot occur without connectivity. But we still don't known what happens in a partially connected structure.

Problems as the previous occurs all the time in a network that is part of the world's daily life, the internet: At any moment a fraction as high as 3 % of all internet routers are non-functional, however we are still able to use it. On the other hand, the failure of some specific routers may cause large areas to become disconnected. It is desirable for us to known why this is so, and how to avoid it.

Another possibility is to be able to use failure to stop unwanted processes in a network such as diseases propagation: the knowledge about which structural damages would stop the diffusion of a disease potentially allows health authorities to design efficient vaccination strategies, i.e. the vaccination can be directed to individuals with specific characteristics in order to cause the largest possible delay of the diffusion process.

Suppose that our initial network $G = \{V, E\}$ has N nodes and M links, and it is functional. If some subset of V and/or E fails or are removed, we still have a network structure. However we don't know if the network is still functional.

This chapter is devoted to analyzing this problem from a purely structural perspective first, and then including qualitative features of the removed elements of G. As the situation we just described is a damaging process, this can be analyzed (at least) using the framework of *percolation*.

4.1 Percolation

The percolation problem can be simple stated as follows: in a connected network of arbitrary topology, each node is either present and functional with probability

p, or absent and non-functional with probability $f = 1 - p$. p is the *ocupation probability*, and f is the *failure probability*. Two nodes are connected if they are neighbors in the underlying network and are both occupied.

At $p = 0$ there are no nodes in the network. Increasing p from zero induces the nucleation and growth of small sub-graphs called *clusters* of increasing size. These clusters are mostly tree-like when $p < p_c$. At some value of the occupation probability, p_c, denoted as the *percolation threshold* a giant cluster appears. As p increases from p_c the giant cluster coalesces with other clusters until it encloses the whole network at $p = 1$, figure 4.1. The creation or dilution of the giant cluster is a *percolation transition*. A network with a giant cluster is said to *percolate*.

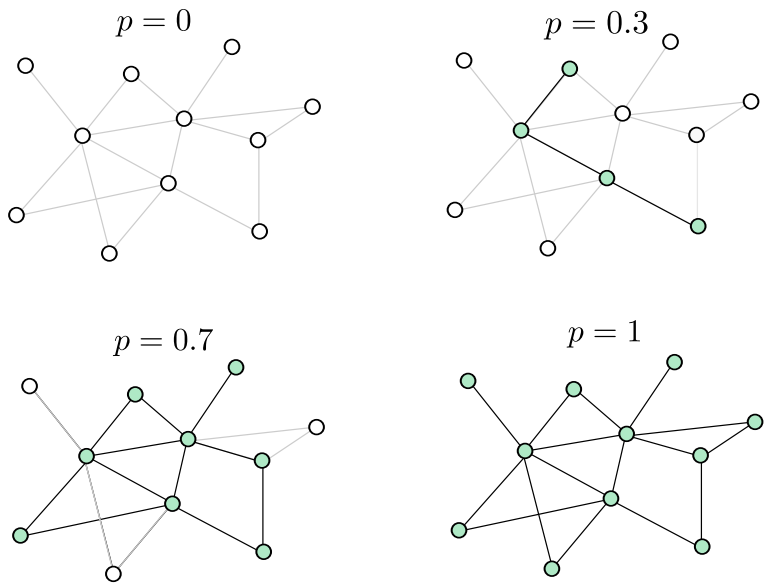

Figure 4.1: Graphical representation of a percolation process. Open circles are absent nodes, and gray links are removed links associated to absent nodes.

4.1.1 Percolation in arbitrary topologies

For simplicity we assume that the original network has degree distribution p_k, was created with the configuration model and is uncorrelated. Let u be the probability that a node is not connected to the giant cluster via a particular neighbor. For a node of degree k, the total probability of not belonging to the giant cluster is u^k. The probability of a random node to not belong to the giant cluster is

$$g_0(u) = \sum_k p_k u^k. \tag{4.1}$$

The fraction of nodes that are in the giant cluster S is equal to the fraction p that have not been removed times the probability of being in the giant cluster

$$S = p\,(1 - g_0(u)) \tag{4.2}$$

The probability u has two contributions: i) the neighbor in the giant component has been removed, with probability $1-p$, and ii) and the probability that none of the neighbors is present but not connects to the giant cluster, pu^k. Then the total probability that we are not connected to the giant cluster via a given neighbor is $(1-p)+pu^k$.

As a neighbor is reached following a link, degree is distributed according to the *excess degree distribution*

$$q_k = \frac{(k+1)p_{k+1}}{\langle k \rangle}, \tag{4.3}$$

and u is defined by the self consistence equation

$$u = \sum_k (1-p+pu^k)q_k = 1-p+p\sum_k u^k q_k = 1-p+pg_1(u). \tag{4.4}$$

The form of $g_1(u)$ depends of the form on p_k. However, we know that $g_1(u)$ is a polynomial with non-negative coefficients (because they are probabilities). All its derivatives of are non-negative for $u \geq 0$. Then, $g_1(u)$ is an increasing function of u and concave upward on the interval $u \in [0,1]$. On the other hand, u is an increasing function of u for any value of u.

Equation (4.4) has a trivial solution in $u = 1$ because $g_1(1) = 1$ (if q_k is correctly normalized). The other non-trivial solution occurs when u and $1-p+pg_1(u)$ meet tangently or intersect at some $u < 1$.

The percolation threshold reads

$$\frac{d}{du}(1-p+pg_1(u))\bigg|_{u=1} = 0, \tag{4.5}$$

yielding

$$p_c = \frac{1}{g_1'(1)}. \tag{4.6}$$

Knowing the definition of q_k

$$g_1'(1) = \frac{1}{\langle k \rangle}\sum_k k(k+1)p_{k+1} = \frac{1}{\langle k \rangle}\sum_k k(k-1)p_k = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}, \tag{4.7}$$

then, the percolation threshold is

$$p_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} = \frac{1}{\kappa - 1}. \tag{4.8}$$

4.1.2 Average component size

In percolation, the order parameter is the probability that an occupied node belongs to the giant cluster, $S = N_G/N$, where N_G is the size of the giant cluster. Small clusters are characterized by the cluster number distribution $n_s(p) = N_s(p)/N$.

As the probability for any node to belong to a cluster of size s is $sn_s(p)$, the occupation probability can be written as the sum of these probabilities. For $p < p_c$

$$p = \sum_s sn_s(p). \tag{4.9}$$

Above the percolation threshold the giant clusters appears and any node has a probability S to belong to it, then

$$p = S + \sum_s{}' sn_s(p). \tag{4.10}$$

where the primed summation excludes the giant cluster.

The cluster number distribution allows to write the conditional probability that an occupied node belongs to a cluster of size s, $sn_s(p)/\sum_s{}' sn_s(p)$.

As we increase the occupation probability (below p_c) $\langle s\rangle$ increases until one of the clusters becomes the giant cluster at the percolation threshold p_c. Further increases in p will cause $\langle s\rangle$ to decrease, because the giant cluster grows by coalescence with the larger small clusters. In fact $\langle s\rangle$ is singular at p_c and this singularity is the fingerprint of a critical phase transition: at p_c the system lacks a characteristic cluster size, giving rise to the scaling

$$n_s(p) = \begin{cases} s^{-\tau} f_+\left(\frac{s}{s_c}\right) & \text{, if } p \geq p_c \\ s^{-\tau} f_-\left(\frac{s}{s_c}\right) & \text{, if } p \leq p_c \end{cases} \text{, with } s_c = |p_c - p|^{-1/\sigma} \tag{4.11}$$

where the scaling functions $f_\pm$ are continuous, equal at 0, and have a fast decay. Exponents τ and σ depend on the dimensionality and other properties of the system. s_c plays the role of size cut-off.

From the scaling of $n_s(p)$ we obtain

$$\langle s\rangle \sim |p_c - p|^{-\gamma} \tag{4.12}$$

$$S \sim (p - p_c)^\beta \tag{4.13}$$

where the exponents are $\gamma = (3-\tau)/\sigma$ and $\beta = (\tau - 2)/\sigma$. Only two critical indices are independent in percolation.

4.2 Percolation in Complex Networks

Percolation is generally considered on topologies which are regular lattices embedded in a D-dimensional space. However, in a random network with N nodes there is no embedding space: any node can (in principle) be connected to any of the other $N-1$ nodes. Thus, in the limit $N \to \infty$ percolation in complex network is infinite-dimensional.

At this point we are in position to define the problem of network tolerance to "failure" or *network resilience*, which we described at the beginning of the present

chapter. It is *inverse percolation*, because we are interested in the characteristics of the system as a function of the failure probability $f = 1 - p$, i.e. we want to see how the network is diluted rater than built.

One final distinction can be made: nodes and edges can be removed at random or in a deterministic way.

4.2.1 Failure Tolerance

Random deletion of nodes (and its incident edges) corresponds to classical *site percolation*, analyzed in section 4.1. This kind of percolation will be an approximation to describe networks' *failure tolerance*, i.e. how resilient networks are to random errors (non-deterministic failures). We specifically focus in two paradigmatic types of networks: random networks and scale-free networks.

A random network with Poisson degree distribution has $\kappa = \langle k \rangle + 1$, then

$$f_c^{rn} = 1 - \frac{1}{\langle k \rangle}, \tag{4.14}$$

which means that in a random network with finite average degree there always exists a finite percolation threshold. See the left side of figure 4.2 .

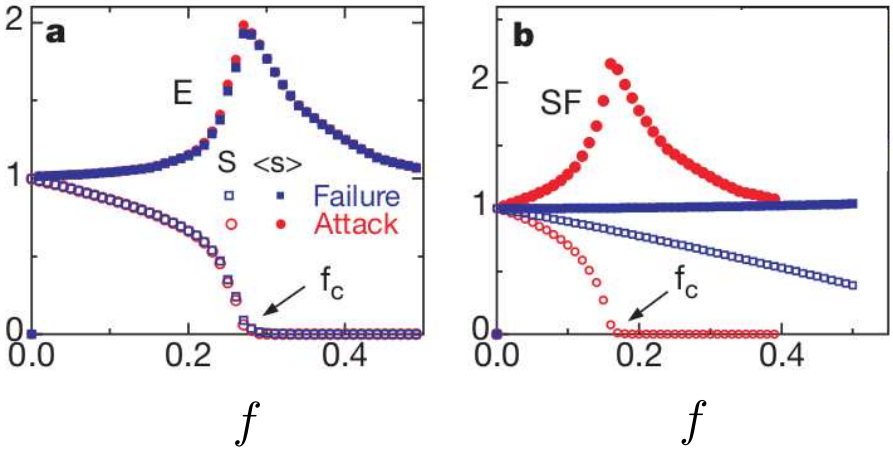

Figure 4.2: Networks under error and targeted removal of nodes. The parameter registered are the size of the giant component S (open symbols) and the average cluster size (filled symbols) as a function of the removed fraction f. In the left side we observe the case of a random network with $N = 10,000$ and $\langle k \rangle = 4$. In the right side the response of a scale-free network with $N = 10000$ and $\langle k \rangle = 4$. The value of γ is not available. Taken from [2].

A scale-free network with $p_k \sim k^{-\gamma}, k = m, m+1, \ldots, k_c(N)$ has [64]

$$\kappa = \frac{2-\gamma}{3-\gamma} \frac{k_c(N)^{3-\gamma} - m^{3-\gamma}}{k_c(N)^{2-\gamma} - m^{2-\gamma}}, \tag{4.15}$$

where $k_c(N)$ is the maximum degree observed in a network of size N

$$k_c(N) = mN^{1/(\gamma-1)}. \tag{4.16}$$

The percolation threshold of a scale-free network varies in accordance to the values of γ

- if $\gamma < 3$, κ diverges as $N \to \infty$: $\kappa \sim N^{(3-\gamma)/(\gamma-1)}$ if $2 < \gamma < 3$, and $\kappa \sim N^{1/(\gamma-1)}$ if $\gamma < 2$. The percolation threshold becomes closer to 1 as N increases.
- if $\gamma > 3$, κ is finite ($\sim m(\gamma-2)/(\gamma-3)$). The percolation threshold is less than 1.

In the range $2 < \gamma < 3$, where many real world scale-free networks lie

$$f_c^{sf} \approx 1 - \frac{3-\gamma}{2-\gamma} m^{2-\gamma} k_c(N)^{\gamma-3}, \tag{4.17}$$

close to 1, even for small networks: for $m = 1, N = 1000, \gamma = 2.5$ one obtains $f_c^{sf} \approx 0.9$. In other words, a scale-free network in the range $2 < \gamma < 3$ will show large sizes of the giant cluster even when a large number of nodes has been (randomly) removed. Se the right side of figure 4.2 [2].

4.2.2 Targeted removal tolerance

Nodes in a network may fail naturally by error. However, other alternative is an intentional (deterministic) removal of nodes, i.e. an *attack*. Such attack could be guided by a *centrality measure* to select target nodes. The simplest centrality measure is node degree, i.e, the most connected node(s) will be removed first, then the second one(s), and so on.

This situation is not so different from that in section 4.1. One simple observation allows us to follow a similar approach.

The initial degree distribution of the network is p_k. After removing a fraction f of the highest degree nodes, the highest degree node remaining has degree $k_c(f)$ implicitly defined by

$$f = \sum_{k_c(f)+1}^{\infty} p_k. \tag{4.18}$$

The probability that one neighbor of any node is removed, $r(f)$, equal the probability that such neighbor has degree larger than $k_c(f)$

$$r(f) = \sum_{k_c(f)+1}^{\infty} \frac{k p_k}{\langle k \rangle} \tag{4.19}$$

The result is a (damaged) network with maximum degree $k_c(f)$, where the neighbor of any node was removed with probability $r(f)$. The degree distribution

of this network in the range $k_{min} \leq k \leq k_c(f)$ is just p_k scaled by the factor $1/(1-f)$. Thus, the giant cluster disappears when [64]

$$r(f_c) = 1 - \frac{1}{\kappa(f_c) - 1}, \tag{4.20}$$

where

$$\kappa(f_c) = \frac{\sum_k^{k_c(f_c)} k^2 p_k}{\sum_k^{k_c(f_c)} k p_k}. \tag{4.21}$$

Using the continuous degree approximation one obtains a value for the heterogeneity parameter $\kappa(f_c)$ similar in form to equation 4.15, so that [64]

$$f^{(2-\gamma)/(1-\gamma)} = 2 + \frac{2-\gamma}{3-\gamma} m (f^{(2-\gamma)/(1-\gamma)} - 1). \tag{4.22}$$

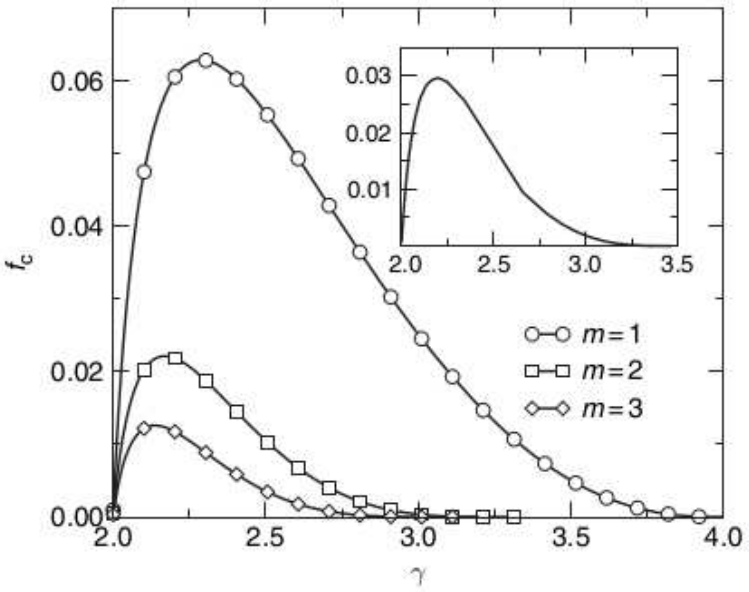

Figure 4.3: Percolation threshold vs. γ for scale-free networks in the continuous approximation. The inset show the same data, from an exact discrete approach. Taken from [3, section 6.5].

The behavior of the removed fraction f_c as a function of γ is show in figure 4.3 [65]. The inset of figure 4.3 show the same data obtained from an exact discrete approach which reveals a qualitatively similar but down scaled behavior of f_c [65]. This result is of great relevance: scale-free (or fat-tail degree distributed networks) appear to be highly fragile to targeted removal of its highest degree nodes, or hubs. Such behavior is notably opposed to that of random removal of nodes in this type of networks. This facts has motivated the assertion that scale-free networks are robust against random failure yet fragile to attack targeted at their highest degree nodes.

It is worth noting that robustness against random failure does not mean "invulnerability" to random failure: although the giant cluster of a scale-free network decreases continuously as the removed fraction increases, through the process the network is still being segmented into an increasing number of small clusters. It is not clear, however, when the functionality will be lost.

Chapter 5

Synchronization

When functional, the components of a networked system are performing *collectively*. If we take connectedness for granted it is clear that the elements are not isolated, they should be able to communicate with one another in some way. Under special circumstances, a subset of the elements follow common dynamical patterns, and a **synchronized** state emerges.

Synchronization is an ubiquitous phenomena in biological, ecological, climatological, sociological, technological and network systems: a large number of electrons march in lockstep in a superconductor allowing a current flow with zero resistance; in the tidal river's mangroves of Malaysia thousands of fireflies flash in unison without any leader. Synchronization is pervasive in nature at every scale [66, 67].

It is part of our every day life, even if we cannot perceive its role: the laser in a today simple cd/dvd/BluRay player is the result of trillions of atoms emitting light waves in synchrony. The same is true for every laser technology such as laser pointers, laser surgery devices, laser levers, lidars, laser sensors, high accuracy optical instrumentation, optic mice, laser scanners, and a very large etcetera.

Synchronization is also importance for our health at the individuals and population level, this because the presence of waves in excitable media requires some level of synchronization

- in cardiac tissue, abnormal patterns of excitation/wave behavior are believed to be caused by pathological conditions such as tachycardia (abnormally fast heart beat) and fibrillation (irregular heart contractions unable to properly pump the blood stream). Without early intervention both conditions are highly lethal.

- in the human brain, synchronization plays an important role in both functioning and dysfunctioning; in the case of human epileptic brain networks, epileptic seizures are usually characterized by an abnormal synchronized firing of neurons related to the epileptic focus[68].

- in the field of epidemiology, synchronization is relevant to describe/forecast the strength of epidemics outbreaks associated with recurrent infectious diseases showing oscillatory behavior [8] such as measles, typhus and cholera

[69]. In this scenario, higher level of synchronization leads to higher-amplitude oscillations in the number of infected individuals.

Clearly, synchronization depends on the nature of the individual oscillators, on the form they interact and the structural pattern of such interaction. The last factor is the one in which we are interested in this thesis.

5.1 General Framework

Fireflies, pacemaker cells, neurons and individuals exposed to a recurrent disease (among many other instances) can be considered as oscillators: all of them display a cyclic behavior, repeating processes over and over again at (more or less) regular time intervals. This is one reason why the study of synchronization in largely devoted to *coupled oscillators*, a collection of oscillators that can influence one another in some physical or chemical way. The central issue in the study of coupled oscillators concerns the emergence of coherent collective behavior, in which the oscillators follow the same dynamical patterns, this is what we call **synchronization**.

Let consider N interacting oscillators. Each oscillator is described by an internal degree of freedom $\phi_i(t)$. This degree of freedom evolves due to i) the internal dynamics of the oscillator and ii) the coupling with other oscillators:

$$\frac{d\phi_i}{dt} = f_i(\{\phi_i\}). \tag{5.1}$$

When isolated, at large times, each oscillator reach either a stable fixed point, a limit cycle or a strange attractor. Two oscillators i and j are coupled through a direct link from j to i, if the evolution equation of i depends on ϕ_j. This definition of coupling leads to a directed coupling network, however, we will review some result of non directed (symmetric) coupling networks.

Various types of synchronization are possible:

- *Complete synchronization*: occurs when all internal variables are equal, i.e. $\phi_i(t) = s(t) \ \ \forall i$.
- *Phase synchronization*: consists in the phase locking of oscillators described by a phase and an amplitude.
- *Generalized synchronization*: two oscillators are in general synchronization if the output of one is equal to certain function of the others' output.

More complex situations such as a sequence of synchronized and non-synchronized behaviors are also possible.

5.2 Identical and linearly coupled oscillators

For a system of linearly coupled oscillators, each unit interacts with its neighbors through a linear superposition of their outputs

$$\frac{d\phi_i}{dt} = F(\phi_i) + \sigma \sum_{j=1}^{N} C_{ij} H(\phi_j), \tag{5.2}$$

where the functions F and H are the same for all oscillators, σ is the interaction strength, and C_{ij} is the coupling matrix. If the coupling between two oscillators depends only on the difference of their outputs

$$\frac{d\phi_i}{dt} = F(\phi_i) + \sigma \sum_{j \in v(i)} \left[H(\phi_i) - H(\phi_j) \right], \tag{5.3}$$

which corresponds to the coupling $C_{ij} = L_{ij}$, where L is the Laplacian matrix of the interaction network.

Let $s(t)$ be the solution of the uncoupled system ($\sigma = 0$). This is also a solution in the case $\sigma \neq 0$ [1], and its stability is studied with the *master stability function* approach: close to the synchronized state $\phi_i = s + \xi_i$, with $\xi_i \ll s$. The synchronized state is stable if the perturbations decrease, otherwise it is unstable. Expanding $F(\phi_i) \approx F(s) + F'(s)\xi_i$ and $H(\phi_i) \approx H(s) + H'(s)\xi_i$, where the primes indicate derivatives w.r.t. s. The evolution of s is given by

$$\frac{d\xi_i}{dt} = F'(s)\xi_i + \sigma \sum_{j} L_{ij} H(s) \xi_j. \tag{5.4}$$

As L is symmetric and has zero-sum columns, it has N real non-negative eigenvalues and one is zero, they are (in ascending order) $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{max}$. Using the eigenvalues ζ_i we can decouple 5.4

$$\frac{d\zeta_i}{dt} = \left[F'(s) + \sigma \lambda_i H'(s) \right] \zeta_j. \tag{5.5}$$

At short times, and assuming that s have negligible variations one obtains

$$\zeta_i(t) = \zeta_i^0 \exp\left(\left[F'(s) + \sigma \lambda_i H'(s) \right] t \right), \tag{5.6}$$

where ζ_i^0 is the initial perturbation. Then, perturbations will either decrease or increase exponentially depending on the sign of $\Lambda_i(\alpha) = F'(s) + \alpha H'(s)$ with $\alpha = \sigma\lambda_i$. The null eigenvalue $\lambda_i = 0$ yields $\Lambda_1 = F'(s)$, and is related to the evolution of an isolated oscillator. The state s is stable only if all the other perturbations decay, i.e. all other Λ_i are negative. Barahona and Pecora [70] showed that for a large class of functions F and H the master stability equation $\Lambda(\alpha) = \max_s (F'(s) + \alpha H'(s))$ is negative within a range of α, $[\alpha_1, \alpha_2]$, as shown in

[1] Any element in the summation of the right side will cancel when counted with swapped indexes, as in $\sum_i \left[\frac{d\phi_i}{dt} - F(\phi_i) \right] = \sum_i \left[\sigma \sum_{j \in v(i)} \left[H(\phi_i) - H(\phi_j) \right] \right]$.

figure 5.1. In these cases the coupling strength should be larger than a minimum, but not arbitrarily high. Otherwise, synchronized states loose stability (we got $\Lambda(\alpha) > 0$). The condition for stable synchronization is then

$$\frac{\lambda_{max}}{\lambda_2} < \frac{\alpha_2}{\alpha_1}, \tag{5.7}$$

i.e. networks exhibits better synchronizability if the eigenratio λ_{max}/λ_2 is as small as possible. Notably, the left hand side of the inequality 5.7 depends only on structural values, while the right hand side depends on F and H, which are properties of the oscillators.

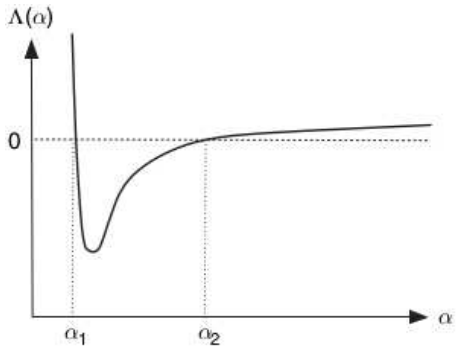

Figure 5.1: Sketch of a typical master stability equation. The synchronized state is stable if for all λ_i, $\sigma\lambda_i \in [\alpha_1, \alpha_2]$. Taken from [3, section 7.2].

5.2.1 Structurally enhanced synchronization

The small-world network model of Watts and Strogatz [25] (seen in subsection 1.2.3) was motivated by something more that obtaining networks with small average path length and high clustering coefficient *per se*: they were interested in ways to enhance synchronization of oscillators by changing the coupling structure. Their model *phenomenologically* associates small-world networks to a "better" propensity to synchronization than lattice couplings. As we have seen, much work was then aimed at analytically describing the properties of small-world networks, from the perspective of structural characterization.

Barahona and Pecora [70] were the first to analytically investigate how the structural properties of small-world networks influences synchronization. Figure 5.1 and inequality 5.7 allow to formulate a very general conclusion: coupling networks with smaller eigenratios λ_{max}/λ_2 will enhance synchronization, independently of the nature of the oscillators they couple.

Barahona and Pecora [70] used a modified small-world model to investigate how the eigenratio behaves. The initial network is a (deterministic) ring of N nodes, each coupled to its $2m$ nearest neighbors. Then they add new randomly wired edges. This way, the *connectance* of the network, $D = 2E/(N(N-1))$, increases.

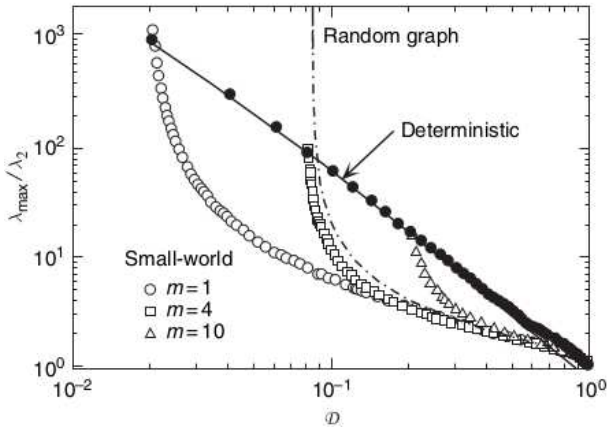

Figure 5.2: Evolution of the eigenratio λ_{max}/λ_2 as a function of the connectance D in small-world networks with $N = 100$ and several values of m. Filled circles correspond to deterministic rings, the dashed line corresponds to random networks. Taken from [3, section 7.2].

From figure 5.2 we see that increasing connectace improves synchronization by lowering the eigenratio value. Small-worlds shown better synchronization propensity than deterministic rings (or lattices). Such evidence sets a firm ground to assert that *the small-world topology is a viable route to enhance synchronization.*

It is tempting to generalize that small average path lengths promotes better communication than larger ones, improving synchronization. Such picture would agree with data in figure 5.2: both, small-world networks and random networks possess the small-world property (i.e. small average path length).

Nishikawa et al. [71] showed that this generalization is not possible. Small average path lengths can also be obtained by growing scale-free networks with degree distribution $p_k \sim k^{-\gamma}$ and small values of γ. Their work pointed out that as γ decreases the eigenratio raises, making synchronization more difficult. Further analysis reveals that this behavior is not characteristic of scale-free networks, it is observed in any network with *hubs*: the eigenratio increases as the number of hubs (which concentrates the same number of edges) is lowered. They concluded that *as a network becomes more heterogeneous, it also becomes less synchronizable yet a small-world*, a fact denoted as *the heterogeneity paradox* [72].

One intuitive explanation for the heterogeneity paradox is the overload of hubs, whose inputs may cancel out when having different phases or frequencies. Motter et al. [72] showed that this paradox can be overcame if a modified coupling strength is used

$$\sigma \to \frac{\sigma}{k_i}, \tag{5.8}$$

which is (in some sense) a coupling homogenization: the more inputs an oscillator has, the weaker they are.

Another work indicates that the eigenratio evolution closely resembles that of

betweeness centrality [73] (*betweeness centrality* is defined as the proportion of information that passes through a given node in communicating other pairs in the network [17, Section 7.4]) and that synchronization is more difficult as the clustering coefficient increases [74], a characteristic trait of small-world networks.

5.3 Integrate-and-fire oscillators

Many systems of biological oscillators present interactions that can be described as episodic or pulse-like: each oscillator accumulates pulsatil perturbations from other oscillators. When a threshold is reached, it fires a pulse which perturbs its neighbors. Then, their perturbation accumulating property is reset, and the process is repeated. An oscillator following this pattern of behavior is dubbed as an *integrate-and-fire oscillator*.

The paradigmatic example of integrate-and-fire systems are neural networks: neurons communicate to each other by firing sudden impulses called spikes. In this sense, the system is *pulse-coupled*.

A simple yet general integrate-and-fire model was developed by Mirollo and Strogatz [75]. Each oscillator i is characterized by a state variable $x_i = U(\phi_i)$, where, $x_i \in [0, 1]$ and ϕ_i is the oscillator's phase. The *Phase Response Function* U is a monotonically-increasing concave function which satisfies $U(0) = 0$ and $U(1) = 1$. When oscillator i reaches the firing threshold at t_i it fires and its state variable is reseted, i.e. $x_i(t_i^+) \to 0$. The state variables of the other oscillators are updated as

$$x_j(t_i^+) = \min(1, x_j(t_i) + \epsilon_{ji}) \qquad j \neq i, \tag{5.9}$$

where ϵ_{ji} is the coupling strength between i and j. The interaction is *excitatory* if $\epsilon_{ji} > 0$, or *inhibitory* if $\epsilon_{ji} < 0$, see figure 5.3.

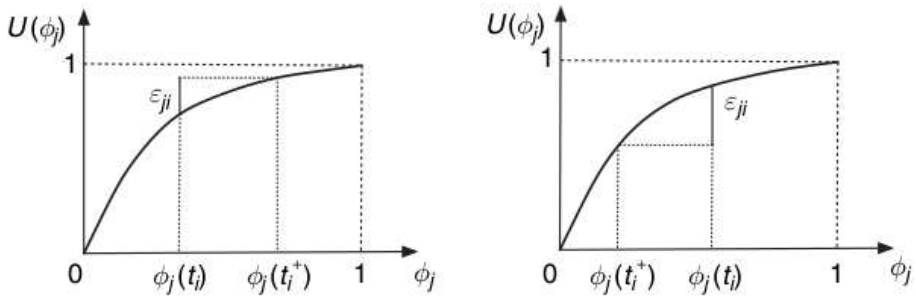

Figure 5.3: Phase dynamics of an integrate and fire oscillator. The situation corresponds to the response of an oscillator j when another oscillator i fires. In the left panel the coupling is excitatory $\epsilon_{ij} > 0$. In the right panel the coupling is inhibitory $\epsilon_{ij} < 0$. Taken from [3, section 7.3].

In terms of the phase variable the update rule is

$$\phi_i(t_i) = 1 \implies \begin{cases} \phi_j(t_i^+) = 0 & \text{if } j = i \\ \phi_j(t_i^+) = \min(1, U^{-1}(U(\phi_j(t_i)) + \epsilon_{ji})) & \text{if } j \neq i. \end{cases} \tag{5.10}$$

In the absence of delay a system of all-to-all connected oscillators, starting in arbitrary initial condition, will reaches a synchronized state with all oscillators firing at synchrony. To assess the transient to synchronization Guardiola et al. [76] defined the quantity

$$S = \frac{1}{N} \sum_i \left(1 - \phi_i(t_j^+)\right), \tag{5.11}$$

measured just after oscillator j has fired. The index j is an arbitrary reference and the election of time t_j^+ ensures that $\phi_j = 0$ in all measures. S allow to measure the *synchronization time* T, defined as the time needed to reach $S = 1$. Guardiola et al. [76] found that oscillators coupled in regular lattices synchronize faster than those coupled in random networks with the same average degree, and this behavior becomes stronger as $\langle k \rangle$ decreases (without going below the percolation threshold of random networks). In small world-networks the synchronization time is an increasing function of the rewiring probability. The conclusion of Guardiola et al. [76] was that an increasing randomness makes synchronization more difficult to reach. They showed that changing the coupling to

$$\epsilon_{ji} \to \frac{\epsilon_{ji}}{k_i} \tag{5.12}$$

removes the synchronization delay due to random induced heterogeneities (a procedure similar in essence to the solution for the heterogeneity paradox [72]).

5.4 The Kuramoto Model

Consider a collection of N nearly identical and weakly coupled limit-cycle [2] oscillators. The system has two timescales. In a short time scale each oscillator converges to its limit cycle at frequency ω_i, so it can be defined solely by a phase θ_i. In a large time scale the phases evolve due to the coupling among oscillators [78].

In order to simplify the analysis suppose each oscillator satisfies:

$$\dot{\theta}_i = \omega_i + \sum_{j=1}^{N} \Gamma_{ij}(\theta_j - \theta_i) \qquad i = 1, \ldots, N \tag{5.13}$$

where Γ_{ij} is an interaction term. Two key aspects demand further simplification: i) the interaction term could have many Fourier harmonics, and ii) the connection topology is unspecified. With the assumption of a zero mean, unimodal and symmetric frequency distribution $g(\omega)$, and an all-to-all interaction function of the form

$$\Gamma_{ij}(\theta_j - \theta_i) = \frac{K}{N} \sin(\theta_j - \theta_i), \tag{5.14}$$

[2] A limit cycle is an isolated periodic solution of the system's dynamics [77].

where $K \geq 0$ is a coupling strength, Kuramoto addressed both problems at the same time [79, Pages 420–422].

The phase dynamics is easier to understand if we consider each oscillator as a point running around the unit circle in the complex plane (see figure 5.4, left). Let define

$$ze^{\mathbf{i}\Theta} = \frac{1}{N}\sum_{j=1}^{N} e^{\mathbf{i}\theta_j}. \tag{5.15}$$

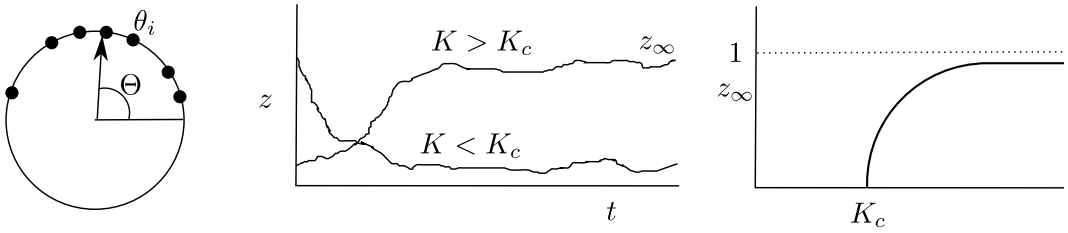

Figure 5.4: Mean-field parameters in the Kuramoto model: Geometric representation of oscillators dynamics, the arrow is $re^{\mathbf{i}\Theta}$ (Left). Qualitative evolution of coherence $z(t)$ as a function of the coupling strength K (Center). Maximum coherence as a function of the coupling strength K for a Cauchy distribution of oscillator frequencies (Right).

