DECISION MAKING IN COMPLEX ENVIRONMENTS : AN ADAPTIVE NETWORK APPROACH

GIANCARLO DE LUCA

Advisor: Prof. Matteo Marsili

Physics Area
SISSA—International School for Advanced Studies

Trieste
30 October 2013

Quotes are a substitute for wit.
— Oscar Wilde

To my family, to my friends, to those who share(d) my life:
for making me a better person

ABSTRACT

In this thesis we investigate decision making in complex environments using adaptive network models.

We first focus on the problem of consensus decision making in large animal groups. Each individual has an internal state that models its choice among the possible q alternatives and we assume that each individual updates its internal state using a majority rule, if it is connected to other individuals, or using a probabilistic rule. In this case, if the individual has no information, the choice shall be totally random, otherwise the probabilistic rule shall have a bias toward one of the q choices, measured by a parameter h_i. The individuals shall also update their neighbourhood adaptively, which is modelled by a link creation/ link destruction process with an effective rate z. We show that the system, if there are no informed individuals, undergoes a I order phase transition at a give value, z^*, between a disordered phase and a phase were consensus is reached. When the number of informed individuals increases, the first order phase transition remains, until one reaches a critical value of informed individuals above which the system is no more critical. We also prove that, for z in a critical range, the removal of knowledgeable individuals may induce a transition to a phase where the group is no able to reach a consensual decision. We apply these results to interpret some data on seasonal migrations of Atlantic Bluefin Tuna.

We, then, build a model to describe the emergence of hierarchical structures in societies of rational self-interested agents. This model constitutes a highly stylised model for human societies. The decision-making problem of the agents, in this situation, is to which other agent to connect itself. We model the preference of agents of that society for connecting to more prominent agents with a parameter β. We show that there exists a sharp transition between a disordered equalitarian society and an ordered hierarchical society as *beta* increases. Moreover, we prove that, in a hierarchical society, social mobility is almost impossible, which captures behaviours that have been observed in real societies.

RIASSUNTO

In questa tesi investighiamo i processi di decision in ambienti complessi usando modelli di reti adattive.

Ci focalizziamo, per prima cosa, sul problema delle decisioni consensuali nei grandi gruppi di animali. Ogni individuo ha uno stato interno che modellizza la sua scelta fra le q possibili alternative e assumiamo che ogni individuo possa aggiornare il suo stato interno usando una regola di maggioranza, se è connesso ad un altro individuo, oppure con una regola probabilistica. In quest'ultimo caso, se l'individuo non ha informazioni la sua scelta sarà completamente casuale, altrimenti, la regola probabilistica avrà uno sbilanciamento verso una delle q possibili scelte misurato dal parametro h_i. Un individuo potrà anche aggiornare il suo vicinato adattivamente, il che è modellizzato da un processo di creazione e distruzione di collegamenti ad un tasso effettivo z. Mostriamo che il sistema, in assenza di individui informati, subisce, at un dato valore z^*, una trasizione di fase del prim'ordine fra una fase disordinata e una fase in cui il gruppo raggiunge il consenso. Quando il numero di individui informati aumenta, la transizione del prim'ordine permane finché non si raggiunge un numero critico di individui informati, al di sopra del quale il sistema non è più critico. Proviamo anche che, per z in un intervallo critico, la rimozioni di individui informati può causare una transizione in una fase in cui il gruppo non è più capace di raggiungere il consenso. Applichiamo questi risultati all'interpretazione di dati sulle migrazioni stagionali del tonno rosso atlantico.

Construiamo, in seguito, un modello per descrivere l'emergenza di stutture gerarchiche in società di agenti razionali ed egoisti. Questo modello costituisce un un modello altamente stilizzato per le società umane. Il problema decisionale degli agenti, in questa situazione, è a quale altro agente connettersi. Modelliziamo le preferenze degli agenti in questa società verso la connessione ad agenti più prominenti con il parametro β. Mostriamo che esiste una transizione brusca tra una società disordinata, equalitaria, e una società ordinata e gerarchica all'aumentare di β. Proviamo, inoltre, che in una società gerarchica la mobilità sociale è quasi impossibile, il che cattura comportamenti che sono stati osservati in società reali.

ACKNOWLEDGEMENTS

I would like to express my very great appreciation to my advisor, Prof. Matteo Marsili, for his help, his support, his patient guidance and for his constructive critiques. I learnt a lot from him in these years.

I would also like to express my gratitude to Dr. Patrizio Mariani and Prof. Brian R. McKenzie: they guided me trough marine ecology with patience and friendliness. Discussing with them helped me a lot.

I would like to thank, as well, Dr. Marco Bardoscia, Dr. Giacomo Livan and Dr. Claudio J.Tessone for their help and contribution to this research project: collaborating with them as been a rewarding and pleasurable experience.

Finally, I would like to thank my friends, my family and those who shared my life in these four years, for having endured and helped me in my bad days and having enjoyed and shared with me my happy moments.

CONTENTS

Part I

INTRODUCTION

1

INTRODUCTION

Decision making has been extensively studied both in social sciences and in biology. Many fundamental questions in economics can be reduced to a decision problem (how much to buy, when etc.); similarly in sociology and political sciences many observed patters stems from a decision problem (why we choose to live in a given neighbourhood, why we vote for a given politician). In biology as well, animals are confronted to the problem of making choices continually, and their evolutionary success depends on how effective their decision abilities are.

Modelling of decision making has a long history in social science, in particular in economics, and a wide literature has been produced on the topic. Putting aside the general problem of modelling complex entities such as human or animals in a mathematically tractable way, the other main problem lies in the complex nature of the environment in which these entities live, and, in particular, their social environment.

This overwhelming complexity, in many cases, has been only partially addressed. In neoclassic economics, for example, the entire theory is built on an object, the utility function, which represent the balance between costs and benefits of making an (economic) choice; all the complexity of the problem is transferred in the choice of an adequate utility function and in a definition of a series of pattern of interaction between the individuals (agents in the language of economics) in the market. In these context, complex *social* interactions have rarely been resolved explicitly.

Similarly in biology, traditional ecological modelling have relied on what, in physical terms, we would call *a mean field* approach, typically writing phenomenological equations for population of a given class (e. g. the famous Lotka-Volterra model). In this case too, *social* complexity as been mainly addressed, essentially writing a stochastic version of these models; possibly inducing some spacial, inhomogeneity.

Only in the last twenty years, with the development of complexity science, and in particular of network science, attempts of resolve explicitly the network of interaction in decision making have been attempted.

In this thesis we focus on modelling decision making in complex environments using what are usually referred as adaptive network models.

We shall present two models of such kind, one which attempts to model consensus decision making in animal groups, and in particular to offer a simplified setting to address migration of large group of animals (chapter 2), and another which focus on decision making in closed societies, in particular on the processes that generate hierarchies in closed societies (chapter 3).

1.1 COMPLEX ADAPTIVE SYSTEMS

The definition of complex adaptive systems[1] is still under debate, as almost any thing in what is called complexity science[2].

We can however use the definition given by Holland [4]:

> Complex Adaptive Systems are systems that have a large numbers of components, often called agents, that interact and adapt or learn.

Animal societies—including human societies—can, clearly, be subsumed in this general category; societies are formed of individuals, these individuals interact with one another, and they adapt their behaviour according to changes either internal to society itself (endogenous changes in the language of sociology) or external influence (exogenous changes).

As it is underlined in [2], complex adaptive systems, are not just *complicated* systems. The fact that a system is characterised by a huge number of possibly different constituents, each of which might have a complicate description, is not sufficient to generate complex pattern and behaviour. Interaction and interdependence is the most relevant aspect: in a complex adaptive system, interaction is such that the collective behaviour is qualitatively and substantially different from that of the single elements. This particular property if usually referred as *emergence* in the literature of complex systems.

This emergent, seemingly *ordered and organised* behaviour does not requires the existence of a centralised planning or of a decisional hierarchy, but it is the consequence of local interaction: this property is usually referred as *self-organisation*. *Self-organisation* and *emergence* are the most characteristic features of Complex Adaptive Systems.

In many cases, as in collective animal behaviour or in human societies, the difference between the behaviour of the group/society and that of the individual that constitute it is so evident, that it has often been questioned whether the usual mechanistic/reductionist approach of exact sciences can be applied. This is particularly true in social sciences.

However, the experience of statistical physics prove that this apparent violation of the reductionist principles of science is, indeed, only apparent. Thanks to the work work of three generation of scientists, a set of rigorous mathematical tools have been developed, which explain why collective properties can be so qualitatively different from microscopical one.

The usual approach of statistical physics is to model the microscopical interaction and then *infer* the global behaviour. This approach can, therefore, be extended l to modelling Complex Adaptive Systems, provided that it is possible to give a meaningful simplified microscopic description of the components and of their interaction

1 For a general review on the topic see [1]; for a more up-to-date review with particular focus on social sciences see [2]; for an historical recognition see [3]

2 The study of complex systems, is by the nature of its problem, intrinsically multidisciplinary. Thus different scientists coming from different backgrounds will often offer a completely different perspective on the topic.

1.1.1 *Agent based models*

Models of Complex Adaptive Systems that use this approach are commonly called *Agent base models*[3].

In this framework one assumes that the system is composed of similar entities called *agents* which interacts with one another over time: It is moreover assumed that each *agent* :

1. is "intelligent and purposeful" ;
2. can interact with other agents using simple rules;
3. can possibly interact with its environment.

The fact that each agent is intelligent and purposeful is usually modelled by the fact that the agent has an internal state (which might simply be the velocity of the agent) and that it updates it according a set of simple rules

The computational complexity of such models usually increases very fast with the number of agents simulations become rapidly inaccessible if large computer clusters are not available.

In general, agent based models are not analytically tractable; however, in some cases, it is possible to build simplified but reasonably descriptive agent based models which are, to many extent, solvable.

1.2 MODELLING SOCIETIES WITH NETWORKS

Modelling decision making in societies requires modelling the underlying social structure. And the *natural language* for these studies is the one of networks.

In principle it would be possible to model social interaction using what we could refer to as *real space* models: for each agent, model the location in space at a given time and all the possible interaction which are possible in that specific location. However social interaction, except possibly for animals in large group, is not simply local in space or in time. This is particularly true for modern human societies in the *the global village*[4] where communication between agents, and thus interaction, is not limited by *spacial proximity*.

In general, a microscopic description that explicitly embodies all the *real space* details, if possible, is exceedingly complicated, and most probably will not lead to an analytically tractable model. In this case realistic details are a distraction and a network description can be seen as a simplified representation of these social systems in which unnecessary details are left away [7] (as Ising model is for magnetisms in physics).

The mathematical theory underlying network models is *graph theory*, a branch of combinatorics and topology, with a long tradition [8].

3 this name was first used in a famous article of Holland and Miller [5] who borrowed this terminology from economics, and was than adopted by the computer scientists doing these set of models)

4 As Marshall McLuhan vividly defined the globalised world [6]

In particular in addressing real networks, the typical approach follows the one pioneered by Erdős-Rényi [9] and Gilbert [10]. Although the aim of those works was to demonstrate the existence of some graphs with given properties, their approach to this problem opened a new fruitful line of research. Erdős and Rényi, in fact, relied on a probabilistic construction to prove their theorems: they defined a statistical ensemble of graphs, and described the *typical* properties of graphs in this ensemble. This has been the main tool used in network science since then.

In the last fifteen years, the study of networks has exploded, fuelled by the development of the Internet and the availability of large data sets. Empirical data showed that Erdős-Rényi model is to oversimplified and triggered a global effort to find more realistic descriptions of real networks (two famous examples of this approach— probably cited in most of network related litterature—are Watts-Strogarts "small world" model [11] and Bárabasi-Albert model [12]). From these studies, two substantially different approaches emerged.

A first approach, is essentially an approach à la Erdős-Renyi. One attempts to describe a graph ensemble with the aim to reproduce the characteristics of networks as obtained from empirical studies. In this sense, we could call this approach a data driven approach.The "dynamical rules" of these systems, if any, do not aim (necessarily) to reflect the true dynamics of the underlying system but are rather chosen in such a way to lead to networks with the desired properties.

The other approach does exactly the opposite. Networks are, in this case, a simplified or stylised description of an underlying system, whose behaviour is not clearly understood and one aims to elucidate with the netwrok model. The dynamic rules are, in this case, a reasonable, albeit simplified, description of the dynamics of the underlying system and the properties of the networks are inferred from the dynamics.

In this thesis we shall adopt the second approach, and provide two examples to which this approach may be fruitfully applied.

In the first case, we tackle the problem of consensus decision making in groups of animals (chapter 2). In this context a network approach provides a mathematically more tractable model with respect to the usual description: there exist the possibility to use a real space model but it is not analytically tractable.

In the second example(chapter 3) we focus on decision making in societies. In this case the network description is a natural modelling framework and not just a mathematically more tractable simplification of a more precise but complex description.

The reader can find an overview of the basic definition of graph theory and the notation used in this thesis in appendix A.1.

1.3 NETWORK DYNAMICAL MODELS

Adaptivity is essentially a dynamical property of a complex adaptive systems. Thus the model that we shall discuss will be essentially network dynamical models. One can generally define a them as stochastic process of a dynamical

system on a graph; to do so one must define a proper phase (or state) space on which the time evolution takes place. A way to do it is to assume that each agent i is a vertex in a graph. We model its internal state with one or more variables that we shall denote collectively with a_i. The network of interaction is given by a graph structure. To keep the discussion simple we shall assume that this structure can be described by and adjacency matrix G.

The phase (or state) space of this system is then given by set Ω of pairs $\omega = (\boldsymbol{a}, \mathsf{G})$ where $\boldsymbol{a}$ is a succinct way to denote the set of all internal states a_i and G is the adjacency matrix of the system.

The stochastic or deterministic rule will essentially specify a sequence of states ω_t indexed by the time.

This general approach is however artificial. In fact when modelling Complex Adaptive Systems the graph represent the network of interaction between agents and thus it is much more natural to describe the evolution of the internal states and that of the network separately and then *infer* general dynamical rules on the phase space.

dynamics on the networks The dynamics of the internal state is often referred as dynamics *on the network*. The main assumption that one makes is that the dynamical update rule is local, i. e. that the evolution of a_i may depend only from the state of a_j of the agents to which a_i is connected. This assumption in modelling Complex Adaptive System is, indeed, consistent with the interpretation network as *network of (social) interaction*.

dynamics of the network The other process that may occur is a change of graph. In this case there is much more freedom in the choice of the update rule. For example one may assume that there may be a *vertex instantiated network update* which models the change of the interaction network originated by an agent (as might be the creation or the interruption break of a friendship), *edge instantiated update* such as the decay of an edge; it is even possible to envisage processes occurring on more general motifs. In general, it is not required that the Dynamics of the network depends on the states of the vertices.

If the time scale of the dynamics over the network is much faster that the dynamics of the network; such a system will behave as stochastic process on a complex topology. On the other hand if the time scale of the network dynamics is much faster, the system will be *well mixed* and thus a mean field approach that model these random interaction is adequate.

However, when these too time scale are comparable, one must take into account the details of network evolution and

1.4 ADAPTIVE NETWORKS

Adaptive networks [13] are network dynamical models in which the network update rules depend on the states of vertices of the system.

An adaptive network model, thus, is essentially a agent based network dynamical system in which the evolution of the properties of the vertices is determined

by the topology of the network whose evolution, is itself determined by the states of the system.

Adaptive networks are a natural framework for model Complex Adaptive Systems and in particular decision making in complex social environments. In fact, it is reasonable to assume that the social neighbourhood of an agent is modified by the choice he makes.

The interest in adaptive networks models is recent, Giving a general overview of adaptive models is beyond the scope of this thesis(see [13] for a general review); however to give an general idea of the

Among all processes that one can consider, *interacting particle systems* [14], also called contact processes, , are of particular relevance.. In this processes the state of an individual (particle, agent) can assume a discrete (and finite) number of state. When two individuals meet one of them shall update its internal state according to the state of the other individual: this mechanism can promote local homogeneity and induces some form of *spreading* of *information*. Contact processes are used in physics, biology or sociology to model—in an extremely stylised manner—as diverse phenomena as the spread of epidemics, opinion dynamics, the spacial competition of different species, etc.

In particular, for contact processes, it is easy to set-up an adaptive network generalisation since the dynamic on the network is rather simple. In fact, adaptive versions of contact processes of fixed discrete topology have been among the first adaptive network models to be developed (i. e. adaptive SIS model [15] or adaptive voter model [16–19])

To clarify some aspects of these stochastic adaptive network model, and, mainly to illustrate some of the standard techniques used to analyse them, we shall focus on the simplest among the state-topology co-evolving adaptive networks: the voter model.

As we shall see in chapter 2, voters model, is not only a toy model, simple enough to illustrate standard techniques, such as moment closure approximation, painlessly ; but it has been extensively used to model opinion dynamics in animal groups. Even though we shall not directly use moment closure approximation in chapter 2— we shall provided an exact solution of the model there discussed,— the other adaptive models for decision making in animal groups do. To give a broader perspective to the reader, we have decided to discuss here at length the adaptive voter model as described in [16] and moment closure approximation.

1.4.1 *Adaptive Voter Models*

The voter model [20] is a prototypical model for opinion dynamics and possibly the simplest of all contact processes. In its first version[21] it was stated as a model for competition of species and received its name *voter model* in [22]. Albeit being an extremely stylised description of any real process, its simplicity and the fact that, it is solvable has made it a popular model to describe opinion dynamics.

Since the adaptive version of this model depend much on the choice of the network updating rule, adaptive voter models refer to an entire class of models.

The voter model

In its original form, the voter model was defined over a two dimensional lattice. In all sites a there is a *voter* that can be in two states $s_i = \pm 1$; at any time step a voter i selects one of its neighbours j and set $s_i = s_j$. This systems is exactly solvable in any dimension, since it can be mapped in a model of random walkers [14].

Let $S = \{s_i\}$ be the set of all the spins. Clearly the *consensus* state defined as the state in which all the spins are either up or down is an absorbing state; all other states are transient and this implies that the system is not ergodic. If the lattice is finite the system will always reach in a finite time the absorbing states and that the stationary distribution will be fully characterised by the probability π of reaching the state the state with all spins up.

The system can be solved intuitively by writing down the transition rates for spin flips [23]. For a finite number of spins, the master equation can be written in terms of the transition rates W_k for the spin flips of the s_k spin .

$$W_k = \frac{1}{\tau}\left(1 - \frac{1}{2d} s_k \sum_{j \mathcal{N}_i} s_j\right) \tag{1.1}$$

where $\mathcal{N}_i$ represent the neighbours of k, τ is a time scale factor and d the dimensionality of the lattice.

The master equation can be solved exactly, since the rates are linear; in particular all the correlation can be calculated exactly for any dimension d. It can be shown, in particular, that the average magnetisation m is conserved by this dynamics. As a consequence the probability of reaching the absorbing state where all $s_i = +1$ is $\pi = \frac{m+1}{2}$. If the lattice is finite, that this suffice to characterise the system. However for infinite lattice, the time to reach consensus may diverge. It has been shown that in fact this depends from dimensionality [23]. Since the system is solvable the asymptotic behaviour of the density of interfaces $n(t)$ can be calculated. For $d \leq 2$ $n(t) \to 0$ which implies that consensus is reached, whereas for $d > 2$ $n(t) \to$ cost. this does not happen.

There have been a lot of extensions, ranging from extending the model to a different dynamics, using different internal states spins, introducing disorder, and changing topology using various network models [20]. In particular in heterogeneous topologies one can loose the conservation of magnetisation property.

Adaptive Voter Model

Among the first adaptive model built to generalise voter model, one can count the work of Holme and Newman [16], and with a similar model, to the work of Gil and Zanette [17, 18]. We shall focus on [16].

In Holme and Newman [16], each node on a network has one opinion g_i out of G possible opinions (Potts spins) at when a node i is selected for update, it picks up one of its neighbours j randomly: with a probability φ the node sets $g_i := g_j$, otherwise, with probability $1 - \varphi$ the link is rewired to another node having the same opinion of node i 1.1. This dynamics tend to reduce the density of *active*

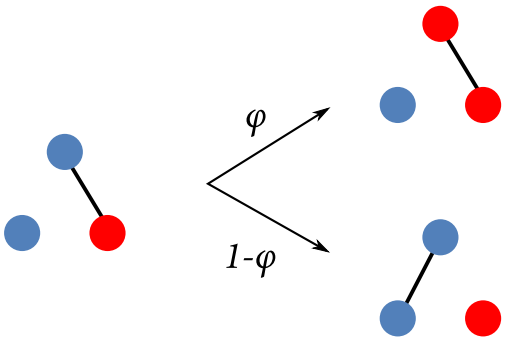


Figure 1.1: An illustration of adaptive voter model dynamics of [16]

links defined as link between node with conflicting opinions. When an upgrade event occurs and an active link is selected, then that link is "inactivated". If $\varphi = 0$ opinions do not change and the final state of the system will be characterised by as many homogeneous connected components as the number of different opinions in the initial system. If, instead, $\varphi = 1$ the network dynamics is frozen and the system will end up in a state with as many homogeneous connected components as the number of connected components in the initial state. More importantly, it can be shown that there is a continuous, out of equilibrium phase transition occurring at a critical value φ_c of the segregation parameter φ.

For an adaptive voter model $G = 2$ starting from a connected graph, this means that aside of the traditional consensus absorbing state, there is another absorbing state characterised by two separated homogeneous communities (*fragmentation*). This phase transition concerns both the topology of the graph (number of connected components) and the distribution of the states: this is a common characteristic of adaptive network models.

Indeed for infinite networks the consensus state might never we reached. A mean field approach has been attempted by [19] to a similar model and it suggests that phase transition for the density of active links exists: below a certain threshold for φ, such density is non-zero and thus the network will never reach a consensus state: above it the network undergoes fragmentation; however for finite size systems due to fluctuations consensus will eventually be reached in any case.

The moment expansion

Vasquez mean field approach [19] can be seen as a special case of a general approach, the so called *moment expansion*. In this approach, one seeks to write down differential equations for the so called moments of the graph, i. e. the densities of particular motifs and subgraphs [24].

The understand how it works let's apply it to the adaptive voter model used by Vasquez.

Let us denote with V the set of agents (vertices) in the model, each of which is identified by an integer number i and let $N = |V|$ be the number of agents in the system; let s_i denote the state of agent i, and $\boldsymbol{s}$ denote the vector whose

i-th component is s_i. Let G be the adjacency matrix of the interaction network. The set $\omega = (\boldsymbol{s}, \mathsf{G})$ define the states of the system and the stochastic rule of voter model define a continuous Markov chain process on these states.

To solve it one should write down the master equation of the system, but solving it is cumbersome. To obtain some hints on the system one can solve its rate equations, i. e. solving the *mean field dynamics*. This approach is usually referred, in this context, as moment expansion.

At any given time t that we shall not explicitly write in the following paragraphs, let us denote by $[s]$ the density of spins in state s, i. e. :

$$[A] = \frac{1}{N}\langle\big|\{i \in V \ : \ s_i = A\}\big|\rangle. \tag{1.2}$$

this quantity is called the zeroth order moment. The first order moments, denoted $[AB]$ are the densities of links between vertices in state A and B, i. e. :

$$[AB] = \frac{1}{2N}\langle\Big|\{(i,j) \in V \times V \ : \ (s_i = A) \wedge (s_j = B) \wedge (g_{ij} = 1)\}\Big|\rangle. \tag{1.3}$$

The second order moments $[ABC]$ are the densities connected subgraphs with three vertices and two links, i. e. :

$$\begin{aligned}[ABC] = \frac{1}{2N}\Big\langle\Big|\{(i,j,k) \in V \times V \times V \ : \ (s_i = A) \wedge (s_j = B) \wedge (s_j = C) \\ \wedge (g_{ij} = 1) \wedge (g_{jk} = 1)\}\Big|\Big\rangle.\end{aligned} \tag{1.4}$$

One can go on defining a whole hierarchy of momenta in a similar way, for and in principle describe the mean field dynamics writing the the infinite system of ODE:

$$\frac{\mathrm{d}}{\mathrm{d}t}[A_1 A_2 \dots A_n] = \mathrm{f}_{A_1, A_2, \dots A_n}([A_1], [A_2,], \dots, [A_1, A_2]], \dots) \tag{1.5}$$

where f are function of the momenta. In general, this system of equation cannot be truncated and the evolution of a given moment depends of higher order moments.