Multiply both sides by $e^{-\mathbf{i}\theta_i}$

$$ze^{\mathbf{i}(\Theta-\theta_i)} = \frac{1}{N}\sum_{j=1}^{N} e^{\mathbf{i}(\theta_j-\theta_i)},$$

now equating imaginary parts

$$z\sin(\Theta-\theta_i) = \frac{1}{N}\sum_{j=1}^{N}\sin(\theta_j-\theta_i).$$

Using the previous results for the interaction term, equation (5.13) takes the form known as the *Kuramoto model*

$$\dot{\theta}_i = \omega_i + Kz\sin(\Theta-\theta_i) \qquad i = 1,\ldots,N \tag{5.16}$$

Oscillators interact through the mean-field quantities z and Θ, know as *coherence* and *mean phase*, respectively. The effective coupling strength is Kz: as z increases more oscillators get into synch. There exists a critical coupling strength K_c such that (see figure 5.4, center)

- If $K < K_c$, starting from any initial condition, the phases become uniformly distributed around the unit circle. The coherence $z(t)$ decays to an $O(N^{-1/2})$ value, as expected for a random set of N points on a circle.

- If $K > K_c$, $z(t)$ grows exponentially, due to nucleation and growth of a small cluster of oscillators, and saturates at some value $z_\infty < 1$, with $O(N^{-1/2})$ fluctuations.

Kuramoto demonstrated that for a Cauchy distribution of oscillation frequencies

$$g(\omega) = \frac{\gamma}{\pi(\omega^2 + \gamma^2)}, \tag{5.17}$$

the super-critical behavior of z_∞ is given by (see figure 5.4, right)

$$z_\infty = \sqrt{1 - \frac{K_c}{K}}. \tag{5.18}$$

5.5 Kuramoto Model with long range interactions

The spatial dependence of the coupling in a system of coupled oscillators is a relevant topic. The limiting cases of interaction are i) the mean-field (all-to-all) interactions of the original Kuramoto model [78], where a super-critical coupling strength allows for a partial synchronization, and ii) the nearest-neighbors interactions in a lattice, where a phase-lock is less probable as the system size increases [80]. In real systems, the interaction signal decays as a function of the distance. For example, the intensity of light and sound decays as a power law of the distance from the source (with exponent 2 in three dimensions).

When the interactions decays as a power law such that the equations for a system of Kuramoto oscillators is

$$\dot{\theta}_i = \omega_i + \frac{K}{\eta} \sum_{j \neq i} \frac{1}{r_{i,j}^\alpha} \sin(\theta_j - \theta_i) \qquad i = 1, \ldots, N \tag{5.19}$$

where

- η is a normalization coefficient, and
- α is the decaying exponent

then one can interpolate between the mean-field coupling case, with $\alpha = 0$, and the lattice coupling, with $\alpha \to \infty$ [81]. In most real systems the coupling does not depends on global information, but only on the phases and the distance between oscillators, then is customary to use $\eta = 1$ in such cases.

Using numerical simulations of equation (5.19) in one-dimensional ($d = 1$) lattices Maródi et al. [82] reported perfect phase order ($z = 1$) for $\alpha < d$, decaying of coherence in the range $d \leq \alpha \leq \frac{3}{2}d$, until z reaches a minimal value $z_\infty \ll 1$ for $\alpha > \frac{3}{2}d$. This behavior was clearer as the number of oscillators increased [82, 81].

In a similar system, Chowdhury and Cross [81] defined the concept of *entrainment*

$$|\Delta_{ij}(t_0 + T) - \Delta_{ij}(t_0)| < 2\pi, \tag{5.20}$$

where $\Delta_{ij} = \theta_i - \theta_j$ and T is an arbitrarily long time, and the *phase correlation function*

$$C_{ij} = \langle \cos(\Delta_{ij}) \rangle \,, \tag{5.21}$$

such that

$$z = \sqrt{\frac{1}{N^2} \sum_{i,j=1}^{N} C_{ij}}. \tag{5.22}$$

The spin-wave analysis they performed found perfect phase ordering for $\alpha \leq d$, entrainment with long range phase ordering for $\alpha < \frac{3}{2}d$, and exponential decay of correlations for $\alpha > \frac{3}{2}d$.

5.6 The Kuramoto model on complex networks

While the all-to-all interactions in the Kuramoto leads to a simplified analysis, it is unrealistic. Interactions tend to be local, defining a network of interaction. The inclusion of this feature in conjunction with random intrinsic frequencies ω_i into the Kuramoto model defines the Sakaguchi-Kuramoto model [83] as

$$\dot{\theta}_i = \omega_i + K \sum_{j \in v(i)} \sin(\theta_j - \theta_i) \qquad i = 1, \ldots, N \tag{5.23}$$

where $v(i)$ denotes the neighborhood of node i.

5.6.1 Small-world effects

When oscillators in the Sakaguchi-Kuramoto model are distributed in a lattice network, synchronization is less probable as N increases [80]. However, even a small amount of shortcuts (via rewiring, as in the original small-world model of Watts and Strogatz [25]) leads to a drop in the average path length, allowing the existence of synchronized states [84]. Hong et al. [4] numerically studied the Sakaguchi-Kuramoto model in a small-world network with coupling

$$K \to \frac{K}{\langle k \rangle}, \tag{5.24}$$

showing that the coherence (averaged over time and realizations of the frequencies ω_i) obeys the scaling

$$z(N, K) = N^{-\beta/\nu} F\left((K - K_c) N^{1/\nu}\right). \tag{5.25}$$

From this scaling relation they obtain that:

- the critical coupling strength, K_c, is a decreasing function of the rewiring probability p (see the left side of figure 5.5),

- K_c diverges as $p \to 0$, and
- the time needed to achieve synchronization, T, is a decreasing function of the rewiring probability p (see the right side of figure 5.5).

Their results indicates the existence of a disorder-synchronization phase transition at $p_c = 0$.

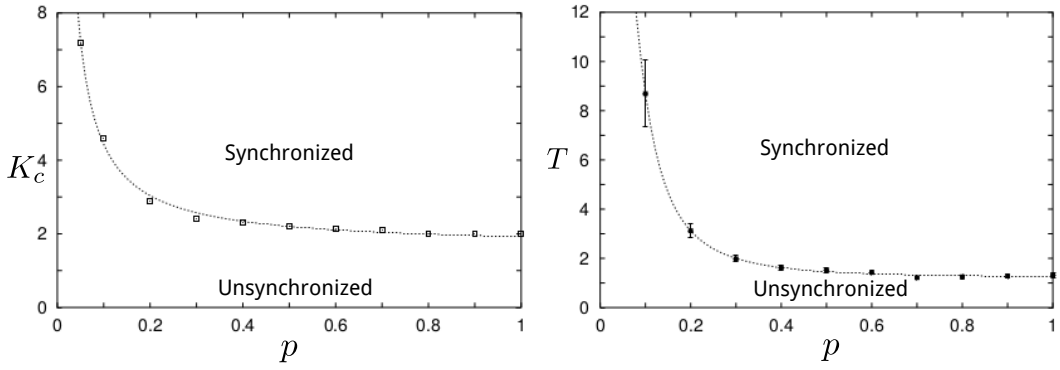

Figure 5.5: Critical coupling strength K_c (left panel) and synchronization time T (right panel) as a function of the rewiring probability p. Adapted from [4].

5.6.2 Degree distribution effects

As we have seen, degree fluctuations in the coupling network can influence synchronization. Ichinomiya [85] analytically studied frequency synchronization in the Sakaguchi-Kuramoto with order parameter

$$z^* = \frac{\sum_i k_i \exp(\mathrm{i}\theta_i)}{\sum_i k_i}, \tag{5.26}$$

showing that for oscillators coupled in uncorrelated random networks with arbitrary degree distribution the critical coupling strength K_c is given by

$$K_c = \frac{2}{\pi g(0)} \frac{1}{\kappa}, \tag{5.27}$$

making an explicit separation between dynamical and topological factors: $\frac{2}{\pi g(0)}$ only depends on the **frequency distribution** $\mathbf{g}(\omega)$, while $\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle}$ is a measure of the heterogeneity level of the degree distribution p_k.

If degree fluctuations are bounded, then κ is finite for large networks as $N \to \infty$, and thus K_c is finite too. For fat-tail distributions (power-laws, shifted power-laws, exponentially truncated power-laws, etc.) when $\gamma < 2$ then $\kappa \to 0$ and $K_c \to 0$ for large networks as $N \to \infty$. They also found that high-degree nodes synchronize faster than low-degree nodes, an effect due to the degree dependence of the order parameter.

Restrepo et al. [86] also obtained equation 5.27 for the Sakaguchi-Kuramoto.

5.6.3 Modularity effects

In general terms, we can see the phase synchronization of N oscillators as an aggregation process in the phase space of phases θ [87]: starting from a random distribution of phases θ_i, the system is partitioned into (at most) N singletons. Pairs, triplets and other higher order clusters appear as oscillators synchronize. In a complex network the process of cluster formation is affected by the community hierarchy of interaction among oscillators [88, 5]. This is relevant because complex networks such as metabolic networks, protein interaction networks, the internet, scientific collaboration networks and food webs [24, 89, 90, 91], among many other, have a nontrivial connectivity hierarchy of modules and/or communities. Even the identification of these structures has promoted the development of several definitions and methodologies [27, 28, 29].

Oh et al. [88] used a modified Sakaguchi-Kuramoto model

$$\dot{\theta}_i = \omega_i + \frac{K}{k_i} \sum_{j \in \langle i \rangle} \sin(\theta_j - \theta_i), \tag{5.28}$$

with a global order parameter

$$\mathbf{M} = \lim_{t \to \infty} \langle z(t) \rangle \,, \tag{5.29}$$

where k_i is the degree of node i, and $\langle \dots \rangle$ is the ensemble average over different configurations, to show that non-modular networks synchronize more easily than modular networks, i.e. the former reach specific values of $\mathbf{M}$ with smaller values of the coupling strength K, and also, the monotonically increasing curve $\mathbf{M}(K)$ becomes steeper as the modularity of the network is decreased by random link swapping.

Arenas et. al [5] used Kuramoto oscillars in modular networks to show that, as synchronization occurs at different time scales, the transient towards synchronization will reveal different topological structures/communities at different scales. They studied the time evolution of oscillator pairs with the (local) order parameter

$$\rho_{ij}(t) = \langle \cos\left[\theta_i(t) - \theta_j(t)\right] \rangle \,, \tag{5.30}$$

where $\langle \dots \rangle$ stand for the average over initial random phases, focusing on the evolution of the *dynamic connectivity* matrix, defined as

$$D_t(T)_{ij} = \begin{cases} 1 & \text{if} \quad \rho_{ij}(t) > T, \\ 0 & \text{if} \quad \rho_{ij}(t) < T \end{cases}, \tag{5.31}$$

given some threshold value T. When T is large enough, $D_t(T)$ represents a set of disconnected synchronized communities/components. As soon as T is decreased, a hierarchical structure of communities begins to emerge.

They noted that the number of disconnected synchronized components, d.s.c. (i.e. the number of null eigenvalues in the spectrum of $D_t(T)$,[92, 93]) is a decreasing function of time. Such behavior is qualitatively similar to that of i versus

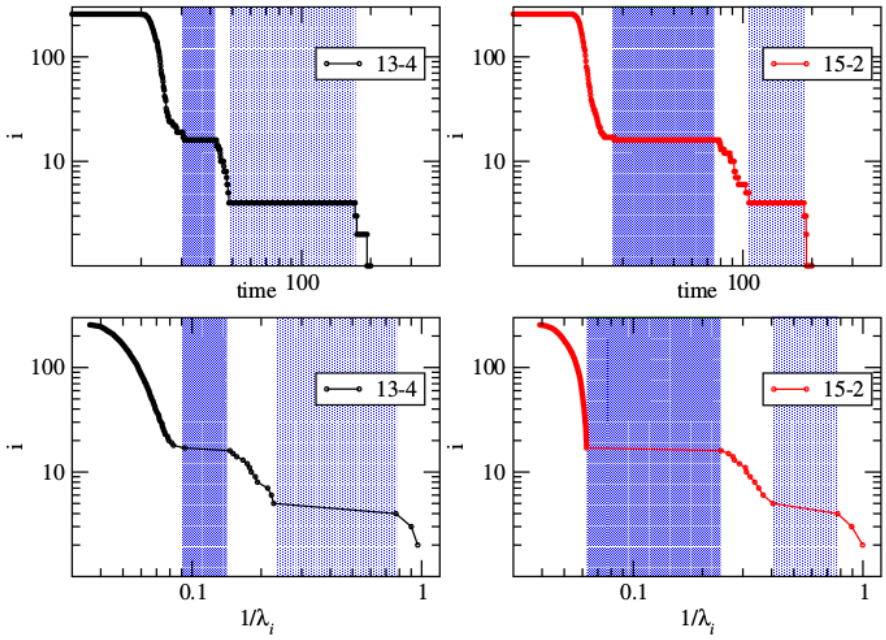

Figure 5.6: Kuramoto oscillators in a modular network G. Network G has $N = 256$ nodes, grouped in 16 compartments for the first level of community. The second level of community groups the first level in four compartments. The internal degree of nodes at the first level is k_{in_1}. The internal degree of nodes at the second level is k_{in_2}. The number of links with any community of the network is k_{out}. For the present data, total degree $k = k_{in_1} + k_{in_2} + k_{out}$ is fixed to 18. Networks are labeled as $k_{in_1} - k_{in_2}$. Time evolution of the number of disconnected synchronized components of the dynamic connectivity matrix (top). Rank index i versus the inverse of the correponding – ascending ordered – eigenvalue in the Laplacian spectrum of the dynamic connectivity matrix (bottom). Taken from [5].

$1/\lambda_i$, where λ_i are the eigenvalues of the Laplacian matrix of $D_t(T)$, ranked in ascending order $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$, see figure 5.6.

Chapter 6

Epidemiological models and complex networks

Epidemiologists, computer scientists, and social scientists share a common interest in understanding and predicting the propagation of entities such as human and animal diseases, computer viruses, knowledge, innovation, etc. These processes of propagation are based on the concepts of spreading and diffusion; epidemiological models are a common framework for them.

Epidemiological models deal with the temporal evolution and location of *infected individuals* within a population. They aim at understanding disease spreading in equilibrium (long time steady state); the existence of a non-zero density of infected individuals; the conditions to observe populations size disease outbreaks, and the emergence of recurrent patterns, among other topics. The understanding of such features provides a reference for the development of health policies to effectively stop epidemic outbreaks.

6.1 Deterministic homogeneous models

The simplest epidemic models assumes a population divided into classes or *compartments*, each corresponding to one stage of the disease. Common labels are

- *Susceptible* compartment, denoted as S, contains individuals who can contract the disease,
- *Infectious/Infected*, denoted as I, contains individuals who have contracted the disease and are contagious,
- *Recovered/Removed/Refractarious*, denoted as R, contains individuals who have recovered from the disease, either temporarily or permanently.

Other compartments could be used to model a more detailed disease evolution, however we will use just the set $\{S, I, R\}$.

The fully mixed model, or mass-action approximation, assumes that within each compartment individuals are identical and have equal chance, per unit time, of coming into contact with each other, meeting completely at random.

Given a population of N individuals, it is possible to obtain the time evolution of the number of individuals in the compartment m at time t, $X^{[m]}(t)$, such that

$$N = \sum_m X^{[m]}(t). \tag{6.1}$$

The numbers $X^{[m]}(t)$ change as individuals transit from one compartment to another, and the transition rates depend on the disease etiology. In compartment models there are two kind of transition rules:

- *Unary transitions*: the spontaneous transition of one individual from compartment m to compartment h. Examples of this kind of transition are
  - the Infectious to Recovered transition $I \to R$, and
  - the Recovered to Susceptible transition $R \to S$.

  The variation in $X^{[m]}$ is
  $$\sum_h \nu_h^m a_h X^{[h]} \tag{6.2}$$
  where a_h is the transition rate from class h and $\nu_h^m = 1, 0,$ or -1 is the change in $X^{[m]}$ due to the spontaneous transition from or to the compartment h.

- *Binary transitions*: transitions due to binary interactions among individuals in different compartments. An example of this kind of transition is the contagion of a susceptible individual in contact with an infectious one: the Susceptible to Infectious transition $S + I \to 2I$.

  The variation in $X^{[m]}$ is
  $$\sum_{h,g} \nu_{h,g}^m a_{h,g} N^{-1} X^{[h]} X^{[g]} \tag{6.3}$$
  where a_h is the transition rate of the interaction, $\nu_{h,g}^m = 1, 0,$ or -1 is the change in $X^{[m]}$ due to the interaction, and $N^{-1}X^{[g]}$ is the probability for each individual of the class h to interact with an individual of the class g.

Different models are characterized by a disease state set and the related transition rule set, but the general deterministic reaction rate equations for $X^{[m]}$ have the form

$$\frac{d}{dt}X^{[m]} = \sum_h \nu_h^m a_h X^{[h]} + \sum_{h,g} \nu_{h,g}^m a_{h,g} N^{-1} X^{[h]} X^{[g]}, \tag{6.4}$$

where $X^{[m]}$ are continuous variables representing the average number of individuals in compartment m.

Paradigmatic epidemiological models worth noting include the SI, SIS, SIR and SIRS models, all sketched in figure 6.1.

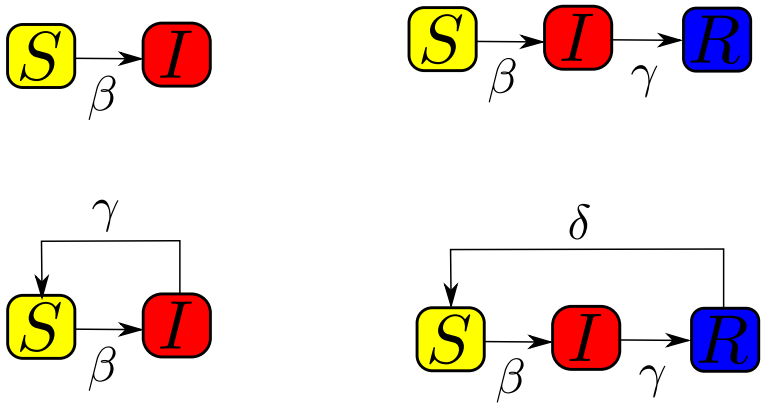

Figure 6.1: Schematic representation of epidemiological models. SI model (top left), SIS model (bottom left), SIR model (top right) and SIRS model (bottom right). These models vary in the number of allowed states from $\{S, I, R\}$, and the number of (constant rate) transitions. β is a binary transition rate, while γ and δ are unary transition rates.

6.1.1 SI model

Consider a population of N individuals. Each one is either infected or susceptible to be infected. Infected individual are said to be in the state I, meanwhile susceptible individuals are said to be in the state S. We refer to $i(t)$ and $s(t)$ as the fractions of individuals in states I or S at time t, respectively.

The disease is transmitted only when an infected individual comes into contact with a susceptible one, and occurs at *spreading rate* β (each individual has, on average, βi contacts per unit time with randomly chosen infected individuals) such that the differential equation for every fraction is

$$\begin{aligned} \frac{ds(t)}{dt} &= -\beta i(t)s(t) \\ \frac{di(t)}{dt} &= \beta i(t)s(t), \end{aligned} \quad , \tag{6.5}$$

where $i(t) + s(t) = 1$, so we can combine (6.5) into a single equation

$$\frac{di}{dt} = \beta(1-i)i. \tag{6.6}$$

This is the logistic growth equation, with solution

$$i(t) = \frac{i_0 e^{\beta t}}{1 - i_0 + i_0 e^{\beta t}}, \tag{6.7}$$

being $i_0 = i(0)$ the initial infected fraction. From this solution we note that the entire population will get infected, it is just a matter of time. The long time result is always the same, a system-size epidemic outbreak.

6.1.2 SIS model

In the SIS model, there are just two states, susceptible and infected. Infected individuals move back into the susceptible at some constant average rate γ. The differential equations for this model are

$$\begin{aligned} \frac{ds(t)}{dt} &= -\beta i(t)s(t) + \gamma i(t) \\ \frac{di(t)}{dt} &= \beta i(t)s(t) - \gamma i(t), \end{aligned} \tag{6.8}$$

where $i(t) + s(t) = 1$, so we can combine (6.8) into a single equation

$$\frac{di}{dt} = (\beta - \gamma - \beta i)i, \tag{6.9}$$

with solution

$$i(t) = \left(1 - \frac{\beta}{\gamma}\right) \frac{Ce^{(\beta-\gamma)t}}{1 + Ce^{(\beta-\gamma)t}}, \tag{6.10}$$

being

$$C = \frac{\beta i_0}{\beta - \gamma - \beta i_0}.$$

In the limit of large populations and small initial infected fractions, the solution is simplified to

$$i(t) = i_0 \frac{(\beta - \gamma)e^{(\beta-\gamma)t}}{\beta - \gamma + \beta i_0 e^{(\beta-\gamma)t}}, \tag{6.11}$$

in this scenario, we are interested in the long time solution for the infected fraction , i.e.

$$i_\infty = \frac{\beta - \gamma}{\beta}. \tag{6.12}$$

Then, there is no system wide epidemic outbreak. The system reaches an *endemic disease state* with a fraction of $i_\infty < 1$ infected individuals. The endemic fraction goes to zero as β approaches γ, and if $\beta < \gamma$ then $i(t)$ will decay exponentially. The point $\beta = \gamma$ marks the point of an *epidemic transition* between a state in which the disease spreads and one in which it dies.

The epidemic transition defines a key parameter in epidemiology, *the basic reproductive number*

$$R_0 = \frac{\beta}{\gamma}, \tag{6.13}$$

which counts the number of secondary infections generated by a primary infected individual.

- if $\beta \leq \gamma$ infected individuals recover faster than susceptible individuals become infected: the infected fraction goes down and the disease dies.
- if $\beta > \gamma$ susceptible individuals become infected faster than infected individuals recover: the infected fraction goes up and the disease spreads.

For its behavior, the SIS model is mainly used for the study of diseases leading to an endemic state, with the prevalence of a constant number of infected individuals.

6.1.3 SIR model

For many real diseases, individuals recover from the infection after a certain time because their immune system fights off the agent causing the disease. The SIR model accounts for this situation.

In the SIR model, there are three states: susceptible S, infected I, and recovered R. Infected individuals move permanently into the recovered state at some constant average rate γ. The differential equations for this model are

$$\begin{aligned} \frac{ds(t)}{dt} &= -\beta i(t)s(t) \\ \frac{di(t)}{dt} &= \beta i(t)s(t) - \gamma i(t) \\ \frac{dr(t)}{dt} &= \gamma i(t), \end{aligned} \tag{6.14}$$

where $i(t) + s(t) + r(t) = 1$. We can rewrite

$$\begin{aligned} \frac{ds(t)}{dt} &= -\beta i(t)s(t) \\ \frac{dr(t)}{dt} &= \gamma i(t), \end{aligned} \tag{6.15}$$

as a single equation

$$\frac{1}{s}\frac{ds}{dt} = -\frac{\beta}{\gamma}\frac{dr}{dt}, \tag{6.16}$$

with solution

$$s = s_0 e^{-\frac{\beta}{\gamma}r}, \tag{6.17}$$

where $s_0 = s(0)$ and we have chosen $r_0 = r(0) = 0$. Now use $i(t) = 1 - r(t) - s(t)$ to write

$$\frac{dr(t)}{dt} = \gamma i(t) = \gamma\left(1 - r - s\right) = \gamma\left(1 - r - s_0 e^{-\frac{\beta}{\gamma}r}\right). \tag{6.18}$$

A measure of the epidemic outbreak size is given by the fraction of recovered individuals at large times, r_∞, when all infected individuals have healed such that $\dot{r} = 0$. For initial condition $s_0 \approx 1$, $i_0 \ll 1$ and $r_0 = 0$ this condition is

$$r_\infty = 1 - e^{-\frac{\beta}{\gamma}r_\infty}, \tag{6.19}$$

which implicitly determines r_∞, and indicates that

- as $\frac{\beta}{\gamma} > 1$ approaches one from above r_∞ goes continuously to zero, and
- for $\frac{\beta}{\gamma} < 1$ we have $r_\infty = 0$.

The SIR model (6.14) has no closed analytic solution. Numerical simulations show that:

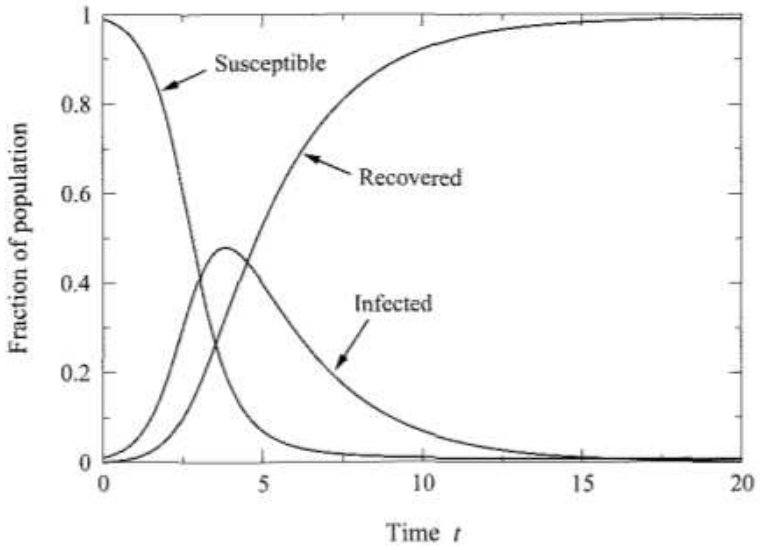

Figure 6.2: Time evolution of a SIR model with parameters $\beta = 1, \gamma = 0.4, s_0 = 0.99, i_0 = 0.01$ and $r_0 = 0$. Taken from [6, section 17.3].

- $s(t)$ decreases monotonically as susceptible individuals are infected,
- $r(t)$ increases monotonically as infected individual get recovered, and
- $i(t)$ goes up at first as susceptible individual get infected, then down again as they recover, and eventually goes to zero as $t \to \infty$.

This behavior is clear from figure 6.2.

6.1.4 SIRS model

The SIRS model extends the SIR model by considering that individuals can get reinfected after recovering. In the SIRS model, there are three states: susceptible S, infected I, and recovered R. Infected individuals move temporarily into the recovered state at some contact average rate γ, then lose immunity at a constant average rate δ. The differential equations for this model are

$$\begin{aligned}\frac{ds(t)}{dt} &= -\beta i(t)s(t) + \delta r(t)\\ \frac{di(t)}{dt} &= \beta i(t)s(t) - \gamma i(t)\\ \frac{dr(t)}{dt} &= \gamma i(t) - \delta r(t),\end{aligned} \tag{6.20}$$

where $i(t) + s(t) + r(t) = 1$.

System (6.20) cannot be solved analytically. Numerical solutions show that the SIRS model has a rich palette of behaviors, depending on the values of the its parameters, including:

- disease persistence in an endemic state,
- disease extinction, and
- oscillations between outbreaks and periods of remission. [6, 69]

The equilibrium of the constant-rate SIRS model (6.20) is

$$s^* = \frac{1}{\beta\tau_i}, \qquad i^* = \frac{\beta\tau_i - 1}{\beta\tau_0}, \qquad r^* = 1 - i^* - s^*, \tag{6.21}$$

where

$$\gamma = \frac{1}{\tau_i}, \qquad \delta = \frac{1}{\tau_r}, \qquad \tau_0 = \tau_r - \tau_i. \tag{6.22}$$

This fixed point is stable and corresponds to an endemic state, reached through damped oscillations [94, 6, 69].

6.1.5 SIRS model with fixed refractory time

Anderson and May [95] noted that treating the duration of a infection, τ_i, as a constant rate, $1/\gamma$, is rarely realistic. It is more common to see recovery from infection to take place after a well defined time. Something similar happens with the recovered to susceptible transition.

These facts motivated the analysis of a SIRS model with fixed infectious and refractory periods, τ_i and $\tau_r = \tau_0 - \tau_i$ respectively. This model is perhaps the simplest mathematical model which captures a wide range of behaviors similar to that of real recurrent diseases such as oscillations and spatial waves. It is modeled by the system

$$\begin{aligned}\frac{di(t)}{dt} &= \beta s(t)i(t) - \beta s(t-\tau_i)i(t-\tau_i)\\ \frac{ds(t)}{dt} &= -\beta s(t)i(t) + \beta s(t-\tau_0)i(t-\tau_0).\end{aligned} \tag{6.23}$$

It has the same equilibrium as the constant-rate SIRS model [69]. Thirty years ago Hethcote et al. [94] showed that the SIRS model with fixed refractory period

presents a *stable* limit cycle for certain parameter values. In the parameter space $(\tau_r/\tau_i, \beta\tau_i)$ a Hopf bifurcation separates solutions into two classes: stable fixed points and stable limit cycles [69].

The SIRS model with fixed refractory period can also be classified as a simple model of excitable media, because it describes a medium that displays the joint properties of wave propagation and refractoriness. There are many systems in which a wave of some sort can pass through a medium, after which the medium cannot support another wave until a suitable length of time, called the refractory time, has passed. Forest fires, the Belousov-Zhabotinsky reaction, neural tissue and cardiac tissue are examples of excitable media [9, chapter 2].

6.2 Epidemics in homogeneous networks

The fully mixed models of epidemiology avoids discussing contact networks at all, however, to understand and predict epidemic outbreaks requires a detailed knowledge of the interactions between individuals. For example, people have contact with only a small fraction of the population (such as family, friends, coworkers, etc), and that fraction is not chosen at random. In this sense, the network of contacts plays a key role in the spread of a disease.

In this section, we outline some fundamental changes introduced by a network of contacts into epidemiological models.

6.2.1 Net infection rate

Each infected individual infects a susceptible neighbor with probability βdt, thus, the total probability of infection for a susceptible individual with n infected neighbors is

$$1 - (1 - \beta dt)^n.$$

Neglecting fluctuations, the average number of infected neighbors for a node of degree k is $n = ki$, then

$$1 - (1 - \beta dt)^n = 1 - (1 - \beta dt)^{ki}.$$

At leading order in $\beta dt \ll 1$ we obtain

$$1 - (1 - \beta dt)^{ki} \approx \beta kidt,$$

so the spreading has changed

$$\beta \to \beta k. \tag{6.24}$$

This *net spreading rate* is explicitly dependent on the individual's *contact number* (i.e. node degree).

In a homogeneous network with $k_i = \langle k \rangle \, \forall i$ the spreading rate is

$$\beta \langle k \rangle . \tag{6.25}$$

Te deterministic equations for the SI, SIS and SIR model maintain their forms, with the substitution $\beta \to \beta \langle k \rangle$.

6.2.2 Epidemic outbreak size

As we have already seen, different mixed models set different upper bounds for the size of an epidemic outbreak. This also happens on networks, with the addition of a direct limitation: the largest outbreak size is restricted by the size of the component where the disease starts. So, in a multicomponent network, outbreak size is limited in accordance to the components size-distribution.

6.3 Epidemics on networks and percolation

As we have seen, the long time behavior of the deterministic SI model is a system-wide epidemic outbreak. On the other hand, the SIS and SIR models present an endemic or epidemic threshold, respectively. The outbreak size in a network is limited by the size of the component where disease starts.

Let us consider the SIR model in a network with arbitrary topology. Infectious individuals remain infected for a constant finite average time

$$\tau_i = \frac{1}{\gamma}, \tag{6.26}$$

then they recover. The probability that a susceptible individual catches the infection from one of their infected neighbors is

$$p = 1 - e^{-\beta/\tau_i} = 1 - e^{-\beta/\gamma}, \tag{6.27}$$

or a total probability

$$p_{total} = 1 - (1-p)^{k_i} \geq p, \tag{6.28}$$

if it has k_i infected neighbors.

Clearly, disease transmission requires the existence of links with a susceptible node in one edge-end and one infected node in the other. Disease only can spread across such links. If we occupy each link with probability p, or not with probability $1-p$, the process is called *bond percolation*. The basic mathematical description of bond percolation is almost the same as that of *site percolation* (seen in subsection 4.1.1). The only difference resides on the equation to obtain the relative size of the giant cluster

$$S = 1 - g_0(u). \tag{6.29}$$

The location of the percolation threshold is the same, namely

$$p_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} = \frac{1}{\kappa - 1}. \tag{6.30}$$

Above p_c the giant cluster appears and an epidemic outbreak of size S becomes possible, however, disease can propagate to any node in a connected cluster where one infected node lies. In the particular case of a random network with Poisson degree distribution the size of the giant cluster is

$$S = 1 - \exp(-p \langle k \rangle S), \tag{6.31}$$

where $\langle k \rangle$ is the average degree of the network when $p = 1$, then $p \langle k \rangle$ is the average degree at any p. Equation 6.31 has the same form as the outbreak size r_∞ of the deterministic SIR model (equation 6.19). Barbor and Mollison [96, chapter 7] demonstrated that the fully mixed (deterministic) SIR model corresponds to the same model on random networks.

For models with reinfection, such as the SIS and SIRS models, the equivalence between epidemic model and bond percolation does not hold. Then, we need to rely on the analysis of the spreading dynamics.

6.4 Epidemics in heterogeneous uncorrelated networks

Relevant networks in epidemiological terms are heterogeneous in their degree distribution p_k, so the homogeneous assumption is no longer valid. A conventional approach to analyze such a case is to consider the *block degree approximation*, i.e. assuming that all nodes with the same degree k are statistically equivalent. Then, nodes are grouped into degree classes, each one having its own density

$$i_k = \frac{I_k}{N_k}, \qquad s_k = \frac{S_k}{N_k}, \qquad r_k = \frac{R_k}{N_k}$$

with global averages related to p_k

$$i = \sum_k p_k i_k, \qquad s = \sum_k p_k s_k, \qquad r = \sum_k p_k r_k$$

6.4.1 SI Model

The evolution equation for the infected fraction in the SI model is given by [7]

$$\frac{di_k(t)}{dt} = \beta(1 - i_k(t))k\Theta_k(t) \tag{6.32}$$

where

- β is the spreading rate,
- $1 - i_k(t)$ is the probability that a node with degree k is not infected,
- k is the degree of nodes in class k, and
- $\Theta_k(t)$ is the density of infected neighbors for nodes of degree k, i.e. the average probability that any given neighbor of a node of degree k is infected. In the fully-mixed approach $\Theta_k(t) = i(t)$ is degree independent.

At least one link of each infected node points to another infected node from which it catched the infection. Then, for uncorrelated networks,

$$\Theta_k(t) = \Theta(t) = \frac{1}{\langle k \rangle} \sum_{k'} (k' - 1) p_{k'} i_{k'}(t), \tag{6.33}$$

is independent of k.

In the initial stage of propagation $i_k(t)$ is small enough to neglect terms of order $\mathcal{O}(i_k^2)$, equation 6.32 reads as

$$\frac{di_k(t)}{dt} = \beta k \Theta(t), \tag{6.34}$$

then

$$\frac{d\Theta(t)}{dt} = \frac{1}{\langle k \rangle} \sum_{k'} (k'-1) p_{k'} k' \beta \Theta(t) = \beta \left(\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \right) \Theta(t) = \frac{1}{\tau} \Theta(t), \tag{6.35}$$

where

$$\tau = \frac{1}{\beta} \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} = \frac{1}{\beta} \frac{1}{\kappa - 1}, \tag{6.36}$$

so $\Theta(t)$ is

$$\Theta(t) = \Theta_0 e^{t/\tau}. \tag{6.37}$$

The initial stage 6.34 can be solved using $i_0(0) = \Theta_0 = i_0$, yielding,

$$i_k(t) = i_0 \left[1 + \frac{k}{\kappa - 1} (e^{t/\tau} - 1) \right], \tag{6.38}$$

where we observe that $i_k(t)$ increases exponentially fast and higher degree classes contribute more to the total infected fraction, $\sum_k p_k i_k$, which is

$$i(t) = i_0 \left[1 + \frac{\langle k \rangle}{\kappa - 1} (e^{t/\tau} - 1) \right]. \tag{6.39}$$

It is remarkable that the *growth time scale* τ is a function of two separate factors: β^{-1}, which only depends on the dynamics through the transition rate $S \to I$; and $(\kappa - 1)^{-1}$, which only depends on the structure, through the *heterogeneity parameter* κ. Some direct conclusions on the form of τ are:

- for random networks with Poisson degree distribution, $\kappa = \langle k \rangle + 1$ so

  $$\tau_{rn} = \frac{1}{\beta \langle k \rangle}$$

  this value coincides with the growth time scale expected for the SI model in homogeneous networks.

- for scale-free networks with degree distribution $p_k \sim k^{-\bar{\gamma}}$ with $2 \leq \bar{\gamma} \leq 3$ we have $\kappa \to \infty$ so

  $$\tau_{sf} \to 0$$

  this value indicates that, on scale-free networks, infection rises almost instantaneously among the population. This happens as a *cascade effect*; the disease first reaches hubs, from where it can spread vary fast to nodes of lower degree [7, 97], as shown in figure 6.3.

In the long time, $t \gg \tau$, the result is always the same: all individuals within the sub-graph where the infection started will be infected.

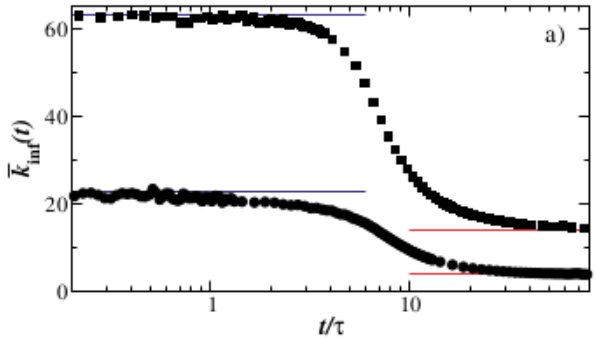

Figure 6.3: Average degree of the newly infected nodes for SI model as a function of re-scaled time by t/τ. Reference lines are drawn at κ for $t/\tau \ll 1$ and m for $t/\tau \gg 1$. Simulation where performed on scale-free networks in the Barábasi-Albert model with $N = 10^4$ and $m = 4(\text{bottom}), 14(\text{top})$. Taken from [7].

6.4.2 SIS Model

The evolution equation of the infected fraction in the SIS model is given by

$$\frac{di_k(t)}{dt} = \beta(1 - i_k(t))k\Theta_k(t) - \gamma i_k(t) \tag{6.40}$$

where

- β is the spreading rate,
- γ is the recovery rate,
- $1 - i_k(t)$ is the probability that a node with degree k is not infected,
- k is the degree of nodes in class k, and
- $\Theta_k(t)$ is the density of infected neighbors for nodes of degree k, i.e. the average probability that any given neighbor of a node of degree k is infected. In the fully-mixed approach $\Theta_k(t) = i(t)$ is degree independent.

An infected node does not have to point to another infected node since the node from which it received the infection can spontaneously become susceptible again.

$$\Theta_k(t) = \Theta(t) = \frac{1}{\langle k \rangle} \sum_{k'} k' p_{k'} i_{k'}(t), \tag{6.41}$$

is independent of k.

In the initial stage of propagation $i_k(t)$ is small enough to neglect terms of order $\mathcal{O}(i_k^2)$, the equation 6.40 reads as

$$\frac{di_k(t)}{dt} = \beta k\Theta(t) - \gamma i_k(t), \tag{6.42}$$

then

$$\frac{d\Theta(t)}{dt} = \frac{1}{\langle k \rangle} \sum_{k'} k' p_{k'} k' \beta \Theta(t) - \gamma \Theta(t) = \beta \frac{\langle k^2 \rangle}{\langle k \rangle} \Theta(t) - \gamma \Theta(t) = \frac{1}{\tau} \Theta(t), \quad (6.43)$$

where

$$\tau = \frac{\langle k \rangle}{\beta \langle k^2 \rangle - \gamma \langle k \rangle} = \frac{1}{\beta \kappa - \gamma}, \quad (6.44)$$

so $\Theta(t)$ is

$$\Theta(t) = \Theta_0 e^{t/\tau}. \quad (6.45)$$

The initial stage 6.34 can be solved using $i_0(0) = \Theta_0 = i_0$ yielding

$$i_k(t) = i_0 \left[1 + \frac{\beta k}{\beta \kappa - \gamma} (e^{t/\tau} - 1) \right], \quad (6.46)$$

where we observe (as in the SI model) that $i_k(t)$ increases exponentially fast and higher degree classes contributes more to the total infected fraction, $\sum_k p_k i_k$, which is

$$i(t) = i_0 \left[1 + \frac{\beta \langle k \rangle}{\beta \kappa - \gamma} (e^{t/\tau} - 1) \right]. \quad (6.47)$$

In this case, the growth time scale is an intertwined function of the transition rates β and γ, and the heterogeneity parameter κ. In order to ensure the existence of an epidemic outbreak with finite threshold we need $\tau > 0$. Such condition reads

$$\frac{\beta}{\gamma} \geq \frac{1}{\kappa}, \quad (6.48)$$

and implies that heavy-tailed degree distribution networks with diverging second moment ($\kappa \to \infty$) in the limit $N \to \infty$ has epidemic threshold equal to zero. Wether or not a finite epidemic threshold exists depends on the contact structure through κ, while the current dynamical regime (sub-threshold or super-threshold) is still an interplay between transition rates β/γ, as in the deterministic homogeneous SIS model.

The endemic state associated with the SIS model can be obtained by imposing the stationary condition $di_k(t)/dt = 0$,

$$i_k = \frac{\beta k \Theta}{\gamma + \beta k \Theta}, \quad (6.49)$$

which makes equation 6.41 yield

$$\Theta = \frac{1}{\langle k \rangle} \sum_k k p_k \frac{\beta k \Theta}{\gamma + \beta k \Theta}, \quad (6.50)$$

a self consistent equation whose solution depends only on the transition rates β and γ. Equation 6.50 has solution if its left hand side $y_1(\Theta) = \Theta$ intersects the

right hand side $y_2(\Theta)$ at $\Theta^* \neq 0$. This happens only if the slope of $y_2(\Theta)$ at $\Theta = 0$ is larger than or equal to 1, i.e.

$$\frac{d}{d\Theta}\left[\frac{1}{\langle k\rangle}\sum_k kp_k\frac{\beta k\Theta}{\gamma+\beta k\Theta}\right]_{\Theta=0} = \frac{\beta}{\gamma}\kappa \geq 1, \tag{6.51}$$

which is the same condition obtained in the early-time approximation, equation 6.48.

Using the continuous degree approximation, it is possible to compute the endemic infected fraction i_∞ for scale-free networks with normalized degree distribution [98, 99]

$$p_k = (\bar{\gamma}-1)m^{\bar{\gamma}-1}k^{-\bar{\gamma}}, \tag{6.52}$$

having average degree $(\bar{\gamma}-1)m/(\bar{\gamma}-2)$ and minimum degree m. We need to solve equation 6.50 close to the epidemic threshold

$$\frac{\beta}{\gamma} = \frac{1}{\kappa} = \begin{cases} \frac{\bar{\gamma}-3}{m(\bar{\gamma}-2)} & \text{if } \bar{\gamma} > 3 \\ 0 & \text{if } \bar{\gamma} \leq 3 \end{cases}. \tag{6.53}$$

- For $2 < \bar{\gamma} < 3$

$$i_\infty \sim \left(\frac{\beta}{\gamma}\right)^{1/(3-\bar{\gamma})}, \tag{6.54}$$

  there is no epidemic threshold, and i_∞ is non-zero for all values of $\frac{\beta}{\gamma}$.

- For $\bar{\gamma} = 3$

$$i_\infty \sim \exp\left(\frac{\gamma}{m\beta}\right), \tag{6.55}$$

  there is no epidemic threshold.

- For $3 < \bar{\gamma} < 4$

$$i_\infty \sim \left(\frac{\beta}{\gamma} - \frac{\bar{\gamma}-3}{m(\bar{\gamma}-2)}\right)^{1/(3-\bar{\gamma})}, \tag{6.56}$$

  the epidemic threshold is given by 6.53 and i_∞ presents a smooth power law behavior.

- For $\bar{\gamma} > 4$

$$i_\infty \sim \frac{\beta}{\gamma} - \frac{\bar{\gamma}-3}{m(\bar{\gamma}-2)}, \tag{6.57}$$

  i_∞ follows the expected behavior for homogeneous networks [98].

6.4.3 SIR Model

The evolution equation of the infected fraction in the SIR model is given by

$$\frac{di_k(t)}{dt} = \beta(1 - i_k(t) - r_k(t))k\Theta_k(t) - \gamma i_k(t) \tag{6.58}$$

where

- β is the spreading rate,
- γ is the recovery rate,
- $1-i_k(t)-r_k(t)$ is the probability that a node with degree k is neither infected nor recovered,
- k is the degree of nodes in class k,
- $\Theta_k(t)$ is the density of infected neighbors for nodes of degree k, i.e. the average probability that any given neighbor of a node of degree k is infected. In the fully-mixed approach $\Theta_k(t) = i(t)$ is degree independent, and

$$\Theta_k(t) = \Theta(t) = \frac{1}{\langle k \rangle} \sum_{k'} (k' - 1) p_{k'} i_{k'}(t), \tag{6.59}$$

is independent of k.

In the initial stage of propagation $i_k(t)$ is small enough to neglect terms of order $\mathcal{O}(i_k^2)$ and $r_k(t) \sim 0$, the equation 6.58 reads as

$$\frac{di_k(t)}{dt} = \beta k \Theta(t) - \gamma i_k(t), \tag{6.60}$$

then

$$\frac{d\Theta(t)}{dt} = \frac{1}{\langle k \rangle} \sum_{k'} (k' - 1) p_{k'} k' \beta \Theta(t) - \gamma \Theta(t) = \beta \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} \Theta(t) - \gamma \Theta(t) = \frac{1}{\tau} \Theta(t), \tag{6.61}$$

where

$$\tau = \frac{\langle k \rangle}{\beta \left(\langle k^2 \rangle - \langle k \rangle \right) - \gamma \langle k \rangle} = \frac{1}{\beta(\kappa - 1) - \gamma}, \tag{6.62}$$

so $\Theta(t)$ is

$$\Theta(t) = \Theta_0 e^{t/\tau}. \tag{6.63}$$

The initial stage 6.34 can be solved using $i_0(0) = \Theta_0 = i_0$ yielding

$$i_k(t) = i_0 \left[1 + \frac{\beta k}{\beta(\kappa - 1) - \gamma} (e^{t/\tau} - 1) \right], \tag{6.64}$$

where we observe (as in the SI and SIS models) that $i_k(t)$ increases exponentially fast and higher-degree classes contribute more to the total infected fraction, $\sum_k p_k i_k$, which is given by

$$i(t) = i_0 \left[1 + \frac{\beta \langle k \rangle}{\beta(\kappa - 1) - \gamma} (e^{t/\tau} - 1) \right]. \tag{6.65}$$

In this case, the growth time-scale is an intertwined function of the transition rates β and γ, and the heterogeneity parameter κ. In order to ensure the existence of an epidemic outbreak with finite threshold we need $\tau > 0$, one must have

$$\frac{\beta}{\gamma} \geq \frac{1}{\kappa - 1}. \tag{6.66}$$

This implies that heavy-tailed degree distribution networks with diverging second moment ($\kappa \to \infty$) in the limit $N \to \infty$ have epidemic threshold zero. Wether or not a finite epidemic threshold exists depends on the contact structure through κ, while the current dynamical regime (sub-threshold or super-threshold) is still an interplay between transition rates β/γ, as in the deterministic homogeneous SIR model.

To obtain the asymptotic number of infected individuals $r_\infty = \lim_{t\to\infty} \sum_k p_k r_k(t)$ we take the rate equation for $r_k(t)$

$$\frac{dr_k(t)}{dt} = \gamma i_k(t), \tag{6.67}$$

which is directly integrable to give

$$r_k(t) = \gamma \int_0^t i_k(t')dt'. \tag{6.68}$$

And the rate equation for $s_k(t)$

$$\frac{ds_k(t)}{dt} = -\beta k s_k(t) \Theta_k(t), \tag{6.69}$$

which is integrable as follows

$$\begin{aligned}
\ln\left(\tfrac{s_k(t)}{s_k(0)}\right) &= \int_{s_k(0)}^{s_k(t)} \tfrac{ds'_k(t)}{s'_k(t)} \\
&= -\beta k \left(\int_0^t \Theta(t')dt' \right) \\
&= -\beta k \left(\tfrac{1}{\langle k \rangle} \sum_k (k-1) p_k \int_0^t i_k(t')dt' \right) \\
&= -\beta k \left(\tfrac{1}{\gamma \langle k \rangle} \sum_k (k-1) p_k r_k(t) \right) \\
&= -\beta k \phi(t),
\end{aligned}$$

with $s_k(0) \simeq 1$, then

$$s_k(t) = \exp\left(-\beta k \phi(t)\right). \tag{6.70}$$

The time evolution equation for ϕ is

$$\begin{aligned}\dot{\phi}(t) &= \Theta(t) \\ &= \tfrac{1}{\langle k\rangle}\textstyle\sum_k (k-1)p_k i_k(t) \\ &= \tfrac{1}{\langle k\rangle}\textstyle\sum_k (k-1)p_k\left[1-r_k(t)-s_k(t)\right] \\ &= \left(1-\tfrac{1}{\langle k\rangle}\right) - \gamma\phi(t) - \tfrac{1}{\langle k\rangle}\textstyle\sum_k (k-1)p_k \exp\left(-\beta k\phi(t)\right).\end{aligned} \tag{6.71}$$

Now r_∞ can be written as

$$r_\infty = \sum_k p_k r_k(\infty) = \sum_k p_k\left[1-i_k(\infty)-s_k(\infty)\right] = \sum_k p_k\left[1-\exp\left(-\beta k\phi(\infty)\right)\right], \tag{6.72}$$

because in the limit $t\to\infty$, all infected nodes have healed. In such conditions, $\phi(t)$ will stop to change, so that $\lim_{t\to\infty}\dot{\phi}(t)=0$. From equation 6.71 one obtains the self consistence equation

$$\phi(\infty) = \frac{1}{\gamma}\left(1-\frac{1}{\langle k\rangle}\right) - \frac{1}{\gamma\langle k\rangle}\sum_k (k-1)p_k\exp\left(-\beta k\phi(\infty)\right). \tag{6.73}$$

Equation 6.73 has a solution if its left hand side $y_1(\phi(\infty))=\phi(\infty)$ intersects the right hand side $y_2(\phi(\infty))$ at $\phi(\infty)^*\neq 0$. This happens only if the slope of $y_2(\phi(\infty))$ at $\phi(\infty)=0$ is larger than or equal to 1, i.e.

$$\frac{d}{d\phi(\infty)}\left[\frac{1}{\gamma}\left(1-\frac{1}{\langle k\rangle}\right) - \frac{1}{\gamma\langle k\rangle}\sum_k (k-1)p_k\exp\left(-\beta k\phi(\infty)\right)\right]_{\phi(\infty)=0} \geq 1, \tag{6.74}$$

which is equivalent to

$$\frac{\beta}{\gamma\langle k\rangle}\sum_k (k-1)kp_k = \frac{\beta}{\gamma}\frac{\langle k^2\rangle-\langle k\rangle}{\langle k\rangle} \geq 1, \tag{6.75}$$

and leads to the same epidemic threshold condition as equation 6.66, namely

$$\frac{\beta}{\gamma} = \frac{1}{\kappa-1}. \tag{6.76}$$

Below the threshold $r_\infty=0$; above the threshold r_∞ takes a finite value.

Using the continuous degree approximation, it is possible to compute r_∞ for scale-free networks with normalized degree distribution $p_k=(\bar{\gamma}-1)m^{\bar{\gamma}-1}k^{-\bar{\gamma}}$. The behavior of $r_\infty(\beta/\gamma)$ is the same as $i_\infty(\beta/\gamma)$ for the SIS model on the same kind of networks [100, 101].

6.4.4 SIRS model in Small-World networks

In previous Chapters from the current Part we have presented the effects of small-world networks over dynamical processes, this chapter will not be an exception. Kuperman & Abramson [8] were the first to study the effects of small-world networks over epidemiological model, specifically, the SIRS model. Their approach was formulated as follows:

- Individuals are located at the nodes of a small-world network built using the Watt-Strogatz model [25], with average degree $2K$ and disorder parameter $p \in [0, 1]$.

- Susceptible nodes can get the disease through contagion by infected neighbors, with probability

$$p_{inf} = \frac{k_{inf}}{k}, \tag{6.77}$$

  where k_{inf} is the number of infected neighbors of node i.

- Infected nodes remain infectious during a period τ_I, after that time they recover.

- Recovered nodes remain recovered during a period τ_R, after that time they become susceptible again.

- A counter

$$\tau_i(t) = 0, 1, \ldots, \tau_0,$$

  where $\tau_0 = \tau_I + \tau_R$, records the location of node i in the cycle of the disease. The state of node i is determined from this counter

  - $i \in S$ if $\tau_i(t) = 0$,
  - $i \in I$ if $\tau_i(t) \in (1, \tau_I)$,
  - $i \in R$ if $\tau_i(t) \in (\tau_I + 1, \tau_0)$.

- The state of node i in the next step depends on i) its current state and ii) the state of its neighbors. The counter evolves following the rules

  - $\tau_i(t+1) = 0$ if $\tau_i(t) = 0$ and $r > p_{inf}$ (no infection occurs),
  - $\tau_i(t+1) = 1$ if $\tau_i(t) = 0$ and $r \leq p_{inf}$ (infection occurs),
  - $\tau_i(t+1) = \tau_i(t) + 1$ if $1 \leq \tau_i(t) < \tau_0$ (infectious + recovered period),
  - $\tau_i(t+1) = 0$ if $\tau_i(t) = \tau_0$ (healing).

  where $r \in [0, 1]$ is a random number from a uniform distribution.

- The synchronization behavior was decribed by the order parameter

$$\sigma = \left| \frac{1}{N} \sum_{i=1}^{N} \exp(\mathrm{i}\theta_i) \right| \tag{6.78}$$

  where $\theta_i = 2\pi(\tau_i - 1)/\tau_0$ is the geometrical *phase* of node i. Only positive phases are counted into the sum, i.e. the sum is over the deterministic part of the disease cycle.

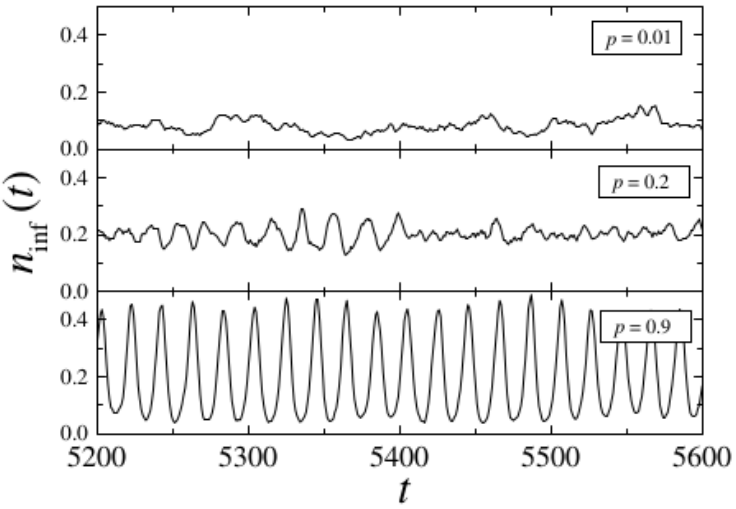

Figure 6.4: Time evolution of the infection fraction $n_{inf}(t)$ for different values of disorder parameter (shown in legends). Simulations were performed with parameters $\tau_I = 4, \tau_R = 9, K = 3, N = 10^4$ and $n_{inf}(t) = 0.1$. Taken from [8].

Their results shows that: at low values of p ($p = 0.01$ top of figure 6.4) the infection is endemic, the infected fraction is low, and present fluctuations; at high values of p ($p = 0.9$ bottom figure 6.4) the infected fraction presents large-amplitude quasi-periodic oscillations, with small fluctuations in amplitude; at some intermediate value of p ($p = 0.2$ middle of figure 6.4) a transition between an endemic and an oscillatory state takes place, oscillatory lapses are intertwined with periods of large fluctuation in n_{inf}.

The transition becomes sharper as system size increases (see figure 6.5) and it moves to lower p as the average degree ($2K$) increases (see figure 6.6).

They also defined an alternative contagion probability

$$p_{inf} = (1 - (1 - q)^{k_{inf}}) \tag{6.79}$$

where $q \in [0, 1]$ is the *infectivity*, and obtained similar results for the range $q < 0.3$.

Particularly, Kuperman & Abramson [8] noted that high values of disorder ($p \geq 0.5$) lead to a situation where large amplitude oscillations arise and the disease dies out. They pointed out that a configuration with $\sigma = 1$ (perfect phase synchronization) naturally leads to disease extinction. This situation is clear: all nodes are either infected or recovered simultaneously and there are no susceptible-infected interactions. In the Chapter 8 we will explore the conditions of extinction in a spatial variation of the model of Kuperman & Abramson.

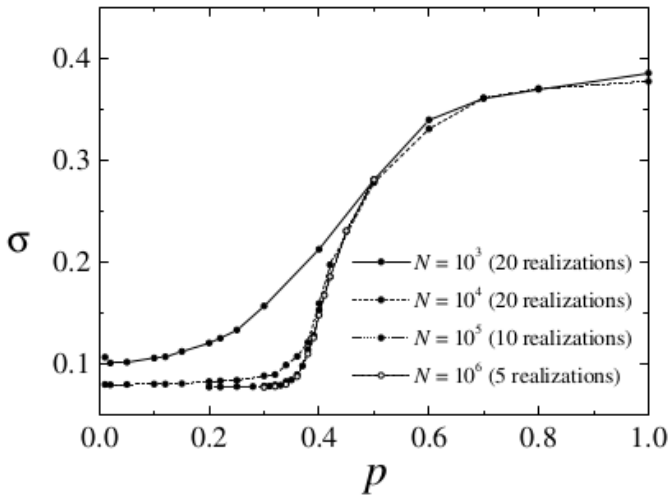

Figure 6.5: Order parameter as a function of the disorder for different network sizes (shown in legends). Each point is a time average over 2000 time steps. Simulations were performed with parameters $\tau_I = 4, \tau_R = 9, K = 3$ and $n_{inf}(t) = 0.1$. Taken from [8].

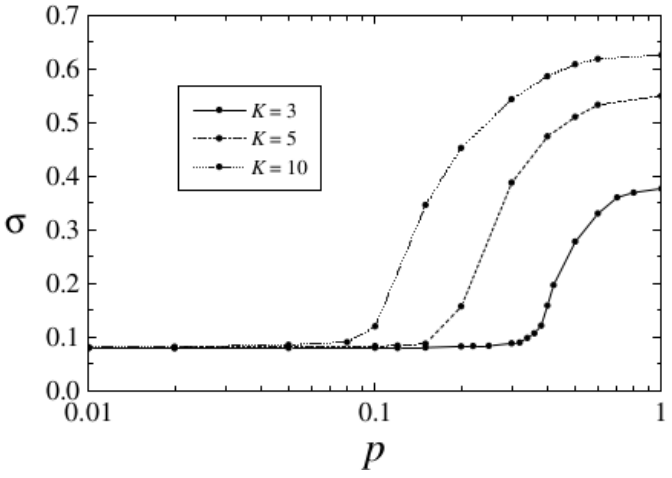

Figure 6.6: Order parameter as a function of the disorder parameter for different average degrees (shown in legends). Simulations were performed with parameters $\tau_I = 4, \tau_R = 9, N = 10^4$ and $n_{inf}(t) = 0.1$. Taken from [8].

6.5 Epidemics in heterogeneous correlated networks

6.5.1 Pair (State) correlations

In the SI model let us define (to avoid confusion with indices) $s_i(t)$ and $x_i(t)$ as the average probabilities that node i is susceptible or infectious at time t. In terms of the adjacency matrix $\mathbf{A}$, with entries A_{ij}, $x_i(t)$ obeys the set of N coupled differential equations

$$\frac{dx_i}{dt} = \beta s_i \sum_j A_{ij} x_j = \beta(1 - x_i) \sum_j A_{ij} x_j, \tag{6.80}$$

$s_i(t)$ obeys similar equations

$$\frac{ds_i}{dt} = -\beta s_i \sum_j A_{ij} x_j = -\beta s_i \sum_j A_{ij}(1 - s_j), \tag{6.81}$$

such that $s_i(t) + x_i(t) = 1$. At early times $x_i(t) \ll 1$ so we can work with the linearized versions of equations 6.80

$$\frac{dx_i}{dt} = \beta \sum_j A_{ij} x_j, \tag{6.82}$$

in matrix form

$$\frac{d\mathbf{x}}{dt} = \beta \mathbf{A}\mathbf{x}. \tag{6.83}$$

The vector $\mathbf{x}$ (of elements x_i) can be written as a linear combination of the eigenvectors of $\mathbf{A}$

$$\mathbf{x}(t) = \sum_r a_r(t)\mathbf{v}_r, \tag{6.84}$$

where the eigenvector $\mathbf{v}_r$ has eigenvalue λ_r. Combining equations 6.83 and 6.84

$$\frac{d\mathbf{x}}{dt} = \sum_r \frac{da_r}{dt}\mathbf{v}_r = \beta \mathbf{A} \sum_r a_r \mathbf{v}_r = \beta \sum_r a_r \lambda_r \mathbf{v}_r, \tag{6.85}$$

and comparing terms, we obtain

$$\frac{da_r}{dt} = \beta \lambda_r a_r, \tag{6.86}$$

with solution

$$a_r(t) = a_r(0) \exp(\beta \lambda_r t). \tag{6.87}$$

Then we have

$$\mathbf{x}(t) = \sum_r a_r(0) \exp(\beta \lambda_r t)\mathbf{v}_r, \tag{6.88}$$

an expression dominated by the term with the largest eigenvalue, λ_N, i.e.

$$\mathbf{x}(t) \sim \exp(\beta\lambda_N t)\mathbf{v}_N. \tag{6.89}$$

As in the fully mixed SI model, the number of infected individuals increases exponentially, with nodes of higher eigenvector centrality becoming infected earlier than those with lower centrality. This "first order" approximation is qualitatively what we expected, but it is not correct: equations 6.80 and 6.81 implicitly assume that the product of the average probabilities $s_i(t)$ and $x_j(t)$ is equal to the average of the probabilities product $s_i(t)x_j(t)$, which in general is not true. If i and j are neighbors (what is needed for transmission) then $s_i(t)$ and $x_j(t)$ are not independent, but correlated.

Let $\langle s_i \rangle$ and $\langle x_j \rangle$ be the average probabilities that node i is susceptible or infected, respectively. In this notation, $\langle s_i x_j \rangle$ is the average probability that node i is susceptible and node j is infected, simultaneously. Equation 6.81 is written as

$$\frac{d\langle s_i \rangle}{dt} = -\beta \sum_j A_{ij} \langle s_i x_j \rangle , \tag{6.90}$$

an equation that cannot be solved due to the unknown term $\langle s_i x_j \rangle$. To obtain i susceptible and j infected we need to start with both susceptibles, then j becomes infected. Note that j must catch the infection from a third (infected) node k, because i is susceptible. This situation is represented by the probability $\langle s_i s_j x_k \rangle$. The total rate as which j becomes infected increases as

$$\beta \sum_{k \neq i} A_{jk} \langle s_i s_j x_k \rangle ,$$

and decreases if either

- i is infected by j (also infected), with probability

$$\beta \langle s_i x_j \rangle ,$$

- or i is infected by $l \neq j$ (also infected), with total probability

$$\beta \sum_{l \neq j} A_{il} \langle x_l s_i x_j \rangle .$$

The rate equation for $\langle s_i x_j \rangle$ is thus

$$\frac{d\langle s_i x_j \rangle}{dt} = \beta \sum_{k \neq i} A_{jk} \langle s_i s_j x_k \rangle - \beta \sum_{l \neq j} A_{il} \langle x_l s_i x_j \rangle - \beta \langle s_i x_j \rangle . \tag{6.91}$$

Equation 6.91 presents the same problema as equation 6.90. The addition of new higher order terms, in this case, three-variable averages. We can continue with our approach, but the succession of equations will never end, it does not *close*. At

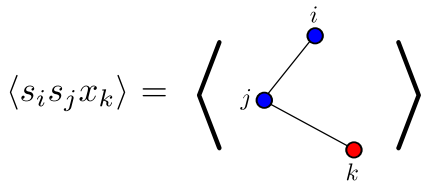

Figure 6.7: Graphical representation of pair correlations. Note that i and j are first order neighbors, also j and k. i and k are second order neighbors, and we assume they are not correlated.

this point, the *moment closure method* is helpful to approximate three-variable averages with (appropriate) combinations of one and two-variable averages.

Using the Bayes Theorem for probabilities

$$\langle s_i s_j x_k \rangle = P(i, j \in S, k \in I) = P(i, j \in S) P(k \in I | i, j \in S). \tag{6.92}$$

Nodes i and j are first order neighbors, also j and k, However, node i and k are second order neighbors, and we assume they are not correlated, then

$$P(k \in I | i, j \in S) = P(k \in I | j \in S) = \frac{P(j \in S, k \in I)}{P(j \in S)} = \frac{\langle s_j x_k \rangle}{\langle s_j \rangle}. \tag{6.93}$$

Combining equations 6.92 and 6.93 we obtain

$$\langle s_i s_j x_k \rangle = \frac{\langle s_i s_j \rangle \langle s_j x_k \rangle}{\langle s_j \rangle}, \tag{6.94}$$

and similarly

$$\langle x_l s_i x_j \rangle = \frac{\langle x_l s_i \rangle \langle s_i x_j \rangle}{\langle s_i \rangle}. \tag{6.95}$$

Equation 6.91 is approximated by

$$\frac{d \langle s_i x_j \rangle}{dt} = \beta \frac{\langle s_i s_j \rangle}{\langle s_j \rangle} \sum_{k \neq i} A_{jk} \langle s_j x_k \rangle - \beta \frac{\langle s_i x_j \rangle}{\langle s_i \rangle} \sum_{l \neq j} A_{il} \langle s_i x_l \rangle - \beta \langle s_i x_j \rangle ,$$

being $\langle s_i s_j \rangle = \langle s_i (1 - x_j) \rangle = \langle s_i \rangle - \langle s_i x_j \rangle$ we obtain

$$\frac{d \langle s_i x_j \rangle}{dt} = \beta \frac{\langle s_i \rangle - \langle s_i x_j \rangle}{\langle s_j \rangle} \sum_{k \neq i} A_{jk} \langle s_j x_k \rangle - \beta \frac{\langle s_i x_j \rangle}{\langle s_i \rangle} \sum_{l \neq j} A_{il} \langle s_i x_l \rangle - \beta \langle s_i x_j \rangle . \tag{6.96}$$

Let define the conditional probability that node j is infected given that i is susceptible, as

$$p_{ij} = \frac{\langle s_i x_j \rangle}{\langle s_i \rangle}, \tag{6.97}$$

the time evolution of this quantity is

$$\begin{aligned}
\frac{dp_{ij}}{dt} &= \frac{d}{dt}\left(\frac{\langle s_i x_j\rangle}{\langle s_i\rangle}\right) \\
&= \frac{1}{\langle s_i\rangle}\frac{d\langle s_i x_j\rangle}{dt} - \frac{\langle s_i x_j\rangle}{\langle s_i\rangle^2}\frac{d\langle s_i\rangle}{dt} \\
&= \frac{1}{\langle s_i\rangle}\left[\beta\frac{\langle s_i\rangle - \langle s_i x_j\rangle}{\langle s_j\rangle}\sum_{k\neq i} A_{jk}\langle s_j x_k\rangle - \beta\frac{\langle s_i x_j\rangle}{\langle s_i\rangle}\sum_{l\neq j} A_{il}\langle s_i x_l\rangle - \beta\langle s_i x_j\rangle\right] \\
&\quad -\frac{\langle s_i x_j\rangle}{\langle s_i\rangle^2}\left[-\beta\sum_j A_{ij}\langle s_i x_j\rangle\right] \\
&= \beta\left(1-\frac{\langle s_i x_j\rangle}{\langle s_i\rangle}\right)\sum_{k\neq j} A_{jk}\frac{\langle s_j x_k\rangle}{\langle s_j\rangle} - \beta\frac{\langle s_i x_j\rangle}{\langle s_i\rangle}\sum_{l\neq j} A_{il}\frac{\langle s_i x_j\rangle}{\langle s_i\rangle} - \beta\frac{\langle s_i x_j\rangle}{\langle s_i\rangle} + \beta\frac{\langle s_i x_j\rangle}{\langle s_i\rangle}\sum_l A_{il}\frac{\langle s_i x_l\rangle}{\langle s_i\rangle} \\
&= \beta(1-p_{ij})\sum_{k\neq i} A_{jk}p_{jk} - \beta p_{ij}\sum_{l\neq j} A_{il}p_{il} - \beta p_{ij} + \beta p_{ij}\sum_l A_{il}p_{il},
\end{aligned}$$

then

$$\frac{dp_{ij}}{dt} = \beta(1-p_{ij})\left[-p_{ij} + \sum_{k\neq i} A_{jk}p_{jk}\right]. \tag{6.98}$$

Using this result, equation 6.90 can be rewritten as

$$\frac{d\langle s_i\rangle}{dt} = -\beta\langle s_i\rangle\sum_j A_{ij}p_{ij}, \tag{6.99}$$

with solution

$$\langle s_i(t)\rangle = \langle s_i(0)\rangle\exp\left(-\beta\sum_j A_{ij}\int_0^t p_{ij}(t')dt'\right). \tag{6.100}$$

Together, equations 6.98 and 6.100 provides a "second order" approximation to the SI model on networks. Note that p_{ij} is not symmetric in i and j, so we need to consider two equations of the form 6.98.