For the adaptive voter model discussed in [16], if we use R (Red) to define the vertices in state 1 and B (Blue) those in state -1 (as in figure 1.1) these moment equations for first two moments truncated to first three moment read:

$$\frac{\mathrm{d}}{\mathrm{d}t}[R] = 0 = \frac{\mathrm{d}}{\mathrm{d}t}[B] \tag{1.6}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}[RR] = \frac{[RB]}{2} + \frac{\varphi}{2}(2[RBR] - [RRB]) \tag{1.7}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}[BB] = \frac{[RB]}{2} + \frac{\varphi}{2}(2[BRB] - [BBR]). \tag{1.8}$$

The first equation is trivially the consequence of the symmetry of the system. The density of same colours link can be obtained either trough direct action of the dynamics, i. e. with probability $\frac{\varphi}{2}$ an active link become an RR and with probability $\frac{1-\varphi}{2}$ and active link is rewired, thus leading to a neat contribution

of $\frac{[RB]}{2}$; the other contribution comes from indirect action of the dynamics: if a central node in a BRB sub graph upgrades to R than two RR are indirectly created, similarly when a central node R in a RRB is upgraded to B a RR node is destroyed, al these events occur with probability $\frac{\varphi}{2}$.

The fixed point of these equations states that the number of active links must satisfy:

$$[RB] = \varphi(2[RBR] - [RRB]) = \varphi(2[BRB] - [BBR]) \tag{1.9}$$

Of course such a set of equations cannot be solved analytically unless we find a way to estimate properly the higher order moments. This is done by moment closure approximation.

Moment closure approximation

There are indeed several possibilities to perform moment closure for these sets of equation [24]. We shall focus on the simplest one, the so called *pair approximation* [15].

In this framework, the second order moments are solved with respect to the first order models $[ABC]$ are approximated by :

$$[ABC] \simeq [AB]\langle q_B\rangle p_{BC} \simeq \kappa[AB]\frac{[BC]}{[B]} \tag{1.10}$$

that is it is proportional to the density of AB links multiplied the average number edges that a node in state B who is already connected to another node has (excess degree) multiplied the probability p_{BC} that a link in state B has a BC link. This approximation is rather crude since it completely ignore topological correlations.

In the case of adaptive voter model, we can make a slight better estimate. Clearly we can write $[BR] + 2[RR] = \langle k_R\rangle[R]$ where $\langle k_R\rangle$ is the average degree of vertices in state $[R]$; thus the probability that a link starting from a node R is of type BR is $\frac{[BR]}{\langle k_R\rangle[R]}$; on the other, if P_k^R be the probability that a node in state R has degree k, the probability that a node in state R at the end another link has l other links attached to it is simply $q_l^R = \frac{(l+1)P_l^R}{\sum_k kP_k^R}$; thus the average number of additional edges that a node in state R who is already connected to another node has is $\langle q_R\rangle = \sum_l lq_l^R \simeq \frac{\langle k_R^2\rangle[R]+\langle k^R\rangle[R]}{\langle k_R\rangle[R]}$. This implies that the moment closure equations for the adaptive voter models take the form:

$$[BRB] = \frac{\langle q_R\rangle}{2\langle k_R\rangle}\frac{[BR]^2}{[R]} \tag{1.11}$$

$$[RRB] = \frac{2\langle q_R\rangle}{\langle k_R\rangle}\frac{[BR][RR]}{[R]} \tag{1.12}$$

$$[RBR] = \frac{\langle q_B\rangle}{2\langle k_B\rangle}\frac{[RB]^2}{[B]} \tag{1.13}$$

$$[BBR] = \frac{2\langle q_B\rangle}{\langle k_B\rangle}\frac{[RB][BB]}{[B]}. \tag{1.14}$$

If we assume a random network approximation we have that $\frac{\langle q_R \rangle}{\langle k_R \rangle} = \frac{\langle q_B \rangle}{\langle k_B \rangle} \simeq 1$ and thus solving the equation (1.6) and (1.9) one obtains that:

$$[RB] = \frac{\varphi\langle k \rangle - 1}{\varphi}[R](1 - [R]). \tag{1.15}$$

This implies that the fragmentation transition occurs at $\varphi_c => \langle k \rangle^{-1}$. Although this description captures the qualitative behaviour of the system, in many cases the agreement with simulations can be quite bad. In particular, if the degree distribution is broad or there are strong topological correlation this approximation schemes may fail. In [24] one can find a review on other, more complicated moment closure approximation schemes, involving either higher moments or non homogeneous approaches (i. e. approaches that take into account also the degree of vertices and not only the density of motifs). On of the bigger limitation of this approximated approach, is the fact that the complexity of moment closure equation and of moment expansion becomes immediately hard to dominate. Voter model can be handled because there are essentially only two choices for the internal state. The number of motifs to take into account when the number of choices becomes higher increases fast and it makes particularly difficult to apply it to systems with a large number of possible choices. On the other hand, in many cases, moment expansion is the only resource to get some analytical insight on the behaviour of the system.

Part II

ADAPTIVE NETOWRK MODELS: TWO CASE STUDY

2

MODELLING COLLECTIVE DECISION MAKING IN ANIMALS

2.1 COLLECTIVE ANIMAL BEHAVIOUR

Collective animal behaviour is a widespread phenomenon in animal ecology. Under many aspects, it represents the most easily accessible (and spectacular) example of collective phenomena in biology [25]; in animal groups, in fact, the the microscopical and macroscopic level are clearly separated— even to an uneducated eye— which is not always the case when dealing with other collective behaviour in biology (such as protein aggregation).

Such impressing and highly complex social behaviours, can be found across many taxa, form insects— that posse limited cognitive abilities—to mammal, including humans— capable of extremely sophisticated social interaction; *emergence* in this context is mainly behavioural, i.e. the fact that the whole group can complete tasks and achieve results (e.g. nests and shelters construction, migrations between geographically separated location, efficient foraging, etc.) that would be impossible for the individuals to achieve [26–28]. These emergent behaviour are so impressive that researchers have used vivid expressions such as *swarm intelligence [29], wisdom of crowds [28] or collective minds [30]* to refer to them.

Collective animal behaviour are studied, as many topics in biology, under two different perspectives: on the one hand one can concentrate on *how* these *collective behaviour* emerge out of individual interactions, on the other hand, one can focus on *why* a given behaviour evolved over another.

These two completely different perspectives (the former referred in [25] as *mechanistic* , the latter as *functional*) are not conflicting but rather complementary, and both have been applied to the study of coordination, synchronisation, decision making, regulation and cooperation in animal groups. Indeed, collective animal behaviour offer a common ground where to confront the results of both approaches and find a link between them.

Mechanistic approach, however, has enjoyed a lot of attention only in recent years; partly due to the large availability of adequate computational resources, partly as a consequence of the growth of number of scientists who apply quantitative methods of statistical physics and complex systems to biology. In fact the conceptual challenges that one faces are similar: how to explain the leap between the collective behaviour of the group and the "microscopical behaviour" of the individuals.[1]

1 The concept of *Collective intelligence* has been investigated since decades and it has also been often popularised to a general public. By example, in his famous book "Gödel, Escher, Bach: an eternal golden braid" [31], D.R. Hofstadter , discussing about ant colonies, discussed the concept of individual in a group: should the colony, as a whole, be considered as a unitary organism and the single individuals be considered as part of it, in the same manner as blood cells are considered

The class of collective behaviour that we shall study in this chapter is collective decision making, in particular, we want to provide a simple model based on stochastic decision making that can reasonably explain how and when consensual decision across a large group animals is possible.

The relevance of these questions is not only limited to theoretical biology but since collective decision making plays a relevant role in the ethology of many species with high economic value— our case study is the migration of Atlantic Bluefin—-it could offer hints to policy-makers having to regulate such economic activities.

2.2 CONSENSUS DECISION MAKING IN ANIMALS AND COORDINATION

Decision making is a crucial aspect in any animal group. Whether we think to bird flocks having to choose the best direction to escape a predator, or a small group of wolves having to coordinate for hunting, we are considering a decision making problem.

With possibly the only exception of social insects, the focus on decision making processes in animal group is recent in ethological and ecological literature [32]; and a lot of both theoretical and experimental work has been done in the last fifteen years.

In a decision making problem there are always variables to consider:

1. information, (How information is spread in the group? Do all the individuals have the same pieces of information? How reliable is information ?)

2. preferences (Are there conflicts in the group about the choices? How are conflicts regulated?)

When it comes to decision making in group one has to distinguish the case in which the decision is truly *collective*, that is, when there is a social predisposition in the species, from the case when each individual decides by itself but its decision are influenced by the choices other individuals have made. In both cases, in fact, we may assist to the emergence of a collective response but modelling is significantly different.

Collective decision making are usually referred to as *Consensus decision making* and are frequently found in problems such as coordinated movement, the choice of travel destination (e.g. during migration) but also hunting or timing.

We can follow L.Conradt's and T.J. Roper's [33], suggestions, and classify consensus decision making problems in animal groups according to two aspects:

1. communication in the group (local vs. global)

2. conflicts of interests

part of an organism and not individuals by themselves? Or, on the contrary, the single Ants should still keep their individuality? In discussing those aspects, Hofstadter used the term *holism* to refer to the fact that the group behave so differently than the individuals by which they are formed. The language is different but the core of the problem is there.

There are thus 4 possible cases:

global communication and no conflict of interest As an example of this situation one can consider that of a small groups of animals having to navigate toward a chosen destination (such as group of primates of mammals); in this case all members of the group can communicate with the others and complex decisional strategies involving complex negotiation may occurs;

global communication and conflict of interest As an example we have synchronisation and choices of travel destination in small groups

local communication and no conflict of interest As an example we can consider large navigating groups (like swarms of locusts);

local communication and conflict of interest As an example we can consider the choice of travel destination in a large group like the choice of migration destination in some fish schools.

This double dichotomy provides a good general framework; of course there are additional factors to be considered:

decision makers there might be a hierarchy, thus only the leaders take decision, or the decision is shared by all the members of the groups; or something in between.

information for large group on the move, information is usually transferred by alignment policies; for smaller group vocalisation might be another possibility; in insects chemical signalling is also a relevant possibility. Moreover a significant process of information pooling between all the constituents may take place, in small group a least;

function group decision may serve as an information pooling strategy correcting the errors of single individuals in elaborating information, or it may optimise the the choice, in case of conflicts, reducing the *consensus coast*. In other cases, as in the case of fish migrations, it may be an evolutionary strategy minimizing individual costs of information gathering.

decision mechanism in large groups self organising decision rules are more likely; in small groups, where communication is global, different policies may be adopted (such as quorum decision, voting)

2.2.1 *Agent based models for consensus decision making*

Traditional ecological modelling relied on deterministic or stochastic processes on subpopulations rather than on individual based models.

Individual based models, in fact, were in most of the cases not solvable analytically nor reasonable approximation scheme were available. It was not until the eighties that the computational power needed for running simulation of agent based models, became widely available;

The first attempt to apply agent based models to the study of collective motion has probably been C. Rynolds' BOIDS programme [34] in 1986; afterwards, this approach has been widely used (the famous Vicsek model [35] is only a simplification of Reynold model) and has now became one of the main tool in addressing mathematical modelling of collective animal behaviour in animal groups.

Agent based models have been thoroughly used in the last fifteen years to model consensus decision making problems. Here we shall not overview the traditional approaches to modelling decision making in animal groups, (see [32] for a review on the topic).

2.3 THE MIGRATION PROBLEM, CONSENSUS DECISION MAKING

Migration is a complex topic in ecology that involves both taxon-specific aspects and general concepts; indeed the term *migration* has itself different meaning that underlines different perspectives, as Daingle and Drake noted [36]. In fact in biology *migration* may refer to a periodic movement of a general population between two or more geographically separated location (e.g. the migration between feeding and spawning sites), or to movement of a single individual between different locations which occurs at scale much larger than the usual movement, or to a general relocation of a spatially distributed population (e.g. human migrations); it may even refer to the locomotory act itself. Historically. *Migration* has been addressed, as many subjects in life sciences, from two apparently contrasting points of view: a *behavioural* perspective where the focus is on the individual and migration behaviours are seen as the consequence of a genetic evolution, and from an *ecological* perspective, and thus seeing it as a phenomenon acting on populations rather than on individuals.

Indeed the term *migration* is used to refer to a variety of different behaviours which require different modelling schemes.

A general definition of *migration* has been attempted by different authors but still there is not always consensus on it, since much depend on the focus on the study; for our scope however we can employ the definition used in [36]:

> Migration is a movement away from the home range that do not cease, at least initially, when suitable resources or home ranges are encountered.

This definition, indeed, helps to discriminate between *migration* and other behaviours involving movement of groups of animals: *foraging* when movement is usually limited in the home range or stops as soon as an adequate food source is found and *ranging* which is an exploratory movement away from home range.

We have to note that, albeit relevant from a conceptual point of view, mechanisms underlying migration and those underlying other collective animal movements may be rather similar.

When speaking about migrations a second relevant distinction is to be made between periodic and *irruptive* migrations.

The former are usually typical of long-lived specie who get an evolutionary advantage from having a range that encompasses different habitats. The latter,

instead, can be found across all taxa but are more frequent among short living species (e.g in insects) and are usually triggered by a deterioration of the habitat (this fact has been shown in insects [37–39], e.g. locusts [40])

Another aspect, useful to classify migratory behaviours, is the the structure of the migratory patterns. In most cases migrations occur between two sites; but there are cases in which the migration has the structure of a round trip (e.g. herrings). One-way migrations are also possible for short lived animals. In some habitats, disordered patterns are also possible; this behaviour is usually called *nomadism*).

From a modern perspective, potential migratory behaviour is encoded into the genotype of an organism: it require the development of sufficient locomotory abilities so as to allow to travel for long distances often through challenging habitats (deserts, oceans), the ability of gathering information from the environment. In case of group migrations, it also requires to develop social abilities, of an efficient way to share information with the other members of the group, and also some form of memory of the migratory route. This migratory phenotype (called in [36] *migration symptom*), which can be environment-dependent (e.g. locusts [40]), will actually generate a migration only if triggered by the environment (in [36] *migration arena*); the mechanisms, by which a population having a given migratory phenotype will actually undergo a migration, encompass the study of collective-decision making (see section 2.2), and different taxon specific aspects (how information is spread, gathered etc.); finally natural selection will select genotypes that allow for phenotypes which are better adapted to the environment. The *functions* of migration may be multiple. The main advantage seems to be *adaptiveness* with respect to unstable habitat: periodic migration may be seen as an adaptation to habitats that change asynchronously over time (e.g. wintering strategies with respect to summer-winter cycle at higher latitude), whereas *irruptive* or *nomadic* migratory behaviour may be seen as an adaptation to unstable habitats like arid regions where rainfalls are erratic and unpredictable. In general, migration behaviours seem to allow for some form of *pre-emption*, that is the ability to detect the deterioration of an habitat before it reaches a critical level.

Migratory patterns for species undergoing a periodic migration, are rather stable and liable. However, there have been observations [41–45] of abrupt changes in migratory patterns of species including the removal of entire migratory routes. Such sudden changes occur on a time scale which is usually too short to be determined by significant changes in the genotype: the mechanism that undergoes group formation, consensus decision making and collective memory is likely to be responsible for such abrupt changes.

2.3.1 *A case study: the Atlantic bluefin tuna*

The Atlantic Bluefin Tuna (ABFT), *Thunnys Thynnus* (Linnaeus 1758), is a pelagic fish, known for its big size, its rich migratory pattern and for its great economic importance [46–50] .

Its size can exceed 3 m and it can weight up to 900 km. During migrations its speed may reach 70 km/h . Its big size, the quality of its flesh and the high value it can reach on fish markets (mainly on Tokyo Tsukiji fish market) have triggered an excessive fishing effort that have seriously reduced the stock (of about 70 – 80% in 30 years [51]); this overexploitation has led IUCN to classify the ABFT as an endangered species in its red list [52].

Habitat

The ABFT is a widely distributed species that live in the pelagic ecosystem of Northern Atlantic Ocean and its adjacent seas (Mediterranean Sea and Gulf of Mexico). It can sustain a large range of sea water temperature (from 3°C to 30°C) while maintaining a constant body temperature. It usually occupy surface and subsurface coastal and open sea waters although extensive tagging experiments [53] have shown that diving down to a depth of 500 – 1000 m is not infrequent.

Migratory pattern

The ABFT exhibits a seasonal south-north group migratory behaviour. Spawning sites are located in temperate-tropical waters (e. g. western Mediterranean Sea, Gulf of Mexico and historically also Black Sea) [46, 54]), but feeding sites used by the largest and oldest individuals are located in northern temperate-boreal waters (eastern Atlantic: Norwegian and North Seas; western Atlantic: coastal and shelf waters of New England, Nova Scotia and Newfoundland;). It seems to exhibit homing behaviour and spawning site fidelity in both the eastern and western spawning areas [51, 53].

Tagging studies [55, 56] also demonstrate that bluefin tunas undergo trans-Atlantic migrations and that a certain degree of mixing between population is possible. These data also demonstrate that migratory routes of ABFT may evolve in time.

The migratory behaviour exhibited by bluefin tuna has evolved to allow the species to benefit from large biomasses of forage species such as herring and mackerel in these regions [48]; this evolutionary behaviour is supported by notable physiological adaptations, including efficient thermal-regulation [53].

The disappearance of Norwegian route

Until the late 50s [41, 46] landings of large adult ABFT on the coasts of Norway were frequent(see Figure 2.11 (a)).

Estimating the correct biomass of ABFT on the coast of Norway is a complicated task: ABFT until the beginning of XX century was only a by-catch of other kinds of fisheries, and reports of landings were not systematic [41], however there are sufficient evidences that the ABFT was commonly found in North Sea and Norwegian Sea. The systematic records of landing of ABFT by ICES started only in 1927.

The Norwegian bluefin tuna fishery started around 1910s as a by-catch or sport fishery but developed into the most important ABFT fishery in northern

Europe. Fisherman until the 30s used to consider ABFT almost as nuisance rather than a resource since the technology available (the main fishery then was that of herrings) was not adequate for an effective commercial bluefin tuna fishery. Starting from the late 20s the fishing effort toward ABFT increased: there was both a development of more efficient catch methods and of the overall fishing capacity that led to a significant increase of the catches. The ABFT by the 50s had become one the most relevant fisheries in northern seas. [41].

Starting from the mid 1950s the catches and landings of ABFT started to decrease([46] and Figure 2.11) whereas the fishing effort remained constant; the decrease was so abrupt that by the beginning of the mid 60s commercial bluefin tuna fishery was almost abandoned.

Starting from the end of the 60s only occasional landing of vagrant have been recorded[45]. As Fig 2.11 shows this abrupt decrease (in particular in 1960) did not correspond to a generalised reduction of spawning site biomass and thus that particular migratory route has been *forgotten* by the tuna fish that used to migrate in the North sea and on the coast of Norway.

Collective memory of migratory pattern

Understanding what may have induced such a dramatic change in the migratory habits of ABFT, is indeed not only important from a theoretical point of view, but has significant economic consequences.

Different hypotheses have been proposed (degradation of habitat, decrease of water temperature, overfishing)[45] but the rather abrupt decline of landing observed requires a different explanation.

In most of the cases, bluefin tuna migration are collective. Thus migration mechanism rely on the ability of the group to reach consensus on one of the possible destinations. We have to consider, moreover, that the main information shared by fish in schools is about speed and direction, thus an effective model for migration must embody to some extent aspects common to all models of collective motions (see [57]).

Indeed the fact that these migrations are a mainly collective is important. Since the ABFT is an apex predator, grouping behaviour are unlikely to occur for protection. There have been suggestions that bluefin tunas engage in some form of cooperative hunting and get hydrodynamic benefit from schooling [58] but these behaviour have been observed mainly in feeding sites.

Grouping for migration, instead, may have other functions. It has been shown [59–61] that grouping may help the fish in a school to gather information from the environment in a way that goes beyond the simple correction of errors (the so called *collective sensing*). This suggests that individual in a group may have different sensing and cognitive abilities: in particular information about feeding and spawning sites might be unevenly distributed among given population (i. e. there might be informed and uninformed individuals): in this perspective, the memory of the migration route may arise from the interaction of individual and is thus an example of *collective memory* [62, 63].

Only recently a completely different approach that does not rely on explicitly considering the spatial dynamics but uses a dynamic network has been suggested [64].

2.4 MODELLING MIGRATION: REAL SPACE AGENT MODEL FOR CONSENSUS DECISION MAKING

Migration in large groups is a consensus decision-making problem, where communication is non local, and there might be conflict of interests. For group on the move, collective motion models are the prototypical models for consensus decision making; many models that have been developed specifically to model migrations [40, 62, 63, 65–67] have, in fact, been heavily inspired by models for collective motion.

Collective motion models are usually real space agent based [57, 68], that rely on these general assumptions [57]:

1. all agents are similar
2. all agents are moving with a nearly constant absolute velocity; however, each agent can modify its direction of motion;
3. all agents interact within a specific interaction range by changing ; their direction of motion, in a way favouring alignment;
4. all agent might be subject to some noise in its dynamics.

Within this framework, in the last 20 years a great deal of work has been done(see [57] for a review).

2.4.1 *Vicsek's Model*

The so called Vicsek's Model [35] is the prototypical model among the class of models described above. In this model, each particle i is self-propelled, i. e. its speed is constant v_o, and its dynamics, described by position$\boldsymbol{x}_i$ and direction $\boldsymbol{d}_i$ is described by the discretised dynamics:

$$\boldsymbol{x}_i(t+\Delta t) = \boldsymbol{x}_i(t) + v_o \boldsymbol{d}_i(t)\Delta t \qquad \boldsymbol{d}_i(t) = \frac{\sum_{j\in\mathcal{N}_i(t)} \boldsymbol{d}_j(t)}{\left|\sum_{j\in\mathcal{N}_i(t)} \boldsymbol{d}_j(t)\right|} + \xi(t) \tag{2.1}$$

where $\xi(t)$ represents a white noise and $\mathcal{N}_i(t) = \{j \mid |\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)| < r\}$.

This simple dynamics exhibits a phase transition with respect the the intensity of the noise, between a *disordered* phase where each fish moves randomly, and a *ordered* phase, where each agent moves toward the same direction.

Beside the fact of being the prototypical model for collective motion, such a simple model highlights an important fact: the group of agent is able to reach a global consensus through local interaction, that is global consensus may be reached, through a self organising mechanism , without assuming global communication within a group.

2.4.2 *Modelling schools of fish*

Vicsek model is indeed too stylised to capture realistic behaviours. During the last 20 years, there has been a great deal of work trying to build more realistic models for collective animal motion. In this contest a great effort has been put into experimental research trying to model the effective *force field* of each animal.

It is known, in fact, that individuals in a group do not align with all the individuals within a certain range but they tend to weight more individual in some specific positions over other; these choices may be induced by the environment, and by the cognitive abilities and species dependent limitations (for fish see [69–71]).