In figure 6.8 we see a comparison of the first- and second-order approximations for the SI model against simulation data in networks (a) with low transitivity (i.e. low cluster coefficient) and (b) with high transitivity. A network with low transitivity is mostly tree-like. The second-order approach is exact because i) the only path between i and k is through j, or ii) there exists another path (through a long cycle) which dilutes any correlation between i and k to the point that it can be neglected. The first-order approach shows good agreement with simulation data, is it just one order below the exact order of correlations, which is length two.

For networks with high transitivity the situation depicted in figure 6.7 does not holds in general: it becomes possible for k to be directly connected to i (if i, j and k form a cycle of length three) an they become correlated. On the other hand, short cycles of order $4, 5, \ldots$ also become possible and the correlations through them may not be diluted to the point of being negligible, so correlation length higher than two may arise. In this case the second-order approach is a relatively good approximation, while the first-order approach shows a poor agreement with the simulation data.

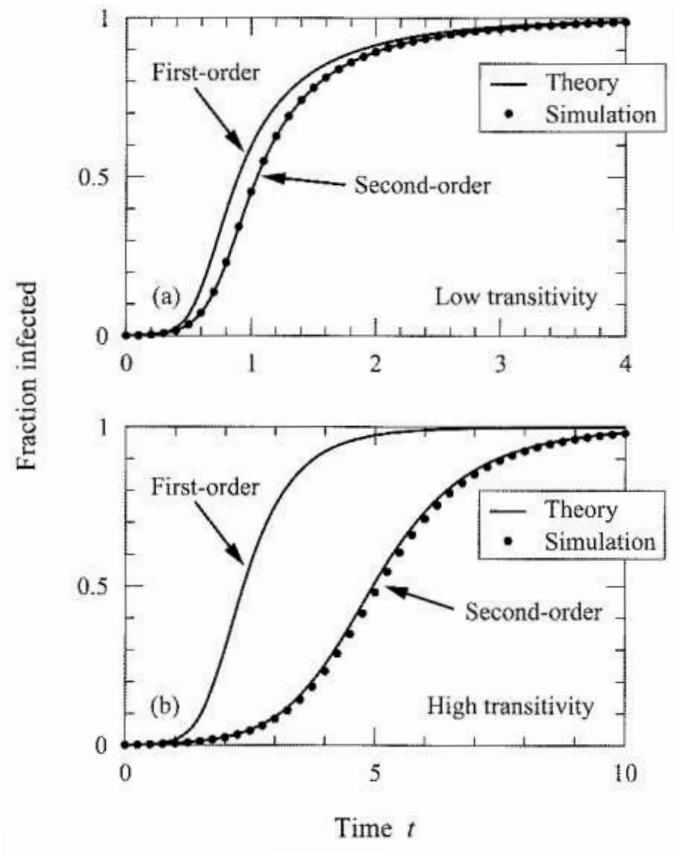

Figure 6.8: First- and second-order approaches in comparison with direct simulations for the SI model. The first- and second-order approaches give similar results in tree-like networks with low clustering coefficient, or *transitivity*. For networks with high transitivity the first-order approach shows a bad agreement with simulation data (dots). Taken from [6, section 17.10].

Chapter 7

Excitable media

Many time evolving systems can be thought of as a dynamical system, with the connotation of a force or perturbation producing a change of state from a resting state to an excited state. If the perturbation needs to be strong enough to produce an excitation, and a given amount of time is needed to restore the system to its resting state, we are dealing with a special class of systems called *excitable media*.

7.1 Characteristics of excitable media

In general terms, any excitable medium shares three qualities:

- In absence of external perturbations, the system possesses a globally attracting equilibrium point, called the *resting state*.
- The phase space if partitioned by a *threshold* into two domains: sub threshold and supra-threshold. A system in the supra-threshold domain is said to be *excited*.
- The transition from the resting state to a excited state, i.e. excitation, only takes place when an external perturbation is greater than the *threshold*. After a small perturbation, that leaves the system in the sub-threshold domains, the system quickly move back to the resting state. After a large perturbation, that leaves the system in the supra-threshold domain, the system performs a large excursion before returning to the resting state.
- It possesses a *refractory period*, i.e. the time needed before the system can be re-excited after a relaxation from the excited state. Before this period, a re-excitation it not possible, even for supra-threshold perturbations.

These are the qualities of a single excitable unit, however, real excitable systems are typically composed of many spatially extended units that display collective dynamics, such as pattern formation and spatiotemporal chaos. One characteristic trait of excitable media dynamics is that two traveling waves of excitation cannot pass through each other. Instead, they annihilate each other. This happens because each excitation front leaves a refractory region behind it, so the

system cannot be re-excited immediately. Clearly, excitable systems composed of a single excitable unit do not display this kind of phenomena.

Spatially extended excitable media are a sub-class of the *reaction-diffussion* systems framework. This framework is a natural description of chemical systems. The language of chemistry is useful to explain the generalities of its coarse-grained mathematical model.The partial differential equation describing their evolution is

$$\frac{\partial}{\partial t}\mathbf{q}(\mathbf{x},t) = \mathbf{D}\nabla^2\mathbf{q}(\mathbf{x},t) + \mathbf{R}(\mathbf{q}), \tag{7.1}$$

where

- $\mathbf{q}$ is the system state vector,
- $\mathbf{D}$ is the diffusion matrix, and
- $\mathbf{R}(\mathbf{q})$ are the non-linear reaction terms.

The first term in the right hand side of equation (7.1) accounts for the components transport, while the second term contains details about the local dynamics of the components; i.e. the reactants are reacting with each other and the products are being transported via diffusion.

Some of the patterns of an extended excitable media can be explained analytically (through equation (7.1) for example) while many others require the use of numerical techniques. The process implies a discretization of the space into a lattice to obtain a (large) system of ordinary differential equations from the system (7.1). In these ODEs, diffusion is implemented by coupling lattice variables to those of its nearest neighbors.

7.2 Examples of excitable media

Excitable media are ubiquitous. Even something as trivial as a toilet flush is an excitable system; it requires a minimum force to be exerted on the lever (threshold) to start the water flow (excitation), and the toilet flush is useless until the cistern is refilled (refractory period).

Excitable media span a broad spectrum of scales, from waves of calcium concentration at the intracellular scale, to plankton growth of up to hundreds of kilometers in the oceans.

7.2.1 Calcium waves in the cell

At the intracellular level, calcium ions (Ca^{2+}) are vital as physiological regulators in the living cell. They signal muscular contractions in the heart and uterus, and act as universal signal messengers. Calcium concentration should be carefully controlled. An excess can induce apoptosis, i.e. the process of cellular self killing. Calcium concentration changes when Ca^{2+} ions form complexes, limiting the amount of free Ca^{2+} ions, and developing calcium waves which often form spatial pattern such as spiral waves.

7.2.2 Waves in the Belousov–Zhabotinsky reaction

The chemical reaction known as the Belousov–Zhabotinsky reaction (B-Z reaction) is perhaps the simplest experimental setup to study a nonlinear system displaying excitations and oscillations, i.e. an excitable system; it has been a convenient system to explore the mechanism for pattern transitions, to prove techniques of spiral waves control, and to explore the dynamics of spiral and scrolls in heterogeneous and anisotropic excitable media [10, Chapter 4].

Figure 7.1: Spiral wave patterns in the Belousov–Zhabotinsky reaction. Taken from [9, section 2.5].

The preparation of the B-Z reaction does not requires any sophisticated equipment nor expertise, and the recipe only contains a couple of reactants [9, Section 2.5]

First Mix

- 67 ml of water
- 2 ml of concentrated sulfuric acid
- 5 gm of sodium bromate

Then, in a glass vessel put

- 6 ml of the previously prepared mix.
- 1 ml of malonic acid solution (1 g in 10 ml)
- 0.5 ml of sodium bromide solution (1 g in 10 ml)

- wait until bromide color to vanish.
- 1 ml of 25 mM of phenanthroline.

Pour the final solution into a covered Petri dish, and once the reaction starts a pattern and spiral waves appears in the surface of the solution. These patterns correspond to spatially varying concentrations.

7.2.3 Waves of electrical activity in the heart

The understanding of pattern formation in excitable media helps in dealing with heart conditions such as cardiac arrythmias and fibrillation.

The heart is an excitable organ; it behaves as a pump by sending deoxygenated blood to the lungs and oxygenated blood to the rest of the body. The pressure of the pump is due to contraction and relaxation of electrically excitable cells in the heart wall. Then waves of excitation propagates along the heart wall.

Under normal conditions the heart beat is set by the sinus node, whose cells are spontaneously activated at a given frequency. This excitation propagates throughout the atria, and to the ventricles after reaching the atrio-ventriular node, so there is a time delay between activation of the atria and the ventricles.

The presence of physical obstacles such as inexcitable scar tissue or transiently inexcitable regions (refractory regions) allows the formation of *reentrant pathways*, i.e. an excitation loop around a core region, when the excitation goes around the obstacle. The typical pattern of reentrant pathways are spiral waves, which persists even after the original refractory region is recovered. The excitation frequency of these spiral waves is likely to be higher than that of the sinus node, thus dominating the heart beat, setting an abnormally high frequency. High frequencies make spiral waves unstable, increasing the probability of wave break up, then forming more spirals until the system reach a state of spatiotemporal chaos. This scenario is know as ventricular fibrillation.

Many people suffering from arrhythmias uses implanted defibrillators to prevent fibrillation, however, the high current that it uses to reset the heart to it resting state can cause scars which in turn promote the development of future arrhythmias. So understanding the dynamics of pattern formation in excitable media can help in the development of safe and effective low current defibrillators.

7.2.4 Waves of spreading depression and persistent activity in the brain

Brain tissue behaves as an excitable medium. i) In the human brain the migraine aura (a flickering sensation that can drastically reduce the field of vision) has been associated with Spreading Depression, a diffusive wave of depolarization resulting from a temporary breakdown of the trans-membrane potential as there occurs a large-scale movement of ions across the membrane. Spreading depression typically lasts for a couple of minutes before recovery to the original electrophysiological

state. ii) Patterns of excitation also appear in neocortical slices of the mammalian brain.

It has been speculated that spiral waves make possible the persistence of cortical oscillations. If such spatial patterns emerge spontaneously from arbitrary initial conditions, then they should be an outcome of the system's internal structure alone.

The brain displays a complex topology. Neurons in the brain have processes that can span large distances (compared to the cell body) and a single neuron can form several thousand connections with other neurons. A majority of the neuron connections may be with neurons at short inter-body distances. However, connections with farther neurons are not rare. Also, neurons are embedded in a matrix of regular topology made from glial cells. As calcium waves mediates glial-glial and glial-neuron communications, the aggregate system of glial cells and neuron forms a network with "small-world" topology [1].

7.2.5 Waves of infection in a population

Mankind has been exposed to infectious diseases since its origin. In the 14th century the *Black Death* killed millions when waves of infection spread across Europe for years. In the 19th century the spread of bubonic plague waves killed millions in India and China. In the period 1918-1919 the *spanish flu* traveled the world in three waves and killed about 500 million people. In 2009 the influenza strain AH1N1v started spreading around the world, killing about 18000 people in 214 countries. Another recent example of disease wave propagation is the spatial pattern of malaria spreading in India reported by Jesan in 2013, see figure 7.2. These are just a few examples of how diseases can display waves and spatial pattern and then be related to excitable media.

In general terms, epidemic spreading implies the propagation of a disease among a population. Such disease often involves the passage of individuals through a finite set of discrete disease states. An individual gets infected by direct interaction/contact with its infected neighbors. If a disease does not confer permanent immunity to the healed individuals, an epidemy could be recurrent. These characteristics, in conjunction with an refractory disease state, define a cellular automaton version for modeling an excitable media.

The SIRS epidemic model is akin to this version, and it is suitable for modeling non-fatal instances of malaria and influenza; an excited cell returns to the initial susceptible (excitable) state passing through other discrete states until recover. Also, the excitation (infection) of any cell depends on its current state and the number of excited (infected) neighbors.

Waves of infection observed in the SIRS model resemble the excitation wave fronts. Understanding the spatial pattern of disease spreading will help in developing better methods of controlling disease propagation.

[1] as seen in Chapter 1

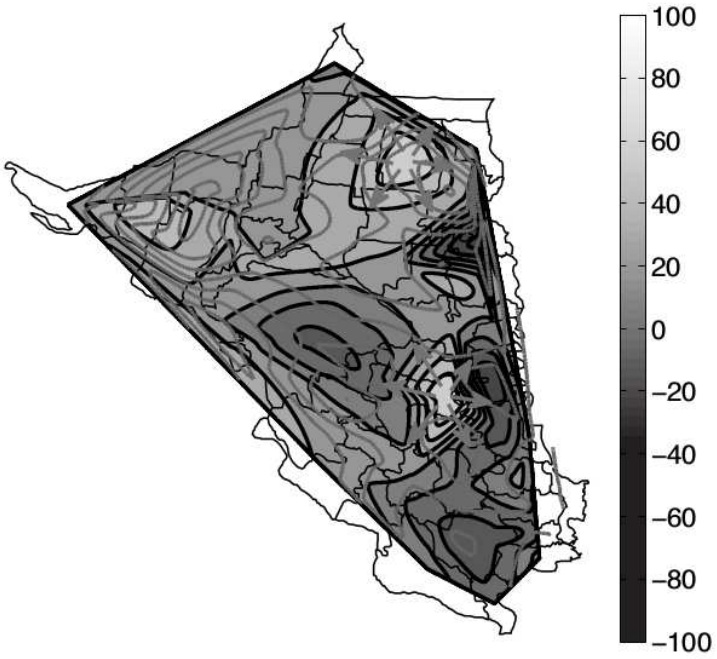

Figure 7.2: Spreading pattern of malaria in northern Bengal, India. Lighter regions mark places or earlier disease activity reported from January 206 to February 2009. Gray scale encodes the mean relative phase angle, i.e. the average phase angle of a given location relative to the spatial average over all regions. Taken from [10, section 4.7].

7.2.6 Waves in Lotka-Volerra model

When species of predators and preys interact they can produce spatiotemporal patterns [10],oscillating population cycles [102], and even chaos [103]. In the simplest case two populations are coupled, then changes in one population affects the other. The most famous model capturing such situation is the Lotka-Volterra model. However, the Lotka Volterra model is generalizable to the case of S species interacting through a complex graph of dominance relations [104].

Dobrinevski and Frey pointed out that i) deterministic rate equations of the general Lotka-Volterra model possess fixed extinction points (i.e. just one species survive) and coexistence points (all species survive), and ii) as stochasticity is introduced into the model, extinctions always occur on a time scale T_{ext} dependent on the system size N and the stability of the fixed points.

One could expect [105]

- *Stable coexistence* with $T_{ext} \propto e^N$ if the deterministic dynamics has a stable attractor in the coexistence region.
- *Unstable coexistence* with $T_{ext} \propto \log(N)$ if the deterministic dynamics approaches the extinction hyper planes for large times, and weak fluctuations are sufficient to make one species go extinct.

- *Neutrally stable coexistence* with $T_{ext} \propto N^{\gamma}$ if the deterministic dynamics possesses a family of neutrally stable, closed orbits, corresponding to the existence of a conservation law.

In particular, the three species cyclic Lotka-Volterra model [106, 107] (also known as the rock-paper-scissors game) and the SIRS model presents some similarities; both have cyclic dynamics through a set of species or disease states.

7.3 The Hodgkin-Huxley model

The work of Alan Hodgkin, Andrew Huxley and Bernard Katz was pioneering to understand the electrical activity of the cell. They described the mechanism for the initiation of the *action potential* in the squid giant axon, and proposed a theoretical model which captures its profile through a RC circuit.

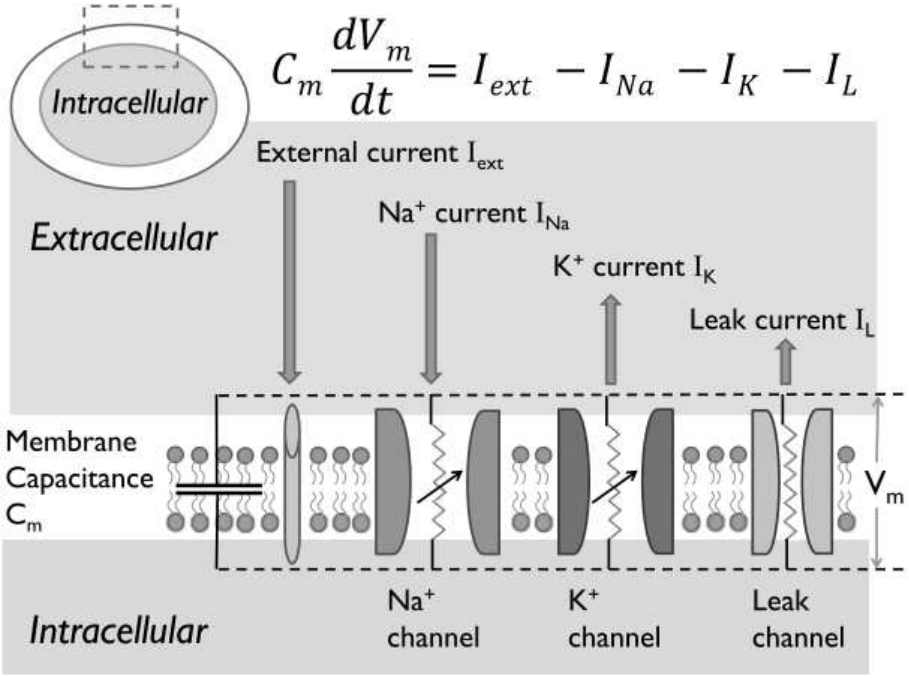

Figure 7.3: Schematic representation of the cell membrane potential as an electrical circuit. The lipid bilayer is represented by a capacitor C_m that maintains the potential $V_m = V$ across the cell membrane. Ion-channels act as voltage dependent rheostats for controlling the inward Na^+ and outward K^+ currents. Taken from [10, section 2.2].

The trans-membrane voltage follow the differential equation

$$C_m \frac{d}{dt} V = I_{ext} - I_{ion}, \tag{7.2}$$

where I_{ext} is the external current that depolarizes the cell membrane, and I_{ion} is the sum of ionic currents

$$I_{ion} = \sum_i I_i, \tag{7.3}$$

each one given as

$$I_i = g_i(V - E_i), \tag{7.4}$$

where E_i is the equilibrium/reversal membrane potential for I_i (the equilibrium potential is reached when there is no net flow of charges across the membrane). The conductance g_i are voltage gated, that can be in ON or OFF state, with a probability modeled as

$$\frac{d}{dt}p_i = \alpha_i(V)(1 - p_i) - \beta_i(V)p_i, \tag{7.5}$$

where $\alpha_i(V)/\beta_i(V)$ is the voltage dependent rate for the OFF $\rightarrow$ ON/ON $\rightarrow$ OFF transition of gate i (this values needs to be determined experimentally). The macroscopic conductance of the ion channel G_{ic} is the product of the individual gate probabilities

$$G_{ic} = \bar{g}_{ic} \prod p_i, \tag{7.6}$$

where $\bar{g}_{ic}$ is the maximum conductance of the ion channel (when all its gates are ON).

Calling the probabilities of the individual gates m, n and h, the total ionic current is

$$I_{ion} = \bar{g}_{Na}m^3h(V - E_{Na}) + \bar{g}_K n^4(V - E_K) + \bar{g}_L(V - E_L) \tag{7.7}$$

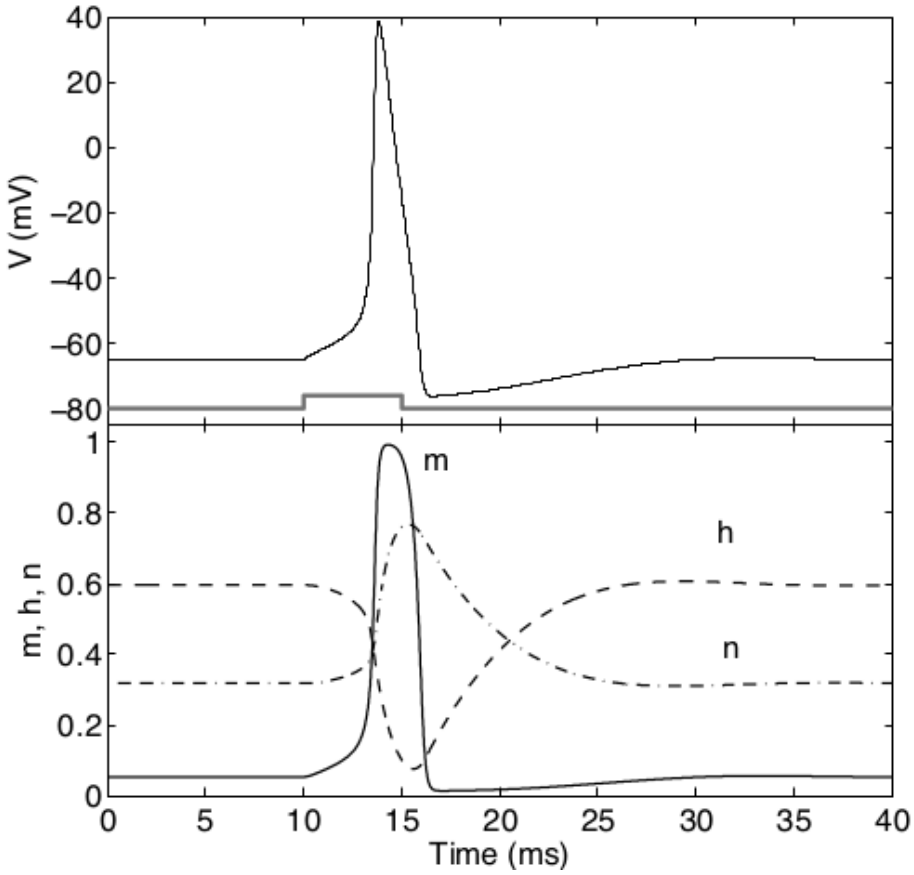

Figure 7.4: Dynamics of the Hodgkin-Huxley model variables. Time evolution of voltage V (top) and gate probabilities (bottom) as a response to $I_{ext} = 4\mu$A applied for 5 ms. Taken from [10, section 2.2].

7.3.1 Integrate-And-Fire neurons in small-world networks

Roxin et al. [11] used integrate-and-fire neurons, with membrane potential

$$\tau_m \frac{d}{dt} V_i = -V_i + I_{ext} + g_{syn} \sum_{j,m} A_{i,j} \delta(t - t_j^{(m)} - \tau_D). \tag{7.8}$$

Neuron i fires when its voltage exceeds 1, then it is reset to 0. If $A_{i,j}$ is nonzero, then neuron i receives a pulse stimulus of amplitude g_{syn}, delayed by τ_D after neuron j fired its m-th spike (at time $t_j^{(m)}$). The synaptic conductance g_{syn} is chosen to satisfy $I_{ext} + g_{syn} > 1$, then, a single input suffices to induce firing. The coupling topology A is a small-world network constructed using the Newman-Watts model [26], in a ring. The long-range links added are *directed*, in order to account for the nonreciprocal character of synaptic connections.

The directed nature of the long-range links allows to overcome extinctions due to wave annihilation: activity is re-injected into already recovered regions. The

probability of survival/failure of the sustained activity is modified by the density of long-range links.

For small values of τ_D (i.e. fast excitation waves) the probability of failure (extinction of the dynamics) is an increasing function of p, with increasing steepness as the system size increases. For larger values of τ_D, $F(t)$ becomes a non-monotonic function of p, making it possible to observe long lived activities.

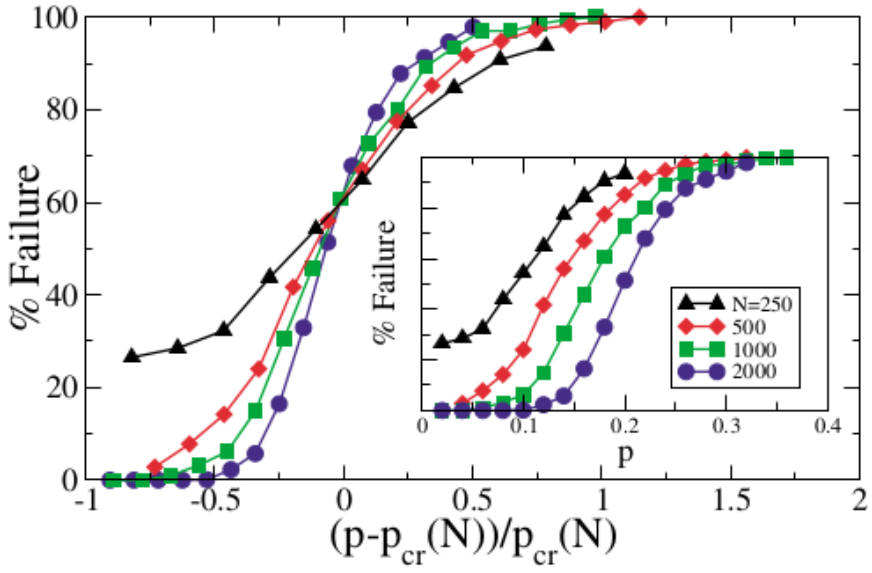

Figure 7.5: Failure (extinction of the dynamics) rates as a function of p, normalized by $p_c(N)$. The inset shows the failure rate as a function of p. Several system sizes are shown. Taken from [11].

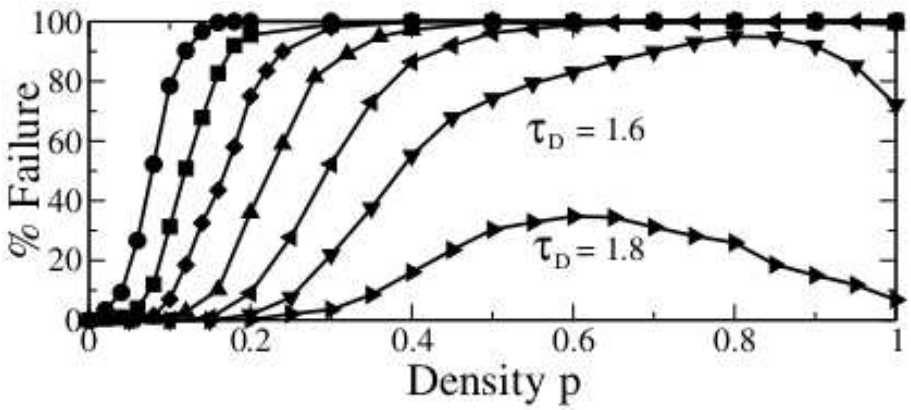

Figure 7.6: Failure (extinction of the dynamics) rates as a function of p for system of size $N = 1000$ and $\tau_D = 0 : 6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8$ (from left to right). Taken from [11].

7.4 The FitzHugh-Nagumo model

One paradigmatic model of excitable unit is the FitzHugh-Nagumo model

$$\frac{du}{dt} = F(u,v) = u(u-u_c)(1-u) - v, \tag{7.9}$$

$$\frac{dv}{dt} = G(u,v) = \epsilon(ku - v) \tag{7.10}$$

where $k \geq 0$, and $\epsilon \ll 1$. The parameter ϵ controls how slow the time evolution of v is in comparison to u. The variable u (which corresponds to V in the Hodgkin–Huxley model) is the *excitation variable*, it represents the rapid depolarization of the cell after a supra-threshold perturbation. The variable v is the *recovery variable*, it represents the recovery of the cell to the resting state. The FitzHugh-Nagumo model was aimed at understanding the electrical signal propagation along a nerve with a two-variable model. In the context of biological cells, this model describes a profile called the *action potential*, originally proposed by Hodgkin-Huxley with a four -variable model.

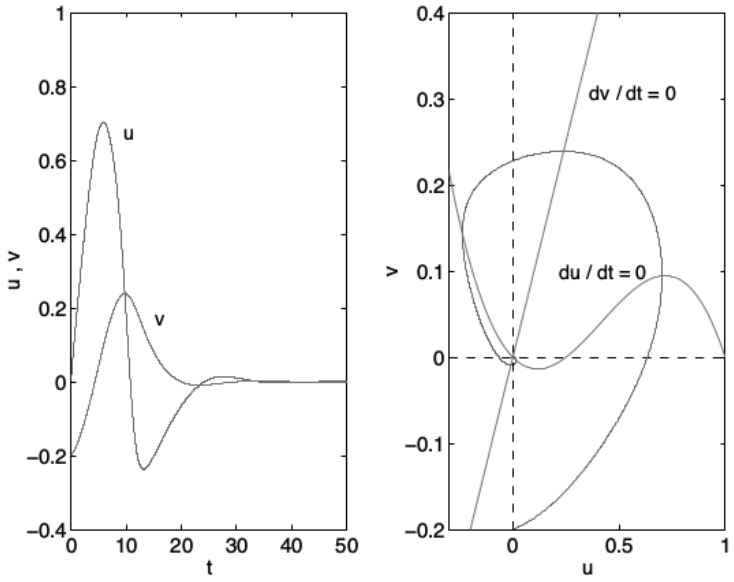

Figure 7.7: Time evolution of variables u and v in the FitzHugh-Nagumo model (left) and the corresponding trajectory in the (u,v) plane (right). Parameter values were $u_c = 0.25, k = 1, \epsilon = 0.1$, and the initial condition were $(0,-0.2)$. Taken from [10, section 1.2].

In equation (7.9) the cubic term sets the system's available fixed points. There are two stable fixed points at $u = 0$ and $u = 1$, and one unstable fixed point at $u = u_c$. The second term is a variable that controls a saddle-node bifurcation in the bistable system that makes unstable one fixed point, at $u = 0$ or $u = 1$, so the system can return to the remaining stable fixed point (resting state).

Figure 7.7 shows both the time evolution and the phase space trajectory of variables u and v in the FitzHugh-Nagumo model. We can observe a large-amplitude departure of u and v from the initial condition at $(0, -0.2)$ before it returns to the resting state at $(0, 0)$. This qualitative behavior is the blueprint of the action potential.

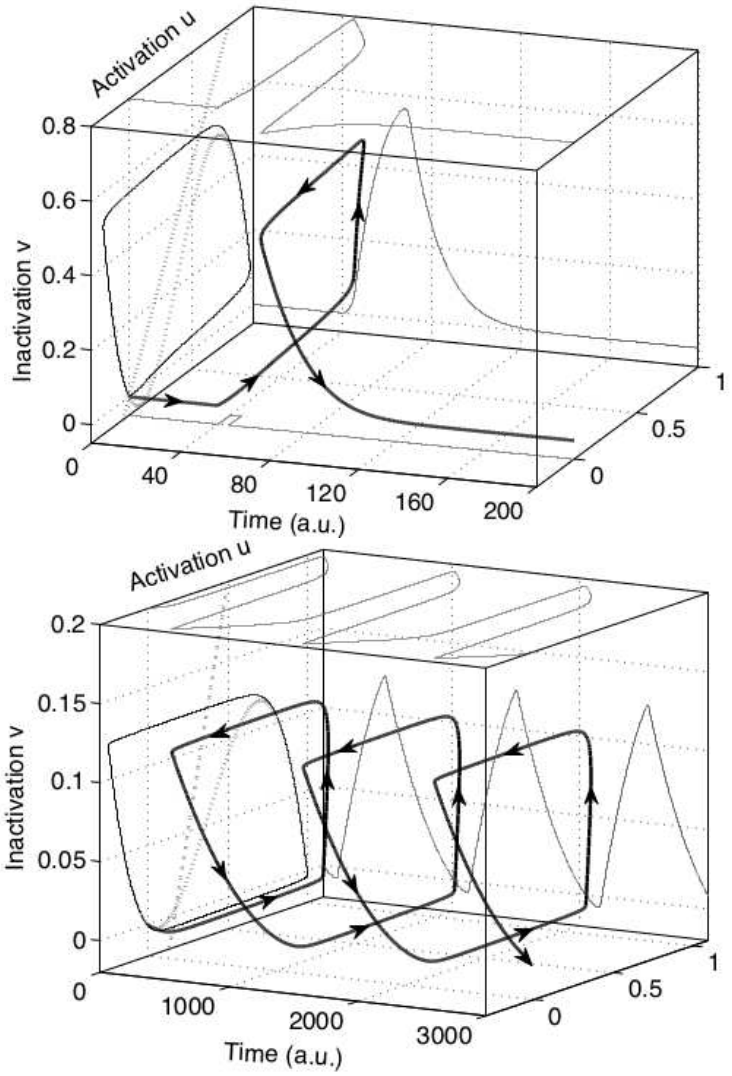

Figure 7.8: Dynamic behaviors of the FitzHugh-Nagumo model. Excitable dynamics (top) and oscillatory dynamics (bottom). Taken from [10, section 2.3].

Depending on the u- and v-nullclines [2] intersection, the system presents qualitatively different behaviors

- If the intersection is located at the left of the u-nullcline minima, then, the dynamics is excitable, see figure 7.8 on top.

[2]The **nullcline** of a dynamic variable x is defined as $\dot{x} = 0$. $\dot{\square}$ stand for "derivative with respect to time".

- If the intersection is located in between the u-nullcline minima and maxima, then, the intersection corresponds to a "limit cycle", and the dynamics is oscillatory, see figure 7.8 on bottom.

7.5 The Chialvo Model

The need for discretization has motivated the development of discrete models of excitability, such as the one proposed by Dante R. Chialvo [108], a two-variable map which evolves in discrete time intervals

$$x_{t+1} = f(x_t, y_t) = x_t^2 e^{y_t - x_t} + k, \tag{7.11}$$
$$y_{t+1} = g(x_t, y_t) = a y_t - b x_t + c, \tag{7.12}$$

where x is a fast *activation/potential* variable, y is a slow *recovery* variable

- k is the external stimulus applied to the system,
- $a < 1$ is a time constant for the recovery dynamics,
- $b < 1$ is the constant for the activation-dependence of the recovery, and
- c is an offset.

Without any external stimulus ($k = 0$), the stable equilibrium of the system is

$$x^* = 0, \qquad y^* = \frac{c}{1-a},$$

otherwise x increases fast, followed by a slow change in y until the system returns to the resting state.

The dynamics of the Chialvo model is clear after a phase plane analysis. Let us define the first difference of the state variables

$$\Delta x = x_{t+1} - x_t, \tag{7.13}$$
$$\Delta y = y_{t+1} - y_t. \tag{7.14}$$

The x-nullcline is the set of values for which $\Delta x = 0$:

$$x^2 \exp(y - x) + k = x \tag{7.15}$$

i.e.

$$y = \ln(x - k) - 2\ln(x) + k. \tag{7.16}$$

The y-nullcline is the set of values for which $\Delta y = 0$:

$$ay - bx + c = y \tag{7.17}$$

i.e.

$$y = \frac{bx - c}{a - 1}. \tag{7.18}$$

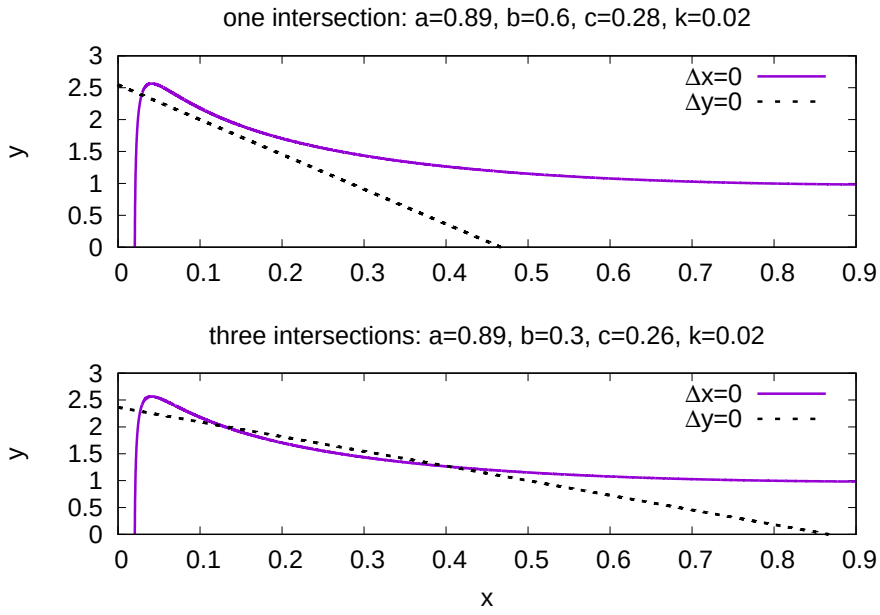

Figure 7.9: Nullclines of the Chialvo model. In dependence of the parameter values, one (top) or three (bottom) intersections are possible.