Since our case study is the migration of ABFT we shall focus on fish. An early attempt to simulate effectively the motion of fish in a school was done by Huth and Wissel in 1992 [72] and followed Reynolds' modelling framework [34]. For each fish they identify three zones:

1. a zone of avoidance for distances smaller that a certain radius r_r, where fish will behave escaping from one another ;
2. an alignment zone, for distances between r_r and r_p, where the behaviour will be similar to Vicsek mode;
3. an attraction zone for distances between r_p and r_a where fish will tend to come closer to another. For distances larger that r_a there will be no interaction.

This model captures better the typical behaviours of fish schools and shoals[73] and is the base of a set of more sophisticated models [74, 75]. Recently, there have been also attempts to solve the inverse problem, that is inferring from the real trajectories of fish in schools the details of dynamic interaction [76].

2.4.3 *Collective memory and the effective leadership model*

General models for collective motion can describe how consensus can be reached without any prior information.

However, for describing a migration process, there is the need to model how agents store information about the "right destination", and possibly encode the fact that information about the "right destination" may be unevenly distributed.

An general agent model for modelling decision making in a large group with unevenly distributed information has been built by I. Couzin, J. Krause Guttal et al. [65, 66] (see also [62, 63]).

In this model, as in Vicsek model, there are N agents and the speed is constant v_o. Thus each agent i is characterised by its position $\boldsymbol{x}_i(t)$ and by the direction of motion $\boldsymbol{d}_i(t)$. Moreover each agent is characterised by two positive parameters $\omega_g^{(i)}$ that measures its gradient detection abilities trait and $\omega_s^{(i)}$ that measure it social detection abilities trait. The space is assumed to be the square $[-1, 1] \times [-1, 1]$ with periodic boundary condition.

Given a distribution of those traits among a population, and the dynamical rule described above, the group might or might not reach migrate efficiently. More importantly, it is possible to apply to treat this simulation as a genetic algorithm, using a simple evolutionary rule [77], and see whether there exists a evolutionary stable solution (ESS).

Rules of Movement

The dynamic of position of is simply defined by

$$\boldsymbol{x}_i(t+\Delta t) = \boldsymbol{x}_i(t) + v_0 \hat{\boldsymbol{d}}_i(t)\Delta t \tag{2.2}$$

where $\hat{\boldsymbol{d}}_i(t)$ represent the unit direction vector.

The upgrade of the directions is more complicated. And can be summarised as follow:

1. Avoid collisions. As in Huth and Wissel model, there exist an avoidance zone defined by a radius r_a. If $\mathcal{N}_i(t) = \{j \ : \ \left|\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)\right| < r\}$ is not empty then the direction is updated to avoid collisions as follow:

$$\boldsymbol{d}_i(t+\Delta t) = -\frac{\sum_{j\in\mathcal{N}_i(t)}(\boldsymbol{x}_j(t) - \boldsymbol{x}_j(t))}{\left|\sum_{j\in\mathcal{N}_i(t)}(\boldsymbol{d}_i(t) - \boldsymbol{d}_j(t))\right|} \tag{2.3}$$

2. If the zone of avoidance is empty, the direction shall be updated as follow:

$$\boldsymbol{d}_i(t+\Delta t) = \frac{\omega_s^{(i)}\hat{\boldsymbol{d}}_g^{(i)}(t+\Delta t) + \omega_g^{(i)}\hat{\boldsymbol{d}}_s^{(i)}(t+\Delta t) + \sigma_r\hat{\boldsymbol{d}}_r^{(i)}(t+\Delta t)}{\left|\boldsymbol{d}_g^{(i)}(t+\Delta t) + \omega_g^{(i)}\hat{\boldsymbol{d}}_s^{(i)}(t+\Delta t) + \sigma_r\hat{\boldsymbol{d}}_r^{(i)}(t+\Delta t)\right|}. \tag{2.4}$$

In this formula the term $\boldsymbol{d}_r^{(i)}$ represents a random direction in $\mathbb{R}^2$, modelling noise and intrinsic random motion; σ_r measures the intensity of this noise.

The term $\boldsymbol{d}_s^{(i)}$ represent the direction update due to social interaction. It is assumed that there is a social interaction radius r_s such that:

$$\boldsymbol{d}_s^{(i)}(t+\Delta t) = \frac{\sum_{j\in\mathcal{S}_i(t)}(\boldsymbol{x}_j(t) - \boldsymbol{x}_j(t))}{\left|\sum_{j\in\mathcal{S}_i(t)}(\boldsymbol{d}_i(t) - \boldsymbol{d}_j(t))\right|} + \frac{\sum_{j\in\mathcal{S}_i(t)}\boldsymbol{d}_i(t)}{\left|\sum_{j\in\mathcal{S}_i(t)}\boldsymbol{d}_i(t)\right|} \tag{2.5}$$

where $\mathcal{S}_i(t) = \{j \mid r_a < \left|\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)\right| < r_s\}$.

Finally, the term $\boldsymbol{d}_g^{(i)}$ represent the direction update due to information. In [65] this was a constant whereas in [66] this is assumed to be of the form:

$$\boldsymbol{d}_g^i(t+\Delta t) = (\cos(\theta), \sin(\theta)) \tag{2.6}$$

where θ satisfies the Ornstein-Uhlenbeck stochastic differential equation:

$$\mathrm{d}\theta = -\omega_g^{(i)}\theta\,\mathrm{d}t + \sigma_g\,\mathrm{d}W_r. \tag{2.7}$$

With this notation it is assumed that there is an environmental gradient along x axis, the goal of migration being to reach $x = 1$.

Fitness and evolutionary algorithm

In order to implement an evolutionary algorithm approach, a proper definition of pay-offs for a given set of traits is needed. Such a fitness function can be obtained by subtracting to the benefit of migration for agent i, $b(i)$ defined as being the distance travelled on x axis the coasts of developing the social traits $c_s(i) = p_s(e^{\frac{\omega_s^{(i)}}{\kappa_s}} - 1)$ and gradient detection traits $c_g(i) = p_g(e^{\frac{\omega_g^{(i)}}{\kappa_g}} - 1)$.

The evolutionary policy is a proportional reproduction [77], where each individual reproduces asexually with a frequency in the next generation proportional to its relative fitness. The offspring has the same traits of the parents plus some random mutation.

Evolutionary Stable Solution

Of course, there are some hidden assumption in these models. In particular social abilities and information gathering abilities are not completely independent: however, a great insight is obtained.

First [65], it shows that a small number of informed individuals can effectively polarise a school without the need of signalling or of a specific hierarchical structure. More importantly, the relative number of informed individuals needed to polarise a school seems to decrease with the number of individuals.

An major result which is obtained in [66] using this approach, is the fact that the ESS is characterised by a population where a small amount of individuals have strong information gathering abilities and weak social abilities (the *leaders*), and a vast majority of the population instead has a strong social interaction abilities and weak information gathering ability (the *followers*).

This suggests that group migration evolved, as for collective sensing, as a global property and memory of migratory routes is likely to be a shared knowledge.

2.5 STOCHASTIC ADAPTIVE MODELS FOR COLLECTIVE DECISION MAKING

If the models previously described can effectively give great insight about the migration mechanism, they cannot be solved analytically.

In a real-space agent model the adaptivity stems from the fact that the neighbourhood of an agent changes over time *due* to to positions and velocities of the other agents; position and velocities are, on the other hand, influenced by the neighbourhood of the agents.

Since, at the end, what is important in the system is who is interacting with whom, it is natural to attempt a simplification of these models using a dynamic network model (Figure 2.1).

Adaptive network

The idea of applying adaptive network models to get intuition about collective decision making in animals has been suggested by Gross and Blasius in 2008 [64].

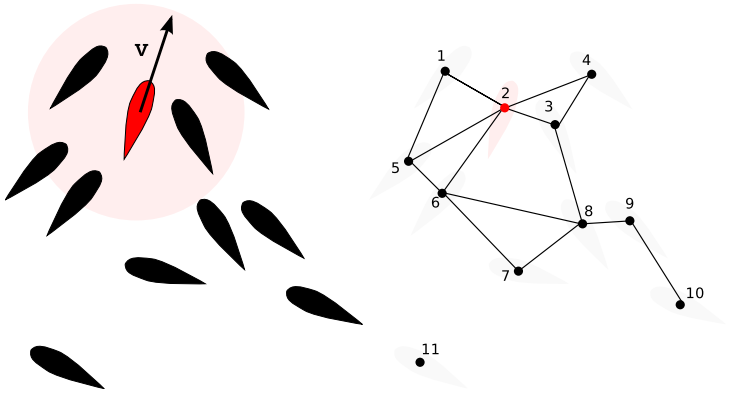


Figure 2.1: An illustration relating real space agent based models and network based models

Later, a particular flavour of an adaptive voter model have been proposed in [78] and [63] to address the problem of collective decision making.

Let us focus on the model described in [78]. In this work the aim is to describe density triggered transition in locusts from a disordered phase where each individual moves almost independently from the others to an ordered one when all individuals move together. In locusts, these transition is not only behavioural (collective motion vs. independent motion) but it correspond to a well known cases of polyphenism (environmental induced changed in phenotypes on a genetically similar individuals); it has been suggested [40] that this behaviour might be induced by cannibalistic behaviour that can be observed, when densities are high, in otherwise herbivores locusts [79].

This work is related to an experiment done by Buhl *et. al.* [80]. In this experiment locusts where put in a annulus and left free to move: the experiment showed that below a certain critical density of individuals, the movement was disordered, whereas above that threshold, the locusts started moving coherently either clockwise or counter clockwise.

The model studied in [78] modelled the individuals locusts as nodes in a network that have two state, left goer L or right goer R. The meeting of opposite going individuals is modelled by a Poissonian $R-L$link creation process with a per node link a_o, and a per $R-L$ edge link destruction rate of d_o. The equal-goers $R-R$ and $R-L$ per node link creation rate is a_e and the per edge link destruction rate d_e. The intrinsic opinion dynamics is modelled through a random opinion switch occurring with a rate q whereas the induced opinion change is modelled as a voter model like update is with a w_2 probability per $R-L$ link plus a 3 node update (i. e. a process that sends LRL to LLL) occurring with probability w_3 per triplet.

This model is not exactly solvable but can be tackled using the moment expansion similarly to what happens for the simple adaptive voter model, discussed

in section 1.4.1, and using a moment closure approximation. Because of the 3 node process one shall need to consider also second order moments

$$[{}^{D}A^{C}_{B}] = \frac{1}{4N}\langle\Big|\Big\{(i,j,k,m)\in V\times V\times V\times V \;:\; (s_i=A)\wedge(s_j=B)\wedge(s_k=C) \\ \wedge(s_m=D)\wedge(g_{ij}=1)\wedge(g_{ik}=1)\wedge(g_{im}=1)\Big\}\Big|\rangle. \tag{2.8}$$

In our case the state can be only $\{R, L\}$ and thus the moment expansion equations truncated to first, for $[R]$ and $[RR]$ order read :

$$\frac{\mathrm{d}}{\mathrm{d}t}[R] = q([R]-[L]) + w_3([RLR]-[LRL]) \tag{2.9}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}[RR] = q([RL]-2[RR])w_2([RL]+2[RLR]-[RRL])+ \\ w_3(3[{}^{R}L^{R}_{R}]-[{}^{R}R^{L}_{L}])+a_e[R]^2-d_e[RR]. \tag{2.10}$$

The moment closure approximation for first order moments as those discussed for voter model (eq. **??** (with $B = L$), for the second moments using the same approach one gets:

$$[{}^{R}L^{R}_{R}] \simeq \kappa_L^2\frac{[LR]^3}{6[L]^2}, \qquad [{}^{R}R^{L}_{L}] = \kappa_R^2\frac{[LR]^2[RR]}{[R]^2} \tag{2.11}$$

If there are no equal-goers link creation ($a_e = d_e = 0$), the system undergoes a supercritical bifurcation (II order phase transition) between an ordered and disordered phase. When there is equal-goers link creation/destruction, the system may undergo a sub-critical bifurcation (I order transition).

In this system there is no information bias and thus it is no suited to model migrations, but a slight modification of this simple model has been proposed by in [63], with information bias. Using the same approximate approach this work prove that the relative number of informed individuals play an important role in the establishment of a consensus, and that, at least when individual have only two alternatives, *uniformed individuals promote democratic consensus*; this, in addition to the results previously found in [66] suggest that an adaptive network model for migrations must embody this information bias.

Both the work in [63, 78] use a voter model approach, with a slight modification to encode the non linearity of state update in animals. A more realistic approach should consider a different internal state upgrade that take into consideration all the neighbours of an individual, or at leas a significant portion of them; there are evidences (e. g. see [81]) that animals update their position using more than one individual at a time.

2.6 A SOLVABLE ADAPTIVE NETWORK MODEL FOR MIGRATION I: DESCRIPTION

The previous models offer a hint of the possibilities offered by adaptive network models; however their solution relies on a description which is essentially a

mean field approach, solved using an approximation model that is known to have issues to correctly describe the system qualitatively. There is, however, the possibility to build an adaptive network model, similar to those discussed in [63, 78] and [82, 83], which can be exactly solved even with the presence of information [84].

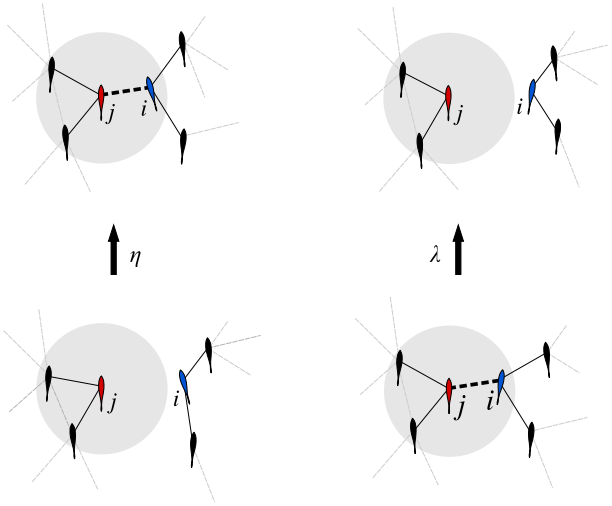


Figure 2.2: Connection between link creation and destruction process and real space models where η is the link creation rate and λ the link destruction rate.

2.6.1 *Informal description of the model*

Let us consider a group with N individuals. In our network model each individual is represented by a node (thus N is the total number of nodes) and each node i has an internal dynamical variable a_i that can take integer values ranging from 1 to q.

Although the mathematical solution does not depend on the specific interpretation of the variable a_i, in the context of migrating groups, a_i might be considered as the direction taken by a single individual to reach the destination site.

A state of the system is defined by the symmetric adjacency matrix of the system g_{ij} and by the set of the internal dynamical states variables a_i. In our model, links are *mutual* and, thus, the adjacency matrix is a symmetric matrix (i.e. for all i, j we have that $g_{ij} = g_{ji}$) such that $g_{ij} = 1$ if there is a link between the nodes i and j, $g_{ij} = 0$ otherwise. The evolution of the system is governed by stochastic dynamics in which both the neighbourhood and the values of the internal dynamical state may vary.

In particular, the choice of the destination a_i is influenced by two factors: *i)* pro-social behaviour, by which an individual keeps the same destination of their neighbours and *ii)* memory, by which an isolated individual preferentially heads toward a destination α_i that is encoded in its memory. The first effect is quantified by the rate η at which individuals form new links with other individuals, This rate should encode both evolutionarily selected traits for pro-social behaviour

and environmental factors, notably the average distance between individuals. In our model we assume that the interactions between individuals heading towards different directions decay much faster than interactions between close individuals heading towards the same direction. This is in agreement with real space dynamical model and is achieved in the network by assuming that link creation can occur only when individuals have the same internal state a_i. Therefore if $a_i \neq a_j$ the link between nodes decays immediately.

On the other hand, memory is quantified by the strength h_{α_i} with which unlinked individuals adopt their preferred choice α_i. The fraction of individuals with a preferred destination α is n^α but we also contemplate a fraction n^0 of "uninformed" individuals, that have no *a priori* preference for any memorized destination. In our model, we use the convention that uninformed individuals have $\alpha_i = 0$ and $h_{\alpha_i} = 0$. Finally, individuals linked and moving in the same direction can also move further apart from each other, which is formally encoded by assuming that links between nodes decay with a constant rate λ. We note that this modelling framework allows to treat sociality, information and preference as independent factors in social animals. Therefore knowledgeable individuals can still display sociality and conformity to the group.

Network dynamics

We model the dynamics of link creation and destruction with stochastic Poissonian processes. At a rate η each individual can make a connection with another individual picked up randomly among all the individuals to which it is not connected. The connection will be established only if the individuals share the same internal dynamical state, i.e. if $a_i = a_j$.

Internal state dynamics

The change of the internal state is a Poissonian process with rate ν. When an internal state update event occurs the individual has two possibilities. If an individual is linked to other individuals(s), it will update its internal state conforming to the state of the majority in its neighbourhood; i.e., the new state a_i' is:

$$a_i' = \underset{x}{\text{argmax}}\Big(\sum_j g_{ij}\delta_{xa_j} \Big) \tag{2.12}$$

This assumption is coherent with what is usually done in modelling group motion that is assuming that an individual tend to "follow" the its neighbours.

On the other hand, if an individual is not linked, it will pick up an internal state at random. More precisely, we assume that each individual has a preferred value of the internal variable, let us call it α_i. In the case of an internal state update event, an unlinked individual will pick up a state according to the following probability distribution:

$$\text{Prob}(a_i = a) = \frac{e^{h_i\delta_{a\alpha_i}}}{q - 1 + e^{h_i}} \tag{2.13}$$

where h_i is a parameter that measures the intensity of the preference. Uninformed individuals, instead, update their direction at random, which is described by Eq. (2.13) with $h_i = 0$.

It is relevant to note that in our model *all individuals* have a social component. For example in groups with conflicting preferences our model suggests that, for some range of the parameters, an informed individual can follow the group and migrate toward a site different from its preferred destination. This approach makes our definition of leaders not only dependent on the amount of information stored but also on the social context in which they live. Therefore, the interaction between personal information and social effects is explicitly resolved in our model and we note it has been suggested to operate in living groups [85, 86].

2.6.2 *The interpretation of the parameters of the model*

The model is described in a formal mathematical way; in this form it can be applied to a large variety of contexts, as long as the parameters have a meaningful interpretation. In the perspective of applying it to modelling consensus decision making in migrating species, we have to give a clearer interpretation in terms of realistic parameters.

As already discussed, migratory behaviours require that an individual or a species has evolved specific locomotory and cognitive abilities that would enable it to undertake a migration. These *genetic* factors evolve on a time scale much slower than the lifespan of each individuals. On the other hand, there are *experiential* factors mainly linked to the environment (e. g. density of fish, habitat quality) and the cognitive elaboration (e. g. memory, learning). These aspects evolve on a significantly slower time scale and can vary during the lifespan of an individual.

The link dynamics

It is rather obvious that the value of η must have a dependence on the density of individuals; most of the real space models, in fact, assume that each individual interacts only with "close" individuals (e.g. closer than a certain radius) which is a reasonable description of the natural behaviour of schooling fishes; thus if the local density of individuals is too low the probability of being close enough to interact with one another is small and thus the link creation rate must also be small; on the contrary the higher the density, the higher the number of "close" individuals and the higher the creation rate must be. Moreover we know that there are physical constraints to the density of fish which mean that the shape of the density induced variation of η has an upper bound. Moreover, social individuals (species) have developed specific social abilities which are encoded in their genetic make-up and thus a social fish is expected to be able to interact with other individuals of its species more effectively than a fish of a non social species: by the rules of our model, this means that at a given local density of

individuals the social species will have a significantly higher link creation rate and lower link destruction rate. We can write:

$$\eta = \eta_s \eta_e(\rho) \tag{2.14}$$

In this case, a variation of local density may induce an even significant variation of $\eta(i)$. The range of these variations, however, may be seen as a *genetically determined quantity*.[2]

We could take into account genetic variability in the population assuming the each individual has its own link update rate $\eta(i)$ but in general these variation should be of limited range, at least for less cognitively advanced species, thus we shall consider them uniform across the population.

Memory of migrating sites and preference parameter

In our model the memory is encoded in two parameters, the α_i and h_i.

We first have to take into account the fact that the ability to remember a migration route as well as gathering information from the environment must be genetically determined, we can assume that individual with no or low information gathering abilities shall have $h_i \sim 0$ whereas individuals with higher information gathering abilities will have and higher value of the In addition to that we have to consider the experiential factor. If, in some species, the memory of an migration route is innate, in most social species, on the contrary, the migration routes are *learned*; thus a juvenile individual shall have little or no experience, an thus a low h_i whereas experienced individuals shall have a higher h_i. An additional aspect that influence h_i is the quality of habitat. In fact, an experienced individual may be less willing to reach a destination α_i if, in previous migrations, it has proven to be unfit (low density preys, overfishing, etc.). We can assume, in principle, that the value of h across the population is noisy. Putting things together:

$$h_i = h_g(i) h_a(i) h_h(i, \alpha_i) + \xi \tag{2.15}$$

where $h_s(i)$ is the social genetic determined factor, where as $h_a(i)$ is the factor linked to age and learning and $h_e(i)$ accounts for the other habitat quality and ξ is a noise factor accounting for genetic and habitat variability.As for the link creation rate the product accounts for the fact that these factors are not independent (e. g. there cannot be difference between juvenile of experienced individuals if there is not a genetic predisposition to remember and gather information about migration routes).

Noise in most cases is irrelevant.

In fact one can assume that the whole population is fractioned in sub-populations, according to their value of h_i.

Firstly, we can use the results of Guttal's and Couzin's work [66] which states that there should be only two sub-population with respect to the factor h_g: an

2 The product accounts for the fact that is an individual is no social $\eta(i) \sim 0$ independently on the density

information able population, having roughly the same $h_s(i) = H_s$, and a purely social component with $h_s(i) = 0$.

Secondly, we might divide the information able sub population according to experience. This is a species dependent factor but it is reasonable to assume that there are a few stages of *experience*; in particular, we can assume that the juvenile have no information bias $h_a(i) = 0$ whereas all the experienced individuals will have the same experienced induced preference $h_a(i) = H_a$.

Finally, we can assume that the habitat induced change preference is a reasonably fair measure of habitat quality and thus, it must depend only on the migration site, that is $h_h(i, \alpha_i) = H_\alpha \delta_{\alpha\alpha_i}$.

Putting evertithing together we have:

$$h_i = \begin{cases} 0, & \text{if } (h_s(i) = 0) \vee (h_s(i) = 0) \quad \textit{(uninformed individuals)} \\ h_\alpha \delta_{\alpha\alpha_i}, & \text{otherwise} \quad \textit{(informed individuals)}. \end{cases} \tag{2.16}$$

2.6.3 *Formal description of the model*

The intuitive description of the model given in the previous section can be formalised in a precise mathematical way.

The processes described above are Continuous Time Markov Chain described on an articulate state space.

The state space

To describe the network structure we have to precisely define the underling graph structure. Since we are not assuming any node volatility we define a set of nodes V such that $|V| = N$, where N is the number of individuals (agents, fish) in the group;, up to isomorphisms we can assume that $V = \{i \in \mathbb{N} \ : \ i \leq N\}$. We shall assume that each agent has a discrete but finite number of internal states q. We can assume, up to isomorphisms, that set of possible internal state of a node is $S = \{i \in \mathbb{N} \ : \ 0 < i \leq q\}$. We define for each node i, a parameter $\alpha_i \in S$ that represent the preferred state, and a real number h_i which represents the strength of that preference. Let, finally, $\Gamma(V)$ be the set of all undirected graphs on the node space V.