Depending on the parameter values, the nullclines intersect at one or three points, see figure 7.9. When there is only one intersection, the intersection point is stable. When there are three intersections: the left one is stable, the middle one is a saddle, and the right one is unstable. Focusing on the case of a single intersection, we can distinguish two types of dynamics

1. Excitable dynamics: if the intersection point lies at the left of the x-nullcline maxima the fixed point is globally attracting. A supra-threshold perturbation to the system results in an action potential. This generally happens at low values of k.

2. Oscillatory dynamics: if the intersection point lies to the right of the x-nullcline maxima, the fixed point loses stability, it bifurcates into a stable fixed point and a stable limit cycle. The limit cycle resembles and envelope for the x-nullcline. Trajectories initiated outside the limit cycle converge to it. Trajectories initiated close to the fixed point spirals toward it.

If more than one intersections exists, the Chialvo model allows for bursting and chaotic solutions.

When dealing with extended excitable media, it is useful to think of this model as a model of forest-fire propagation (one with large time and space scale): in such case, the variable x represents the "local temperature" of a given tree, while variable y represents its "local height". In a full grown forest, a tree burns if its temperature is raised above a certain threshold and will increase the temperature

of its neighbors by diffusion. Fire propagates if heat-transference conditions are adecuate.

7.5.1 Chialvo model in small-world networks

To simulate extended excitable media consider a two-dimensional array of $N \times N$ Chialvo maps with parameters $a = 0.89, b = 0.6, c = 0.28, k = 0.02$, coupled diffusively as [108]

$$x_{t+1}^{i,j} = (1 - D) f(x_t^{i,j}, y_t^{i,j}) + \frac{D}{4} \sum_{q=1,-1} f(x_t^{i+q,j+q}, y_t^{i+q,j+q}),$$

where $D = 0.2$ is the diffusion coefficient [3]. The array is located at the vertices of a small-world network constructed with the Newman-Watts model [26]. As the shortcut probability p increases, three dynamical regimes are observable [10, chapter 8]

- $0 < p < p_c^l$: after a transient period (characterized by multiple circular waves), the state of the system is completely covered by one or more self-sustained spiral waves, see figure 7.10. Defining *activity* as the fraction of maps with $x > 0.9$, the transient shows oscillatory activity, which settles to a stable value when spiral waves emerge.

  The probability per unit time for spiral waves to be created increases with the system size N and the shortcut probability p, see figure 7.11.

- $p = p_c^l$: a large fraction of the system simultaneously oscillates between excited and refractory.

- $p > p_c^u(N)$: self-sustained activity ceases and the system is absorbed into $x^{i,j} = 0, \forall i, j$, ans the dynamics get extincted, see figure 7.13.

Here p_c^l is a lower critical probability, and $p_c^u(N)$ is a system-size dependent upper critical probability.

[3] Only the excitation like variable x is diffusively coupled. Consider a forest-fire propagation model; the idea of "temperature" diffusion is natural, however, it would be weird to talk about "height" diffusion.

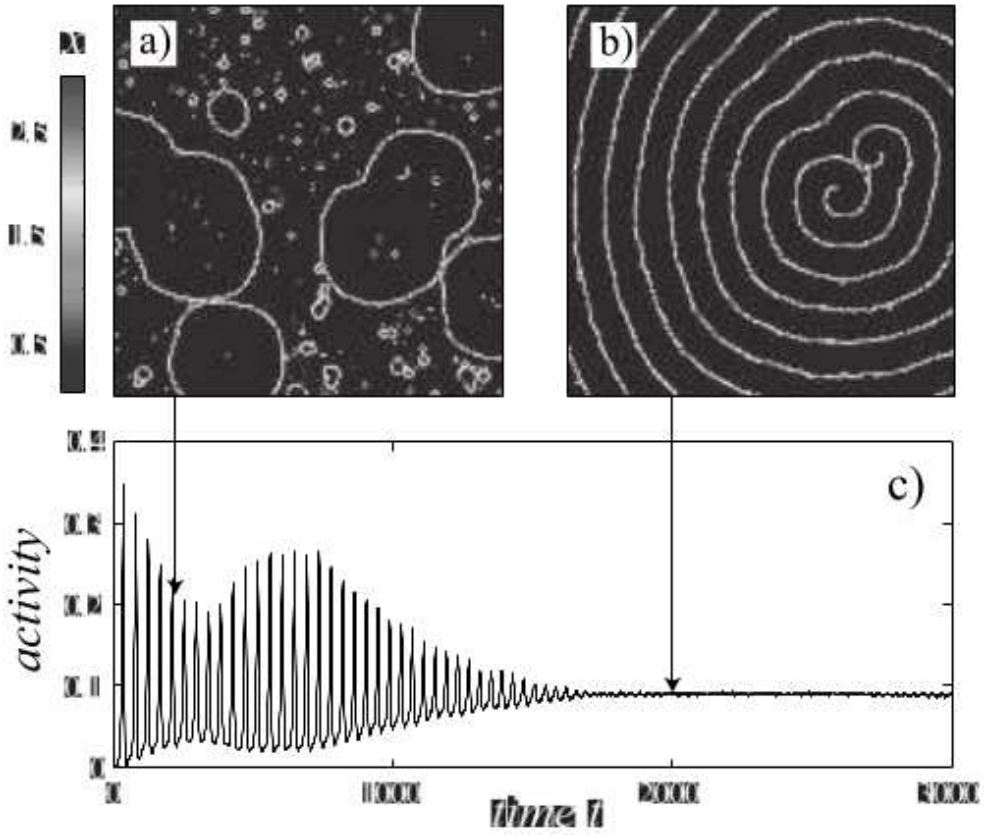

Figure 7.10: Regime $0 < p < p_c^l$ of the Chialvo model in small-world networks. a) circular waves in the transient. b) self-sustained spiral waves. c) Time behavior of activity. Taken from [10, section 8.3].

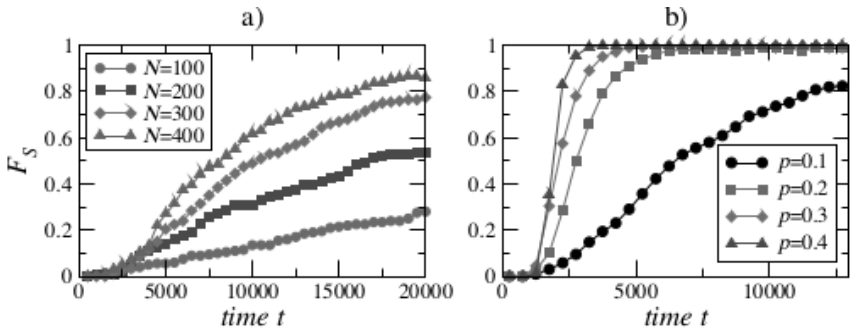

Figure 7.11: Fraction of system with spiral waves, F_S as a function of a) N for $p = 0.05$ and b) p with $N = 300 \times 300$. Taken from [10, section 8.3].

7.6 Effects of complex topologies

Complex topologies constitute a source of heterogeneity in excitable media, and play a role in the collective dynamics. The small-world topology is of special interest in this aspect. As the fraction of long range links is increased in a small

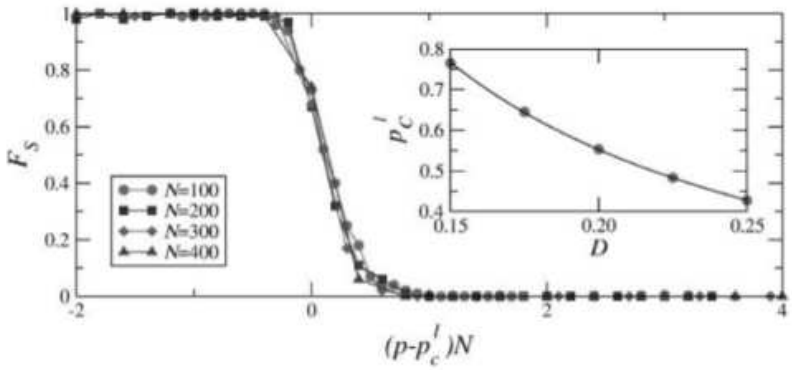

Figure 7.12: Fraction of system with spiral waves, F_S as a function of $(p-p_c^l)N$. The inset shows the dependence of p_c^l with the diffusion coefficient D. Taken from [10, section 8.3].

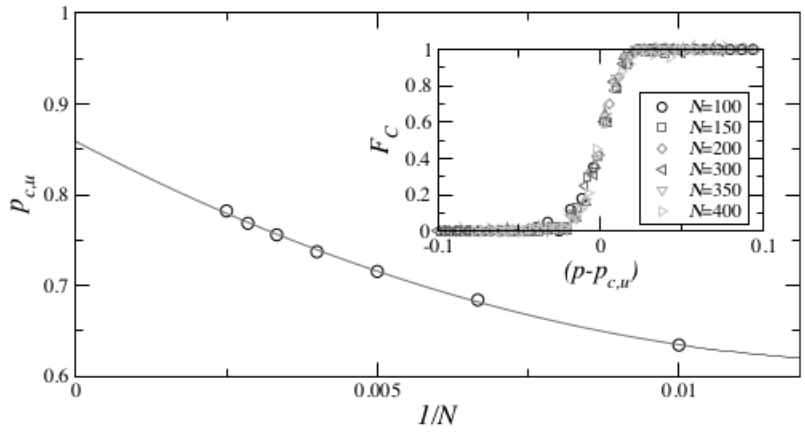

Figure 7.13: Dependence of p_c^u with the system size N. Te inset shows the traction of ceased systems after a time $t = 20,000$, F_C as a function of $(p - p_c^u)$. Taken from [10, section 8.3].

world topology, the following is observed

- a transition between different epidemic patterns in a one-dimensional SIRS model [8],

The topology of an excitable media can also shed light onto our understanding of several diseases related to the brain.

- In order to observe similar patterns to that of epileptic seizures and bursts in a neuronal model of excitable media, it is necessary to take into account a small world topology of connections via glia surrounding [109].
- It has been found that the brains of patients with schizophrenia, depression, or bipolarity have lower glia to neuron ratios when compared to normal subjects. Under normal conditions the glia to neuron ratio is approximately 10:1 [110].

This evidence supports the hypothesis that the topology of connections in the units of an excitable media influences its dynamical behavior and determines which dynamical patterns may be observed. Changing the internal structure allows transitions between different dynamical regimes and then different excitation patterns.

Chapter 8

SIRS model on a network with long-range interactions

Mankind has been exposed to infectious diseases from its dawn. In the middle of an epidemic outbreak questions such as *when will this be over?* or *how long will this last?* may be common. Today, disease spreading is reinforced by the increasing mobility of people around the world, via terrestrial, maritimal and aerial transportation. This mobility represents effective long range interactions between people in the origin and the destination of a moving individual. As a result, diseases could persist due to the existence of several locations for the dynamics.

For a recurrent disease, the time needed for it to become extinct is a random variable of great interest. This information is encoded by the survival probability or *persistence*, $F(t)$, i.e. the fraction of active systems at time t in a large number of realizations of the dynamics [111].

Several works have been aimed to analytically derive expressions for the extinction time on SI, SIS and SIR models, either by birth-death or branching processes [111, 112, 113, 114]. These studies rely on the fully mixed assumption. In epidemiology, the fully mixed assumption gave way to the more realistic **contact network**. The SI, SIS and SIR models over complex (contact) networks have been largely studied in order to determine structural influences on the existence of an epidemic threshold [7, 97, 98, 99, 100, 101] , and the effects of fluctuations [112].

Not much work has been aimed at investigating extinction times on recurrent models such the SIR model with removal and renewal [115] or the SIRS model [116]. Notable is the work of van Herwaarden & Grasman [115] who have proposed a size dependent scaling relation for the expected extinction time in stochastic recurrent epidemics. However, it is unclear what is the effect of spatial embedding and long-range interactions over such scaling.

This chapter contains the second part of the original research presented in this thesis, which is devoted to the study of extinction times of recurrent diseases with the implementation of a SIRS model. To build a realistic model of recurrent diseases, we keep in mind that: systems of interacting units possess an underlying

topology of interactions, i.e. a network; in real world networks, both node degree and link lengths can be distributed accordingly to physical restrictions; the small-world topology in a network has direct influence on its dynamics; the SIRS model with fixed refractory period is a simple instance of excitable media, and; due to pattern formation, synchronization is relevant for understanding the dynamics of excitable media.

We will account for these facts in the analysis of the SIRS model with fixed refractory period developed over small-world networks embedded on a d dimensional lattice with long-range connections following a power-law decay. The focus of our analysis will be synchronization measures and their relation to extinction times of the disease dynamics. We will also explore how the extinction time is influenced by the link-length distribution, $p(l)$, and the system size, $N = L^d$.

The main results of this chapter can be summarized as follows:

1. **Extinctions are always preceded by a transient of hyper-synchronization: Stage 3**.

   The functional relation of the squared coherence z^2 and phase variance σ^2 is common at extinction time: $z^2 = \exp(-\sigma^2)$. This relation suggests that, at extinction time, phase fluctuation is a random variable with approximately Gaussian distribution (subsection 8.3.2).

   This Gaussian (unimodal) distribution of phase fluctuation is always reached through a short transient of hyper-synchronization before extinction (*Stage 3*, in sub-subsection 8.3.3.2). Such hyper-synchronization transient is only visible in time measured from extinction, *time-to-extinction*.

2. **The mean lifetime distribution $F(t)$ may be obtained by mapping the extinction to a First-Passage Time Problem, where the minima of the dynamics frontier size is considered as a random walker biased toward zero**.

   This approach reproduces very well our data in the mean field domain, i.e. $\alpha < d$ (subsection 8.3.5). Agreement with data is striking for small systems. For large systems and $\alpha > d$ the mean lifetime distribution becomes even broader than in the mean field, then differences between data and predictions becomes evident.

3. **A network with power-law link-length distribution has an effective size $\Lambda \equiv L/\langle l \rangle$ [1] and mean lifetime $t_{1/2}(L, \alpha)$. The logarithm of the mean lifetime amplification with respect to $\alpha = 0$, $\ln(y) \equiv \ln(t_{1/2}(L, \alpha)/t_{1/2}(L, 0))$, is a function of the form $\ln(y) = A\ \Lambda^{\beta}$, with $\beta \simeq 3.5$.**

   This allows us to collapse multiple data into a single curve, and to estimate expected mean lifetimes for systems that would be computationally expensive to simulate (subsection 8.3.7).

[1] L is the linear-length of the system, and $\langle l \rangle$ its average links length

Additionally, we observe that for a system of size $N = L^d$, the mean lifetime of the dynamics scales asymptotically as $\ln(t_{1/2}N^{1/2}) = QN$, where Q is a constant that depends on dynamical parameters. This result is consistent with the results of van Herwaarden & Grasman [115] for the SIR model with removal and renewal. The asymptotic slope of $\ln(t_{1/2}N^{1/2})$ increases with α (subsection 8.3.6). The system reported by van Herwaarden & Grasman is similar to the model we used, in the case $\alpha = 0$, i.e. in the mean field domain.

8.1 Model definition

8.1.1 Building networks with power law length distribution

The first step is to assign nodes/sites to the vertices of a hyper-cubic d-dimensional lattice with helicoidal boundary conditions. We want to build a network such that the probability for a pair (i, j) to be connected decays as $p_{i,j} \propto r_{ij}^{-\alpha}$, where r_{ij} is the distance between the pair (i, j). To achieve this, we need to generate long-range links with length distribution

$$p(l) = Al^{-\alpha+d-1} \qquad l_{min} \leq l \leq l_{max}. \tag{8.1}$$

This probability is the product of the quantity of sites at distance l from a reference l^{d-1}, times the probability that two sites at distance $r_{ij} = l$ be connected, $l^{-\alpha}$. Each link is generated from a radius l, with distribution $p(l)$, by inversion [52] (see Appendix A).

Then, select a random site i and connect it to other site j at a distance l. We repeat this procedure until we have assigned links_per_site $\cdot N$ links.

8.1.2 SIRS model specification

In order to initialize the SIRS system, we start by assigning a uniformly distributed last-infection time $t_i \in [0, \tau_0 = \tau_i + \tau_r]$ to every node i, and set the simulation time $t = \tau_0$.

The state of node i is determined by the simulation time t and its last infection time t_i, as follows

- i is infected and infectious if $0 \leq t - t_i < \tau_i$,
- i is refractory if $\tau_i \leq t - t_i < \tau_0$,
- i is susceptible if $t - t_i \geq \tau_0$,

where τ_i, τ_r and τ_0 are the *infected period*, *refractory period* and *active period*, respectively.

The SIRS dynamics is defined by the transitions

- $S \to I$: a susceptible individual with k_i infectious neighbors get infected with probability

$$p_{inf} = 1 - (1 - p_0)^{k_i}$$

where $p_0 \in [0, 1]$ is the *link infectivity*.

- $I \to R$: infection lasts τ_i time steps. After that time an infectious individual becomes refractory.

- $R \to S$: refractory individuals stay so for τ_r time step. After that time a refractory individual becomes susceptible again.

Instead of going over the entire system at each time step, we set up a list containing the labels of all **infectible nodes**. This list is the union of: a) susceptible nodes having at least one infected neighbor, plus b) refractory nodes with at least one infected neighbor, which were themselves infected at an early enough time to become susceptible before their infected neighbor heals. For each infected node i, all its neighbors j having

$$t_i - t_j > \tau_r, \tag{8.2}$$

are infectible (see figure 8.1). Infectible nodes constitute the *surface* where the non-deterministic part of the dynamics takes place ($S \to I$). Simulation ends when the list of infectibles is empty.

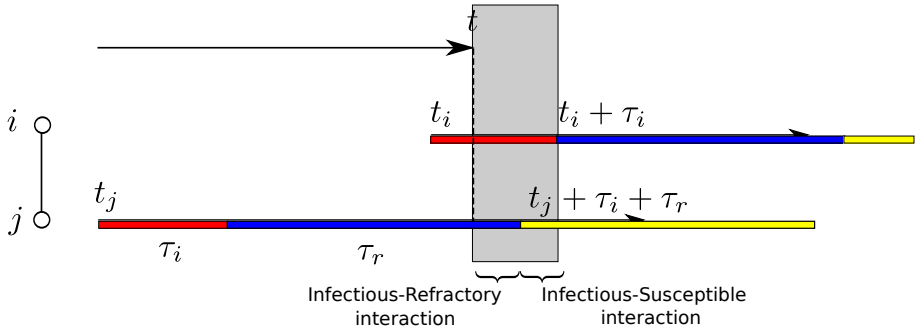

Figure 8.1: Graphical definition of infectible nodes. Consider two neighboring nodes i and j, such that i is infectious. j may be refractory or susceptible. A $S \to I$ transition is possible only if the gray time band if non-zero, i.e. $t_i + \tau_i - (t_j + \tau_i + \tau_r) = t_i - t_j - \tau_r > 0$, this leads to the relation $t_i - t_j > \tau_r$, as given by equation 8.2. In such case we say that node j is infectible. Note that node j will become susceptible before node i heals.

A queue contains infectible sites. We consider them in turn and, if they do not get infected, resend them to the queue. If all their infected neighbors have already healed, we remove them from the list. Otherwise if infected at this time, we send their infectible neighbors to the queue. One sweep through the list constitutes a time-step.

8.1.3 Definitions and notation

Our primary interest is to know the conditions of *extinction*, i.e. a system configuration which avoids disease spreading and turns off the dynamics. Empirically, *extinction means an empty list of infectibles*, and it happens at **extinction time**, t_{ext}. We are particularly interested in how the decay exponent α affects t_{ext}.

For a disease, the extinction time is distributed according to the survival probability or **persistence**, $F(t)$, defined as the fraction of active systems at time t in a large number of realizations of the dynamics [111]. In practice, we measured persistence at time t as the fraction of measurements done at time t with respect to the number of simulation runs performed. Any system that become extinct before t will not contribute to the measurements count.

Persistence allows to implicitly define another time value of interest, the **mean lifetime** $t_{1/2}$, as

$$F(t_{1/2}) = \frac{1}{2}. \tag{8.3}$$

Unlike extinction time, mean lifetime is not a per simulation measure. It is defined within a set of simulations performed under the same conditions. Its interpretation is probabilistic: at $t_{1/2}$ one half ot the dynamics will be extincted. For any $t > t_{1/2}$ the probability of extinction, $1 - F$, will be greater than $1/2$.

Mean lifetime increases fast with the system size N and the decay exponent α. We decided to measure at time t if it is in the "quasi"-geometric sequence $[\tau_0, \lfloor a\tau_0 \rfloor, \lfloor a\lfloor a\tau_0 \rfloor \rfloor, \dots]$, with $a = 1.1$. This policy saves memory, but also reduces time-resolution as the number of measurements increases.

High time-resolution is particularly necessary close to extinction time. A second measurement policy was to maintain a circular array to save the last 10^4 time steps, and save it to disk at extinction time. This allows us to observe behavior backward from extinction time. We denote time measured backwards from extinction as **time-to-extinction**, and it is defined as

$$t' = t_{ext} - t. \tag{8.4}$$

Forward-time measurements has "geometric" step [2] with low time-resolution, while measures in time-to-extinction has unitary step, the highest possible resolution in our setup.

At each measurement time, we register simulation time t and calculate: the fraction of infected nodes i, the fraction of susceptible nodes s, the fraction of refractory nodes r, the fraction of infectible nodes x, the squared coherence z^2, the phase variance σ^2, and the mean phase Θ.

Calculated measures were saved for each simulation run. Then, they were averaged over 1000 simulation runs under the same conditions, i.e, using the same set of values for the settable parameters in our algorithm. We used one starting configuration per network for each set of parameters. All our measurements

[2] steps are taken at increasingly separated intervals, such as 1,2,4,8,...

were averaged over networks at running time. Simulations with multiple starting configuration per network yield similar results.

Link infectivity p_0 is fixed to 0.75, a large enough value to ensure extinctions. Other parameters we maintain fixed are the infectible period $\tau_i = 4$, and the refractory period $\tau_i = 8$.

8.1.4 Acceleration of the dynamics

We set a maximum simulation time of 10^8 times steps. Many simulations did not finish in this time and were orders of magnitude longer. Simulation series in α for systems of moderate size as $N = 40 \times 40$ required weeks of cpu time; larger systems required months without even completing the same range in α. Simulations in one and two dimensions yield similar results.

Long extinction times motivated us to implement annealing within our model. We fix an *annealing probability* p_{ann} such that, at each time step, a total of $p_{ann} \times$ links_per_site links will be reallocated (both edge-ends), with the same length distribution used to build the original network. By using annealing we were able to reduce the mean lifetime, see figure 8.19. Dynamics over annealed networks present the same qualitative behavior than static networks. Just the times scales are reduced, such that, in a fixed amount of wall time, we can simulate larger systems as the annealing probability increases.

8.1.5 Normal distribution over a circle

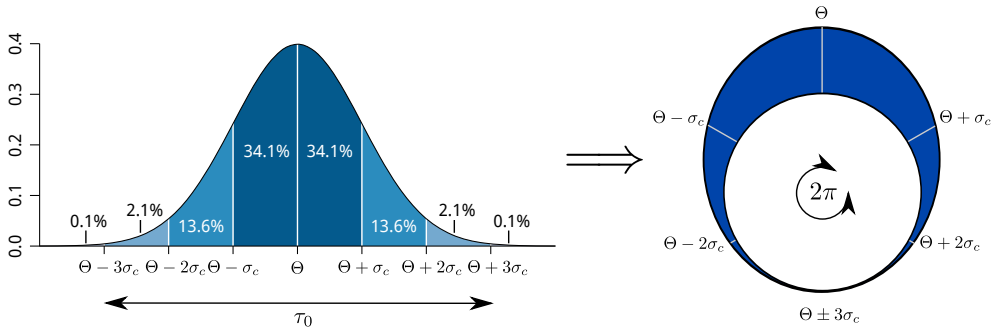

Figure 8.2: Normal distribution $N(\Theta, \sigma^2)$ folded over a circle.

If a variable θ is normally distributed with mean Θ and standard deviation σ, 99.73 % of their probability is within three standard deviations from the mean. Denote this length as $\tau_0 = 6\sigma$, and bend it as a circle (see figure 8.2). Then, we define the constant

$$\sigma_c = \frac{\tau_0}{6} = \frac{\pi}{3} \simeq 1.047 \text{ rad.} \tag{8.5}$$

σ_c^2 will be useful as a phase variance reference when dealing with coherence measures in the next sections.

8.2 Dynamic Analogies

8.2.1 SIRS nodes as oscillators

For the purpose of analysis, every node in the system will be considered as an oscillator running in discrete time through the Susceptible state (S), the Infectious state (I), and the Refractory state (R). As stated before, the Infectious state and the Refractory state possess fixed periods, τ_i and τ_r, respectively. There is no fixed susceptible period; the time a node stays in the Susceptible state is a random variable ($\tau_s \geq 1$), because the $S \to I$ transition is a probabilistic process. This means that our nodes are *non-identical oscillators*, they oscillate at different frequencies ω_i. Non-identical oscillators are generally analyzed in the framework of the Kuramoto model [78].

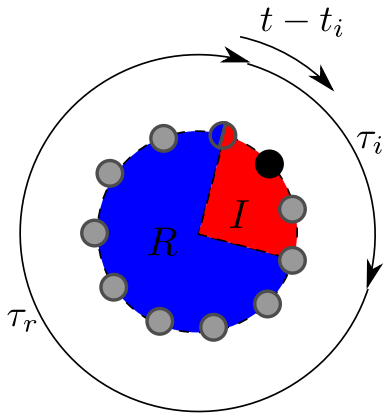

Figure 8.3: Active period circle. The black dot marks the actual location of $t - t_i$ in the active period. Gray dots mark possible locations for $t - t_i$. The red-and-blue dot marks the position where the oscillator enters (red) and leaves (blue) the active period. t_i is the last infection time of oscillator i. The red region represents the Infectious state, while the blue region represents the Refractory state.

We focus on the active period $\tau_0 = \tau_i + \tau_r$, to define a phase θ, i.e. only count active nodes (those in either the Infectious state or the Refractory state). Figure 8.3 shows a graphical representation of the transition of oscillators through the *active period circle*, which depicts the definition of the phase of node i

$$\theta_i = \frac{t - t_i}{\tau_0}, \tag{8.6}$$

used to define an order parameter as done by Kuperman & Abramson [8]

$$ze^{\mathrm{i}\Theta} = \frac{1}{N_a}\sum_{j=1}^{N_a} e^{\mathrm{i}\theta_j} = \frac{M}{N_a}, \tag{8.7}$$

where z is the *coherence*, N_a is the number of active sites, and

$$M = \sum_{j=1}^{N_a} e^{\mathrm{i}\theta_j} = \sum_{j=1}^{N_a} (\cos\theta_j + \mathrm{i}\sin\theta_j)\,.$$

Therefore,

$$M^2 = MM^* = \sum_{i,j=1}^{N_a} (\cos\theta_j + \mathrm{i}\sin\theta_j)(\cos\theta_j - \mathrm{i}\sin\theta_j) = N_a + 2\sum_{i>j}^{N_a} (\cos\theta_i\cos\theta_j + \sin\theta_i\sin\theta_j)\,.$$

8.2.2 SIRS and SIR model with removal and renewal as models for recurrent epidemics

van Herwaarden & Grasman [115] used the WKB approximation on the SIR model with removal and renewal close to the deterministic stable equilibrium, to show that the asymptotic expected extinction time is

$$\ln\left(t_{1/2}N^{1/2}\right) = QN. \tag{8.8}$$

Where Q is a positive constant dependent of dynamic parameters.

The model used by van Herwaarden & Grasman [115] is similar to our model in the case $\alpha = 0$.

8.3 Analysis of simulation results

8.3.1 Types of extinctions

Generically, a disease stops spreading if there are no interactions between susceptible and infected individuals. Let denote the phase space's vertices as

$$\mathbf{I} \equiv (s=0, i=1, r=0), \qquad \mathbf{S} \equiv (s=1, i=0, r=0), \qquad \mathbf{R} \equiv (s=0, i=0, r=1). \tag{8.9}$$

We can distinguish three extinction cases:

a) Extinctions with high coherence: there are susceptible sites but the density of infected individuals is zero. This case is composed by all states in the border line $\mathbf{S} \leftrightarrow \mathbf{R}$ $(s \geq 0, i=0, r=1-s)$. See figures 8.4 and 8.5.

   Coherence is high because all active nodes, if any, aggregate into the refractory range within the active period circle.

b) Extinctions with low coherence: there are infectious sites but the density of susceptible individuals is zero. This case is composed by all states in the border line $\mathbf{I} \leftrightarrow \mathbf{R}$ $(s=0, i \geq 0, r=1-i)$.

   Coherence is low because all active nodes may be dispersed in the full active period circle.

c) Extinctions by isolation: the density of susceptibles and infectious are non zero ($s \neq 0, i \neq 0, r \neq 0$), but they are disconnected from each other, by means of a refractory barrier. This type of extinction does not happen at the any border of the phase space of densities. It is not clear what levels of coherence to expect in this case.

Previous works has reported both extinctions with high levels of coherence [8, 117] and extinctions by isolation [117] in the stochastic SIRS model. We have no report of extinctions with low levels of coherence so far.

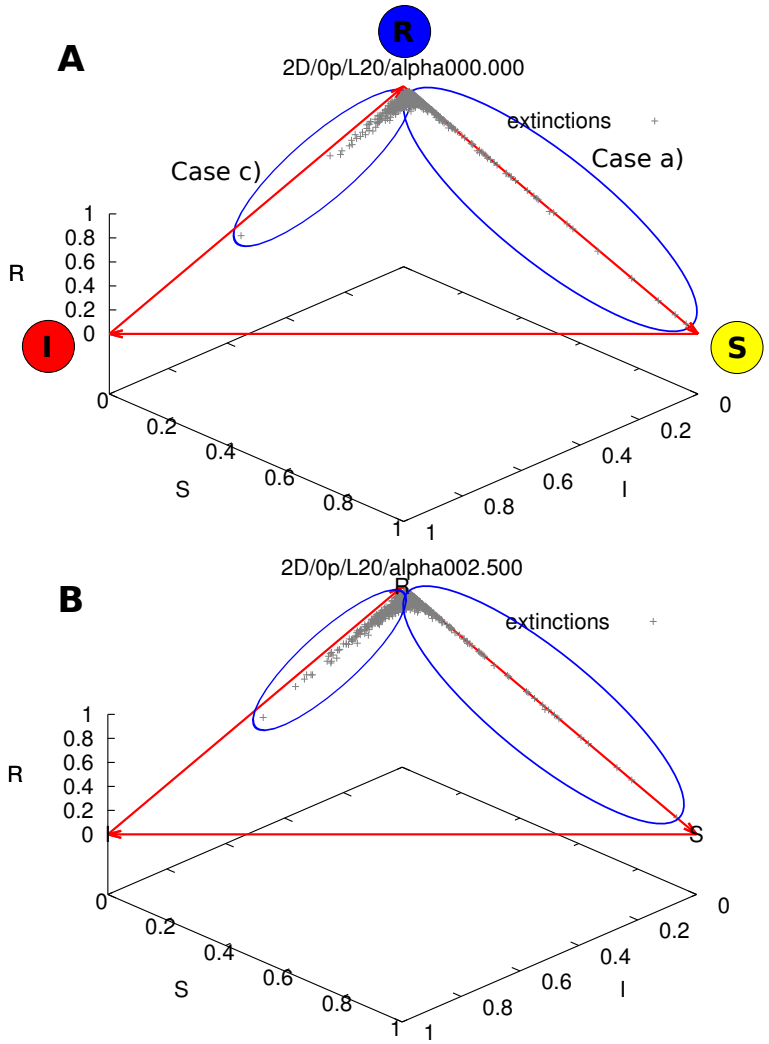

Figure 8.4: Extinction states in phase space for 1000 networks in dimension $d = 2$ with $L = 20$ and $p_{ann} = 0$. Two values of α/d are shown; $\alpha/d = 0$ (A) and $\alpha/d = 1.25$ (B). For non-zero annealing Case c) extinctions are absent, and Case a) extinctions ore more uniformly distributed over the $\mathbf{S} \leftrightarrow \mathbf{R}$ border.

In figure 8.4 our simulations shows only extinctions with high coherence (Case a)) and extinctions by chance (Case c)). None of the extinctions recorded lies

in the border line $\mathbf{I} \leftrightarrow \mathbf{R}$ (Case b)), i.e. we do not observe extinctions with low coherence. Also, most extinction happens near the vertex **R**, indicating that extinctions present high levels of refractoriness.

For non-zero annealing Case c) extinctions are absent, and Case a) extinctions ore more uniformly distributed over the $\mathbf{S} \leftrightarrow \mathbf{R}$ border. In this sense, rewiring through annealing allows the infection to overcome refractory barriers.