A state of the system ω will be defined by :

1. A graph $\mathcal{G}_\omega \in \Gamma(V)$, that is a couple $\mathcal{G}_\omega = (V, \mathsf{G}(\omega))$ where $\mathsf{G}(\omega)$ is the adjacency matrix of the graph in the state ω. To refer to the ij element of the matrix $\mathsf{G}(\omega)$ we shall use $g_{ij}(\omega)$. Since the interaction is mutual, we shall use a undirected graph model and thus $\mathsf{G}(\omega)$ shall be symmetric, i. e. $g_{ij}(\omega) = g_{ij}(\omega), \quad \forall\, (i, j) \in V \times V$;

2. A ordered N-tuple $\boldsymbol{a}(\omega) \in S^N$ whose i-th element, that we shall write $a_i(\omega)$, represents the internal state of the i-th node; we shall—improperly— refer to it a the *state vector*.

Thus a state ω can be defined by the ordered couple $(\mathcal{G}^{(\omega)}, \boldsymbol{a})$ but since the nodes are fixed, one can just define the state ω as the tuple formed by the elements of

the adjacency matrix $g_{ij}(\omega)$ and the elements of the state vector. Finally let $\mathcal{S}$ denote the set of all the states.

The description of the process and the master equation

By the intuitive description we have given in section 2.6.1 it is clear the the system can be describe by a Continuous Time Markov Chain (CTMC; for a general introduction see [87]).

Let Ω_t be the random variable describing the state of the system at time t; Due to Markov property all the properties of the system are completely determined by the transition probabilities:

$$P(\omega', t|\omega, s) = \text{Prob}(\Omega_t = \omega'|\Omega_s = \omega) \tag{2.17}$$

and the probability $P(\omega', s) = \text{Prob}(\Omega_s = \omega)$; for convenience we may collect all the transition probabilities in a single matrix valued function $\mathsf{P}(t, s)$ such that $\mathsf{P}_{\omega\omega'}(s, t) = P(\omega', t|\omega, s)$.

More precisely in a CTMC the probability must satisfy, for all ω, ω':

$$P(\omega', t + \delta t|\omega, t) = \delta_{\omega\omega'} + \mathsf{Q}_{\omega\omega'}\delta t + o(\delta t^2). \tag{2.18}$$

The non diagonal elements of the Q matrix (or transition matrix), since the probabilities are positive numbers, must be positive $\mathsf{Q}_{\omega\omega'} \geq 0$. Due to conservation of probability, moreover, it is easy to check that :

$$\sum_{\omega'} \mathsf{Q}_{\omega\omega'} = 0 \tag{2.19}$$

which implies that $\mathsf{Q}_{\omega\omega} = -\sum_{\omega' \neq \omega} \mathsf{Q}_{\omega,\omega'}$.

By definition the transition probabilities satisfy the so called *Kolmogorov forward and backward equations*:

$$\partial_t \mathsf{P}(s, t) = \mathsf{P}(s, t)\mathsf{Q} \tag{2.20}$$

$$\partial_s \mathsf{P}(s, t) = -\mathsf{Q}\mathsf{P}(s, t). \tag{2.21}$$

with the initial condition $\mathsf{P}(t, t) = \mathsf{I}$.

Since the system is, by definition, translationally invariant (the Q matrix is a constant) we can simply set $s = 0$ and use only Kolmogorov forward equation. It is thus straightforward to see that:

$$\mathsf{P}(t) = \mathsf{P}(0, t) = \mathrm{e}^{\mathsf{Q}t}. \tag{2.22}$$

Kolmogorov equations has the advantage to describe the system independently of the initial state of the system; however, one is generally interested in the quantities $P(\omega, t) = \text{Prob}(\Omega_t = \omega)$. The equation for $P(\omega, t)$ can be easily obtained. Let $\boldsymbol{p}(t)$ be the column vector of probabilities such that $\boldsymbol{p}_\omega(t) = P(\omega, t)$, and let $\boldsymbol{p}(0)$ be the column vector of initial probability distribution, defined as $p_\omega(0) = P(\omega, 0)$. Clearly, $\boldsymbol{p}(t) = \mathsf{P}^T(t)\boldsymbol{p}(0)$. Rewriting the Kolmogorov forward equation in terms of probabilities instead of the transition probabilities we obtain the *master equation* usually used in natural sciences :

$$\partial_t \boldsymbol{p}(t) = \mathsf{Q}^T \boldsymbol{p}(t). \tag{2.23}$$

To describe the system in the long run one looks at the stationary distribution $\boldsymbol{\pi}$ which is defined by:

$$\boldsymbol{\pi} = \mathsf{P}^T(s,t)\boldsymbol{\pi}. \tag{2.24}$$

This immediately implies, using equation (**??**), that the stationary state is defined by:

$$\mathsf{Q}^T\boldsymbol{\pi} = 0 \tag{2.25}$$

It can be shown [87] that, if the state space has a finite cardinality, a CTMC *has always at least one stationary distribution*. The study of CTMC amounts to the study of the structure of the matrix Q and of the state space.

The form of Q matrix

In our case, there are 3 main processes that are involved in the definition of the Q matrix.

link creation With a rate η each node i can establish a link with another node j picked up randomly among the others. The link is established only if the nodes share the same internal state. This process can connect only a state to another state that differ by the fact of having an additional an edge between node in the same internal state, i. e. two states ω and ω' such that $\exists(i,j) \in V \times V \ : \ (g_{ij}(\omega) = g_{ji}(\omega) = 0) \wedge \ (g_{ij}(\omega') = g_{ji}(\omega') = 1) \wedge (\forall\,(h,k) \in V \times V \ : \ (h \neq i) \wedge (k \neq j),\ g_{hk}(\omega) = g_{hk}(\omega') = g_{kh}(\omega) = g_{kh}(\omega'))$ and $\boldsymbol{a}(\omega) = \boldsymbol{a}(\omega')$. The transition rate is clearly:

$$\mathsf{Q}_{\omega\omega'} = \frac{2\eta}{N-1}\delta_{a_i(\omega')a_i(\omega')_j} \tag{2.26}$$

link destruction Each link has a destruction rate λ.

This process can connect a state only to another state that differ from it by the fact of having removed an edge between node in the same internal state,i. e. two states ω and ω' such that $\exists(i,j) \in V \times V \ : \ (g_{ij}(\omega) = g_{ji}(\omega) = 1) \wedge \ (g_{ij}(\omega') = g_{ji}(\omega') = 0) \wedge (\forall\,(h,k) \in V \times V \ : \ (h \neq i) \wedge (k \neq j),\ g_{hk}(\omega) = g_{hk}(\omega') = g_{kh}(\omega) = g_{kh}(\omega'))$ and $\boldsymbol{a}(\omega) = \boldsymbol{a}(\omega')$. The transition rate is clearly:

$$\mathsf{Q}_{\omega\omega'} = \lambda \tag{2.27}$$

internal state update Each node can update its internal state at a rate ν.

If the node is linked when an internal state upgrade event occurs, it will conform to its neighbourhood. When an internal state upgrade event occurs for a linked individual i, thus, the new internal state of that individual a_i' is chosen using a *majority rule*, that is :

$$a_i' \in \underset{x}{\operatorname{argmax}}\Big(\sum_j g_{ij}\delta_{xa_j}\Big) \tag{2.28}$$

Thus this process can only link to states that differ by the value of the internal variable of a linked node,i. e. two states ω, ω' such that: $\exists (i,j) \in V \times V \ : \ (g_{ij}(\omega) = 1) \wedge (a_i(\omega') \neq a_i(\omega)_j)$, and $\mathsf{G}(\omega) = \mathsf{G}(\omega')$[3]

We thus have:

$$\mathsf{Q}_{\omega\omega'} = \nu \frac{\chi_{\underset{x}{\operatorname{argmax}}\left(\sum_{j\in V} g_{ij}(\omega)\delta_{xa_j(\omega)} \right)}(a_i(\omega'))}{\left| \underset{x}{\operatorname{argmax}}\left(\sum_{j\in V} g_{ij}(\omega)\delta_{xa_j(\omega)} \right) \right|} \tag{2.29}$$

where $\chi_A(x)$ is the indicator function of the set A.

When, instead, a node is not linked, at a rate ν it will undergo a random transition to a state a; with the probability:

$$\text{Prob}(a) = \frac{\mathrm{e}^{h_i \delta_{a\alpha_i}}}{q - 1 + \mathrm{e}^{h_i}} \tag{2.30}$$

where $\alpha_i \in S$ is a parameter measuring the state in which the node i would like to update whereas the $h_i \in \mathbb{R}$ measure the strength of this bias.

This update rule can allow for transitions between two states that differ by the value of an internal state of an unlinked node, i. e. two states ω and ω' such that $\exists i \in V \ : \ \forall j \in V, \quad g_{ij}(\omega) = 0 \wedge a_i(\omega) \neq a_i(\omega')$ and $\mathsf{G}(\omega) = \mathsf{G}(\omega')$ and thus:

$$\mathsf{Q}_{\omega\omega'} = \nu \frac{\mathrm{e}^{h_i \delta_{a_i(\omega')_i \alpha_i}}}{q - 1 + e^{h_i}} \tag{2.31}$$

All other non diagonal transition rates are zero.

2.7 A SOLVABLE ADAPTIVE NETWORK MODEL FOR MIGRATION II: INVARIANT DISTRIBUTION

We know that a finite state CTMC has always a stationary distribution. However, according the the structure of the process the might be more than one stationary state. To understand how many stationary states this system admits and their properties we have first to analyse the class structure of the state space.

2.7.1 *The class structure of the state space*

Let us denote by:

$$\mathscr{E} = \{\omega \in \mathcal{S} \ : \ \forall (i,j) \in V \times V \ : \ (g_{ij}(\omega) = 1), \qquad (a_i(\omega)_i = a_j(\omega))\} \tag{2.32}$$

that is the subset of state space that encompasses all the states that do not have links between nodes in a different internal state.

3 In principle this rule allows for transition between ω to ω itself, but in CTMC $\mathsf{Q}_{\omega\omega'}$ is determined by normalisation

Let us denote by $\mathscr{T} = \mathscr{S} \setminus \mathscr{E}$ the set of states that have links between nodes in a different internal state. By definition $\mathscr{S} = \mathscr{E} \oplus \mathscr{T}$.

The first step to characterise the behaviour of the system is to describe the structure of the state space with respect to the dynamics. We shall show:

1. First, we prove that $\mathscr{E}$ is closed, that means that, starting from a state in $\mathscr{E}$, it is impossible to reach any state of $\mathscr{T}$
2. Then we prove that the dynamics is *irreducible* on $\mathscr{E}$, i.e. that it is possible to reach any state in $\mathscr{E}$ starting from any state in $\mathscr{E}$; this is enough in this system to prove that the dynamics is *ergodic* when restricted to $\mathscr{E}$.
3. We prove that the states of $\mathscr{T}$ are *transient*,i. e. that from any state in $\mathscr{T}$ we can reach a state in $\mathscr{E}$ in a finite time. Thus any state in $\mathscr{T}$ is not relevant in the long run, i. e. for any $\omega \in \mathscr{T}$, $\lim_{t\to\infty} P(\omega, t) = 0$.

We know, since the state space is finite, that this system admits at least one stationary distribution. The fact that the state space admits a unique ergodic subspace ensures that the stationary distribution is unique.

PROPOSITION 2.1. *$\mathscr{E}$ is closed* □

Proof. Let $\omega \in \mathscr{E}$. By definition,

$$g_{ij}(\omega) \neq 0 \implies a_i(\omega) = a_j(\omega'); \tag{2.33}$$

in order to generate a transition to a state $\omega' \in \mathscr{T}$, one should either create a link between two nodes h, k such that $a_h(\omega') \neq a_k(\omega')$ and $g_{hk}(\omega) = 0$, but this is clearly prohibited by eq. (2.26) since $\delta_{a_k(\omega')a_k(\omega')} = 0$; or change the internal state of a node i such that $\exists j \in V \ : \ g_{ij}(\omega) = 1$ which is impossible because of eq. (2.33) (the upgrade process following (2.28) will not change the state of that node since its neighbours have share al the same internal states). This proves the thesis, since, for state that differ by more than a link or by the value of more than one internal state, $\mathsf{Q}_{\omega\omega'} = 0$. □

PROPOSITION 2.2 (ERGODICITY WITH RESPECT TO $\mathscr{E}$). *The system is ergodic when restricted to $\mathscr{E}$* □

In order to prove this proposition we need first to prove this technical lemma

LEMMA 2.1. *Given an arbitrary state $\omega \in \mathscr{E}$, there exists a finite sequence of states $S_L = \{\omega\}_0^L$ in $\mathscr{E}$ such that:*

i. $\omega_0 = \omega$;

ii. ω_L *is a state with empty network and all internal states equal to 1, i. e.* $(G)(\omega_L) = 0$ *and* $a_i(\omega_L) = 1 \ \forall i \in V$;

iii. $\mathsf{Q}_{\omega_n\omega_{n+1}} > 0$ *and* $\mathsf{Q}_{\omega_{n+1}\omega_n} > 0$ *for all* $\omega_n \in S_L$.

□

Proof. For any state ξ, let M_ξ indicate the number of edges of the network in the state ξ (i. e. $M_\xi = \frac{1}{2}\sum_{i,j\in V} g_{ij}(\xi)$).

Let us first define recursively a sequence of states $\{\beta_n\}_0^{L_1}$ in in $\mathscr{E}$, such that $\beta_0 = \omega$ and, for any β_n:

(a-1) $\exists (a,b) \in V\times V$, such that $g_{ab}(\beta_n) = 1$ and $g_{ab}(\beta_{n+1}) = g_{ba}(\beta_{n+1}) = 0$;

(a-2) $\forall (i,j) \in V\times V \ : (i,j) \notin \{(a,b),(b,a)\}, \quad g_{ij}(\beta_{n+1}) = g_{ij}(\beta_n)$;

(a-3) $\boldsymbol{a}(\beta_n) = \boldsymbol{a}(\beta_{n+1})$.

Clearly, because of equations (2.26) (2.27) and (2.18), $\mathsf{Q}_{\beta_n\beta_{n+1}} = \lambda > 0$ and, $\mathsf{Q}_{\beta_{n+1}\beta_n} = \frac{2\eta}{N-1} > 0$ since $\beta_n \in \mathscr{E}$ by construction.

Moreover, because of property (a-1), we have that $M_{\beta_{n+1}} = M_{\beta_n} - 1$, and, since, $M_{\beta_0} = M_\omega > 0$, the sequence is finite and of length $L_1 = M_\omega$. Thus β_{M_ω} is a state with no edges and all the internal variables equal to those of ω.

Let us define recursively another sequence $\{\gamma_n\}_0^{L_2}$ such that $\gamma_0 = \beta_{M_\omega}$ $\gamma_{L_2} = \omega_L$ and for all γ_n and γ_{n+1} with $n < L_2$ it is true that:

(b-1) $\exists k \in V$ such that $a_k(\gamma_n)_k \neq 1$;

(b-2) $a_k(\gamma_{n+1}) = 1$;

(b-3) $\forall i \in V \ : \ (i \neq k), \quad a_i(\beta_{n+1})i = a_i(\gamma_n)$.

Let $J_\omega = \sum_i (1 - \delta_{a_i(\omega)1})$ be the number of nodes having an internal state different from 1. Clearly, because of property (b-2), $J_{\gamma_{n-1}} = J_{\gamma_n} - 1$. Moreover because of equation (2.31), $\mathsf{Q}_{\gamma_n,\gamma_{n+1}} = \nu\frac{e^{h_k\delta_{1\alpha_k}}}{q-1+e^{h_k}} > 0$ and $\mathsf{Q}_{\gamma_{n+1},\gamma_n} = \nu\frac{e^{h_k\delta_{a_k(\omega)\alpha_k}}}{q-1+e^{h_k}}$. The length of the sequence is $L_2 = J_{\gamma_0} \leq N$ and thus the sequence is finite. Thus the sequence:

$$\omega_n = \begin{cases} \beta_n \text{ if } n \leq M_\omega \\ \gamma_{n-M_\omega} \text{ if } n > M_\omega \end{cases} \tag{2.34}$$

of length $L = M_\omega + J_{\gamma_0}$ satisfies the conditions of the thesis. □

Proof of proposition 2.2. We only need to show that the system dynamics can link, in a finite time, any two states in $\mathscr{E}$ (i.e. that Q is irreducible when restricted to $\mathscr{E}$, see [87, p. 126]). Since the number of states is finite, this immediately implies that that all the states in $\mathscr{E}$ are positive recurrent, i. e. have a finite the average recurrence time.

To prove this, it is sufficient to show that, given and two states $\omega, \omega' \in \mathscr{E}$, there is a positive integer $L < \infty$ such that there exists a finite sequence of states $\{\omega_n\}$ that satisfies:

(a-1) $\omega_0 = \omega$

(a-2) $\omega_L = \omega'$

(a-3) $\mathsf{Q}_{\omega_n\omega_{n+1}} > 0$ and $\mathsf{Q}_{\omega_{n+1}\omega_n} > 0$

Let $\bar{\omega}$ be the state such that $\forall i,j \in V \times V, \quad g_{ij}(\bar{\omega}) = 0$ and $\forall i \in V \ : \ a_i(\bar{\omega})_i = 1$. By applying Lemma 2.1 first to ω and $\bar{\omega}$,then to ω' and $\bar{\omega}$ and finally merging the two sequences of states obtained, we obtain a sequence satisfying properties a. This proves the thesis. □

PROPOSITION 2.3. *If $\omega \in \mathscr{T}$, then ω is transient.* □

Proof. Let $\omega \in \mathscr{T}$, let $M_\omega > 0$ be the number of edges of the network in the state ω . Let $\bar{\omega}$ be the state such that $\mathsf{G}(\bar{\omega}) = 0$ and $\boldsymbol{a}(\bar{\omega}) = \boldsymbol{a}(\omega)$. By construction, thus, $\bar{\omega} \in \mathscr{E}$. Since $\mathscr{E}$ is closed, to prove the proposition it is sufficient the prove that any state ω communicates with $\bar{\omega}$ on the jump process induced by the stochastic dynamics of our system (see [87, p. 111, p. 115]); this amounts to prove that there exists a finite sequence of states $S_L = \{\omega_n\}_0^L$ such that:

(a-1) $\omega_0 = \omega$;

(a-2) $\omega_L = \bar{\omega}$;

(a-3) $\mathsf{Q}_{\omega_n\omega_{n+1}} > 0 \ \forall \omega_n \in S_L$;

Let us construct such a sequence. Let ω_n , $0 < n < L$, be defined as follow:

(b-1) $\exists a, b$ such that $g_{ab}(\omega_n) = 1$ and $g_{ab}(\omega_{n+1}) = 0$;

(b-2) $\forall (i,j) \in V \times V \ : \ (i,j) \notin \{(a,b),(b,a)\}, \quad g_{ij}(\omega_{n+1}) = g_{ij}(\omega_n)$;

(b-3) $\boldsymbol{a}(\omega_n) = \boldsymbol{a}(\omega_{n+1})$.

Since $\omega_0 = \omega$ and $\omega \in \mathscr{T}$, the network in ω_0 must have at least one edge. Let M_{ω_0} be the number of edges in the network of state ω_0. If $M_{\omega_0} = 1$ the sequence consists only of two states and $\omega_1 = \bar{\omega}$. If $M_{\omega_0} > 1$ then there shall be exactly $L = M_{\omega_0} + 1$ states in S_L (due to property (b-1)). Clearly $\mathsf{Q}_{\omega_n\omega_{n+1}} = \lambda$ (due to (2.27)) which proves the the thesis □

2.7.2 *Stationary distribution and detailed balance solution*

Once we have proved that the stationary distribution is unique, we can attempt to obtain the stationary distribution.

The stationary $\boldsymbol{\pi}$ satisfies by definition

$$\mathsf{P}(s,t)\boldsymbol{\pi} = \boldsymbol{\pi} \tag{2.35}$$

and:

$$\pi(\omega) = \lim_{t\to\infty} P(\omega, t) \tag{2.36}$$

In [87] one can find the proof that, on an ergodic subspace (like $\mathscr{E}$), indeed, for a finite size CTMC the system will eventually converge to the stationary distribution for any given initial condition.

Derivation of detailed balance distribution

We look for distribution that satisfy detailed balance (DB), which means that, for any couple of states ω, ω', we must have:

$$\pi(\omega)\mathsf{Q}_{\omega\omega'} = \pi(\omega')\mathsf{Q}_{\omega\omega'} \tag{2.37}$$

Clearly any distribution satisfying detailed balance is in particular stationary .

To obtain explicitly such distribution, let us assume that DB apply, we obtain $\boldsymbol{\pi}$ by repeatedly applying (DB) starting from a state with non empty network. Let $M_\omega = \frac{1}{2}\sum_{i,j\in V} g_{ij}(\omega)$ be the number of edges of the graph in the state ω.

Let ω and ω' in $\mathscr{E}$ be such that $M_\omega = M_{\omega'} + 1$, and they differ only by an edge between the nodes i, j. In this case only link creation/destruction processes may occur and we get:

$$\pi(\omega) = \frac{\mathsf{Q}_{\omega'\omega)}}{\mathsf{Q}_{\omega\omega')}}\pi(\omega') = \left(\frac{2\eta\delta_{a_i(\omega)a_j(\omega)}}{\lambda(N-1)}\right)^{g_{ij}(\omega)}\pi(\omega') \tag{2.38}$$

Applying recursively the previous equation one gets, for any $\omega in \mathscr{E}$:

$$\pi(\omega') = \pi(\bar{\omega})\prod_{i,j\in V,\, i<j}\left(\frac{2\eta\delta_{a_i(\omega)a_j(\omega)}}{\lambda(N-1)}\right)^{g_{ij}(\omega)} \tag{2.39}$$

where $\bar{\omega})$ is such that $\mathsf{G}(\bar{\omega}) = 0$ and $\boldsymbol{a}(\bar{\omega}) = \boldsymbol{a}(\omega)$, i. e. the probability measure of a state with an empty network and all the internal states equal to those in ω.

THE FORM OF $\pi(\bar{\omega})$ Let us define $\boldsymbol{\alpha}$ the vector of the values α_i in equation (2.31). Let ω_0 be such that $(G)(\omega_0) = 0$ and $\boldsymbol{a}(\omega) = \boldsymbol{\alpha}$, that is the state in which the network has no edges and all internal states equal to the preferred ones. Let us define a vector $\boldsymbol{s}^{(i)}$ as

$$s_j^{(i)} = \begin{cases} \alpha_j & \text{if } j > i \\ a_j(\omega) & \text{if } j \le i \end{cases} \tag{2.40}$$

Clearly $\boldsymbol{\alpha} = \boldsymbol{s}^{(0)}$ and $\boldsymbol{a}(\bar{\omega}) = \boldsymbol{s}^{(N)} = \boldsymbol{a}(\omega)$. If we define the state ω_i as being the state with $\mathsf{G}(\omega_i) = 0$ and $\boldsymbol{a}(\omega_i) = \boldsymbol{s}^{(i)}$ The transition rate between ω_{i+1} and ω_i correspond to a *spin flip* of the i-th $a_i(\omega_{i+1})$ to α_i we have:

$$\mathsf{Q}_{\omega_i\omega_{i+1}} = \nu\frac{e^{h_i\delta_{a_i(\omega)\alpha_i}}}{q-1+e^{h_i}} \tag{2.41}$$

and thus applying recursively DB we get:

$$\pi(\omega_{i+1}) = \frac{\mathsf{Q}_{\omega_i\omega_{i+1}}}{\mathsf{Q}_{\omega_{i+1}\omega_i}} = e^{h(\delta_{a_i(\omega)\alpha_i}-1)}\pi(\omega_i) \tag{2.42}$$

The complete invariant measure becomes:

$$\pi(\omega) = \pi(\omega_0)e^{-\sum_i h_i}\prod_{i,j\in V, j<i} e^{\sum_i h_i\delta_{a_i(\omega)\alpha_i}}\left(\frac{2\eta\delta_{a_i(\omega)a_j(\omega}}{\lambda(N-1)}\right)^{g_{ij}(\omega)} \tag{2.43}$$

where $\pi(\omega)$ represents the probability for state characterized by an empty network and each of the internal states equal to the preferred state to occur. Nevertheless the quantity $\pi(\boldsymbol{\omega_0})e^{-\sum_i h_i}$ does not depend on ω and can thus be determined by normalization and then the invariant measure (2.43) can be written as:

$$\pi(\omega) = \frac{1}{\mathscr{Z}} \prod_{j<i} e^{\sum_i h_i \delta_{a_i(\omega)\alpha_i}} \left(\frac{2\eta \delta_{a_i(\omega)a_j(\omega)}}{\lambda(N-1)} \right)^{g_{ij}(\omega)} \tag{2.44}$$

where $\mathscr{Z}$ is the normalization constant (the partition function in the language of statistical physics).

A formal proof that the system satisfies detailed balance

To prove that our system satisfy detailed balance one should check that there are no unbalanced loops in the system. However, to explicitly carry on this calculation is cumbersome.

In the previous section we have derived the detailed balance invariant distribution. It should be clear we have indeed exploited all non zero transition rates to obtain it. Moreover π in equation (2.44) is positive, finite for any ω and normalisable, since $\sum_\omega \pi(\omega) \leq |\Omega| \pi(\omega^*)$ where ω^* is a state on which $\pi(\omega^*)$ is maximal (since the Ω is finite such state must exists). This proves, since our system admits a unique invariant distribution, that the system must satisfy detailed balance (and in particular all cycles must be balanced).

We can, redundantly, write a properly formalised proof, and the simplest way to do it, is to assume that one has *guessed* the form of the distribution (2.44) and *verify* that the distribution (2.44) satisfies detailed balance. The uninterested reader can jump directly to next section.

THEOREM 2.1. *The system admits a unique invariant distribution:*

$$\pi(\omega) = \frac{1}{\mathscr{Z}} \prod_{j<i} e^{\sum_i h_i \delta_{a_i(\omega)\alpha_i}} \left(\frac{2\eta \delta_{a_i(\omega)a_j(\omega)}}{\lambda(N-1)} \right)^{g_{ij}(\omega)} \tag{2.45}$$

that satisfies detailed balance. □

Proof. Since the states are finite, the system always have at least one invariant distribution $\boldsymbol{\pi}$ (see [87, p. 120] and [88, p. 37]). Moreover since $\mathscr{T}$ is transient, $\boldsymbol{\pi}|_{\mathscr{T}} = 0$. We have proven that the system is ergodic on $\mathscr{E}$, this implies (see [87, p. 118]) that $\boldsymbol{\pi}|_{\mathscr{E}} = 0$ is unique.

The function $\pi(\omega)$ written (2.45) is positive and finite on any state ω and since Ω is finite, normalisable; thus it defines a distribution on the state space.

To prove the thesis, since detailed balance implies stationarity, and we have already proven the stationary distribution in this system is unique, we only need to check that $\pi(\omega)$ satisfies detailed balance on all couple of states ω and ω' for which either $Q_{\omega\omega'} > 0$ or $Q_{\omega'\omega} > 0$ or both (if both $Q_{\omega\omega'} = 0$ and $Q_{\omega'\omega} = 0$ detailed balance is trivially satisfied).

There are three different situations to take into account:

(a-1) $\omega, \omega' \in \mathscr{T}$;

(a-2) $\omega \in \mathscr{T}, \omega' \in \mathscr{E}$ (or symmetrically $\omega' \in \mathscr{T}, \omega \in \mathscr{E}$) ;

(a-3) $\omega, \omega' \in \mathscr{E}$.

In case (a-1) detailed balance is trivially satisfied because $\pi(\omega) = 0$. In fact, since $\omega \in \mathscr{T}$, in the state ω $\exists i, j : g_{ij} = 1$ and $\delta_{a_i a_j} = 0$ thus at least one of the factor in eq (2.44) is zero.

In case (a-2) we have that $\mathsf{Q}_{\omega\omega'} > 0$ but $\mathsf{Q}_{\omega'\omega} = 0$ detailed balance is satisfied for any value of $\pi(\omega')$, because $\pi(\omega) = 0$ (since $\omega \in \mathscr{T}$).

In case (a-3), there are only two classes of states that for which either $\mathsf{Q}_{\omega\omega'} > 0$ or $\mathsf{Q}_{\omega'\omega} > 0$:

(b-1) ω and ω' differs at most by an edge that is $\exists a, b \in V \times V : g_{ab}(\omega) = 1 \wedge g_{ab}(\omega') = 0$ in ω', and $\forall (i, j) \in V \times V : (i, j) \notin \{(a, b), (b, a)\}, g_{ij}(\omega) = g_{ij}(\omega')$ and $\boldsymbol{a}(\omega) = \boldsymbol{a}(\omega')$: ω';

(b-2) ω and ω' differ only by the value of the internal state of one node, that is $\forall i, j \in V \times V \; g_{ij}(\omega) = g_{ij}(\omega')$ and $\exists n \in V, a_n(\omega) \neq a(\omega')$ and $\forall i \in V : i \neq n a_i(\omega) = a_i(\omega')$.

In case (b-1) we have that $\pi(\omega) = \frac{2\eta}{\lambda(N-1)}\pi(\omega')$ thus

$$\pi(\omega)\mathsf{Q}_{\omega\omega'} = \frac{2\eta}{\lambda(N-1)}\pi(\omega') \cdot \lambda = \frac{2\eta}{(N-1)}\pi(\omega') = \mathsf{Q}_{\omega'\omega}\pi(\omega'). \tag{2.46}$$

In case (b-2) we have that $\pi(\omega) = \mathrm{e}^{h_n(\delta_{\alpha_n a_n(\omega)} - \delta_{\alpha_n a_n(\omega')})}\pi(\omega')$ and thus

$$\pi(\omega)\mathsf{Q}_{\omega\omega'} = \mathrm{e}^{h_n(\delta_{\alpha_n a_n(\omega)} - \delta_{\alpha_n a_n(\omega')})}\pi(\omega')\mathrm{e}^{h_n \delta_{\alpha_n a_n(\omega')}} = \mathrm{e}^{h_n \delta_{\alpha_n a_n(\omega)}}\pi(\omega') = \pi(\omega')\mathsf{Q}_{\omega'\omega}. \tag{2.47}$$

The theorem is proven. □

2.7.3 *Describing the system in terms of populations*

In the previous sections we derived the form of the stationary distribution. However we need to express it in another form to be able to treat it analytically.

Moreover, in the perspective of applying it to the description of migration in a group of animal, we would like to be able to obtain information about migrating population, since, in principle, they are the measurable quantities (the network structure itself, is still hard to deduce, even though there have been attempts, in some context to use tagging methods to experimentally study network dynamics [89]).

As discussed in the previous section, there is a natural sub-population structure define by how able the individuals are to gather information (the h_i parameter). A description of the system in terms of these sub-population instead that in term of network is a natural step. Moreover, such a description allows us to take the *thermodynamic limit* more easily.

The distribution of node and edges population

To proceed further we make some simplification. In particular we assume that $h_i = h_{\alpha_i}$ for the informed individuals and that $h_i = 0$ for the uninformed ones. The rationale for this choice has been already discussed in section 2.6.2.

With this choice it is possible to naturally partition the entire population $N = |V|$ in sub-populations. In particular let us define

$$\hat{N}_a^\alpha(\omega) = \sum_{i \in V} \delta_{a_i(\omega)a}\delta_{\alpha_i\alpha} \tag{2.48}$$

i. e. the number of nodes that have an internal state a while they would prefer it to be α, and

$$\hat{N}_a^0 = \sum_{i \in V} \delta_{a_i(\omega)a}\delta_{h_i 0} \tag{2.49}$$

the population of uninformed individuals that have an internal state α.

Since h_i and α_i are parameters of the system, thus fixed at the beginning this partition must satisfy the following constraints:

$$\sum_a \hat{N}_a^\alpha(\omega) = N^\alpha, \tag{2.50}$$

where N^α is the number of individuals in the whole population that have information about directionα

$$\sum_a \hat{N}_a^0(\omega) = N^0 \tag{2.51}$$

where N^0 is the number of individuals that have no information. Clearly:

$$N^0 + \sum_{\alpha \in S} N^\alpha = N \tag{2.52}$$

For sake of simplification we shall use the following notation: $\hat{N}_a(\omega) = \hat{N}_a^0(\omega) + \sum_\alpha \hat{N}_a^\alpha(\omega)$ which denotes the number of nodes in state which have and internal state a.

Let also define a partition of the edges of the system, in particular let

$$M_{ab} = \frac{1}{1+\delta_{ab}} \sum_{(i,j) \in V \times V} g_{ij}(\omega)\delta_{a_i(\omega)a}\delta_{a_j(\omega)b} \tag{2.53}$$

be the number of edges between nodes of type a and b. Clearly, $\hat{M}_{ab}(\omega) = \hat{M}_{ba}(\omega)$ due to the symmetry of G.

The stochastic dynamics of our system defines an induced process on the sub populations $\hat{N}_{a,t}^\alpha = hat N_a^\alpha(\Omega_t)$ and $\hat{M}_{ab,t} = \hat{M}_{ab}(\Omega_t)$ which can be studied using the knowledge we have on process Ω_t.

A state of this induced process is characterised by a sequence of positive integers $\{N_a^\alpha\}_{a \in S, \alpha \in S \cup \{0\}}$ satisfying $\sum_a N_a^\alpha = N^\alpha$ and a sequence of positive integers $\{M_{ab}\}_{a \in S, b \in S, b \geq a}$ such that $M_{ab} \leq N_a N_b$ if, $a \neq b$ and $M_{aa} \leq \frac{N_a(N_a-1)}{2}$. We can define the probabilities for the induced process easily; but since we have derived the invariant distribution $\pi(\omega)$ in eq.(2.45) we shall restrict to that case.

If $\omega \in \mathscr{E}$, it is easy to check, using, the definition, that $\forall\,(a,b) \in V \times V \;:\; a \neq b, \quad \hat{M}_{ab}(\omega) = 0$; since $\pi(\omega) = 0$, if $\omega \in \mathscr{T}$. Without loss of generality we shall restrict ourselves to $\mathscr{E}$ and thus we shall assume that al non diagonal M_{ab} are zero. We shall use the short-cut notation $\boldsymbol{N}$ to refer to all N_a^α and $\boldsymbol{M}$ to refer to all diagonal elements M_{aa} . The induced stationary distribution reads:

$$\pi(\boldsymbol{N}) = \sum_{\omega\in\mathscr{E}} \pi(\omega) \prod_{a\in S} \delta(\hat{N}_a^0(\omega) - N_a^0)\delta(\hat{M}_{aa}(\omega) - M_{aa}) \prod_{\alpha\in S} \delta(\hat{N}_a^0(\omega) - N_a^0). \quad (2.54)$$

To calculate it let us define a subspace of $\mathscr{E}$ as follow:

$$\begin{aligned}\Omega(\boldsymbol{N},\boldsymbol{M}) = \{\omega \in \mathscr{E} \;:\; \forall a \in S, \alpha \in S \cup \{0\},\ (\hat{N}_a^\alpha(\omega) = N_a^\alpha) \\ \wedge\, (\hat{M}_{aa}(\omega) = M_{aa})\}\end{aligned} \quad (2.55)$$

It is easy to check that $\pi(\omega)$ is the same for all states in $\Omega(\boldsymbol{N},\boldsymbol{M})$; to see it, we note that each link carries a weight of $\frac{2\eta}{\lambda(N-1)}$ and thus, since all states in $\Omega(\boldsymbol{N},\boldsymbol{M})$ have the same edge population they must all have the same network contribution:

$$\prod_{a\in S} \left(\frac{2\eta}{\lambda(N-1)}\right)^{M_{aa}} \quad (2.56)$$

Moreover, noting that:

$$\begin{aligned}\sum_i h_i \delta_{a_i(\omega)\alpha_i} &= \sum_j h_\alpha \sum_{a\alpha} \delta_{a_i(\omega)a}\delta_{a\alpha}\delta_{\alpha\alpha_i} = \sum_{a\alpha} h_\alpha \delta_{a\alpha} \sum_i \delta_{a_i(\omega)a}\delta_{\alpha\alpha_i} \\ &= \sum_{a,\alpha} h_\alpha \delta_{a\alpha} N_a^\alpha(\omega)\end{aligned} \quad (2.57)$$

we can directly rewrite that weight as :

$$\pi(\boldsymbol{N},\boldsymbol{M}) = \frac{|\Omega(\boldsymbol{N},\boldsymbol{M})|}{\mathscr{Z}} \mathrm{e}^{\sum_a h_a N_a^a} \prod_{a\in S} \left(\frac{2\eta}{\lambda(N-1)}\right)^{M_{aa}}. \quad (2.58)$$

It is easy also easy to calculate the number of states in $\Omega(\boldsymbol{N},\boldsymbol{M})$; for a given distribution of a_i (and thus N_a nodes in state a) the number of possible different choices for the network with M_{aa} edges $a-a$ corresponds to the ways that one can extract M_{aa} edges out of the maximal possible number of $a-a$ edges $\frac{N_a(N_a-1)}{2}$ which is $\binom{\frac{N_a(N_a-1)}{2}}{M_{aa}}$; the possible choices of distributions of a_i corresponding to a given distribution $\boldsymbol{N}$ are also easy to obtain and correspond to: $\frac{N!}{\prod_a N_a^0! \prod_{a\alpha} N_a^\alpha!}$. Thus we have:

$$|\Omega(\boldsymbol{N},\boldsymbol{M})| = \frac{N!}{\prod_a N_a^0! \prod_{a\alpha} N_a^\alpha!} \prod_{a\in S} \binom{\frac{N_a(N_a-1)}{2}}{M_{aa}} \quad (2.59)$$

and thus the population stationary distribution reads:

$$\pi(\boldsymbol{N},\boldsymbol{M}) = \frac{1}{\mathscr{Z}} \frac{N!}{\prod_a N_a^0! \prod_{a\alpha} N_a^\alpha!} \mathrm{e}^{\sum_a h_a N_a^a} \prod_{a\in S} \binom{\frac{N_a(N_a-1)}{2}}{M_{aa}} \left(\frac{2\eta}{\lambda(N-1)}\right)^{M_{aa}}. \quad (2.60)$$

The node population distribution

We can sum in eq.(2.60) over $\boldsymbol{M}$ and obtain the stationary node population distribution, $\pi(\boldsymbol{N})$, that we shall call simply *population distribution*:

$$\pi(\boldsymbol{N}) = \frac{1}{\mathscr{W}} \frac{N!}{\prod_a N_a^0! \prod_{a\alpha} N_a^\alpha!} \mathrm{e}^{\sum h_\alpha N_a^\alpha \delta_{a\alpha}} \prod_a \left[1 + \frac{2\eta}{\lambda(N-1)}\right]^{\frac{N_a}{2}(N_a-1)} \tag{2.61}$$

Where $\mathscr{W}$ is the normalization constant.

The measure is defined over the multisymplex defined by:

$$0 \leq N_a^\alpha \leq N^\alpha \tag{2.62}$$

and

$$\sum N_a^\alpha = N^\alpha \tag{2.63}$$

2.7.4 *The free energy of the system and the thermodynamic limit*

The $\pi(\boldsymbol{N}, \boldsymbol{M})$ and $\pi(\boldsymbol{N})$ describe the system at finite N; however since we are interested in properties of migration of *large* groups we can use the thermodynamic description of this system (i. e. the $N \to \infty$ limit).

Let us denote the densities with $n_a^\alpha = \frac{N_a^\alpha}{N}$ and consequently $n_a = \sum_\alpha n_a^\alpha$ and $n^\alpha = \sum_a n_a^\alpha$. We want, first, to address the problem whether there is a function $\mathrm{F}(\boldsymbol{n}; \frac{\eta}{\lambda}, h)$, that plays the role played by the *free energy* in usual statistical physics, such that:

$$\pi(N) = \frac{1}{\mathscr{Z}} \mathrm{e}^{-N\left[\mathrm{F}(\boldsymbol{n}; \frac{\eta}{\lambda}, h) + O(1/N)\right]}. \tag{2.64}$$

To do so, we expand $\pi(\boldsymbol{N})$ in a large N limit, using Stirling approximation($\log(N!) \simeq N\log(N) - N$, $\log(1+x) \simeq x$ for small x and normalization :

$$\sum_a n_a^\alpha = n^\alpha. \tag{2.65}$$

We then obtain :

$$\begin{aligned} \log(\pi(\boldsymbol{N})) &= N\log(N) - N - \sum_{a\alpha} (N n_a^\alpha \log(N) + N n_a^\alpha \log(n_a^\alpha) - N n_a^\alpha) \\ &\quad + N \sum_{a\alpha} h_\alpha \delta_{a\alpha} n_a^\alpha + N \frac{\eta}{\lambda} \sum_a \left(\frac{N}{N-1} (n_a)^2 + \frac{N}{N-1} \frac{1}{N} n_a \right) \\ &= -N \left[\sum_{a\alpha} (n_a^\alpha \log(n_a^\alpha) - \sum_{a\alpha} h_\alpha n_a^\alpha \delta_{a\alpha} - \frac{\eta}{\lambda} \sum_a (n_a)^2 + O(1/N) \right] \\ &= F(\boldsymbol{n}, \boldsymbol{h}) \end{aligned} \tag{2.66}$$

Thus in the thermodynamic limit the invariant measure will converge to a sum of Dirac delta measures centred in the global minimum (or minima) of the free energy (because of normalization all other minima will be exponentially depressed).

2.7.5 *Edge distribution and average degree*

In order to obtain the edge distribution, one has to sum over $\boldsymbol{N}$ in equation (2.60), but the calculation is not easy. However, on can take advantage of the fact that the probability measure that describes the system is concentrated on its minima for simplifying the joint probability (2.60).

In fact, when the system is in one of these minima $\boldsymbol{N}^*$, we can consider the node population variables to be *frozen* around $\boldsymbol{N}^*$ and thus instead of the joint probability distribution (2.60) one can consider the conditional probability $\pi(\boldsymbol{M}|\boldsymbol{N})$ to calculate the relevant quantities.

The conditional probability may be calculated as follow:

$$\pi(\boldsymbol{M}|\boldsymbol{N}^*) = \frac{\pi(,\boldsymbol{N}^*,\boldsymbol{M})}{\sum_{\boldsymbol{M}}\pi(,\boldsymbol{N}^*,\boldsymbol{M})} = \frac{\prod_{\alpha=1}^{q}\binom{\frac{N_a^*}{2}(N_a^*-1)}{M_{aa}}\left(\frac{2\eta}{\lambda(N-1)}\right)^{M_{aa}}}{\prod_{a=1}^{q}\left[1+\frac{2\eta}{\lambda(N-1)}\right]^{\frac{N_a^*}{2}(N_a^*-1)}} \tag{2.67}$$

2.7.6 *Average Degree*

The first and most important quantity that can be calculated from the degree distribution is the average degree.

By definition the average degree is equal to $\langle k\rangle = \frac{2\langle\sum_s M_{aa}\rangle}{N}$ where the numerator is the average number of edges.

We can easily calculate :

$$\begin{aligned}\langle M_{aa}\,|\boldsymbol{N}^*\rangle &= \frac{2}{N_a^*}\sum_{\boldsymbol{M}} M_{aa}\pi(\boldsymbol{M}|\boldsymbol{N}^*) = \frac{\sum_{M_{aa}} M_{aa}\binom{\frac{N_a^*}{2}(N_a^*-1)}{M_{aa}}\left[\frac{2\eta}{\lambda(N-1)}\right]^{M_{aa}}}{\left[1+\frac{2\eta}{\lambda(N-1)}\right]^{\frac{N_a^*}{2}(N_a^*-1)}} \\ &= \frac{\eta}{\lambda(N-1)}N_a^*(N_a^*-1)\frac{1}{1-\frac{2\eta}{\lambda(N-1)}}\end{aligned} \tag{2.68}$$

The total average degree reads:

$$\begin{aligned}\langle k\rangle \simeq \langle k\mid \boldsymbol{N}^*\rangle = \frac{2\langle\sum_a M_{aa}\rangle}{N} = \frac{2}{N}\sum_a \frac{\eta}{\lambda(N-1)}N_a^*(N_a^*-1)\frac{1}{1-\frac{\eta}{\lambda(N-1)}} = \\ \frac{2\eta}{\lambda}\sum_a (n_a^*)^2 + O(1/N) \simeq z\sum_a (n_a^*)^2\end{aligned} \tag{2.69}$$

As we immediately note, for a given value of z due to normalisation constraints we have that $\frac{z}{q} \le \langle k\rangle \le z$. It is clear that, in finite systems, the previous result is true only for $z \ll N$: the degree of a node in a finite graph, in fact, cannot exceed $N-1$; when doing simulations with finite N, one should expect to see, as z approaches N relevant finite size effects: i. e. the average degree reaches a plateau.

2.8 A SOLVABLE ADAPTIVE NETWORK MODEL FOR MIGRATION III: STABLE STATES AND PHASE DIAGRAM

As we mentioned in the previous sections, the invariant distribution is peaked around the minima of the free energy, thus, in order to analyse the system we proceed to minimise the Free Energy of the system:

$$\mathrm{F}(\boldsymbol{n};\frac{\eta}{\lambda},\boldsymbol{h}) = \sum_a n_a^0 \log n_a^0 + \sum_{a\alpha} n_a^\alpha \log(n_a^\alpha) - \sum_{a\alpha} h_\alpha n_a^\alpha \delta_{\alpha a} - \frac{\eta}{\lambda}\sum_a (n_a)^2 \quad (2.70)$$

subject to the constraints:

$$n^\alpha = \sum_a n_a^\alpha \quad (2.71)$$

and

$$n^0 = \sum_a n_a^0 \quad (2.72)$$

We introduce then Lagrange multiplier $\beta_\alpha - 1$ for the first constraints and $\beta_0 - 1$ the second one and impose the First Order Condition (FOC):

$$\nabla\left\{\mathrm{F}(\boldsymbol{n};\frac{\eta}{\lambda},\boldsymbol{h}) - \sum_\alpha(\beta_\alpha - 1)\left[-n^\alpha + \sum_a n_a^\alpha\right] - (\beta_0 - 1)(n^0 - \sum_a n_a^0)\right\} = 0 \quad (2.73)$$

obtaining

$$\log(n_a^0) - \frac{2\eta}{\lambda}n_a - \beta_0 = 0 \quad (2.74)$$

and

$$\log(n_a^\alpha) - h_\alpha\delta_{a\alpha} - \frac{2\eta}{\lambda}n_a - \beta_\alpha = 0 \quad (2.75)$$

and thus

$$\begin{cases} n_a^0 = \mathrm{e}^{\frac{2\eta}{\lambda}n_a}\mathrm{e}^{\beta_0} \\ n_a^\alpha = \mathrm{e}^{h_\alpha\delta_{a\alpha}+\frac{2\eta}{\lambda}n_a}\mathrm{e}^{\beta_\alpha} \end{cases} \quad (2.76)$$

If we define :

$$Q = \sum_a \mathrm{e}^{zn_a} \quad (2.77)$$

we can write

$$n^0 = \mathrm{e}^{\beta_0}Q \quad (2.78)$$

and

$$n^\alpha = \mathrm{e}^{\beta_\alpha}\mathrm{e}^{\frac{2\eta}{\lambda}n_\alpha}(\mathrm{e}^{h_\alpha} - 1) + \mathrm{e}^{\beta_\alpha}Q \quad (2.79)$$

Equation (2.78) allows us to eliminate one Lagrangian multiplier:

$$e^{\beta_0} = \frac{n^0}{Q} \tag{2.80}$$

Equation (2.79) instead allows us to eliminate the quantity

$$e^{\beta_\alpha} = \frac{n^\alpha}{(e^{h_\alpha} - 1)e^{\frac{2\eta}{\lambda} n_\alpha} + Q} \tag{2.81}$$

We then have also the normalization constraint:

$$\sum_{i>0} n^i = 1 - n^0. \tag{2.82}$$

We can then write the FOC for our system in the following way

$$\boxed{n_a = e^{\frac{2\eta}{\lambda} n_a} \left[\frac{n^0}{Q} + W + \frac{(e^h - 1)n^a}{(e^h - 1)e^{\frac{2\eta}{\lambda} n_a} + Q} \right]} \tag{2.83}$$

where

$$Q = \sum_{i=1}^{q} e^{\frac{2\eta}{\lambda} n_i} \tag{2.84}$$

and

$$W = \sum_{i=1}^{q} \frac{n^i}{(e^{h_i} - 1)e^{\frac{2\eta}{\lambda} n_i} + Q} \tag{2.85}$$

2.8.1 *Stable states without information: collective movement*

When no information is available in the group, the system reduces to an adaptive network model in which group coordination only depends on the rates at which links are created or destroyed [82, 83].