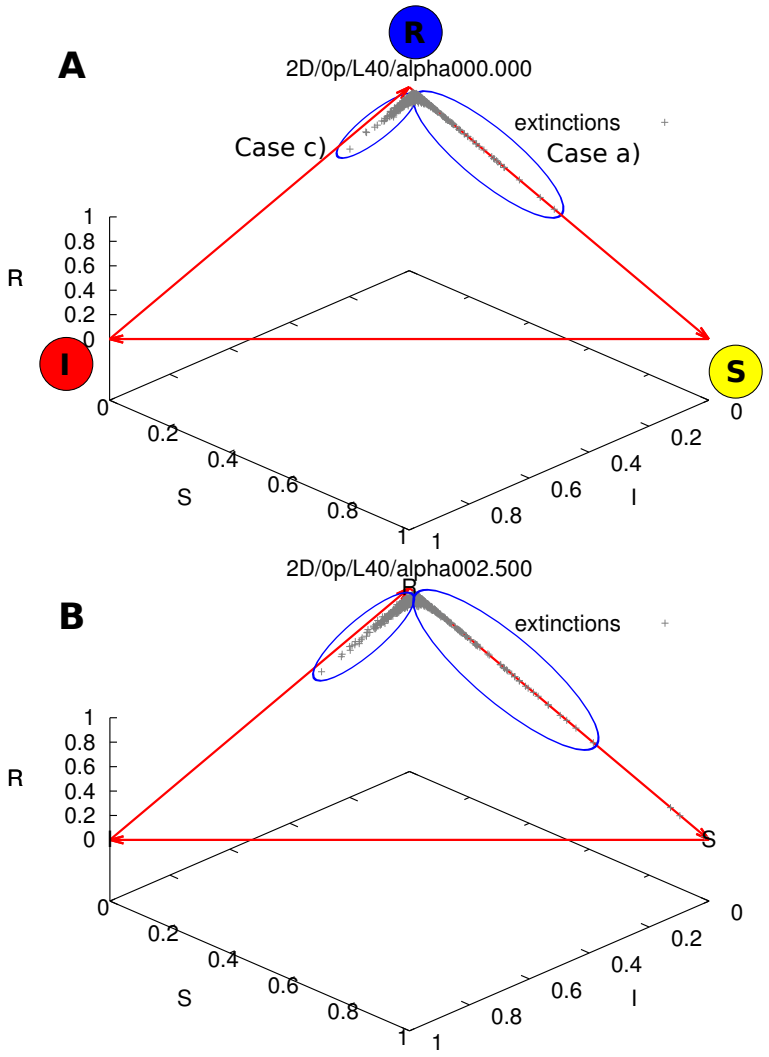

Figure 8.5: Extinction states in phase space for 1000 networks in dimension $d = 2$ with $L = 40$ and $p_{ann} = 0$. Two values of α/d are shown; $\alpha/d = 0$ (A) and $\alpha/d = 1.25$ (B). For non-zero annealing Case c) extinctions are absent, and Case a) extinctions ore more uniformly distributed over the $\mathbf{S} \leftrightarrow \mathbf{R}$ border.

We can summarize these results, saying that, in the context of extinctions

- at high values of infectivity, only extinctions with high coherence (Case a)) and extinctions by chance (Case c)) are possible,

- the border line $\mathbf{S} \leftrightarrow \mathbf{I}$ is repulsive,
- the border line $\mathbf{R} \leftrightarrow \mathbf{S}$ is absorbing, and
- the border line $\mathbf{I} \leftrightarrow \mathbf{R}$ is forbidden,
- as the system size increases, extinctions concentrates more closely to the vertex $\mathbf{R}$.

In accordance with the description of the three possible extinction cases we can differentiate them in the phase space (see figure 8.4). However, we don't know if extinctions follow some pattern(s), regardless of the case. The following sections will address this topic.

8.3.2 Phase dispersion distribution at extinction time

When infectious activity dies, $\delta\theta = \theta - \Theta$ approach a normal distribution $P(\delta\theta) = N(\Theta, \sigma^2)$. For the sake of simplicity (and without loss of generality) we can assume zero mean, and arbitrary standard deviation σ

$$P(\delta\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\delta\theta^2}{2\sigma^2}\right). \tag{8.10}$$

From equation 8.7, the average value of M^2 is

$$\left\langle M^2 \right\rangle = N_a + 2\sum_{i>j}^{N_a} \int_{-\infty}^{\infty} \frac{\cos(\delta\theta_i)\cos(\delta\theta_j)}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\delta\theta^2}{2\sigma^2}\right) \mathrm{d}\delta\theta,$$

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{\cos(\delta\theta)}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\delta\theta^2}{2\sigma^2}\right) \mathrm{d}\theta &= \int_{-\infty}^{\infty} \frac{\cos\sigma u}{\sqrt{2\pi}} e^{-u^2/2} \mathrm{d}u \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \left(\frac{e^{\mathrm{i}\sigma u} + e^{-\mathrm{i}\sigma u}}{2}\right) \mathrm{d}u \\
&= e^{-\sigma^2/2},
\end{aligned}$$

$$\left\langle M^2 \right\rangle = N_a + 2\sum_{i>j}^{N_a} e^{-\sigma^2} = N_a + N_a(N_a - 1)e^{-\sigma^2},$$

$$z^2 = \frac{\left\langle M^2 \right\rangle}{N_a^2},$$

so at extinction we get

$$z^2 = \frac{1 + (N_a - 1)e^{-\sigma^2}}{N_a} \sim e^{-\sigma^2} \qquad N_a \gg 1. \tag{8.11}$$

From the stochastic nature of our model we don't expect neither perfect synchronization nor stability [118, 119].

In order to test the validity of equation 8.11 we need to measure σ for a zero mean $\delta\theta$ distribution, as follows: being $\delta\theta_i \in [-\pi, \pi]$ we measured the actual mean Θ to define a new unsigned phase fluctuation

$$\delta\hat{\theta}_i = \min\left(|\theta_i - \Theta|, \pi - |\theta_i - \Theta|\right), \tag{8.12}$$

then we measured the variance as

$$\sigma^2 = \frac{1}{N_a}\sum_{i=1}^{N_a}(\delta\hat{\theta}_i)^2. \tag{8.13}$$

z^2 where calculated from the module of the right hand side in equation 8.7.

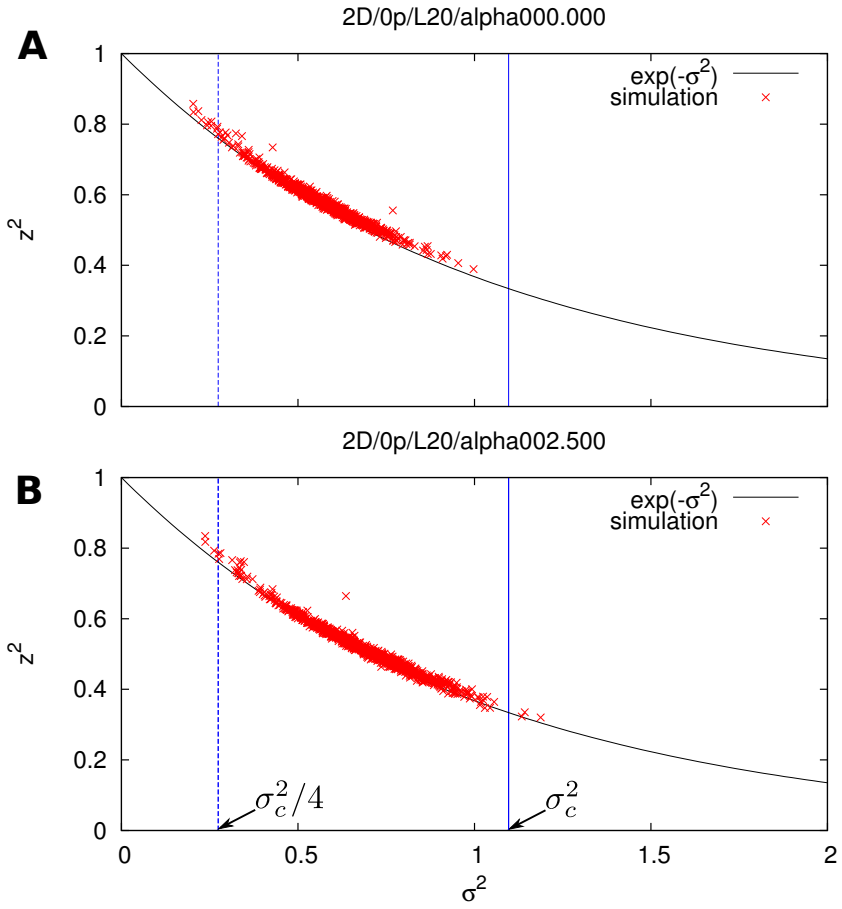

Figure 8.6: z^2 vs. σ^2 corresponding to 1000 extinctions for networks in dimension $d = 2$ with $L = 20$ an $p_{ann} = 0$. Two values of α/d are shown; $\alpha/d = 0$ (A) and $\alpha/d = 1.25$ (B). Vertical lines indicates the location of σ_c^2 (solid line) and $\sigma_c^2/4$ (dashed line). Most extinctions happens to the left of σ_c^2. Some outliers extinctions appears to the right of σ_c^2 only for small system sizes and large values of α.

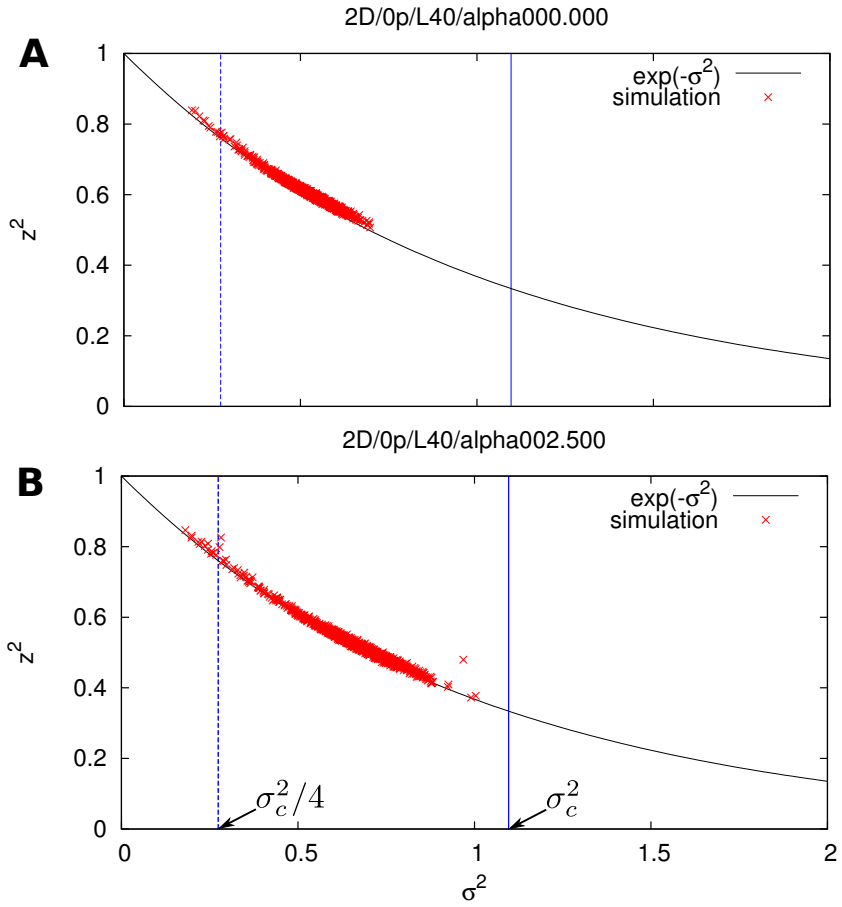

Figure 8.7: z^2 vs. σ^2 corresponding to 1000 extinctions for networks in dimension $d = 2$ with $L = 40$ and $p_{ann} = 0$. Two values of α/d are shown; $\alpha/d = 0$ (A) and $\alpha/d = 1.25$ (B). Vertical lines indicates the location of σ_c^2 (solid line) and $\sigma_c^2/4$ (dashed line). In this series of simulations all extinctions happens to the left of σ_c^2.

Figures 8.6 and 8.7 shows the measured behavior of $z^2 = f(\sigma^2)$, averaged over realizations of the dynamics, against the prediction of equation 8.11. The correspondence suggests that active nodes have decorrelated phase fluctuations at extinction, i.e. their phase fluctuations are closely described by a normal distribution over the active period circle. Extensive simulations yield similar results for all values of α, L and p_{ann} we used for networks in dimension $d = 1, 2$. These results validates the hypothesis of phase fluctuations decorrelation at extinction.

From figures 8.6 and 8.7 we specifically note that, in the plane σ^2 – z^2

- the centroid of the extinction distribution lies *on* the curve $z^2 = e^{-\sigma^2}$,

- the tails of the extinction distribution lie *above* the curve $z^2 = e^{-\sigma^2}$,

- all extinctions, *independently of their types*, are contained inside the region

[3],

$$\sigma^2 \leq \sigma_c^2 \wedge z^2 \geq e^{-\sigma^2}, \tag{8.14}$$

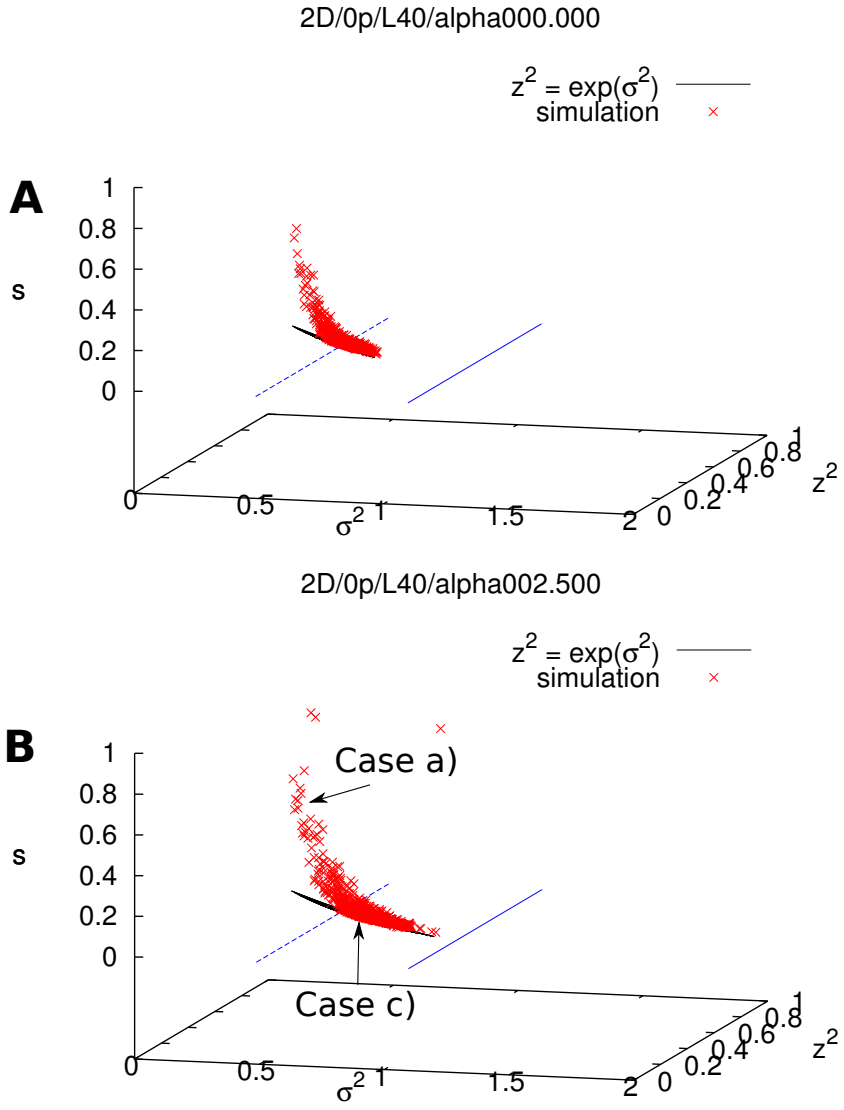

Figure 8.8: Susceptible fraction s vs. z^2 and σ^2 corresponding to 1000 extinctions for networks in dimension $d = 2$ with $L = 40$ and $p_{ann} = 0$. Two values of α/d are shown; $\alpha/d = 0$ (A) and $\alpha/d = 1.25$ (B). The lines at the bottom plane indicates the location of σ_c^2 (solid line) and $\sigma_c^2/4$ (dashed line).

Figure 8.8 plots the susceptible fraction $s = s(z^2, \sigma^2)$. This view allows to identify types of extinctions. Case a) extinctions lie *above and to the left* of the curve $z^2 = e^{-\sigma^2}$, with s varying in a wide range of values. Case c) extinctions

[3]Some outlier extinctions appears to the right of σ_c for large values of α and small system sizes

lie *on* the curve $z^2 = e^{-\sigma^2}$, with s varying in a narrow range of small values. Extinctions come closer to the curve $z^2 = e^{-\sigma^2}$ as the system size increases.

8.3.3 Dynamic behavior before extinction

8.3.3.1 Densities in phase space

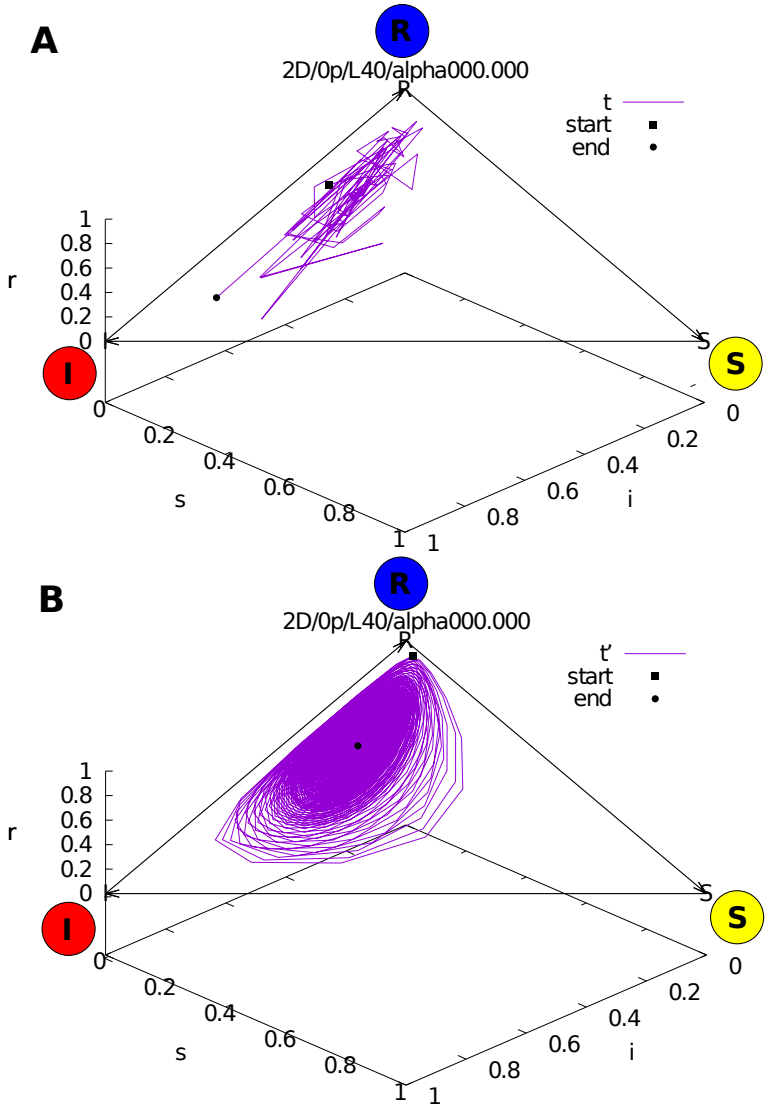

Figure 8.9: Densities (i, s, r) in phase space corresponding to 1000 networks in dimension $d = 2$ with $L = 40$, $p_{ann} = 0$ and $\alpha/d = 0$. Forward time averages (A) and time-to-extinction averages (B).

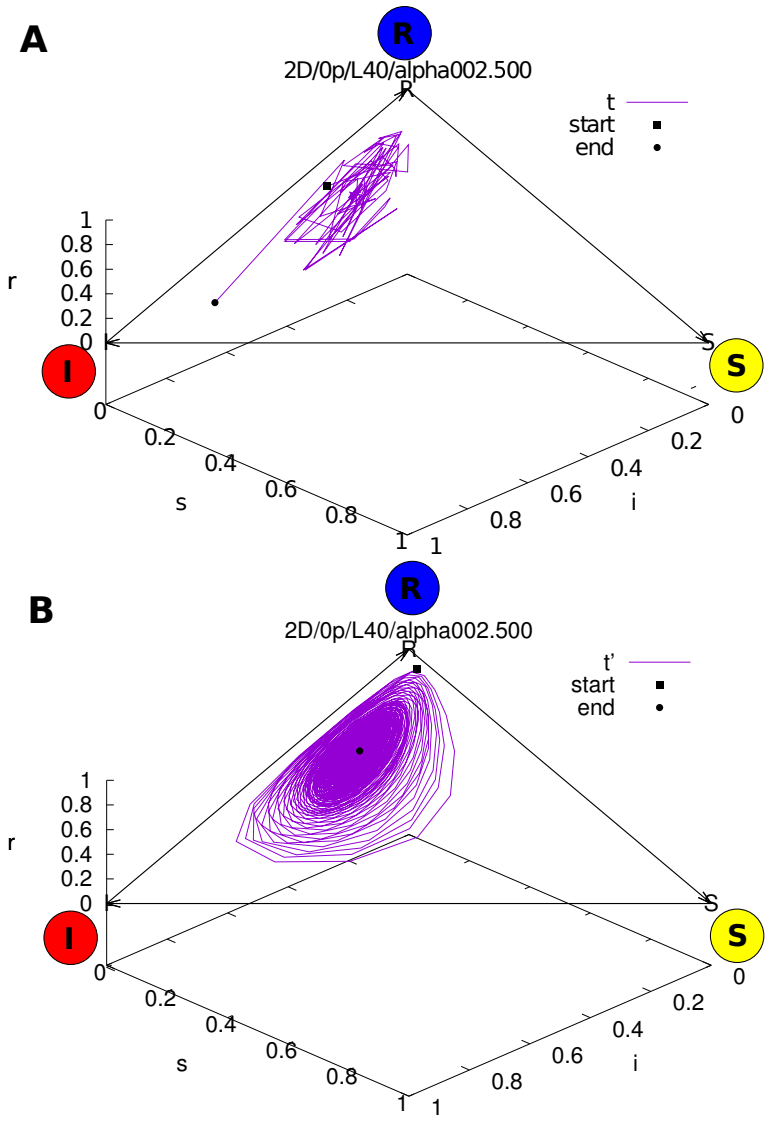

Figure 8.10: Densities (i, s, r) in phase space corresponding to 1000 networks in dimension $d = 2$ with $L = 40$, $p_{ann} = 0$ and $\alpha/d = 1.25$. Forward time averages (A) and time-to-extinction averages (B).

Figures 8.9 and 8.10 show the phase space of fractions (i, s, r) inter-simulation averaged in (forward) time (with geometric step, i.e. the new measurement time is approximately 1.1 times the previous measurement time) A) and in time-to-extinction (with unitary step) B).

In forward time, (i, s, r) starts at a random location (■), and ends close to the borders (•). The latter observation is dubious: forward-time measures becomes less representative of the simulations population as we get closer to the last forward measure (at the last measurement we may get only one alive system, but it is possible to get more than one). To avoid this problem we need to observe all simulations at the same time before they get extincted. This is exactly why we

take time-to-extinction measures.

In time-to-extinction, the point (i, s, r) is characterized by oscillations of continuously decreasing amplitude from extinction (■). The oscillation path is more densely cluttered close to the border $\mathbf{I} \leftrightarrow \mathbf{R}$ than to the border $\mathbf{S} \leftrightarrow \mathbf{R}$. The average extinction (■) does not happen on any border, but close to the vertex $\mathbf{R}$. At large time-to-extinction oscillations disappear and we just observe a static point with fluctuations (•).

8.3.3.2 Coherence and phase variance

Figures 8.11 and 8.12 present the inter-simulation averages of phase variance σ^2, squared coherence z^2, and persistence F for 1000 networks in dimension $d = 2$ with $\alpha/d = 0, 1.25$.

We observed three stages in the synchronization dynamics for large system sizes over realizations (see figure 8.12):

Stage 1 *Synchronization*: at the start of a simulation ($t = \tau_0$) z^2 has a low value, because phases are uniformly distributed over the active period circle (see figure 8.3). Coherence then increases during a transient period (of approximately 10^3 time steps for the parameters we used), due to a continuous aggregation of oscillators [87] in the active period circle.

Stage 2 *Sustained coherence*: coherence reaches a sustained average value with small fluctuations. This value is maintained until shortly before extinction time.

Stage 3 *Hyper-synchronization*: **average coherence increases many cycles before extinction time. This second transient, similar in duration to the first transient, was always present independently of the parameters used in the simulations. Stage 3 is only visible for measurements averaged vs. time-to-extinction. Therefore all extinctions are by hyper-synchronization.**

For small system sizes the behavior differs:

- for $\alpha/d < 1$ Stage 2 is absent, while
- for $\alpha/d > 1$ Stage 2 is short in comparison to Stage 2 for large systems.

As general patterns from figures 8.11 and 8.12 we observe that

i Persistence starts to decrease at the beginning of Stage 2. This means that extinctions are more likely (persistence is lower) as soon as coherence stabilizes.

ii Systems with lower levels of sustained coherence have longer activity. At fixed N, those systems correspond to networks with high α/d values.

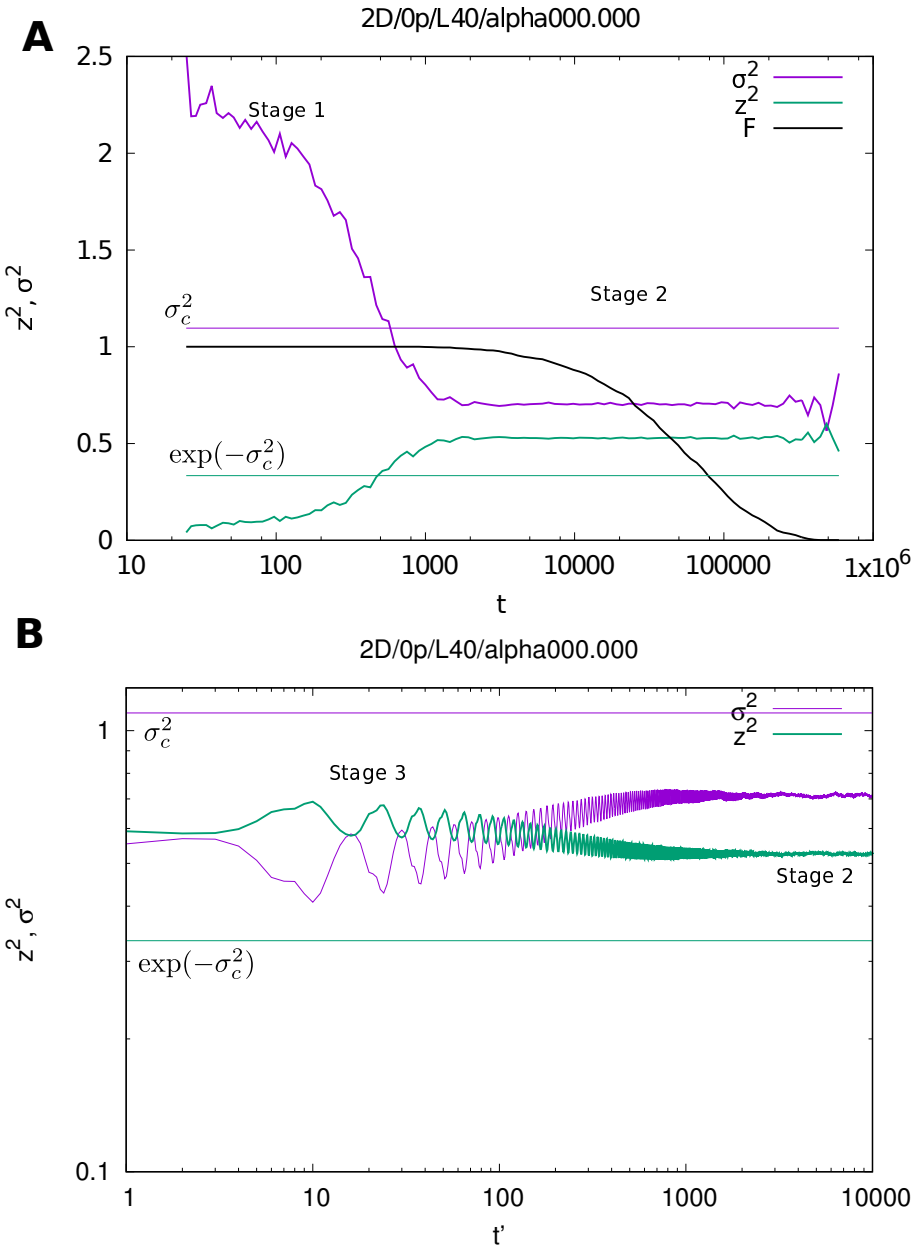

Figure 8.11: Time behavior of z^2, σ^2 and persistence for networks in dimension $d = 2$, with $L = 40$, $\alpha = 0.00$, and $p_{ann} = 0$. In forward time (A) and in time-to-extinction (B). In (B) the y axis is logarithmic to better appreciate variations with respect to t'. Horizontal doted lines indicate the location of σ_c^2 (upper) and $\exp(-\sigma_c^2)$ (lower) on each plot.

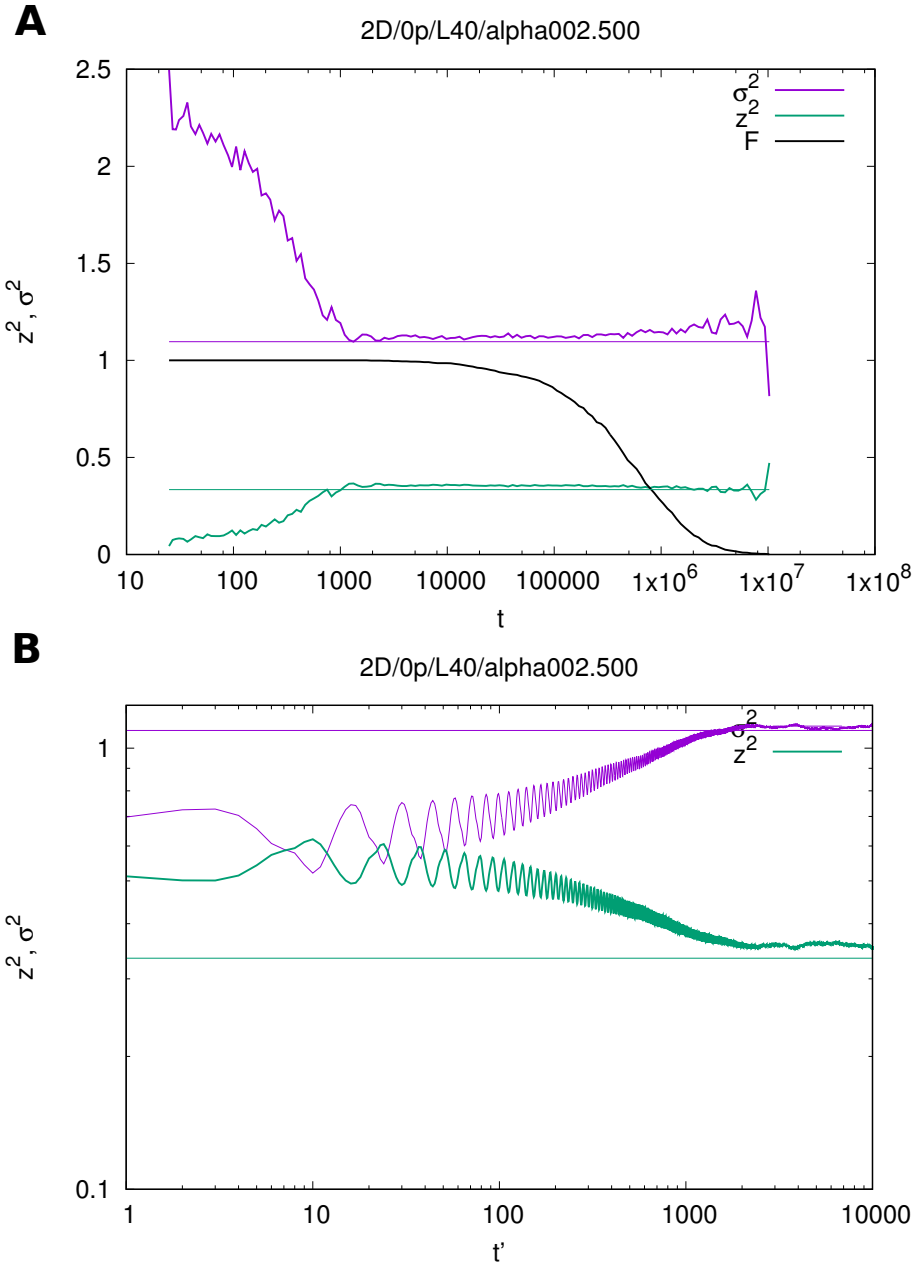

Figure 8.12: Time behavior of z^2, σ^2 and persistence for networks in dimension $d = 2$, with $L = 40$, $\alpha = 2.50$, and $p_{ann} = 0$. In forward time (A) and in time-to-extinction (B). In (B) the y axis is logarithmic to better appreciate variations with respect to t'. Horizontal doted lines indicate the location of σ_c^2 (upper) and $\exp(-\sigma_c^2)$ (lower) on each plot.

On large systems, large values of α display the same behavior shown in figure 8.12, with larger time scales, up to 10^8.

For large system sizes, in addition to z^2 and σ^2, three quantities remain constant (up to small fluctuations) in the first half of Stage 2: the active fraction $i + r$, the susceptible fraction s, and the infectible fraction x. These averages and

their fluctuations vary with α. See figure 8.13.

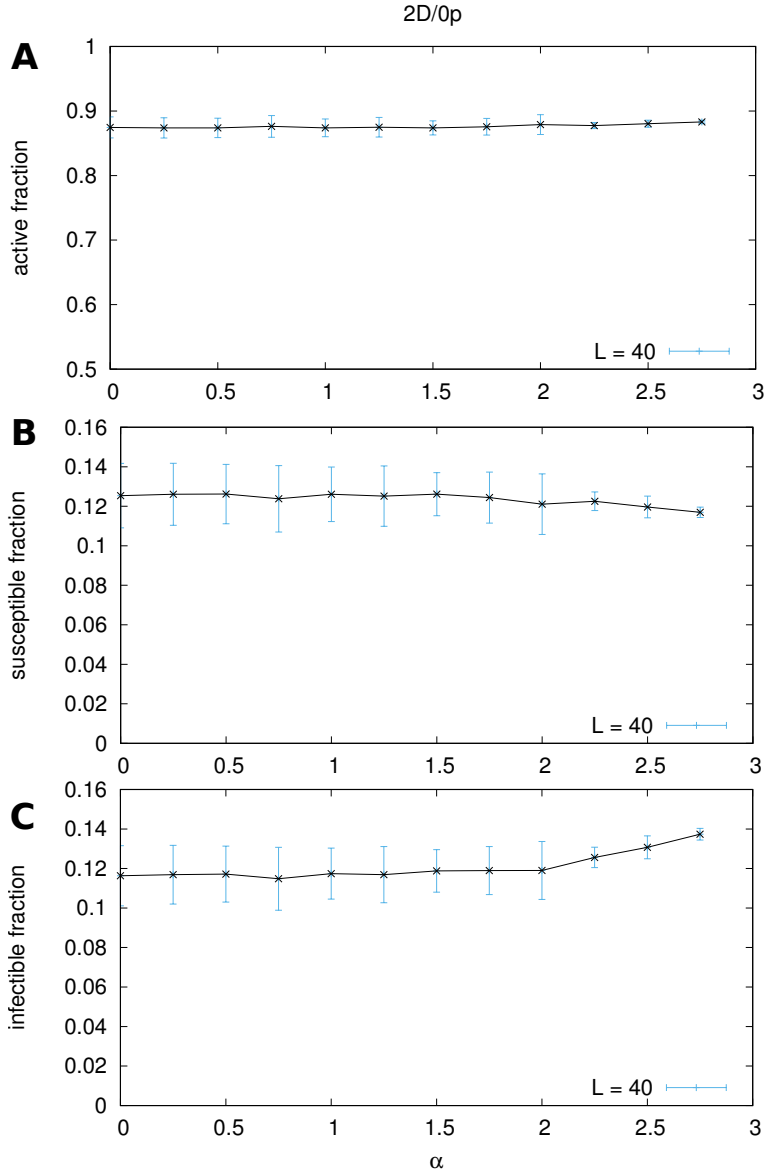

Figure 8.13: Time averages of active fraction (A), susceptible fraction (B) and infectible fraction (C) vs. α for networks in dimension $d = 2$, with $L = 40$ and $p_{ann} = 0$. These quantities are almost constant for $\alpha \leq 2$, while change slowly for $\alpha > 2$.

8.3.4 Characteristics of Stage 2

For large systems, Stage 2 spans almost all the dynamics lifetime. This stage corresponds to the "stable" part of the dynamics, contained in between two transients, Stage 1 and Stage 3.

We take the time averages of constant quantities[4] in the persistence range $1 > F > 0.5$ for $L = 20, 40$ and $\alpha = 0.00, 2.00, 2.25, 2.50, 2.75$. This persistence range corresponds to the first half of Stage 2, when the probability to survive (F) is larger than the probability of extinction ($1 - F$). This range was arbitrarily selected so as to involve Stage 2 mostly.

8.3.4.1 Coherence and phase variance

Figure 8.14 presents the time averages of z^2 vs. σ^2. It shows that these averages came close to the line where most extinctions lied, $z^2 = e^{-\sigma^2}$. At all levels of σ^2, z^2 is larger than $e^{-\sigma^2}$. In fact, the quantity $z^2 - e^{-\sigma^2}$ presents small amplitude-varying oscillations around a small mean value.

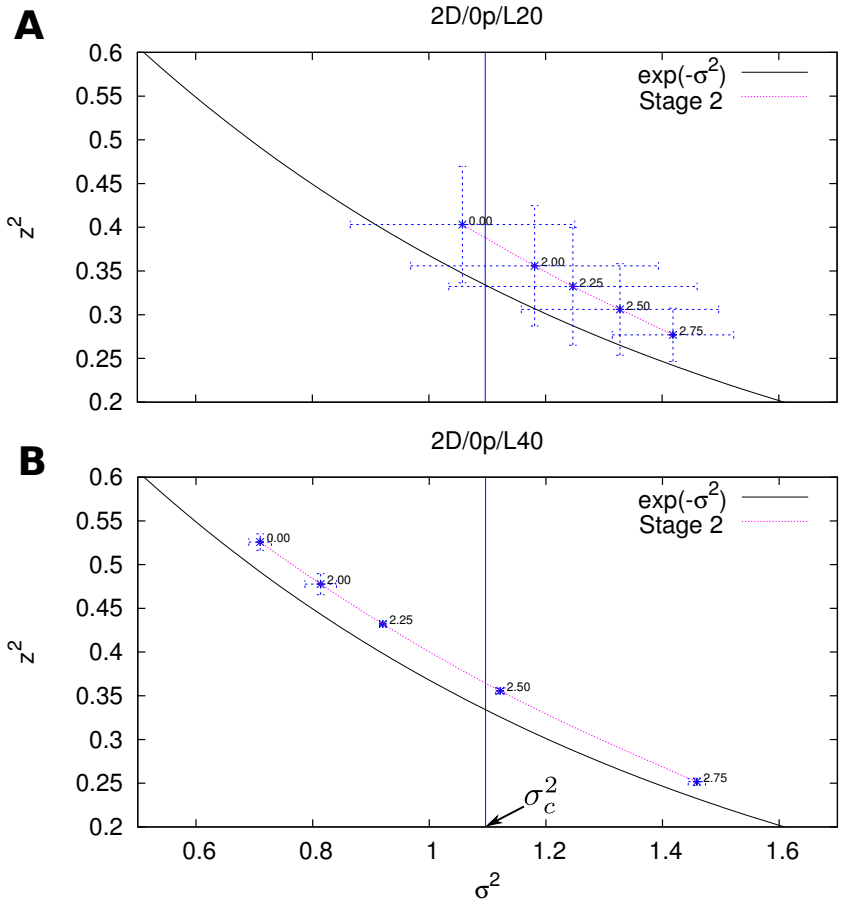

Figure 8.14: z^2 vs. σ^2 corresponding to self-sustained oscillations for networks in dimension $d = 2$ with $p_{ann} = 0$. Values of α are shown for every data point; $L = 20$ (A) and $L = 40$ (B). Vertical lines indicate the location of σ_c^2 (solid line) and $\sigma_c^2/4$ (dashed line).

In general we note that:

[4] Up to small fluctuations

- larger values of α are associated with lower sustained coherence,
- at larger values of α, the time average of σ^2 presents smaller fluctuations, and
- larger system sizes allow for stable coherence in a wider range of σ^2.

A visual analysis of phase-fluctuations distributions over the active period shows the patterns depicted in figure 8.15: an alternance between unimodal and bimodal distributions of phase-fluctuations.

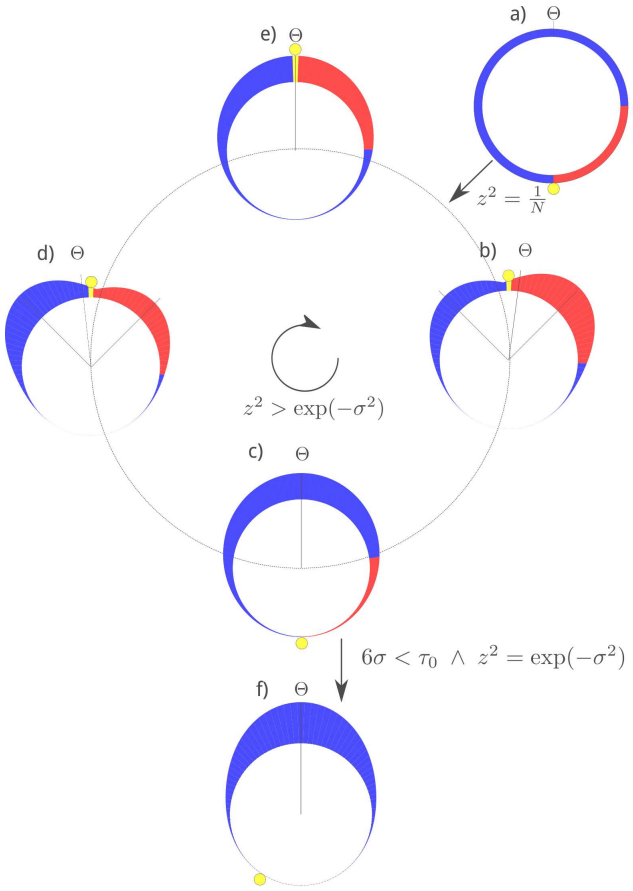

Figure 8.15: Density distribution of nodes over the active period circle. States are color coded; S = yellow, I = red, R = blue. The yellow circle and band indicates the extremes of the phase range that exists in an equivalent linear distribution.

8.3.5 Modeling the distribution of extinction times

In figure 8.16 we present the average infectible fraction. The temporal behavior of this average allows to distinguish the same three stages discussed in the previous section, with another interpretation: In Stage 1, the average infectible fraction

shows large fluctuations. In Stage 2, it is constant up to small fluctuations. In Stage 3 (in time-to-extinction), the oscillation minima decay exponentially.

In order to survive, the infection needs to travel from one infected node to a susceptible neighbor. The infection moves through the network, with shape and size deformations. Infectible fraction is a measure of the dynamic's surface, i.e. it accounts for those nodes that can become active (infectious) at the next simulation step. In this scenario we can rethink the dynamic stages as

Stage 1: Surface size stabilization.

Stage 2: Surface size maintenance.

Stage 3: Surface size destabilization and extinction.

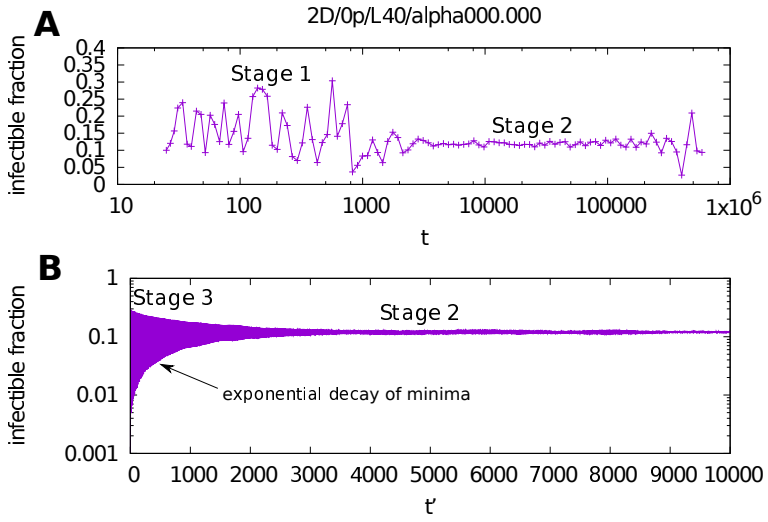

Figure 8.16: Time evolution of the infectible fraction for 1000 networks in dimension $d = 2$ with $L = 40$, $p_{ann} = 0$, and $\alpha = 0$ in forward time t (A), and in time-to-extinction t' (B). The three stages discussed in the previous section are distinguishable here. It worth noting that, in Stage 3, the mean oscillation minima departs from zero exponentially. Increasing the infectible fraction is equivalent to increasing the dynamics surface. Larger surfaces provides more "reactive" nodes to the dynamics. Disease spread is enhanced with large infection fluctuations. This condition leads to faster extinctions.

In figure 8.17 we observe the behavior of infectible fraction minima x_{min}. Data were taken from the same time-to-extinction data on figure 8.16. From the analysis of x_{min} we observe that these values decay exponentially with t' as

$$x_{min} = x_{min,\infty}\left(1 - e^{-at'}\right). \tag{8.15}$$

Figure 8.17 shows this model applied to two sequences of infectible fraction minima. The correspondence is notable, and consistent in L and α.

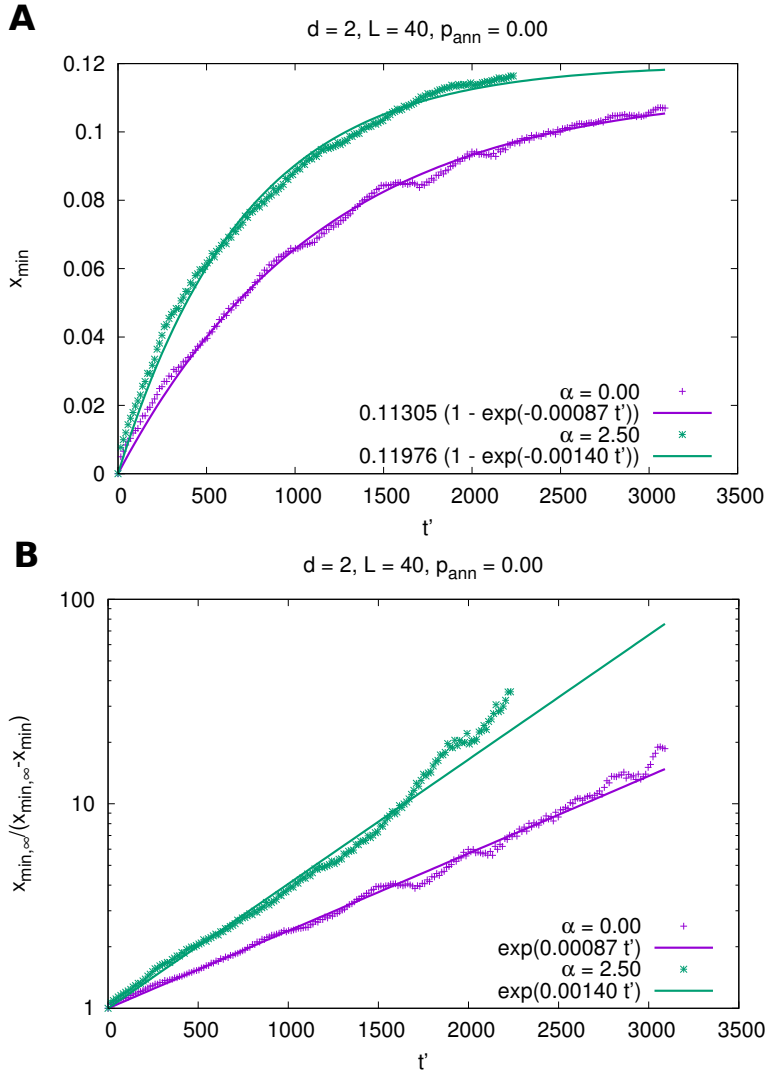

Figure 8.17: (A) Temporal behavior of the minima x_{min} of infectible fraction for 1000 networks in dimension $d = 2$ with $L = 40$ and $p_{ann} = 0$. Two values of α are shown: $\alpha = 0$ and $\alpha = 2.5$. Infectible fraction minima behave as $x_{min} = x_{min,\infty}\left(1 - e^{-at'}\right)$. In (B) we plot $x_{min,\infty}/(x_{min,\infty} - x_{min})$ to make clearer the exponential behavior of minima. Parameters $x_{min,\infty}$ and a vary with L and α.

It is worth noting that the infectible fractions we observe in figures 8.16 and 8.17 are inter-simulation averages. In forward time they are measured at the same simulation time, in time-to-extinction they are measured at the same time before extinction. The infectible fraction minima recorded in a single simulation is permanently varying in time, not as smooth as shown in figure 8.17. These minima will follow a noisy path toward zero.

If we think about this noisy path as a random walk, persistence $F(t)$ is obtained by mapping the extinction dynamics to a First-Passage-Time problem as follows: the system is considered as a random walker with initial position $x = x_0 =$

$x_{min,\infty} > 0$. It performs a random walk with non-zero drift $\mu < 0$ toward lower positions until it is absorbed at the origin $x = 0$ (i.e. extinction).

Let $W(x,t)$ be the probability of the random walker to be in position x at time t, $W(x,t)$ is solution of the Fokker-Plank equation

$$\frac{\partial W(x,t)}{\partial t} = -\mu \frac{\partial W(x,t)}{\partial x} + \frac{\sigma_0^2}{2} \frac{\partial^2 W(x,t)}{\partial x^2}. \tag{8.16}$$

As the random walker is absorbed at $x = 0$, $W(x,t)$ is subject to the boundary condition

$$W(0,t) = 0 \qquad \forall\, t. \tag{8.17}$$

Solving (8.16) and (8.17) by the method of images yields [120, 121]

$$W(x,t) = \frac{1}{\sqrt{2\sigma_0^2 t}} \left[\exp\left(-\frac{(x+\mu t + x_0)^2}{2\sigma_0^2 t} \right) - \exp\left(-\frac{(x+\mu t - x_0)^2}{2\sigma_0^2 t} \right) \exp\left(-\frac{2\mu x_0}{\sigma_0^2} \right) \right], \tag{8.18}$$

from which $F(t)$ is derived as

$$F(t) = \int_0^\infty W(x,t)dx, \tag{8.19}$$

$$F(t) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{t+t_0}{\sqrt{2K^2 t}} \right) \right] - \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{t-t_0}{\sqrt{2K^2 t}} \right) \right] \exp\left(-\frac{2t_0}{K^2} \right), \tag{8.20}$$

where

$$t_0 = \frac{x_0}{\mu}, \qquad K = \frac{\sigma_0}{\mu},$$

are related to the average and variance of the extinction probability $P(t)$ given by

$$P(t) = -\frac{\partial}{\partial t} F(t) = \frac{\sigma_0^2}{2} \frac{\partial}{\partial x} W(x,t) \Big|_{x=0}, \tag{8.21}$$

as

$$\langle t \rangle = t_0, \qquad \sigma_{t_{ext}}^2 = \langle t^2 \rangle - \langle t \rangle^2 = \frac{1}{2} t_0 K^2. \tag{8.22}$$

We now test this model by comparison with numerical results. Figure 8.18 presents the fits of equation (8.20) to persistence data corresponding to 1000 networks with $L = 20$ and $\alpha/d = 0.00, 1.00, 1.25$, for both static networks and annealed networks with $p_{ann} = 0.05$, in dimensions $d = 2$. The correspondence of equation (8.20) to simulation data is clear in the cases shown. However, from extensive simulations, we observed that equation (8.20) stops being a good approximation for persistence data $\alpha/d > 1$. Such behavior is accentuated in large systems.

The limited validity of equation (8.20) arises from the fact that survival probability becomes broader due to the scaling of mean lifetime with α and N increases (see next section). Then, our 1000 simulations per parameter-set results insufficient to construct a numerical approximation of $F(t)$. For small system sizes and

small values of α Stage 2 is comparable in length to Stage 1 and Stage 3 (these are approximately equal in length). In other cases this is not true; Stage 2 becomes by far the longest stage. The scaling length of Stage 2 alone is not analyzed in depth in this thesis.

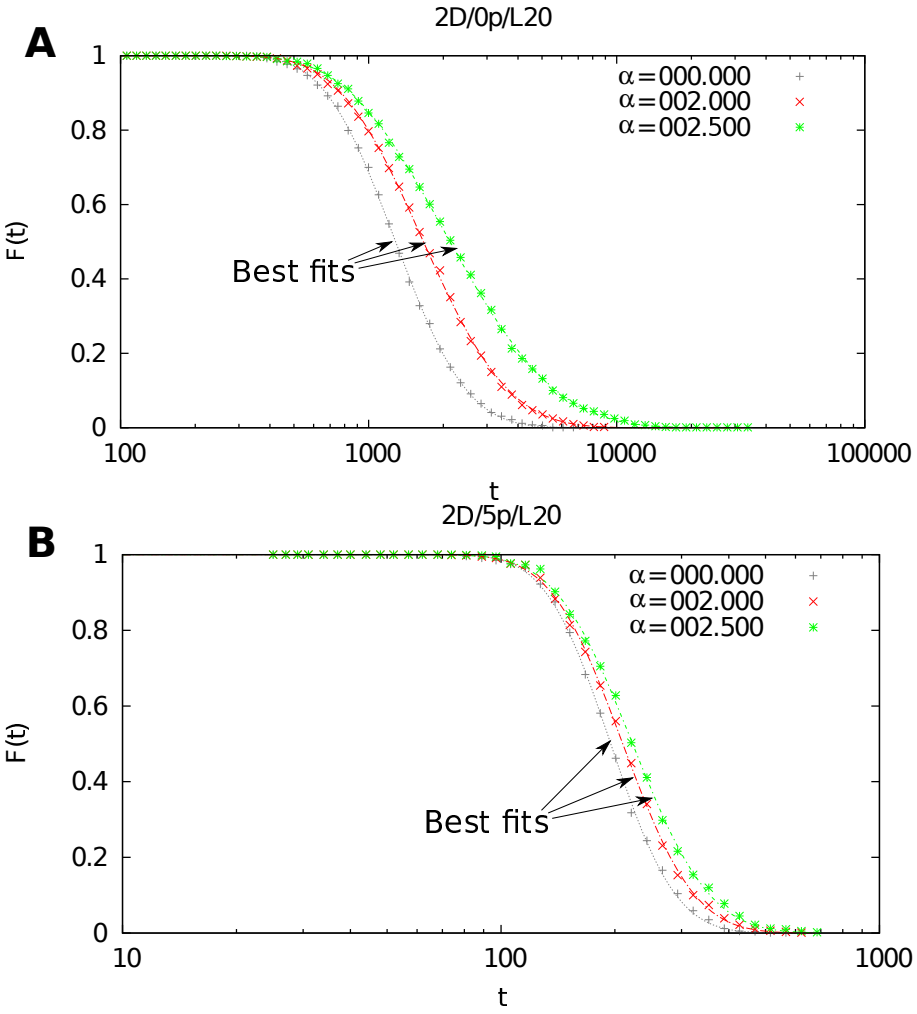

Figure 8.18: Persistence fits for data of networks with $L = 20$ in dimension $d = 2$. Three values of α/d are shown in each sub-figure; $\alpha/d = 0.00, 1.00, 1.25$. Static case with $p_{ann} = 0.00$ (A), and annealed networks with $p_{ann} = 0.05$ (B). **Symbols represent simulation data, while dotted lines of the same color are their respective best fits of equation** (8.20). The correspondence is striking. Data fits in $d = 1$ yield similar results.

8.3.6 Mean lifetime scaling with system size $N = L^d$

The mean lifetime $t_{1/2}$ is implicitly defined by

$$F(t_{1/2}) = \frac{1}{2}.$$

Mean lifetimes were interpolated using the best fits of equation 8.20 to survival data. Remember, all our measurements happen in a "geometric" time-sequence.

To forecast how the mean lifetime $t_{1/2}$ increases with the system size we need to identify the kind of stability of the deterministic SIRS model underlying our model [115, 104]. As we used a SIRS model with fixed refractory period, whose attractors are either a stable fixed point or a stable limit cycle [94, 69], we expect $\ln\left(t_{1/2}N^{1/2}\right)$ to scale exponentially with the system size [115].

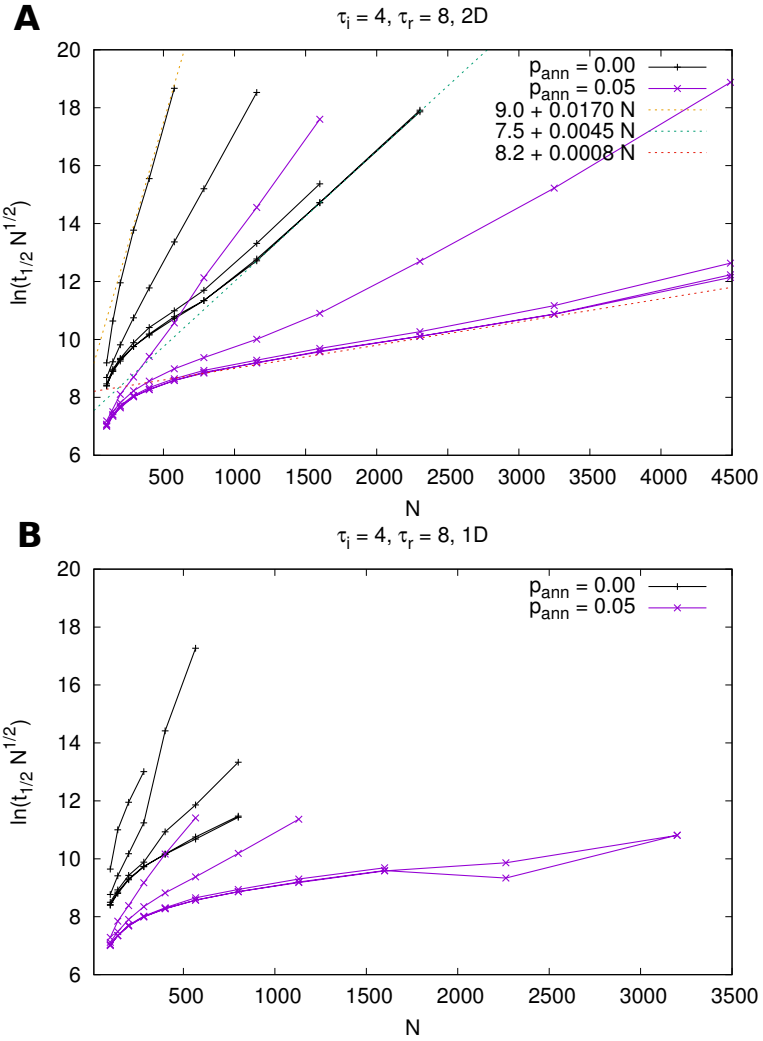

Figure 8.19: $\ln\left(t_{1/2}N^{1/2}\right)$ as a function of the system size N for networks in $d = 2$ (A) and $d = 1$ (B). Several values of α/d are shown in each sub-figure, from left to right; $\alpha/d = 2.0, 1.5, 1.0, 05, 0.0$. Different levels of annealing were considered. $\ln\left(t_{1/2}N^{1/2}\right)$ is asymptotically linear with N. Annealing reduces the asymptotic slope of $\ln\left(t_{1/2}N^{1/2}\right)$.

Figure 8.19 shows approximately this behavior for different values of α and L. The behavior of $\ln\left(t_{1/2}N^{1/2}\right)$ shows higher slopes for higher values of α/d. This results make clear that, as N or α/d increases we need more simulation runs

in order to get reliable persistence statistics, because $F(t)$ becomes broader. It also offers an explanation to why we were able to simulate just small systems: $\ln\left(t_{1/2}N^{1/2}\right)$ is asymptotically linear in N. Simulation series in α for systems of moderate size as $N = 40 \times 40$ required weeks of cpu time; larger systems required months without even completing the same range in α. Simulations in $d = 1$ yield similar results.

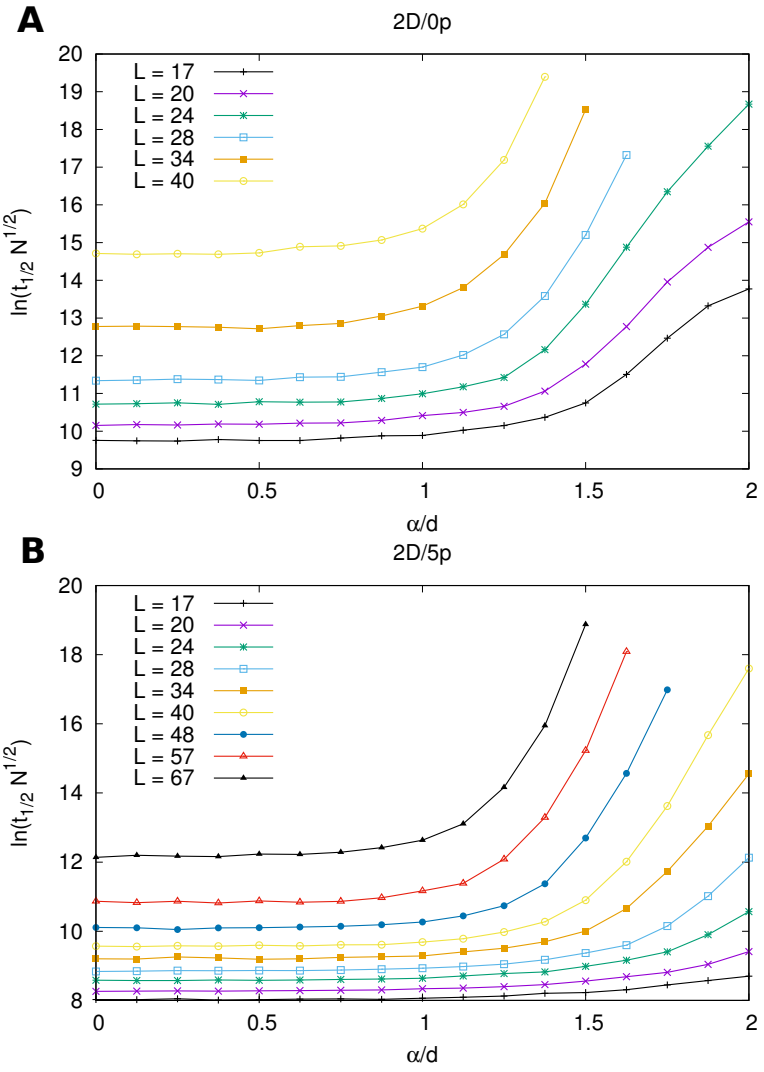

Figure 8.20: $\ln\left(t_{1/2}N^{1/2}\right)$ as a function of the system size α for networks in $d = 2$. Two levels of annealing were considered; $p_{ann} = 0.00$ (A), and $p_{ann} = 0.05$ (B). $\ln\left(t_{1/2}N^{1/2}\right)$ increases with α, and this is enhanced by the system size N. Annealing reduces growth of $\ln\left(t_{1/2}N^{1/2}\right)$. Simulations in $d = 1$ yield similar results.

As the system size increases, the range of α values we were able to simulated becomes narrower. Figure 8.20 allows to see why this is so: $\ln\left(t_{1/2}N^{1/2}\right)$ increases with α, and this growth is enhanced by the system size N. Annealing reduces

growth of $\ln\left(t_{1/2}N^{1/2}\right)$.

As previously stated, large mean life times motivated us to implement annealing within our model to reduce mean lifetime, see figure 8.19. Dynamics over annealed networks present the same qualitative behavior than static networks. Table 8.1 presents a resume of the parameters we used in our simulation. In all of them the infectible period is $\tau_i = 4$, the refractory period is $\tau_i = 8$, and the link infectivity is $p_0 = 0.75$.

	$p_{ann} = 0.00$		$p_{ann} = 0.01$		$p_{ann} = 0.05$	
			$d = 1, \Delta\alpha = 0.25$			
L	α	cpu time	α	cpu time	α	cpu time
100	0.00–3.00	0:44:50.69	0.00–3.00	0:33:41.75	0.00–3.00	0:31:34.18
141	0.00–3.00	0:48:49.30	0.00–3.00	0:35:18.89	0.00–3.00	0:32:12.76
200	0.00–3.00	0:56:15.82	0.00–3.00	0:36:43.52	0.00–3.00	0:33:25.19
283	0.00–3.00	1:45:59	0.00–3.00	0:40:01.99	0.00–3.00	0:35:36.79
400	0.00–3.00	8:40:28	0.00–3.00	0:45:48.87	0.00–3.00	0:40:20.41
566	0.00–3.00	162:01:0	0.00–3.00	1:33:10	0.00–3.00	1:11:34
800	0.00–1.50	–	0.00–3.00	9:54:39	0.00–3.00	7:16:04
1131	0.00–1.25	–	0.00–3.00	313:38:02	0.00–3.00	195:52:42
1600	0.00–1.25	–	0.00–1.50	–	0.00–1.75	–
2263	0.00–0.75	–	0.00–1.25	–	0.00–1.50	–
3200	–	–	0.00–1.25	–	0.00–1.25	–
4525	–	–	0.00–1.00	–	0.00–1.25	–
			$d = 2, \Delta\alpha = 0.25$			
L	α	cpu time	α	cpu time	α	cpu time
10	0.00–4.00	1:00:04	0.00–4.00	0:55:02.75	0.00–4.00	0:35:57.25
12	0.00–4.00	0:46:27.42	0.00–4.00	0:39:18.36	0.00–4.00	—
14	0.00–4.00	1:13:56	0.00–4.00	0:58:05.17	0.00–3.25	0:37:06.33
17	0.00–4.00	2:00:15	0.00–4.00	1:01:58	0.00–4.00	0:38:37.49
20	0.00–4.00	6:25:48	0.00–4.00	1:01:58	0.00–4.00	0:43:12.81
24	0.00–4.00	131:00:42	0.00–4.00	1:27:20	0.00–4.00	0:49:09.06
28	0.00–3.25	–	0.00–4.00	3:50:08	0.00–4.00	1:15:39
34	0.00–3.00	–	0.00–4.00	74:55:33	0.00–4.00	4:58:13
40	0.00–2.75	–	0.00–3.75	–	0.00–4.00	95:22:57
48	0.00–1.25	–	0.00–3.25	–	0.00–3.50	–
57	–	–	0.00–2.75	–	0.00–3.25	–
67	–	–	0.00–2.50	–	0.00–3.00	–

Table 8.1: Completed simulations in $d = 1$ (top) and $d = 2$ (bottom), with cpu time usage per series.

8.3.7 Mean lifetime collapse

From the figures in the previous section (i.e. 8.19 and 8.20) it is clear that $\ln\left(t_{1/2}N^{1/2}\right)$ depends on N, α and p_{ann}. A general scaling relation is desirable. Such relation is the aim of this subsection.

When α is varied, with L fixed, the average link-length varies as

$$\langle l\rangle = \frac{d-\alpha}{d+1-\alpha}\left(\frac{\left(\lfloor\frac{L}{2}\rfloor+1\right)^{d+1-\alpha}-1}{\left(\lfloor\frac{L}{2}\rfloor+1\right)^{d-\alpha}-1}\right). \tag{8.23}$$

The result is that we have changed the **effective geodesic size of the system** i.e., the ratio of the system length to the average link length

$$\Lambda = \frac{L}{\langle l\rangle} = \frac{1}{\lambda}. \tag{8.24}$$

Asymptotically in L we get

$$\Lambda \propto \begin{cases} L^0 & \text{if } \alpha \leq d, \\ L^{\alpha-d} & \text{if } d < \alpha \leq d+1, \\ L^1 & \text{if } \alpha > d+1. \end{cases} \tag{8.25}$$

Λ is the infection-chain length the disease needs to traverse the linear dimension of the system. A system with small Λ requires a short correlation length to establish a system-wide synchronized cluster, making extinctions more likely. As a result, disease is active on a shorter time, when compared to a system with larger Λ.

For a fixed L, increasing α from zero results in an increase of $\ln\left(t_{1/2}N^{1/2}\right)$ (see figure 8.20). Let $t_{1/2} = t_{1/2}(L,\alpha)$, the mean lifetime's amplification factor is defined as

$$y \equiv \frac{t_{1/2}\left(L,\alpha\right)}{t_{1/2}\left(L,0\right)}. \tag{8.26}$$

Figure 8.21 show a data collapse of $\ln(y)$ vs. Λ: $\ln(y)$ is an increasing function of Λ. Such behavior is consistent in dimension $d = 2$ for different values of size lengths L, decay exponent α, and annealing probabilities p_{ann}.

The amplification of $t_{1/2}$ with Λ

$$t_{1/2}\left(L,\alpha\right) = t_{1/2}\left(L,0\right)\exp(A\Lambda^{\beta}) \qquad \beta \simeq 3.5, \tag{8.27}$$

is notable because $\Lambda = \Lambda(L,\alpha)$ is a purely geometrical quantity ($d = 2$), meanwhile $t_{1/2}$ is supposed to depend on dynamic parameters too [115]. Moreover, annealing appear to only affects the factor A, reducing it.