In this case, only n_a^0 are present in the FOC (2.73) and thus $n_a = n_a^0$. The Free Energy now is completely symmetric with respect to the change of colour $\mathbb{Z}_q$ symmetry.

Solution of First Order condizion

To simplify the notation, we call $z = \frac{2\eta}{\lambda}$ and $x_a = z n_a$, the FOC can be put in the form:

$$x_a e^{-x_a} = \frac{z}{Q}, \tag{2.86}$$

and the normalisation constraint becomes:

$$\sum_a x_a = z. \tag{2.87}$$

Equation (2.86) has to possible solutions in term of the W Lambert function (see appendix B.1 for the details).

$$x_- = -\mathrm{W}_0(-zQ^{-1}) \tag{2.88}$$

and

$$x_+ = -\mathrm{W}_{-1}(-zQ^{-1}). \tag{2.89}$$

It is easy to show that $x_- < 1$ and $x_+ > 1$ and Q is a parameter to be determined auto-consistently using normalisation constraint.

Since all x_a satisfy the same equation, a general solution of the first order equation shall be obtained by specifying the number L_+ colours that are in the + solution and the number L_- of colours that are in − solution. Of course $L_+ + L_- = q$ The normalisation will enforce:

$$L_+ x_+ + L_- x_- = z \tag{2.90}$$

and thus, using normalisation, we can eliminate the Q constant from equations and reduce the FOC to:

$$x_- \mathrm{e}^{-x_-} = \frac{z - L_- x_-}{q - L_-} \mathrm{e}^{-\frac{z - L_- x_-}{q - L_-}}. \tag{2.91}$$

According to the value of z this equation shall have one or more solutions.

It is easy to check that the solution where all $x_a = \frac{z}{q}$ always exists (and corresponds to $L_- = q$ for $z < q$ and to $L_+ = q$ otherwise). It is possible to prove (see [82, 83] for the complete details) that it is a minimum of the free energy only for $z < q$.

It is possible to prove that the the solution with $L_+ = 1$ is the only other minimum of the free energy (see [82, 83] for the complete details).

The average degree of the system can be obtained from the population distribution using equation (2.69). Another quantity which is:

$$\sigma = \frac{q \sum_{i \in S} n_i - 1}{q - 1} \tag{2.92}$$

which we call *school efficiency*. It measures how polarised is the population toward one direction and, thus, how effectively the group is able to reach consensus: if the group is uncoordinated than $\sigma = 0$, if all individuals share the same internal state $\sigma = 1$. Because of (2.69) these two quantities are linked $\langle k \rangle = z\sigma + (1 - \sigma)\frac{z}{q}$. Such kind of relations are the hallmark of adaptive networks and derive directly from state topology co-evolution typical of these systems. This also clarifies the characteristics of this system. Low or non extant coordination is linked to a network structure with a low density and no giant component. High σ,if $z > 1$, is linked to the emergence of a giant component, and a higher network density.

The behaviour of the system in absence of information may be summarised as follow. Below a certain threshold $\check{z}$ only one local minimum exists which corresponds to a symmetric solution $\sigma = 0$ (Figure 2.3 a); there the network has a low density a no giant component with $\langle k \rangle < 1$. In this case the group does

not reaches consensus (in the context of migration problem, the group does not migrate). At $\check{z}$, a new bundle of q local minima appears with $\sigma > 0$. There the network has a higher density, $\langle k \rangle > 1$, and a *giant component* emerges. In this case a large portion of the population succeeds in reaching consensus (in the context of migration problem, the group start migrating).

Between $\check{z}$ and $\hat{z}$ both solutions coexist and individuals can coordinate with one another or not. Above $\hat{z}$ the only local minima are for $\sigma > 0$ while the low density solution $\sigma = 0$ becomes unstable. There is an intermediate point z^* below which global minimum of the free energy is the low density solution, and above which the system has q global minima corresponding to high density solutions. At this point, in the thermodynamic limit, the system undergoes a first order transition, whereas for finite N there system is bistable between $[\check{z}, \hat{z}]$.

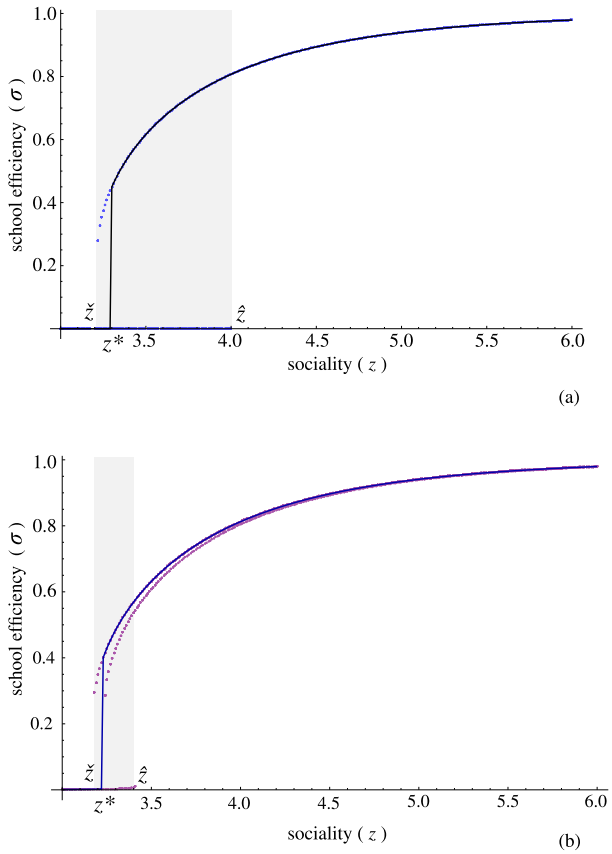


Figure 2.3: Critical group dynamic: school efficiency, σ as function of the social parameter z in (a) non informed group $n^1 = 0$ and (b) informed group $n^1 = 0.05$, $h = 0.5$. The dotted lines correspond to all the stable solutions of (**??**), the shadowed areas identify the coexistence region whereas the solid lines corresponds the equilibrium solution.

2.8.2 *Stable state with information with $n^1 = 1 - n^0$: migration without conflict of interests*

The first generalisation is to study the problem of consensus decision making without conflict of interest. We study the case with q destinations, a density of informed individual $n^1 = 1 - n^0$ and a preference h about destination 1. The equation (2.73) again can be solved numerically to obtain prediction on collective behaviour.

In this case equation (2.83) take the simplified form:

$$\begin{cases} x\mathrm{e}^{-x} = z\frac{n^0}{Q} + z\mathrm{e}^{h}\frac{1-n^0}{(\mathrm{e}^{h}-1)\mathrm{e}^{x}+Q} \\ y_i\mathrm{e}^{-y_i} = zP = z\frac{n^0}{Q} + z\frac{1-n^0}{(\mathrm{e}^{h}-1)\mathrm{e}^{x}+Q} \quad \text{q-1 times} \end{cases} \tag{2.93}$$

with the conditions

$$x + \sum_{i=0}^{q-1} y_i = z \tag{2.94}$$

$$Q = \mathrm{e}^{x} + \sum_{i=0}^{q-1} \mathrm{e}_i^{y} \tag{2.95}$$

and

$$P = \frac{n^0}{Q} + W = \frac{n^0}{Q} + \frac{1-n^0}{(\mathrm{e}^{h}-1)\mathrm{e}^{x}+Q}. \tag{2.96}$$

Numerical solution of this system using naively the standard methods is problematic; however it is possible to reduce the system to a one dimensional equation which can be solved numerically (the details of the this calculation can be found in appendix B.2). The stability as well cannot be addressed exactly as in the case in with no information, however it is possible to check the stability of the numerical solutions using standard methods (see appendix B.2).

Phase diagram

Information has two main effects on the system (Figure 2.3 b). Firstly, it breaks the symmetry between the q high density stable solutions found in the $n_0 = 1$ case, by selecting the solution with the preferred destination $\alpha = 1$ as the most likely (the global minimum of the free energy). The $q-1$ solutions corresponding to migration toward other destinations *remain stable*, but are much less likely to be selected by the population. In the thermodynamic limit, only the solution with $\alpha = 1$ will occur.

Secondly, the coexistence region between high and low density solutions $[\check{z}, \hat{z}]$ is reduced in the case of informed migration (Figure 2.3). In fact this region becomes smaller as the number of informed individuals increases (Figure 2.4).

Eventually, there exists a critical value of n_c^1 at which the region collapses into a point and the phase transition becomes of second order. Above that critical

value the system has a smooth crossover between low and high density states, as z increases, and a single solution is found. We can, thus, plot an exact phase diagram for this region for the two parameter z and n^1 2.4. As one notice the shape of the coexistence region varies with h.

The edges of coexistence region

Even though solving equation (B.21) allows us to calculate all the solutions of the FOC and obtain an exact phase diagram„ this process can be quite computationally intense (especially as q increases). We can however attempt to calculate the boundaries of the coexistence region in more effective way.

Let specialise to the case when all $y_i = y$, since we know from solution of (B.5) that these are the solution relevant to the description of the system. If we use the same notation used in section 2.8 and me do some algebraic manipulation on equation (B.5) we can reduce the system to the solution of two equations:

$$z = x + (q-1)y \tag{2.97}$$

$$n^1 = \frac{\left(e^h + (q-1)e^{-x+y} - q\right)\left(x - e^{-x+y}y\right)}{\left(-1 + e^h\right)\left(x + (-1+q)y\right)}. \tag{2.98}$$

The edge of the coexistence region corresponds to the birth (or death) of two new branches of solutions. In particular for the lower edge $\check{z}$ corresponds to the birth of two high density branches $x_+(z, n^1, h)$ (stable) and $\hat{x}_+(z, n^1, h)$ (unstable) such that $x_+(\check{z}, n^1, h) = \hat{x}_+(\check{z}, n^1, h)$ whereas the high edge of the coexistence region corresponds to the death of the stable low density branch $x_+(\check{z}, n^1, h)$ and of the unstable high density branch$\hat{x}_+(\check{z}, n^1, h)$.

However we now think to z in equation (2.97) as a dependent variable $z = z(x, y(x)) = z(x)$ where $y(x)$ is implicitly defined by equation (2.98) ; it is easy to

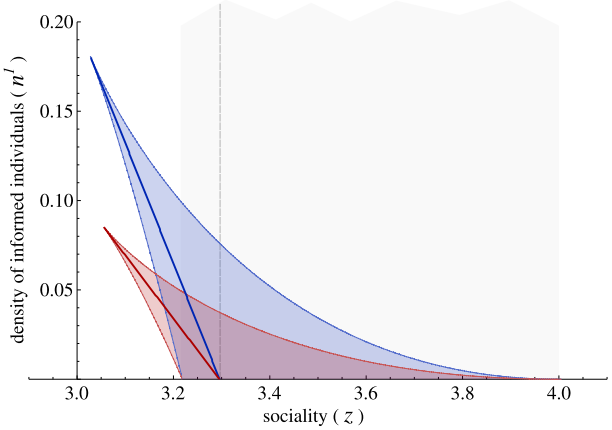


Figure 2.4: Phase diagram of the system with $q = 4$ possible directions. The grey area corresponds to a preference parameter $h = 0$ (no preferences) and the dashed line is the critical line. The blue area corresponds to $h = 0.5$ and the thick blue line represents the corresponding critical line. The red area corresponds to $h = 1$ and the thick red line represents the corresponding critical line as well.

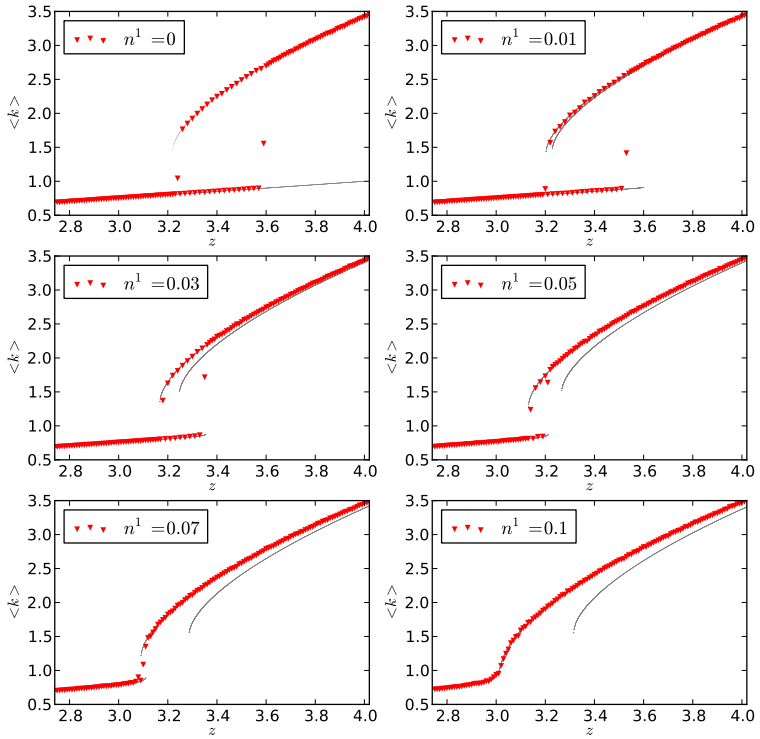


Figure 2.5: In this figure we check our theoretical result with a simulation using Gillespie Algorithm [90]. The red triangle are the simulated data. Simulations are done with $q = 4$, $N = 5000$, $\nu = 30$ and $\lambda = 10$. We first let the system thermalise, at $z = 2.75$, and then increase z by 0.01, take data for$\Delta T = 250$, and repeat until we $z = 4.01$. At this point we decrease z by 0.01 and repeat the procedure until we return to $z = 2.75$.

understand that $\check{z}$ and $\hat{z}$ correspond to a stationary points of $z(x)$ and we can take advantage of this intuition to find them.[4]

We thus impose the first order condition for z

$$\frac{\mathrm{d}}{\mathrm{d}x}z = 1 + (q-1)\frac{\mathrm{d}}{\mathrm{d}x}y = 1 - (q-1)\frac{\partial_y \hat{n}^1(x,y)}{\partial_x \hat{n}^1(x,y)} = 0 \tag{2.99}$$

where $\hat{n}^1(x,y)$ is defined as the right side and y is such that $\hat{n}^1(x,y) - n^1 = 0$.

Equation (2.99) does not depend explicitly on n^1 and z (but it depends on h); To solve this equation we impose normalisation $y = \frac{z-x}{q-1}$. If for some z equation (2.99) admits a set of solutions labelled by z ($x(z)$, $\frac{z-x(z)}{q-1}$), by construction, must lie on the edge for some value of n^1. Since we have assumed $\hat{n}^1(x,y) - n^1 = 0$ we easily derive that $n^1 = \hat{n}^1(x(z), \frac{z-x(z)}{q-1}) = n^1(z)$, which defines parametrically the edges of the coexistence region $(z, n^1(z))$ (cf. figure 2.6).

4 To visualise it better one can think to $y = \sqrt{x}$. The point $x = 0, y = 0$ corresponds to the birth of the two branches of the square root function, but it can also be seen as the minimum of the function $x = y^2$

The critical point

We can expand further the previous reasoning to get the critical point, that is the point in which the transition becomes of second order, it is easy to visualise that this happens when $\frac{\mathrm{d}^2}{\mathrm{d}x^2}z = 0$. In our case this correspond to the condition where the unstable high density branch $\hat{x}_+$ vanishes and thus the two stable branches $x_+ = x_-$ join.

In addition to equations (2.97), (2.98) and (2.99), one gets an additional equation form the requirement that the second derivative of z vanishes, i.e..

$$\frac{1}{q-1}\frac{\mathrm{d}^2}{\mathrm{d}x^2}z = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}y = \frac{\partial_x^2\hat{n}^1(x,y) + 2\partial_{xy}^2\hat{n}^1(x,y)\frac{\mathrm{d}}{\mathrm{d}x}y + \partial_y^2\hat{n}^1(x,y)\left(\frac{\mathrm{d}}{\mathrm{d}x}y\right)}{\partial_y\hat{n}^1(x,y)} = 0. \tag{2.100}$$

We can use (2.99) and set $\frac{\mathrm{d}}{\mathrm{d}x}y = -\frac{1}{q-1}$ in equation (2.100). Solving (2.99) and substituting in equation (2.100) we obtain an equation that does not depend neither on h nor on n_1. As before, we can use normalisation and solve; for the value of z for which equation (2.99) admits, a solution, and curve parametrized by z $(z, n^1(z), h(z))$ (cf. figure 2.6).

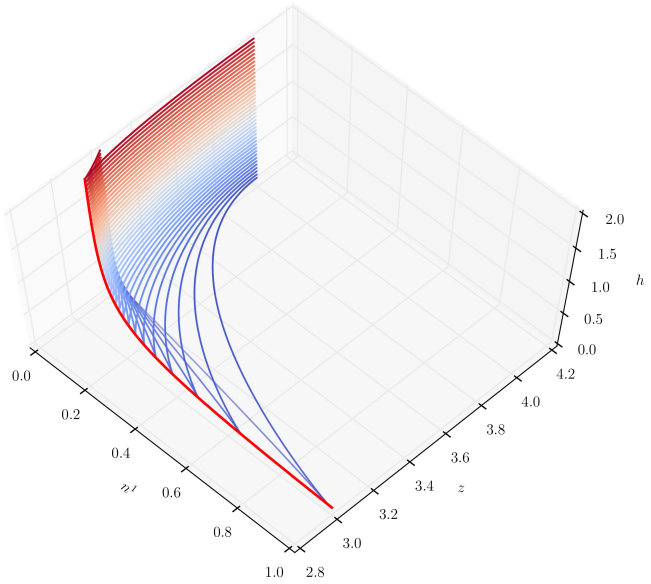


Figure 2.6: 3 dimensional phase diagram obtained solving numerically equations (2.99), (2.98) and (2.98), for $z \in [2.9, 4]$ with a minimal step $\Delta z = 0.0001$; $h \in [0.1, 2]$ and a resolution $\Delta h = 0.05$ (higher resolution are possible but the resulting graph becomes difficult to handle using standard methods.). The red line is the line representing II order critical points line numerically obtained solving, in addition to the previous equations, also equation (2.100), in this case h is parametrically determines, whereas we have used $\Delta z = 0.00001$.

Such curve, for $q = 4$, starts at $h_c \simeq 0.0986122$ and $n^1 \simeq 1$; this means that the if the strength of preference is low, the system will always have a coexistence region. On the other limit, the curve approaches asymptotically $z \simeq 3.0986122$; the value of n^1 from the numerical solution seams to decreases slowly to 0 when $h \to \infty$.

Hysteretic cycles

From the tridimensional phase diagram in figure 2.6 we get another important feature of these system: when z is below the value at which the system has a first order transition when $n^0 = 1$, the system may show a n^1-h hysteretic cycle.

The behaviour of the solution as the parameters h and $n^1 = 1 - n_0$ vary, at fixed z, is depicted in Figure 2.7. For low values of z (Figure 2.7a) we observe a smooth crossover from low to high density solutions as h and/or n_1 increase whereas when z is in a critical region 2.7b). The presence of a sharp transition with coexistence in a broad range of parameters is a robust feature of this model.

This provides us with a vivid picture of how we expect the collective behaviour of the population to change when the parameters z, h and n_0 change. Adapting this picture to the observed behaviour of populations (e. g. of migrating fish) provides hints on the likely underlying causal effects. In brief, when z is large, i.e. for individuals with a marked pro-social behaviour, we expect abrupt transitions when either the density n^α of individuals with a given preference, or the intensity h_α of that preference varies in such a way as to cross the boundaries in the phase diagram (Figure 2.4).

When both preference and information are decreased in the informed migration problem, abrupt transition from efficient group formation to collapse of migration efficiency are visible. We note that this hysteresis cycle is consistent with observed stock collapses of migratory fish populations [42]. When the migratory population is described using a social parameter z close to the critical point, then the interplay between the memory for a given destination, h, and the fraction of the individuals informed, n^1 about this destination can produce an abrupt transition in the migration of the species.

In the case of a school migrating in direction 1, a decrease of the value of h and n^1 over years e.g. due for example to overfishing of both individuals and prey in the migration site, can force the system to cross the critical line reaching eventually low values of both h and n^1. When in this condition, an increase in the value of h might occur due for example to better habitat conditions or food availability, for those few vagrant fish that might still be present in the area. However this increase alone cannot bring the system back to the original state because the system may not cross again the critical line. Thus the group may not migrate in direction 1 even though previous habitat conditions are re-established.

2.8.3 *Stable state with information and conflict of interests*

Conflict of interests, in our model, is characterised by the presence of more than one group of informed individuals: the simplest case being, a fraction n^1 of

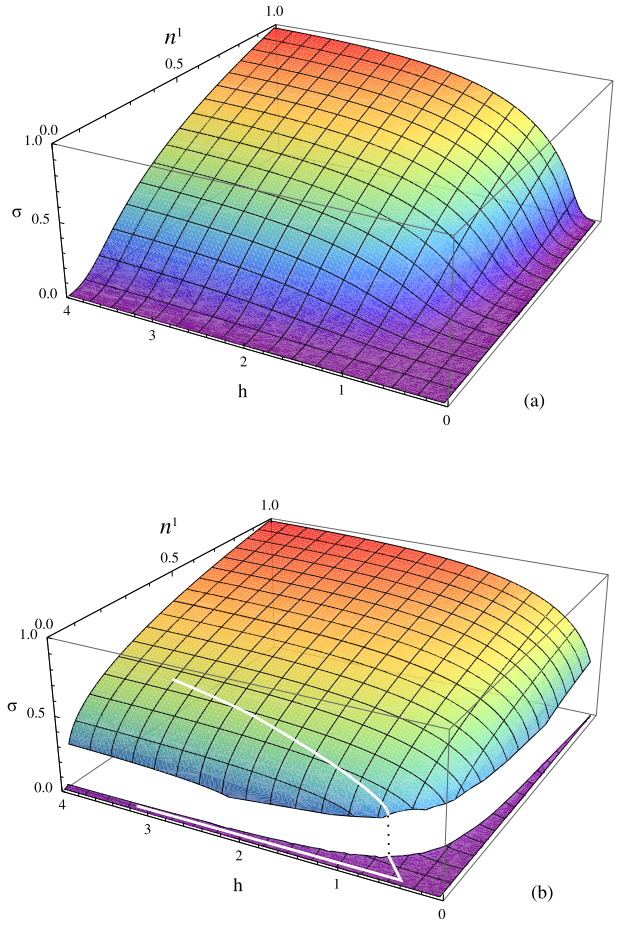


Figure 2.7: School efficiency σ as function of the fraction of informed individuals n^1 and strength of the preference h when the social parameter z is (a) in a non critical region $z = 2.5$ and (b) in a critical region $z = 3.1$ The white line in panel (b) is a schematic illustration of the hysteresis mechanism for a bluefin tuna population starting with high n^1 and h, then decreasing n^1 and h (overfishing of both preys and predators) and subsequently increasing h (increase of population of preys).

the individuals willing to reach location 1 and a fraction of n^2 willing to reach location 2.