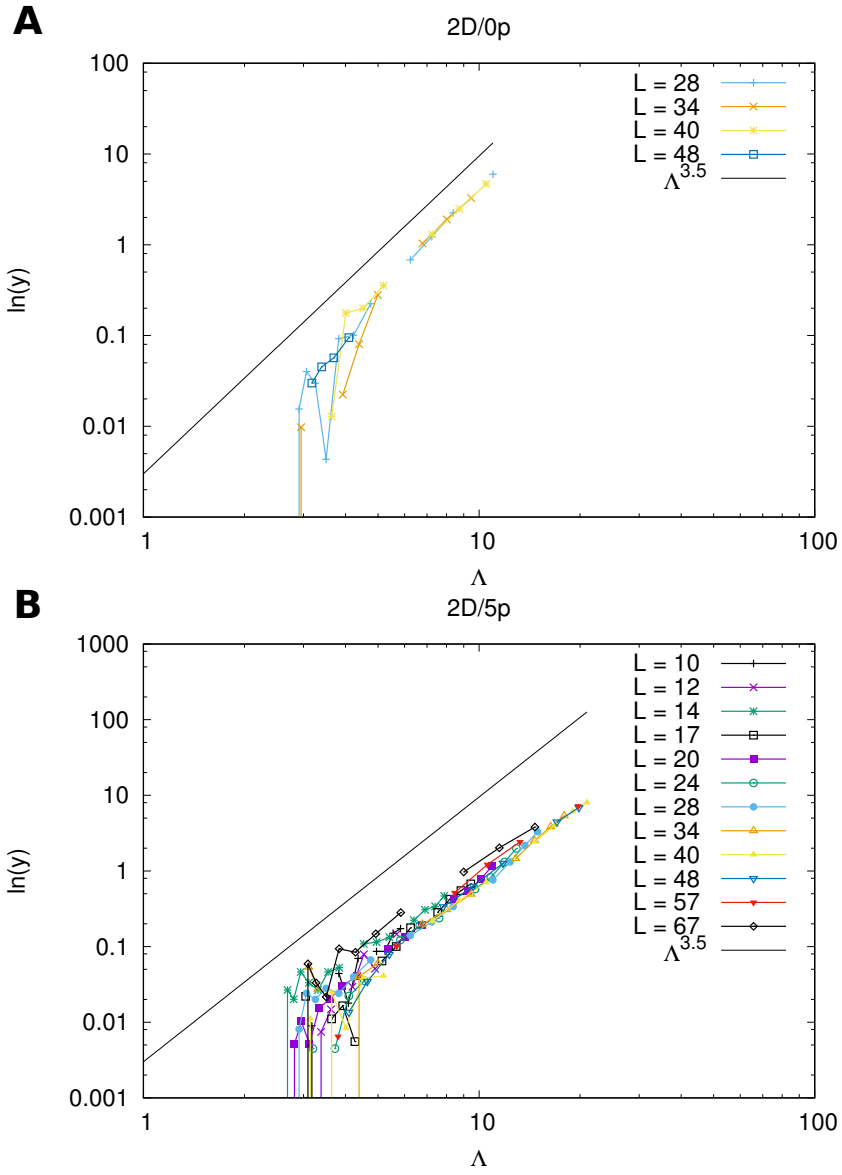

Figure 8.21: $\log(y)$ vs. Λ for networks with $p_{ann} = 0.00$ (A) and $p_{ann} = 0.05$ (B). Values of $\log(y)$ for different values of L collapse into a common behavior when plotted against Λ. The common behavior resembles a power-law $\log(y) = A\Lambda^{\beta}$ with exponent $\beta \simeq 3.5$. The power-law scaling is clearer for annealed dynamics data (B). Annealing appears to only affect the factor A, reducing it.

The scaling in equation 8.27 is important for two reasons: i) It incorporates the effect of the system length L, the decay exponent α and the annealing probability p_{ann} over the mean lifetime $t_{1/2}$ in a single functional relation. To our knowledge, this relation does not exist in the current literature yet. ii) It allows us to estimate mean lifetimes for networks with parameters (L, α) by measuring mean lifetimes in networks with $(L, 0)$. This is a desirable situation, because $\ln\left(t_{1/2}N^{1/2}\right)$ is asymptotically linear in N, and the slope of such behavior is higher for larger

values of α.

8.4 Conclusions and Discussion

We analyzed a SIRS model with fixed infectious and refractory period over networks with power-law link-length distribution

$$p_{ij} \propto r_{ij}^{-\alpha}.$$

When focus on extinctions of the disease dynamics, we found that

1. At extinction time, squared coherence z^2 and phase variance σ^2 satisfy the relation $z^2 = e^{-\sigma^2}$, which means that phase fluctuations are unimodal, and are represented by a Gaussian distribution, approximately.

2. **In time-to-extinction, the unimodal distribution of phase fluctuations is reached after a short period of hyper-synchronization preceding extinction, Stage 3. This behavior was consistent in all the parameters sets we used.** The time length of this transient presented little sensitivity to parameter changes.

3. The prior to hyper-synchronization behavior of Stage 2 is notable close to the curve $z^2 = e^{-\sigma^2}$ (see figures 8.9 and 8.10). In two dimensions, this behavior is enhanced by large system sizes, and stabilized for $\alpha/d > 1$ (see figure 8.13). Then, for the SIRS spatial model with long-range interactions, large and locally-connected contact networks [5] promote long lasting disease activity.

   The long lasting activity in Stage 2 may result from the coexistence of two synchronized populations whose phase lag is stable up to fluctuations. This stable phase lag prevents hyper-synchronization extinction by maintaining the system in a state of partial synchronization. Further studies are necessary in order to determine which mechanism drives the inter-population coalescence leading to extinction.

   In general, disease's lifetime is dominated by the time length of Stage 2, which is a random variable that needs to be analyzed separately.

4. Infectible fraction x represents the dynamic frontier of the disease. In time-to-extinction this frontier oscillates. The minima of this oscillation, x_{min}, vanish at extinction time. The approach to zero is exponential, $x_{min} = x_{min,\infty}(1 - e^{-at'})$.

   Conceiving the extinction of the disease dynamics as a First-Passage Time problem for a random walker (associated with infectible fractions minima) with starting position x_0 and drift velocity $\mu < 0$, the survival probability $F(t)$ for $\alpha/d \leq 1$ is closely

[5] not necessarily lattice networks

described by

$$\boxed{F(t) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{t+t_0}{\sqrt{2K^2 t}}\right)\right] - \frac{1}{2}\left[1 + \text{erf}\left(\frac{t-t_0}{\sqrt{2K^2 t}}\right)\right]\exp\left(-\frac{2t_0}{K^2}\right),}$$

where $t_0 = \frac{x_0}{\mu}, K = \frac{\sigma_0}{\mu}$, and σ_0 is the diffusion term in the Fokker-Planck equation.

This approach is highly accurate in the range $\alpha < d$. In such conditions the average link length is comparable to the system length L, i.e. this is the mean field domain. For $\alpha > d$ this approach is only an approximation. In very small systems this is a good approximation. Disagreement between data and predictions becomes more evident when the systems size increases.

5. The asymptotic scaling of $\ln\left(t_{1/2}N^{1/2}\right)$ is linear with the system size N for the parameters we used. This suggests that, the qualitative relation between the mean lifetime scaling and the stability of the deterministic SIR model with removal and renewal [115] also applies to the SIRS model.

   The role played by attractor stability of the underlying deterministic model requires a deeper analysis from the context of *nonlinear dynamics*, a subject that is beyond the scope of this work.

6. **We can estimate the mean lifetime of a system in dimension $d = 2$ with parameters (L, α) by measuring the less expensive case $(L, 0)$ as**

$$\boxed{t_{1/2}(L,\alpha) = t_{1/2}(L,0)\exp(A\Lambda^{\beta}) \qquad \beta \simeq 3.5,}$$

   where A is a factor that depends on the annealing.

   $\Lambda = \Lambda(L,\alpha)$ is a purely geometrical quantity ($d = 2$), meanwhile $t_{1/2}$ is supposed to depend on dynamic parameters [115]. Annealing appear to only affects the factor A, which decreases with the annealing probability.

   This scaling relation is important for two reason: i) It incorporates the effect of the system length L, the decay exponent α and the annealing probability p_{ann} over the mean lifetime $t_{1/2}$ in a single functional relation. To our knowledge, this relation does not exist in the current literature yet. ii) It allows us to estimate mean lifetimes for networks with parameters (L, α) by measuring mean lifetimes in networks with $(L, 0)$. This is a desirable situation, because $\ln\left(t_{1/2}N^{1/2}\right)$ is asymptotically linear in N, and the slope of such behavior is higher for larger values of α.

   For the SIRS spatial model with long-range interactions, mean lifetimes does not scales directly with the linear size of the system. It does scales (in a non-linear way) with the system length relative to the average link length.

This shows that, more than absolutes, what matters is how big a system is for the average interaction length.

Long-lasting activity allows to remark some similarities with other instances of excitable media. In the context of epidemiology, short lived disease activity is desired. However, in other circumstances, sustained activity of spatially extended excitable media is of fundamental importance. In ecology, considering the S, I and R states of the SIRS model as species in the three species Lotka-Volterra model, sustained activity is synonymous of ecological diversity, i.e. coexistence of species [104, 105, 106]. Extinction is something to fight against, with exceptions such as plague eradication. Reichenbach et al. [107] shown that moderate mobility promotes diversity; to much mobility leads to extinction. This is similar to our results that short-ranged "quenched" interactions on locally connected networks ($\alpha/d > 1, p_{ann} = 0$) promotes sustained activity, while long-ranged "annealed" interactions on globally connected networks ($\alpha/d < 1, p_{ann} > 0$) leads to extinctions.

Notably, the stochastic three species Lotka-Volterra model presents mean extinctions times that scale as e^N when the deterministic underlying dynamics has a stable attractor in the coexistence region [104, 105].

It is still required to explore how τ_i, τ_r and $\langle k \rangle$ affects the results we had observed. By the large amount of simulations already done, we expect only quantitative changes for the behaviors previously shown.

Appendices

Appendix A

Generating power-law distributed radius

$$1 = A\int_{l_{min}}^{l_{max}} p(l)dl = A\int_{l_{min}}^{l_{max}} l^{-\alpha+d-1}dl = A\left.\frac{1}{d-\alpha}l^{d-\alpha}\right|_{l_{min}}^{l_{max}} = A\frac{l_{max}^{d-\alpha}-l_{min}^{d-\alpha}}{d-\alpha},$$

$$A = \frac{d-\alpha}{l_{max}^{d-\alpha}-l_{min}^{d-\alpha}}.$$

Let $s \in [0,1]$ be a uniform random variable, then

$$1-s = \int_{l}^{l_{max}} p(l')dl' = A\left.\frac{1}{d-\alpha}l'^{d-\alpha}\right|_{l}^{l_{max}} = A\frac{l_{max}^{d-\alpha}-l^{d-\alpha}}{d-\alpha} = \frac{l_{max}^{d-\alpha}-l^{d-\alpha}}{l_{max}^{d-\alpha}-l_{min}^{d-\alpha}},$$

$$(1-s)\left(l_{max}^{d-\alpha}-l_{min}^{d-\alpha}\right) = l_{max}^{d-\alpha}-l^{d-\alpha},$$

$$s\left(l_{max}^{d-\alpha}-l_{min}^{d-\alpha}\right)+l_{min}^{d-\alpha} = l^{d-\alpha},$$

$$l = \left(s\left(l_{max}^{d-\alpha}-l_{min}^{d-\alpha}\right)+l_{min}^{d-\alpha}\right)^{\frac{1}{d-\alpha}},$$

Let us define $l_{max} = \lfloor\frac{L}{2}\rfloor + 1$, $l_{min} = 1$,

$$\text{konstant} = \begin{cases} \left(\lfloor\frac{L}{2}\rfloor+1\right)^{d-\alpha}-1 & \text{if } \alpha \neq d \\ \lfloor\frac{L}{2}\rfloor+1 & \text{if } \alpha = d, \end{cases} \tag{A.1}$$

$$\text{dexponent} = \begin{cases} \frac{1}{d-\alpha} & \text{if } \alpha \neq d \\ 0 & \text{if } \alpha = d, \end{cases} \tag{A.2}$$

then

$$l = \begin{cases} (s\cdot\text{konstant}+1)^{\text{dexponent}} & \text{if } \alpha \neq d \\ \text{konstant}^s & \text{if } \alpha = d. \end{cases} \tag{A.3}$$

Appendix B

Average link-length

The networks with power-law link-length distribution we build have average link length

$$\langle l \rangle = \frac{d-\alpha}{d+1-\alpha}\left(\frac{\left(\lfloor\frac{L}{2}\rfloor+1\right)^{d+1-\alpha}-1}{\left(\lfloor\frac{L}{2}\rfloor+1\right)^{d-\alpha}-1}\right). \tag{B.1}$$

We can define the quantity

$$\lambda = \frac{\langle l \rangle}{L} \tag{B.2}$$

as a global measure of the average single link coverage on a medium of length L.

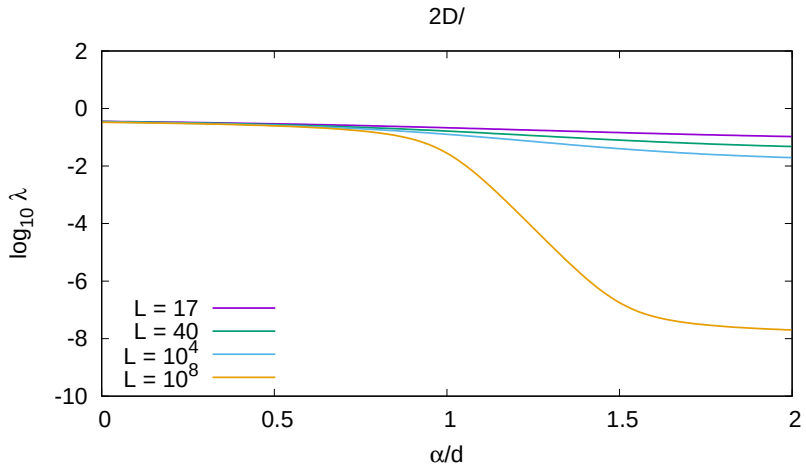

Figure B.1: λ vs. α/d for different values of L in $d = 2$. As stated by equation B.3, for $\alpha \leq d$ we observe that λ varies little. In the range $d < \alpha \leq d+1$ it is clear that λ decays with α. For $\alpha > d$ we can observe a "saturation" of λ, i.e it becomes independent of α.

For finite systems λ decreases continuously from a maximum bounded by $\frac{d}{2(d+1)}$ at $\alpha/d = 0$. The limit case $\alpha = 0$ corresponds to a network in which nodes are connected to each other irrespective of their distance of separation, i.e. the network is a small-world network. In the limit of large L

$$\lambda \propto \begin{cases} \frac{d-\alpha}{d+1-\alpha} L^0 & \text{if } \alpha \leq d, \\ \frac{\alpha-d}{d+1-\alpha} L^{d-\alpha} & \text{if } d < \alpha \leq d+1, \\ \frac{\alpha-d}{\alpha-(d+1)} L^{-1} & \text{if } \alpha > d+1. \end{cases} \tag{B.3}$$

In this limit, λ is a finite fraction of L if $\alpha \leq d+1$. For $\alpha > d+1$, λ decay with the system size. See figure B.1.

Appendix C

Directories and datafiles

d = dimension
a = annealing probability (percent) DONE=(0,1,5)
L = Linear Length of system DONE=(2D: 12,14,17,20,24,28,34,40,48,57,67,80.
1D: 100,141,200,283,400,566,800,1131,1600,2263,3200,4525)

./dD/ap/LL/**state.vs.time*** contents:

Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Value	sim time	I	S	R	infectible fraction	z^2	-	-	$F(t)$	σ^2	Θ

./dD/ap/LL/**state.vs.time.reverse*** contents:

Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Value	reverse time	I	S	R	infectible fraction	z^2	-	-	$F(t)$	σ^2	Θ

./dD/ap/LL/**extinction.state*** contents:

Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Value	extinction time	I	S	R	infectible fraction	z^2	σ^2	-

./dD/ap/**persistence.parameters.dD.ap.LL.out** contents:

Column	**1**	**2**	**3**	**4**	**5**	**6**
Value	α	L	t_o	K	$t_{1/2}$	t_{mp}

./**constant.averages.dD.ap.LL.out** contents:

Column	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Value	α	$\langle s \rangle$	σ_s	$\langle i+r \rangle$	σ_{i+r}	$\langle x \rangle$	σ_x	$\langle z^2 \rangle$	σ_{z^2}	$\langle \sigma^2 \rangle$
Column	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
Value	σ_{σ^2}	$\langle \Theta \rangle$	σ_Θ	$t_{1/2}$	N_a	$\ln(y)_{meas}$	$\ln(y)_{pred}$	λ	σ_λ	θ_0

where

- $x \equiv$ infectible fraction
- $N_a = \frac{1-\exp(-\sigma^2)}{z^2-\exp(-\sigma^2)}$
- $z^2 = \frac{\cos^2\theta_0}{\cos^2\sqrt{\theta_0^2-\sigma^2}}$

Bibliography

[1] Romualdo Pastor-Satorras and Alessandro Vespignani. *EVOLUTION AND STRUCTURE OF THE INTERNET A Statistical Physics Approach*. Cambridge University Press, first edition, 2004.

[2] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Error and attack tolerance of complex networks. *Nature*, 406:378–382, 200.

[3] Alain Barrat, Marc Barthélémy, and Alessandro Vespignani. *Dynamical Processes On Complex Networks*. Cambridge University Press, 2008.

[4] H. Hong, M. Y. Choi, and B. J. Kim. Synchronization on small-world networks. *Physical Review E*, 65:026139(5), 2002.

[5] A. Arenas, Díaz-Guilera, and C. J. Pérez-Vicente. Synchronization reveals topological scales in complex networks. *Physical Review Letters*, 96(11):114102(4), 2006.

[6] M.E.J. Newman. *Networks an introduction*. Oxford University Press, 1st edition, 2010.

[7] Marc Barthélemy, Alain Barrat, Romualdo Pastor-Satorras, and Alessandro Vespignani. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of Theoretical Biology*, 235(2):275–288, 2005.

[8] Marcelo Kuperman and Guillermo Abhramson. Small world effects in an epidemiological model. *Physical Review Letters*, 86(13):2909(4), 2001.

[9] Daniel Kaplan and Leon Glass. *Understanding Nonlinear Dynamics*. Springer-Verlag New York, 1995.

[10] Sitabhra Sinha and S. Srighar. *Patterns in excitable media: Genesis, Dynamics and Control*. CRC Press Taylor & Francis Group, 2015.

[11] Alex Roxin, Hermann Riecke, and S. A. Solla. Self-sustained activity in a small-world network of excitable neurons. *Physical Review Letters*, 92(19):198101(4), 2004.

[12] W. Deng, W. Li, X. Cai, and Q. A. Wang. The exponential degree distribution in complex networks: Non-equilibrium network theory, numerical simulation and empirical data. *Physica A*, 390:1481–1485, 2011.

[13] Rodrigo Huerta-Quintanilla, Efrain Canto-Lugo, and Dolores Viga-de Alva. Modeling social network topologies in elementary schools. *Plos One*, 8(2):e55371, 2013.

[14] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[15] L. A. N. Amaral, A. Scala, M. Barthélémy, and Stanley H. E. Classes of small-world networks. *PNAS*, 97(21):11149–11152, 2000.

[16] Stefano Mossa, Mark Barthélémy, H. Eugene Stanley, and Luís A. Nunes Amaral. Truncation of power law behavior in "scale-free" network models due to information filtering. *Physical Review Letters*, 88(13):138701(4), 2002.

[17] Estrada Ernesto. *The Structure of Complex Networks - Theory and Applycations*. Oxford University Press, first edition, 2011.

[18] Fernando Vega-Redondo. *Complex Social Networks*. Cambridge University Press, 2007.

[19] S.N. Dorogovtsev and J.F.F. Mendes. *Evolution of Networks From Biological Nets to the Internet and WWW*. Oxford University Press, 2003.

[20] S.N. Dorogovtsev and J.F.F. Mendes. Evolution of networks. *Advances in Physics*, 51(4):1079–1187, 2002.

[21] M. E. J. Newman. Assortative mixing in networks. *Physical Review Letters*, 89:208701(4), 2002.

[22] Romualdo Pastor-Satorras, Alexei Vázquez, and Alessandro Vespignani. Dynamical and correlation properties of the internet. *Physical Review Letters*, 87(25):258701, 2001.

[23] Alexei Vázquez, Romualdo Pastor-Satorras, and Alessandro Vespignani. Large-scale topological and dynamical properties of the internet. *Physical Review E*, 65(6):066130(12), 2002.

[24] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Olvai, and A.-L. Barabási. Hierarchical organization of modularity in metabolic networks. *Science*, 297:1551–1555, 2002.

[25] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of *small-world* networks. *Nature*, 393:440–442, 1998.

[26] M.E.J. Newman, C. Moore, and D.J. Watts. Mean-field solution of the small-world network model. *Physical Review Letters*, 84(14):3201–3204, 2000.

[27] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69:026113(15), 2004.

[28] M. E. J. Newman. Detecting community structure in networks. *Eur. Phys. J. B*, 38:321–330, 2004.

[29] Leon Danon, Albert Díaz-Guilera, Jordi Duch, and Alex Arenas. Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(09):P09008(10), 2005.

[30] P. Erdös and A. Rényi. On random graphs. *Publ. Math. Debrecen*, 6:290–297, 1959.

[31] P. Erdös and A. Rényi. On the evolution of random graphs. *Mathematikai Katutó Intézet Közleményei*, 5:17–61, 1960.

[32] M. Molloy, B. J Reed, and S. Redner. A critical point for graphs with a given degree sequence. *Random Structures and Algorithms*, 6:161–180, 1995.

[33] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distribution and their applications. *Physical Review E*, 64(026118):17, 2001.

[34] Stefan Bornholdt and Heinz Georg Schuster. *Handbook of Graphs and Networks : From the Genome to the Internet*. WILEY-VCH, 2003.

[35] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47–97, 2002.

[36] A. Barrat and M. Weigt. On the properties of small-world network models. *The European Physical Journal B*, 13:547–560, 2000.

[37] Hawoong Jeong, S. P. Masont, A. L. Barabási, and Z. N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411:41, 2001.

[38] Kwang-Il Goh, Eulsik Oh, Hawoong Jeong, Byungnam Kahng, and Doochul Kim. Classification of scale-free networks. *PNAS*, 99(20):12583–12588, 2002.

[39] M. E. J. Newmann. The structure of scientific collaboration networks. *PNAS*, 98(2):404 – 409, 2001.

[40] Konstantin Klemm and Víctor M. Eguíluz. Growing scale-free networks with small-world behavior. *Physical Review E*, 65(5):057102(4), 2002.

[41] Albert-László Barabási, Erzsébet Ravasz, and Tamás Vicsek. Deterministic scale-free networks. *Physica A: Statistical Mechanics and its Applications*, 299(3-5):559 – 564, 2001.

[42] Mark Buchanan, Guido Caldarelli, Paolo De Los Rios, Francesco Rao, and Michelle Vendruscolo. *Networks in Cell Biology*. Cambridge University Press, 2010.

[43] Albert-László Barabási and Réka Albert. Topology of evolving networks: local events and universality. *Physical Review Letters*, 85(24):5234–5237, 2000.

[44] Fragkiskos Papadopoulos, Maksim Kitsak, M. Angeles Serrano, Marian Boguna, and Dmitri Krioukov. Popularity versus similarity in growing networks. *Nature*, 489:537–540, 2012.

[45] M. E. J. Newman. The structure and functions of complex networks. *Society for Industrial and Applied Mathematics*, 45(2):167–256, 2003.

[46] A.L. Barabási, R. Albert, and H. Jeong. Mean-field theory for scale-free random networks. *Physica A*, 272:173–187, 1999.

[47] P. L. Krapivsky, G. J., S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Physical Review Letters*, 85(21):4629–4632, 2000.

[48] P. L. Krapivsky, G. J. Rodgers, and S. Redner. Degree distributions of growing networks. *Physical Review Letters*, 86(23):5401–5404, 2001.

[49] Santo Fortunato, Alessandro Flammini, and Filippo Menczer. Scale-free network growth by ranking. *Physical Review Letters*, 96:218701(4), 2006.

[50] S.N. Dorogovtsev, J.F.F. Mendes, and A. N. Samukhin. Structure of growing networks with preferential linking. *Physical Review Letters*, 85(21):4633–4636, 2000.

[51] Duncan S. Callaway, John E. Hopcroft, Jon M. Kleinberg, M. E. J. Newman, and Steven H. Strogatz. Are randomly grown graphs really random? *Physical Review E*, 64(4):041902(7), 2001.

[52] Aaron Clauset, Cosma Rohilla Shalizi, and M.E.J. Newman. Power-law distributions in empirical data. *SIAM*, 51(4):661–703, 2009.

[53] S.N. Dorogovtsev and J.F.F. Mendes. Scaling properties of scale-free evolving networks: Continuous approach. *Physical Review E.*, 63(5):056125(19), 2001.

[54] David M. Pennock, Gary W. Flake, Steve Lawrence, Eric J. Glover, and C. Lee Giles. Winners don't take all: Characterizing the competition for links on the web. *PNAS*, 99(8):5207–5211, 2002.

[55] Benjamin Shargel, Hiroki Sayama, Irving R. Epstein, and Yaneer Bar-Yam. Optimization of robustness and connectivity in complex networks. *Physical Review Letters*, 90(6):068701(4), 2003.

[56] Mattew O. Jackson. *Social and economic networks*. Princeton University Press, 2008.

[57] R. Albert, I. Albert, and G.L. Nakarado. Structural vulnerability of the north american power grid. *Phys. Rev. E*, 69(2):025103(4), 2004.

[58] R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas. Self-similar community structure in a network of human interactions. *Phys. Rev. E*, 68:065103(4), 2003.

[59] W.B. Deng, L. Guo, W. Li, and X. Cai. Worldwide marine transportation network: Efficiency and container throughput. *Chin. Phys. Lett.*, 26(11):118901(4), 2009.

[60] John P. J. Pinel. *Biopsychology*. Pearson, eighth edition, 2013.

[61] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature*, 10:186–198, 2009.

[62] Valentino Braitenberg and Almut Schüz. *Cortex:Statistics and Geometry of Neuronal Connectivity*. Springer-Verlag Berlin Heidelberg, second edition, 1998.

[63] Béla Bollobás, Robert Kozma, and Dezsö Miklós. *Handbook of Large-Scale Random Networks*. János Bolyai Mathematical Society and Springer-Verlag, first edition, 2008.

[64] Reuven Cohen, keren Erez, Daniel ben Avraham, and Havlin Shlomo. Breakdown of the internet under intentional attack. *Physical Review Letters*, 86(16):3682–3685, 2001.

[65] S. N. Dorogovtsev and J. F. F. Mendes. Comment on "breakdown of the internet under intentional attack". *Physical Review Letters*, 87(21):219801, 2001.

[66] Steven H. Strogatz. *Sync: the emerging science of spontaneous order*. Hyperion, 1st edition, 2003.

[67] Alex Arenas, Albert Díaz-Guilera, Kurths Jurgen, Yamir Moreno, and Chansong Zhou. Synchronization in complex networks. *Physics Reports*, (469):93–153, 2008.

[68] Klaus Lehnertz, Bialonski Stephan, Horstmann Marie-Therese, Krug Dieter, Rothkegel Alexander, Staniek Matthäus, and Wagner Tobias. Synchronization phenomena in human epileptic brain networks. *Journal of Neuroscience Methods*, 183:42–48, 2009.

[69] S. Gonçalves, G. Abramson, and M.F.C. Gomes. Oscillations in sirs model with distributed delays. *The European Physical Journal B*, 83(1):363–371, 2011.

[70] Mauricio Barahona and Louis M. Pecora. Synchronization in small-world systems. *Physical Review Letters*, 89(2):054101(4), 2002.

[71] Takashi Nishikawa, Adilson E. Motter, Ying-Chen Lai, and Frank C. Hoppensteadt. Heterogeneity in oscillator networks: Are smaller worlds easier to synchronize? *Physical Review Letters*, 91(1):014101(4), 2003.

[72] Adilson E. Motter, Changsong Zhou, and Jürgen Kurths. Network synchronization, diffusion, and the paradox of heterogeneity. *Phys. Rev. E*, 71(1):016116(9), 2005.

[73] H. Hong, Beom Jun Kim, M. Y. Choi, and Hyunggyu Park. Factors that predict better synchronizability on complex networks. *Phys. Rev. E*, 69(9):067105(4), 2004.

[74] Luca Donetti, Pablo I. Hurtado, and Miguel A. Muñoz. Entangled networks, synchronization, and optimal network topology. *Physical Review Letters*, 95(18):188701(4), 2005.

[75] Renato E. Mirollo and Steven H. Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM Journal of Applied Mathematics*, 50(6):1645–1662, 1990.

[76] X. Guardiola, A. Díaz-Guilera, M. Llas, and C. J. Pérez. Synchronization, diversity, and topology of networks of integrate and fire oscillators. *Phys. Rev. E*, 62(4):5565–5570, 2000.

[77] Lawrence Perko. *Differential equations and dynamical systems*. Springer New York, third edition, 2001.

[78] Yoshiki Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer, 1st edition, 1984.

[79] Huzihiro Araki. *Lecture Notes in Physics 39*. Springer-Verlag New York Heidelberg Berlin, 1975.

[80] Steven H. Strogatz and R. E. Mirollo. Phase-locking and critical phenomena in lattices of coupled nonlinear oscillators with random intrinsic frequencies. *Physica D*, 31:143–168, 1988.

[81] Debanjan Chowdhury and M.C. Cross. Synchronization of oscillators with long-range power law interactions. *Physical Review E*, 82(1):011605(12), 2010.

[82] Máté Morádi, Francesco d'Ovidio, and Tamás Vicsek. Synchronization of oscillators with long range interaction: Phase transition and anomalous finite size effects. *Physical Review E*, 66(1):011109(4), 2002.

[83] Hietsugu Sakaguchi, Shigeru Shinimoto, and Yoshiki Kuramoto. Local and global self-entrainments in oscillator lattices. *Progress of Theoretical Physics*, 77(5):1005–1010, 1987.

[84] D. J. Watts. *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton Univ. Press, Princeton, NJ, 1999.

[85] Takashi Ichinomiya. Frequency synchronization in random oscillator network. *Physical Review E*, 70:026116(5), 2004.

[86] J. G. Restrepo, E. Ott, and B. R. Hunt. Onset of synchronization in large networks of coupled oscillators. *Physical Review E*, 71(3):036151(12), 2005.

[87] Kevin O'Kefee, P. L. Krapivsky, and Steven H. Strogatz. Synchronization as aggregation: Cluster kinetics of pulse-coupled oscillators. *Physical Review Letters*, 115:064101(5), 2015.

[88] E. Oh, K. Rho, H. Hong, and B. Kahng. Modular synchronization in complex networks. *Physical Review E*, 72:047101, 2005.

[89] Alexander W. Rives and Timothy Galitski. Modular organization of cellular networks. *PNAS*, 100(3):1128–1133, 2003.

[90] Kasper Astrup Eriksen, Ingve Simonsen, Sergei Maslov, and Kim Sneppen. Modularity and extreme edges of the internet. *Physical Review Letters*, 90(14):148701(4), 2003.

[91] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *PNAS*, 99(12):7821–7826, 2002.

[92] Norman Biggs. *Algebraic Graph Theory*. Cambridge University Press, 1st edition, 1974.

[93] Alex Pothen, Horst D. Simon, and Kang-Pu Liou. Partitioning sparse matrices with eigenvectors of graphs. *SIAM*, 11(3):430–452, 1990.

[94] Herbert W. Hethcote, Harlan W. Stech, and P. Van Den Driessche. Nonlinear oscillations in epidemic models. *SIAM Journal of Applied Mathematics*, 40(1):1–9, 1981.

[95] Roy M. Anderson and Robert M. May. *Infectious Diseases of Humans:Dynamics and Control*. Oxford University Press, 1991.

[96] *Stochastic Processes in Epidemic Theory*. Springer-Verlag Berlin Heidelberg, 1990.

[97] Marc Barthélemy, Alain Barrat, Romualdo Pastor-Satorras, and Alessandro Vespignani. Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Physical Review Letters*, 92(17):178701(4), 2004.

[98] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic dynamics and endemic states in complex networks. *Physical Review E*, 63(6):066117(8), 2001.

[99] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Physcal Review Letters*, 86(14):3200–3203, 2001.

[100] Yamir Moreno, Romualdo Pastor-Satorras, and Alessandro Vespignani. Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B*, 26(4):521–529, 2002.

[101] Robert M. May and Alun L. Lloyd. Infection dynamics on scale-free networks. *Physical Review E*, 64(6):066112(4), 2001.

[102] Ilkka Gilg, Oliver an Hanski and Benoît Sittler. Cyclic dynamics in a simple vertebrate predator-prey community. *Science*, 302:866–68, 2003.

[103] Peter Turchin and Stephen P. Ellner. Living on the edge of chaos: populations dynamics of fennscandian voles. *Ecology*, 81(11):3099–3116, 2000.

[104] Alexander Dobrinevski and Erwin Frey. Extinction in neutrally stable stochastic lotka-volterra models. *Physical Review E*, 85(5):051903(12), 2012.

[105] Jonas Cremer, Tobias Reichenbach, and Erwin Frey. The edge of neutral evolution in social dilemmas. *New Journal of Physics*, 11:093029(15), 2009.

[106] Tobias Reichenbach, Mauro Mobilia, and Erwin Frey. Coexistence versus extinction in the stochastic cyclic lotka-volterra model. *Physical Review E*, 74(5):051907(11), 2006.

[107] Tobias Reichenbach, Mauro Mobilia, and Erwin Frey. Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature*, 448:1046–1049, 2007.

[108] Dante R. Chialvo. Generic excitable dnamics on a two-dimensional map. *Chaos, Solitons & fractals*, 5(3/4):461–479, 1995.

[109] Suhita Nadkarni and Peter Jung. Spontaneous oscillations of dressed neurons: A new mechanism for epilepsy? *Physical Review Letters*, 91(26):268101(4), 2005.

[110] Rebecca A. Brauc, M. Adnan El-Masri, Joseph C. Parker Jr., and Rif S. El-Mallakh. Glial cell number and neuron/glial cell ratios in postmortem brains of bipolar individuals. *Journal of Affective Disorders*, 91:87–90, 2006.

[111] Ingemar Nåsell. On the time to extinction in recurrent epidemics. *J.R. Statist. Soc. B*, 61:309–330, 1990.

[112] Alex Kamenev and Baruch Meerson. Extinction of an infectious disease: A large fluctuation in a nonequilibrium system. *Physical Review E*, 77(6):061107(4), 2008.

[113] Ira B. Schwartz, Lora Billings, Mark Dykman, and Alexandra Landsman. Predicting extinction rates in stochastic epidemic models. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(01):P01005, 2008.

[114] Miguel González, Rodrigo Martínez, and Maroussia Slavtchova-Bojkova. *Workshop on branching processes and their applications - Chapter 17, pages 241–256*. Springer-Verlag Berlin Heidelberg, 2010.

[115] Onno A. van Herwaarden and Johan Grasman. Stochastic epidemics: major outbreaks and the duration of the endemic period. *Journal of Mathematical Biology*, 33:581–601, 1995.

[116] Quan-Xin Liu, Rong-Hua Wang, and Zhen Jin. Persistence, extinction and spatio-temporal synchronization of sirs spatial models. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(07):P07007, 2009.

[117] Seung Ki Baek. Oscillatory behaviors of an epidemiological model on small-world networks. *Journal of the Korean Physical Society*, 50(1):320–326, 2007.

[118] Hidetsugu Sakaguchi. Cooperative phenomena in coupled oscillator systems under external fields. *Progress of Theoretical Physics*, 79(1):39–46, 1988.

[119] Steven H. Strogatz and Renato E. Mirollo. Stability of incoherence in a population of coupled oscillators. *Journal of Statistical Physics*, 63(3/4):613–635, 1991.

[120] Redner Sidney. *A guide to first passage processes*. Cambridge University Press, 2001.

[121] A. Molini, P. Talkner, G.C. Katul, and A. Porporato. First passage time statistics of brownian motion with purely time dependent drift and diffusion. *Physica A*, 390:1841–1852, 2011.