Alas solving numerically the FOC, in this case, proved to be harder than the previous case: the numerical methods (multidimensional Newton method, Iterative method) have show instability for large range of z.

To get some insight on the behaviour of the system in this case, we can look at the behaviour of the system in two limiting cases $z \to 0$ and $z \to \infty$. In the first case we can do a perturbative expansion in z (see Appendix B.3 for the details). At $z = 0$ the solution is determined only by the update process, and thus, will be essentially given by the average of population with respect to eq. (2.13):

$$n_{a,0} = \frac{n^0}{q} + \sum_\alpha \frac{n^\alpha e^{h_\alpha \delta_{a\alpha}}}{e^{h_\alpha} - 1 + q}. \tag{2.101}$$

The higher order of perturbative expansion do not give any insight except for the fact that they are finite. This assures that at least in a neighbourhood of $z = 0$ this solution should exists. When $n^0 \to 1$ this solution tend to the symmetric solution and for z small enough the it corresponds to a low density network configuration.

For the other limiting case the calculation are cumbersome (see Appendix B.4 for the details). It is easy to convince oneself, from the structure of the free energy, that the stationary points around which expanding are those of the quadratic part. And thus the minimum around which to expand is one of those having the form:

$$n_{a,\infty} = \delta_{a,\alpha} \tag{2.102}$$

Expanding asymptotically the free energy around those points one gets

$$F \simeq -\frac{z}{2} - n^\alpha h_\alpha + \left[n^0 \log(n^0) + \sum_{i \neq 0} n^i \log(n^i) \right] + z e^{-z} \left(- n^\alpha h_\alpha e^{h_\alpha} + \sum_i n^i h_i e^{h_i} \right) + o(e^{-z}) \tag{2.103}$$

From this expansion we get that for large z the group prevailing should be the one with higher $n^\alpha h_\alpha$.

2.8.4 *Intermediate region*

To describe what happens in between, we can only rely on simulations. Because of the presence of various local minima, the system shall have many metastable states, especially when more than two population have the comparable values of $n^i h_i$. It may be necessary to do simulation with networks with a large number of vertices which may take long time to equilibrate.

If we restrict to the case with two informed subpopulations we have essentially three cases to check:

1. $n^1 < n^2$ and $h_1 < h_2$;
2. $n^1 < n^2$ and $h_1 > h_2$ but $n^1 h_1 < n^2 h_2$

3. $n^1 < n^2$ and $h_1 > h_2$ but $n^1 h_1 > n^2 h_2$

In in the first case we are considering an less strongly opinionated minority; in the second case a minority with a stronger preference than the majority, in the third case we are considering a strongly opinionated minority versus a weakly opinionated majority.

We know from the expansions, that for large z we should expect the group with higher $h_\alpha n^\alpha$ to prevail but we cannot blindly use this result for intermediate values of z. Another aspect that we would like to check is what is the role of n^0. In [63] it is suggested, in the case with $q = 2$ that higher value of n^0 should enforce democratic decision i.e. the group prevailing should always be the one with higher n^α irrespectively of the strength of h_α. Numerical evidences (see Figure 2.8 and 2.9) seem to suggest that, for $q > 2$ even for intermediate values z the group with the higher value of $n^a h_a$ dominates.

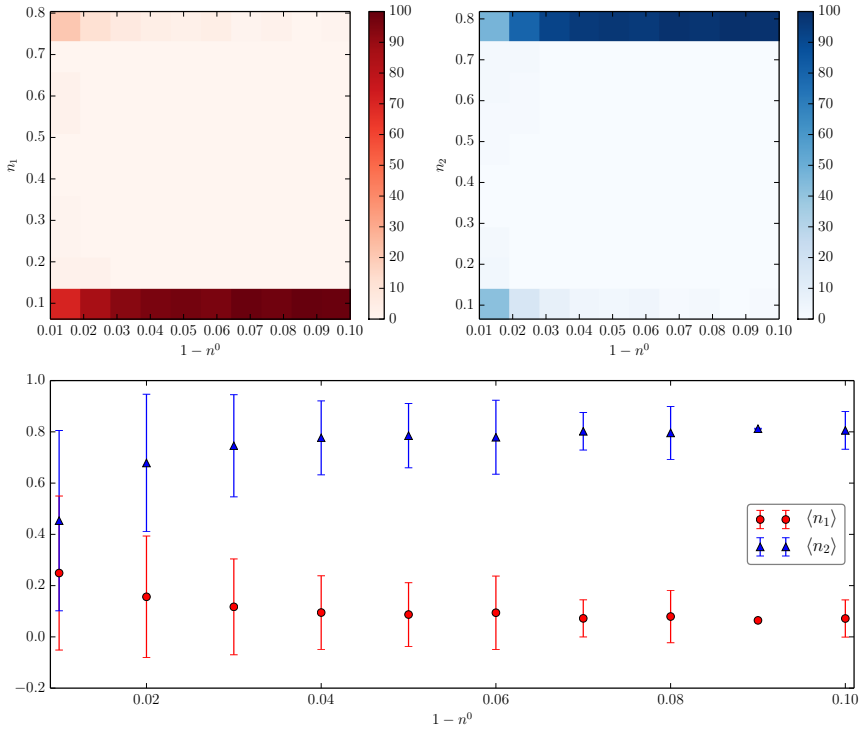


Figure 2.8: Bottom panel: On the y-axis the value of n_1 averaged on 100 realisation. on x-axis $1 - n^0$. Simulation parameters: $N = 1000$ $n1/n2 = 1.5$,$h_1 = 0.5$, $h_2 = 1.5$ $z = 3.4$ and $v = 30$ Each run lasts $T = 1000$. Top panels: two 2D histograms with the distribution of values of n_1 and n_2 in the different realisations. In this case $n^1 > n^2$ but $h_1 < h_2$ and $n^1 h^1 < n^2 h^2$. The system is in a frustrated state and the group prevailing it 2 as we expect. The presence of other (meta)stable where one of the other group is prevailing is the origin of the noise (top panels).

As already mentioned, when there are two or more classes have values of $h_\alpha n^\alpha$ close to the maximum, the system has long lasting metastable states that make the numerical investigation complicated. This is particularly true when n_0 is high, i.e. when the number of uninformed individuals is high; in this case the

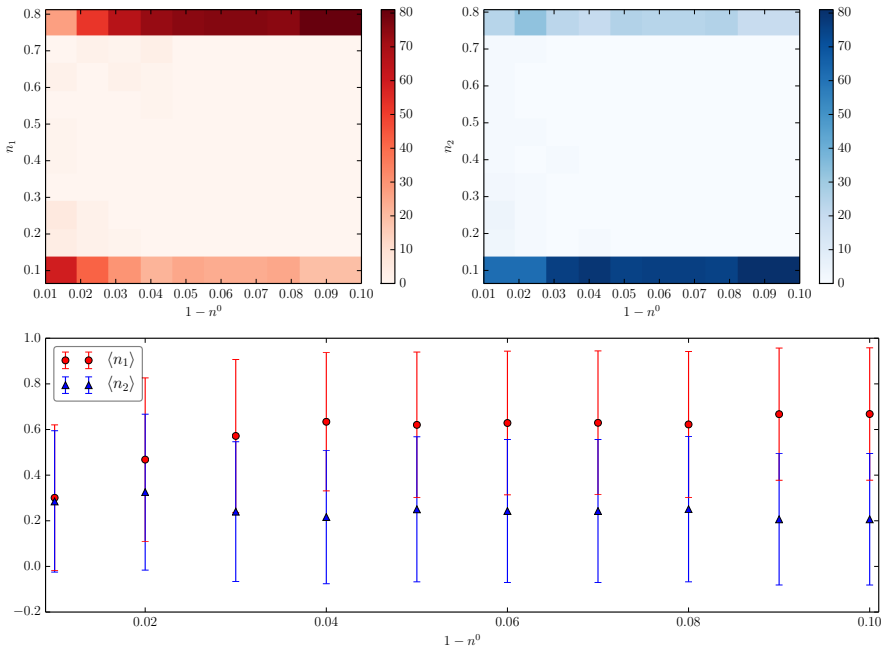


Figure 2.9: Bottom Panel: on the y-axis the value of n_1 averaged on 100 realisation. On x-axis $1 - n^0$. Simulation parameters: $N = 1000$ $n1/n2 = 4$,$h_1 = 0.5$, $h_2 = 1.5$ $z = 3.4$ and $v = 30$ Each run lasts $T = 1000$. Top panels: two 2D histograms with the distribution of values of n_1 and n_2 in the different realisations. In this case $n^1 > n^2$ but $h_1 < h_2$ and $n^1 h^1 > n^2 h^2$. The system is in a frustrated state and the group prevailing it 1 as we expect. The presence of other (meta)stable where one of the other group is prevailing is the origin of the (large noise).

state with maximal n^i, has a long survival time, especially for small values of the preference parameters.

The infinitely strongly opinionated minorities

The results of our expansions should be used with caution if strength of preferences becomes infinite. In fact, when $h_\alpha \to \infty$ we have that the ergodic subspace is reduced to the state having at least $n_\alpha \geq n^\alpha$ and in particular , $n_a^\alpha = n^\alpha$, since the state update process becomes deterministic.

In particular, if the system has no uninformed individuals, the node dynamics will be frozen and the system will be become fragmented in as many independent subgroups as the non-zero n^α.

On the other hand, if there are uninformed individuals, the system restricted to them behaves as in [82, 83] but, the dense solution will always favour the majority. It is easy to rewrite the free energy of the system in this case:

$$F = \sum_{a \in S} n_a^0 \log(n_a^0) - z \sum_{a \in S} n^a n_a^0 - \frac{z}{2} \sum_{a \in S} (n_a^0)^2. \tag{2.104}$$

It is clear that the FOC are of the same form of those with $n^0 = 1$ except for the fact that the normalisation constant reads:

$$\sum_{a\in S} n_a^0 = n^0 = 1 - \sum_{a\in S} n^i < 1. \tag{2.105}$$

It is also clear, form the form of the free energy, that the in the thermodynamic limit, the high density solution corresponding to the internal state with higher n^α is global minimum of the free energy . In this sense *uninformed individuals favour democratic consensus decision making* as stated in [63].

2.9 A SOLVABLE ADAPTIVE NETWORK MODEL FOR MIGRATION IV: MEAN FIELD DYNAMICS

To get a better insight on the behaviour of the system we can can use the mean field dynamics.

Mean Field Dynamics is an approximation method to treat stochastic processes in the limit in which fluctuations are depressed (in our case $N \to \infty$, which is similar to a non homogeneous moment expansion. As for moment expansion it consists in writing down a set of ordinary differential equations for the averages of some stochastic variables, in our case :

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{N}_{a,t}^\alpha \rangle = \lim_{\Delta t\to 0} \frac{\langle \hat{N}_{a,t+\Delta t}^\alpha - \hat{N}_{a,t}^\alpha \rangle}{\Delta t} \tag{2.106}$$

2.9.1 *Notation*

Let us the following notation

$$N_{ak}^\alpha(t) = \langle \hat{N}_{akt}^\alpha \rangle \tag{2.107}$$

to denote the average number of nodes in state a having preference α and degree k, where:

$$\hat{N}_{akt}^\alpha = \hat{N}_k^\alpha(\Omega_t) = \sum_{i\in S} \delta_{k,\sum_j g_{ij}(\omega)}. \tag{2.108}$$

In the following we shall also write

$$n_{ak}^\alpha(t) = \frac{N_{ak}^\alpha(t)}{N} \tag{2.109}$$

When t is not explicitly written, $t = \infty$, is to be understood. We shall also say *an* (a, α)*-node* to refer to a node with internal state a and preference α.

2.9.2 *Derivation*

Let's analyse the the possible contribution to the Mean Field Equation First for $k > 0$.

Case $\Delta N^{\alpha}_{ak,t} = +2$

This situation can be realized only if two (a, α)-nodes in with degree $k-1$ are selected to form a link or a link between two nodes (a, α)-node with degree $k+1$ decays. This happens with probability:

$$p(+2) = \eta N^{\alpha}_{ak-1,t} \frac{N^{\alpha}_{ak-1,t}}{N-1} + \lambda L_{a,k+1} = N\eta (n^{\alpha}_{a,k-1}(t))^2 (1 + \frac{1}{N-1}) + \lambda \frac{L_{a\alpha,k+1}}{N} \tag{2.110}$$

where $L_{a\alpha,k}$ represents the number of edges between (a, α) nodes with degree $k+1$

Case $\Delta N^{\alpha}_{ak,t} = +1$

This situation can be realised when:

1. an (a, α)-node of degree $k-1$ and one with degree $k' \neq k, k-1$ are selected to form a link;
2. an (a, α)-node of degree $k-1$ and a node with internal state a but preference different form α is selected to form a link;
3. when an link, incident on (a, α)-node degree $k+1$ and an (a, α)-node of degree $k \neq k+1$' decays.

This happens with probability:

$$\begin{aligned} p(+1) &= 2\eta N^{\alpha}_{ak-1,t} \frac{\left(\sum_{\beta \neq \alpha} \sum_k N^{\beta}_k\right) + \left(\sum_{k' \neq k-1,k} N^{\alpha}_{ak',t}\right)}{N-1} \\ &+\lambda (N^{\alpha}_{a,k+1}(k+1) - 2L_{a\alpha,k+1}) = N \Big[\eta n^{\alpha}_{a,k-1}(t) \Big((n_a(t) - n^{\alpha}_{a,k-1}(t) \\ &-n^{\alpha}_{a,k}(t)\Big) (1 + \frac{1}{N-1}) + \lambda (k+1) n^{\alpha}_{a,k+1}(t)\Big] - 2\lambda \frac{L_{a\alpha,k+1}}{N} \end{aligned} \tag{2.111}$$

where we have used the fact that $(k+1)N^{\alpha}_{ak+1} = 2L_{a,\alpha,k+1}+$ number of edges only one end of which is incident in a (a, α)-node of degree $k+1$.

Case $\Delta N^{\alpha}_{ak,t} = -1$

1. an (a, α)-node of degree k and one with degree $k' \neq k, k-1$ are selected to form a link;
2. an (a, α)-node of degree k and a node with internal state a but preference different form α is selected to form a link;
3. when an link, incident on (a, α)-node degree k and an (a, α)-node of degree $k \neq k$ decays.

This happens with probability:

$$\begin{aligned} p(-1) =& 2\eta N^{\alpha}_{ak,t} \frac{\left(\sum_{\beta\neq\alpha}\sum_k N^{\beta}_k\right)+\left(\sum_{k'\neq k-1,k} N^{\alpha}_{ak',t}\right)}{N-1} \\ &+\lambda(N^{\alpha}_{a,k}(k)-2L_{a\alpha,k}) = N\Big[\eta n^{\alpha}_{a,k}(t)\Big(n_a(t)-n^{\alpha}_{a,k-1}(t) \\ &-n^{\alpha}_{a,k}(t)\Big)(1+\frac{1}{N-1})+\lambda k n^{\alpha}_{a,k}(t)\Big]-2\lambda\frac{L_{a\alpha,k}}{N} \end{aligned} \tag{2.112}$$

where $L_{a\alpha,k}$ represents the number of edges incident between (a,α)-nodes with degree k and we have used the fact that $kN^{\alpha}_{ak} = 2L_{a,\alpha,k}+$ the number of edges only one end of which is incident in a (a,α)-node of degree k.

Case $\Delta N^{\alpha}_{ak,t} = -2$

This situation can be realized only if two (a,α)-node of degree k are selected to form a link of if one edge incident on two (a,α)-node of degree k decays. This happens with probability:

$$p(+2) = \eta N^{\alpha}_{ak,t}\frac{N_{ak,t}}{N-1}+\lambda L_{a\alpha,k} = N\eta(n^{\alpha}_{a,k}(t))^2(1+\frac{1}{N-1})+\lambda\frac{L_{a\alpha,k}}{N} \tag{2.113}$$

Case $k = 0$

In this case besides the graph dynamics which accounted in the previous contribution we have to consider the colour dynamics too.

CASE $\Delta n^{\alpha}_{a0,t} = 1$ DUE TO COLOUR DYNAMICS Now according to (2.28) the probability that a class gains one element due to colour dynamics is simply :

$$p_{\text{col}}(+1) = \frac{\nu}{e^{h_\alpha}+q-1}\sum_{b\neq a} e^{h_\alpha\delta_{a\alpha}}N^{\alpha}_{b0,t} = N\frac{\nu}{e^{h_\alpha}+q-1}e^{h_\alpha\delta_{a\alpha}}\sum_{b\neq a} n^{\alpha}_{b0}(t) \tag{2.114}$$

CASE $\Delta n^{\alpha}_{a0,t} = -1$ DUE TO COLOUR DYNAMICS Now according to (2.28) the probability that a class gains one element due to colour dynamics is simply :

$$p_{\text{col}}(-1) = \frac{\nu}{e^{h_\alpha}+q-1}N^{\alpha}_{a0,t}\sum_{b\neq a} e^{h_\alpha\delta_{b\alpha}} = Nn^{\alpha}_{a0}(t)\frac{\nu}{e^{h_\alpha}+q-1}\sum_{b\neq a} e^{h_\alpha\delta_{b\alpha}} \tag{2.115}$$

The complete equations for $k > 0$

Putting everything together we obtain the following infinite set of Ordinary Differential Equation (all n are here functions of the time)

$$\begin{cases} \dot{n}^{\alpha}_{a,k} = \lambda(k+1)n^{\alpha}_{ak+1}(t)-(2\eta n_a(t)+\lambda k)n^{\alpha}_{ak}(t)+2\eta n_a(t)n^{\alpha}_{ak-1}(t) \\ \dot{n}^{\alpha}_{a,0} = \lambda n^{\alpha}_{a1}(t)-2\eta n_a(t)n^{\alpha}_{a0}(t)+\frac{\nu}{e^{h_\alpha}+q-1}\left[e^{h_\alpha\delta_{a\alpha}}\sum_b n^{\alpha}_{b0}(t)-n^{\alpha}_{a0}(t)\sum_b e^{h_\alpha\delta_{b\alpha}}\right] \end{cases} \tag{2.116}$$

where $n_a = \sum_{\alpha\in\mathcal{S}\cup\{0\}} n^{\alpha}_a$ and we assume that $h_0 = 0$. In particular the contribution coming from $L_{a\alpha,k}$ erases out.

Summing over k one immediately gets the equations for the population densities:

$$\dot{n}_a^k(t) = \frac{\nu}{e^{h_\alpha} + q - 1}\left[e^{h_\alpha \delta_{a\alpha}} \sum_b n_{b0}^\alpha(t) - n_{a0}^\alpha(t) \sum_b e^{h_\alpha \delta_{b\alpha}}.\right] \tag{2.117}$$

The structure of equations (2.116) makes clear that in this case the network dynamics depends only on the state and not on the parameter α.

2.9.3 *Stationary degree distribution*

We study the stationary state of the system. Solving the recurrence equation can be done using a simple ansatz, i. e. that $n_{a,k}^\alpha = C_a^\alpha \frac{(\mu_a^\alpha)^k}{k!}$.

By direct substitution one gets:

$$n_{a,k}^\alpha = n_{a,0}^\alpha \frac{\left(\frac{2\eta n_a}{\lambda}\right)^k}{k!} \tag{2.118}$$

where $n_a^\alpha = \lim_{t\to\infty} n_a^\alpha(t)$ and $n_a = \lim_{t\to\infty} n_a(t)$.

Summing over k and solving for n_a^α we get $n_{a,0}^\alpha = n_a^\alpha e^{-\frac{2\eta n_a}{\lambda}}$ and thus substituting in (2.118) we get:

$$n_{a,k}^\alpha = n_a^\alpha \frac{\left(\frac{2\eta n_a}{\lambda}\right)^k}{k!} e^{-\frac{2\eta n_a}{\lambda}}. \tag{2.119}$$

Summing over α we obtain finally:

$$n_{a,k} = n_a \frac{\left(\frac{2\eta n_a}{\lambda}\right)^k}{k!} e^{-\frac{2\eta n_a}{\lambda}} \tag{2.120}$$

which is the degree distribution within a sub-population of nodes in a state. Where n_a and n_a^α are quantity that need to be found self-consistently.

To do that we first used eq (2.117). If $n_0^\alpha = \sum_{b\in S} n_{b,0}^\alpha$ immediately obtain:

$$n_{a,0}^\alpha = \frac{e^{h_a \delta_{a\alpha}}}{e^{h_a} + q - 1} n_0^\alpha. \tag{2.121}$$

Substituting in equation (2.118), summing over k and α, and using normalisation constraints, we obtain:

$$n_o^\alpha = \frac{n^\alpha(e^{h_\alpha} + q - 1)}{(e^{h_\alpha} - 1)e^{-\frac{2\eta n_a}{\lambda}} - \hat{Q}} \tag{2.122}$$

where $\hat{Q} = \sum_{b\in S} e^{-\frac{2\eta n_a}{\lambda}}$. Thus, substituting first in (2.121) then in (2.118) and summing over k we get:

$$n_a^\alpha = e^{\frac{2\eta n_a}{\lambda}} \frac{e^{h_a \delta_{a\alpha}} n^\alpha}{(e^{h_\alpha} - 1)e^{\frac{2\eta n_a}{\lambda}} - \hat{Q}} \tag{2.123}$$

These equation are formally equivalent to the first order conditions in equation (2.76), thus the quantities n_a^α corresponding to the stationary solution of MFD satisfy the same equations than the extremal points of the free energy. This of course had to be expected, and it is a check that the MFD is a good description of the system for large N.

This allows us to write degree distribution p_k for this system as:

$$p_k = \sum_a n_a \frac{\left(\frac{2\eta n_a}{\lambda}\right)^k}{k!} \mathrm{e}^{-\frac{2\eta n_a}{\lambda}}. \tag{2.124}$$

As a check we can calculate immediately the average degree and trivially retrieve the expression of equation (2.69).

We know that the system has a stationary distribution, and it is known that, in the large N limit, deterministic MFD equations describes correctly the stochastic system[91]; yet, we should be cautioned that only the asymptotically stables solution of MFD offer a good approximation of the system; however we can use our knowledge obtained solving the stochastic dynamics to select the solutions of (2.123) that give rise to a stable state.

Having the exact form of the degree distribution enables us to investigate many of the properties of the system (such as giant component, small components, etc.). However, the degree distribution depends on the values of n_a which are available only numerically, and the general study is indeed quite articulated. Nonetheless, it should come with no surprise that for strongly polarised solutions of the FOC—i. e. $n_\alpha \simeq 1$, $n_{i\neq\alpha} \simeq 0$—the system has a giant component; in this case, in fact, we have that the degree distribution is almost Poissonian with average $\frac{2\eta}{\lambda} n_\alpha > 1$. On the other hand we the stable symmetric solution corresponds to a disordered phase when almost all nodes are disconnected, or in small components. In Figure 2.10 we provide two snapshot of the system in these situation.

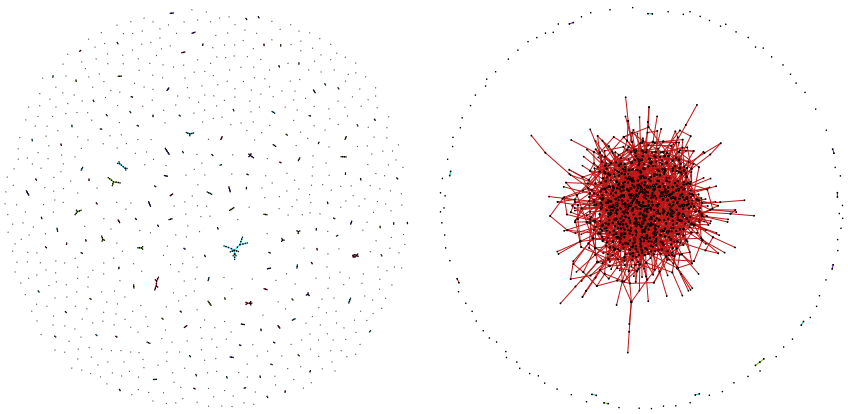
2.10 A SOLVABLE ADAPTIVE NETWORK MODEL FOR MIGRATION V: THE ROLE OF SEGREGATION AND UPDATE POLICY

In our model we have made strong assumptions about the segregation policy and the internal state update rule. One may wonder to what extend the behaviours of our system are dependent on these assumptions.

2.10.1 *The role of segregation*

We want here to discuss what happens when we allow link creation between individuals in a different internal states to occur with probability $\varepsilon \neq 0$, i. e. to relax the strict segregation policy of our model.

When this happens, the system is no longer exactly solvable, at least using the methods we have used here; Even an mean field approach is hard. In fact, if the network dynamics is only marginally touched, the internal update using majority rule becomes relevant, and describing it in mean field is complicated.



(a) (b)

Figure 2.10: Snapshots of the network of the system for two different values of $z = \frac{2\eta}{\lambda}$. Simulations are done with $N = 1000$, $n^0 = 0.99$, $n^1 = 0.01$, $h_1 = 1$ $v = 0$, $\lambda = 10$ and are run for $T = 1000$. Panel (a) corresponds to $\eta = 10$ and thus $z = 2$, and shows the typical "low density" disordered phase corresponding to a non migrating phase. Panel (b) correspond to $\eta = 20$ and thus $z = 4$; in this case the system is polarised in direction ($\alpha 1$, red color in the plot) and there exists a giant component which corresponds to the group migrating.

We can however get some insight noting that when $\varepsilon = 1$ the network dynamics decouples from the dynamics of the internal states of the nodes. In this case we could write directly the mean field equation for the network dynamic ignoring the internal state of the system,i. e. for the quantities $N_k = \sum_{a,\alpha} N^{\alpha}_{a,k}$ and obtain a Poissonian degree distribution:

$$p_k = \mathrm{e}^{-\frac{2\eta}{\lambda}} \frac{\left(\frac{2\eta}{\lambda}\right)^k}{k!}. \tag{2.125}$$

The system would thus has an average degree $\langle k \rangle = z$ and at $\langle k \rangle = 1$ a giant component emerges. In the giant component, due to the majority rule, the nodes will soon reach consensus and thus the system will exhibit a non zero value of σ. From this, we can get a intuitive and clear picture of what happens. The probability ε interpolates between a regime in which there is a first order phase transition for the average degree and σ and a regime where the average degree is smooth whereas σ undergoes a second order phase transition. This means think that the coexistence region between a low and high density phase that exists at $\varepsilon = 0$ shall shrink increasing the value of η until, for some $\varepsilon < 1$, it vanishes. From that point onward the system shall approach smoothly the Erdős-Rényi like regime. Of course if the value of n^1 and h do not permit the presence of a coexistence region, the system will undergo a smooth transition. Numerical simulations support the qualitative picture we described above.

2.10.2 *The role of the upgrade policy*

One may also wonder what happens when another internal state update policy is used (i. e. vote model like). Indeed as long as the segregation policy is the same of our model, any update rule that enforce local homogeneity shall lead to the same results: the structure of the state space is, in fact, the same.Nevertheless, we might expect that the non equilibrium properties of the system shall depend on the chosen policy; and that a majority rule should lead to equilibrium faster that other rules (such as voter model like rules).

If we allow for link creation between nodes in different internal states, we may expect to see differences between models with different update rules; however all these rules behaves similarly for states in which the neighbourhood of a node is homogeneous; one expects that these states are the only stable states for these system and thus qualitatively we should not see much difference with respect to the system using majority rules.

2.11 DISCUSSION AND CONCLUSIONA

The solvable model that has been discussed provide simple yet sufficiently descriptive framework to address consensus decision making and in particular it offer a new insight on the migration process.

The migration game

The migration process can be described as an emergent property of the population undertaking a group formation game: when the spatial density of individuals is locally low, each individual moves independently, and the system is in a sparse network configuration with a value of z below the lower edge of the coexistence region. In this state uninformed individuals cannot migrate whereas informed individuals can undertake a *solitary* migration towards their preferred destination. Due to external stimuli the density may increase and so does the value of z, driving the system toward the coexistence region. In this region even though the local density of individuals high, a sparse network configuration with individuals moving independently is still stable but an alternative denser and stable network configuration also appears. When the system reaches the upper edge of this region, further increasing the density, the sparse network state becomes unstable, the system then undergoes a transition toward the high density network state and the group starts a migration toward the preferred destination. When the local density of individuals decreases in the group, also z decreases and the system is driven back to the coexistence region. The the denser configuration remains stable until the system reaches the lower edge of the coexistence region: at this point, individuals stop migrating and the system switches back into the sparse configuration.

Migration without information: the locusts case

When in the group there not any informed individual, once the group is formed, it will go in a random direction (among the q possible ones). A nice illustration of this behaviour is offered by desert locust *Schistocerca gregaria* Forsskål 1775. This locust is known as one of the most destructive animal pests, and desert locusts swarms (which can cover hundreds of square kilometres and contain billions of individuals) exhibit triggered transition between a disordered phase where each individual moves independently from the others to and ordered phase where individuals swarm. Swarms of locusts are known for the unpredictability of their directions, and for their sudden change of directions, consistent with the presence of bi-stability region and several equally probable directions . These facts have been illustrated in a nice experiment [80], where these three regions have been clearly identified.

The repeated migration game and seasonal migrations

The group formation game described above repeated each year, offer a possible interpretative framework for seasonal migrations. One can think that each year individuals, repeat this migration game, with, however, different parameters. In particular experience of the previous year(s) may drive changes in the preference term h, hence in the memory of migratory species. Likewise changes in this or in the relative density of informed individuals can occur when the migratory population is affected by external stimulus , e.g., overfishing, habitat degradation, demographic fluctuations.

In particular, because of the presence of (h, n^1)- hysteretic cycle, concurring overfishing and habitat degradation may prevent a group to successfully migrate, even if the habitat quality is re-established.

Conflicting Preferences

In the case with conflicting preferences the situation is more complicated. We know that, for large value of z, the group shall migrate toward the direction α for which the product $n^\alpha h_\alpha$ is maximal, whereas in the limit of small z, the sparse configuration is the only stable one.

As numerical evidences suggest, this should hold true also for smaller value z and offer us a nice picture: when for some reason either n^i or h^i of the dominant group are reduced to a point where $n^i h_i$ is no longer maximal the system, during next year shall undergo a migration toward another site. This mechanism might be a strong asset for a group having different feeding sites since it allows the group to respond rapidly to a deterioration of one of them. This is the so called *pre-emption* often observed in animal migrations: the ability to change a migration pattern in response to habitat degradation prior that the quality of the habitat becomes not compatible with the needs of an individuals any more.

In our model we did not consider neither the demography nor the *discovery* of new routes. In normal condition demography equilibrates, whereas the discovery of new routes is an uncommon event in most of the species.

Collective memo ry and breakdown of social traditions

Breakdown of social traditions, due to selected fishing on older informed individuals, has been hypothesized to have contributed to stock collapses in several large commercially important fish populations [42, 43].

Our sketch of the migration game suggests that social dynamics may lead to such collapses and that the integrity of migration pathways and spatial distributions of migratory predators might be particularly vulnerable to perturbations such as fishing or habitat degradation. Fishing out informed individuals and their prey can exacerbate the loss of collective memory up to the point where a migratory pathway is suddenly interrupted.

We can assume that each year young individuals join the group: among them a fraction is able to gather information and remember a migratory route whereas the rest has a purely social behaviour. The "information-gathering-able" individuals behave as uninformed individuals ($h = 0$) but *can learn* a new migratory route during the first migration(s). If the group does not succeed in starting migration, or migrates toward a different location, the young "information-gathering-able" individuals will not learn the traditional migration route of the group and the social traditions of the group will not be transmitted to the new generations. The loss of collective memory in the group will then force the system to cross the critical line and the migration toward the destination site will stop.

2.11.1 *The Atlantic Bluefin tuna case*

Overfishing and habitat degradation, and subsequent interruption of socially transmitted might offer a plausible explanation for the changes of the migratory path of ABFT.

The main prey ABFT on the coast of Norway is herring *Clupea harengus* Linnaeus 1758[5]. During the 50s-70s they have also been extensively fished which resulted in a fast collapse of the stocks([51, 92, 93]. cf. figure 2.11). Since the sea temperature has been relatively constant during this period (See figure 2.11 (c)), is is reasonable to consider the herring biomass in the Norwegian see as good habitat quality indicator (thus an indirect measure of h). Overfishing of large ABFTs and of their main prey correspond, in the language of our model, to a reduction of both the relative number of informed individuals and of the preference parameters.

Since the 60s, the herring populations have recovered to moderate-high levels [92, 93], but bluefin tuna are still extremely rare and apparently have not migrated to these areas since the disappearance several decades ago [51] (figure 2.11 (b)). These hysteretic dynamics are consistent with a fishing-induced removal of predators having preference for migration to these regions and a fishing induced decline in habitat quality which then leads to the collapse of group formation and a sudden change in migratory path (cf. Figure 2.7).

5 Mackerels are also consistently preyed by ABFT in northern Atlantic but their numerical relevance on the Norwegians coast is lesser

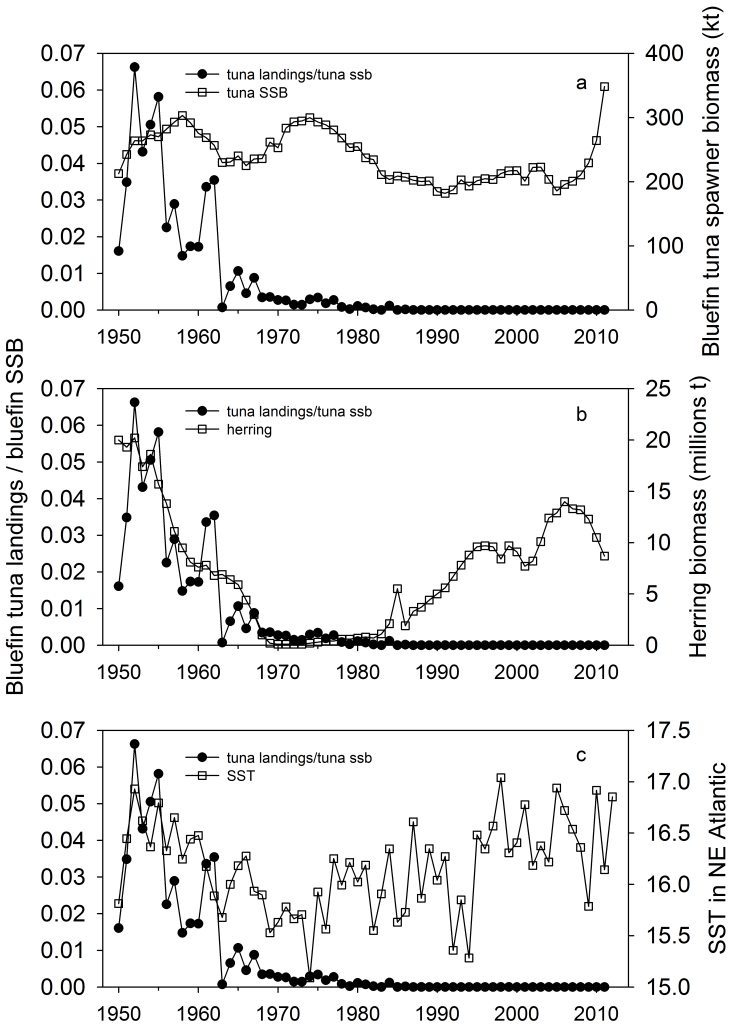


Figure 2.11: A graph showing Tuna landings on the coast of Norway with respect to the estimated spawning site biomass (a), with respect to the herrings biomass in the Norwegian and north sea (b) and with respect to the sea surface temperature (c)

2.11.2 *Conclusions*

As we have sees, the usual approaches to tackle the decision making in animal groups rely on real space model and cannot be approached analytically. However we have illustrated another possible approach to these problems. Instead of describing all the microscopical details of the system as done in standard real space models, it is possible to implicitly account for the change of the neighbourhood of an individual in a group using an adaptive network description.

Early works using this approach have focused on simplified settings involving only to possible choices for the individuals in the group and using approximate methods to get some insights. We discussed at length in this chapter another model, that with the only simplifying assumption to enforce total segregations, i. e. the fact that only individuals choosing to go in the same direction can effectively interact, is analytically solvable.

In particular, it is possible to provide a plausible mechanism for migration dynamics and to show that partial removal of knowledgeable individuals may be sufficient to interrupt the transmission of social traditions in groups of animals.

The model discussed, however, is very general. In fact, as long as one can give meaning to the Potts variable and to the update rule of the models the results can apply. It is possible to envisage different generalisation of the model proposed. One particularly natural one is to study an adaptive version of Axelrod Model [20] for cultural interaction. Preliminary studies seem to suggest that a phenomenology similar to the one discussed in this chapter can occur in this context, however in this case it is not possible to apply detailed balance and thus an analytic solution is much harder to obtain.

Another important generalisation might be to attempt to model the effective topology of the groups. However, with few exception such as [81] or [89], empirical studies are still too rare to obtain significant data on the topology of animal groups; the development of tagging experiments in the future might allow us to have explicit data on the network evolution, enabling us to build more realistic models.

3

DECISION MAKING OF RATIONAL AGENTS: THE SOCIAL CLIMBING GAME

Societies are changing entities that evolve over time. Explaining their structure and how social structure[1] change over time (the so called *social change*) is one of the main end of sociology, and an important part of other social sciences (as anthropology, history and marginally even economics).

As often in social sciences, each discipline, and within each discipline, each theory stresses different characteristic(s) as the most fundamental cause(s) of social change; it is however commonly accepted that societies are subject to two kind of *forces* that generate their change: *exogenous forces* exerted from the outside of the society itself— natural events such as famine, epidemics, climatic change; historical events such as wars, migrations, political instability; and to some extent economic changes—and *endogenous forces* that are built within society itself.

If the former have always been acknowledged to be able to induce social change, many sociological theories have been more reluctant to recognise the latter as a relevant or at least independent factor in social change; some theories, in fact, assume the fact that in the absence of external forces, the society should remain stable. Nowadays, however, there are empirical evidences that endogenous forces play a relevant role in shaping social structure[2] and modelling endogenous forces is a relevant part of modern sociology.

In modelling animal societies in chapter 2, we thought the individuals as *simple* agents that make choices mainly adapting to their social neighbourhood and, for informed individuals, taking into account information gathered on one possible outcome of their decision making problems. The evolutionary pressure selects the best strategies; thus evolution makes the *rational decision*.

In modelling human societies, however, assuming simple conformity rules is reductive. Human societies should be rather thought as the product of the interaction of many *self-interested* individuals making *rational* decisions.

In this chapter we shall focus on one specific aspect of social structure in human societies: hierarchy. In particular we want to investigate, with a simple, stylised, *microscopic* model, how do social elites emerge in a society of self-interested rational decision makers.

3.1 MODELLING RATIONAL DECISION MAKING: UTILITY FUNCTIONS

Modelling decision making by *self-interested rational* has been studied since the beginning of the XVIII century.

1 which is often called, with a lot semantic significance, *social order* in sociological literature.

2 A nice illustration of this can be found in Stanley Lieberson empirical theory of taste [94]

The first to formally tackle the problem using the concept of *utility function*, which measures the perceived balance between costs and benefit for a given choice, has been Daniel Bernoulli in a famous paper in 1738 [95]. Since then, the concept has been widely used in economic science. The typical decision being modelled in economy is the the choice of an agent in a market: given a commodity c what quantity of this commodity to buy x_c at a given price p_c (the problem of the consumer), or vice-versa how much to produce of commodity c, x_c, and at what price to sell it p_c, or at what price p_c to sell a given quantity x_c (the problem of the firm). In neoclassical economics, each agent i is *self-interested and rational* and makes decisions optimising a function u_i, called *utility function*, defined as the balance of benefits $B(x_c, p_c)$ and costs$C(x_c, p_c)$ of buying quantity x_c of commodity c at a price p_c (or vice-versa producing x_c and selling it at p_c). This utility takes the form:

$$u_i = B(x_c, p_c) - C(x_c, p_c) \tag{3.1}$$

In this context, this optimisation process will determine a relationship between price and quantity—the so called (aggregate) demand and supply curve—and eventually define an optimal choice for both consumers and firms.

The concept of utility, which proved to be an effective framework to study rational decision making in economics, albeit with some caveats[3], has subsequently been adopted to model *rational decision making* in more general contexts.

In a general probabilistic framework for rational decision making, a *utility function* can be seen [96] as a map between the different alternative choices and a real numbers that *ranks* the states of the world with respect of different levels of happiness of the agent.

Although mapping the whole world into a function may seem an extremely simplifying assumption, it is not. A famous theorem,proven by Von Neumann and Morgenstern in 1944 [97], states that, as long as it is possible to define a rational concept of preference between the different alternatives, it is possible to define an (expected) utility function u_i for the agent i whose maximisation shall single out the best rational choice for agent i.

3.1.1 *Rational choice in complex environment: random utility model*

In many situations, however, agents exhibit some degree of randomness in their decisions. The same outcome might induce the agent to take different choices, which in principle seem to conflict with the assumption of self-interested and rational agents. However, D. McFadden [98] has shown that it is possible to extend the formalism of utility function in a more general context.

Let us restrict to the case in which the the possible alternatives $s_1, \dots, s_n$ are finite.

When modelling a decision-making problem in a complex environment one might not be able to describe the full utility with respect to which and agent is optimizing; in particular it is not always possible to describe perfectly the

3 The most relevant of which is the fact that utility function by itself is not a measurable quantities, and thus

states of the world, or all the possible variables that an agent effectively takes into account to make its decision.

To model this complexity, thus, we assume that the utility function of the agent has this form:

$$U_i(s, \omega) = \beta u_i(s) + \eta_i(s, \omega), \tag{3.2}$$

where $u_i(s)$ represent the *a priori* observable part of the *full* utility while $\eta_i(s, \omega)$ is a random function that accounts for the unobservable part of the utility and ω collectively represent the possible other choices variable on which the agent is optimising.

The factor β thus measures how well the *observed* utility captures the decision making process of the agent. Alternatively, one may may think to β as an *implicitly measure* of the relative importance the factors that are accounted by u_i in the effective decision making process of agent i i. e. how much do agent i cares about the factors accounted in u_i when making its decisions.

In this framework an agent shall still optimise with respect to the full utility$U_i(s, \omega$, however the outcome

$$s^*(\omega) \in \arg\max_s(\beta u_i(s) + \eta_i(s, \omega)) \tag{3.3}$$

will be stochastic variable.

As a result of this stochasticity, agent i will appear to take decisions with a probability distribution function $\text{Prob}(s^*(\omega) = s)$ induced by the distribution of $\eta_i(s, \omega)$.

To retrieve it one must specify the distribution of $\eta_i(s, \omega)$: let $P_{\eta_i}(s, x) = \text{Prob}(\eta_i(s, \omega) < x)$, one can write:

$$P(s^*(\omega) = s) = \text{Prob}(s^*(\omega) = s) = \int_{-\infty}^{\infty} \mathrm{d}P_{\eta_i}(s, x) \prod_{r \neq s} P_{\eta_i}(s, x + \Gamma(u(s) - u(r)). \tag{3.4}$$

The Probit Model

A first simplification that one can make is that the $\eta_i(s, \omega)$ are independent identically distributed variables normally distributed. In this case $P_{\eta_i}(s, x) = \text{erf}(x)$.

The Logit Model

In McFadden derivation [98], instead assumes that $\eta_i(a, \omega)$ are distributed with a Gumbel distribution $P_{\eta_i(s)}(s, x) = \exp(\exp(-x))$; with this assumption it is possible to obtain a closed form for $P(s^*(\omega) = s)$

$$P(s^*(\omega) = s) = \frac{\mathrm{e}^{\beta u_i(s)}}{\sum_s \mathrm{e}^{\beta u_i(s)}}. \tag{3.5}$$

The assumption that η_i a identical distributed with a Gumbel distribution is indeed quite general. In fact, it can be proven (see [99, p. 33] for an explicit

derivation) that if the agents are optimising with respect to a complex function $U_i(\boldsymbol{s}, \boldsymbol{b}) = u_i(\boldsymbol{s}) + v_i(\boldsymbol{b}|\boldsymbol{b})$ where $v_i(\boldsymbol{b}|\boldsymbol{s})$ is a an unobserved contribution to the utility that depends on a vector $\boldsymbol{b}$ of unobserved choices, and one assumes that $v_i(\boldsymbol{b}|\boldsymbol{s})$ are random independent and identically distributed, $\max_{\boldsymbol{b}}(U_i(\boldsymbol{s}, \boldsymbol{b})) = u_i(\boldsymbol{s}) + \eta_i(\omega)/\beta$, where $\eta_i(\omega)$ are i.i.d. with a Gumbel distribution.

3.2 THE EMERGENCE OF SOCIAL ELITES

The emergence of social elites has interested social scientists ever since Pareto's observation of persistent inequalities in our societies [100]. Inequality is acceptable if it results from differences of individuals in terms of their capabilities, but not if it results, in one way or another, from discrimination[4]. Not only discrimination conflicts with ethical principles that all individuals are *a priori* equal and should have access to the same opportunities. It also damages societies in terms of efficiency [101] as it hampers social mobility, preventing society from promoting individuals to positions in the social hierarchy that are consistent with their capabilities.

3.2.1 *The social climber problem*

We want to build a model that accounts for the emergence and persistence of hierarchies only relying on *endogenous* forces. In particular, it is a well established fact that social prominence bears a series of appealing benefits to individuals (wealth, social influence, etc.). It is likely that the endogenous force responsible for the emergence of hierarchies is the desire of each individual in the society to increase his/her social prominence.

We shall call the agents of this model, who seek to become more important in the society in which they live, *social climbers*.[5]

This inner dynamics which induces agents to desire to become more important inside a society will manifest itself in a process of adaptive change of each individual's neighbourhood. In particular individuals will attempt to connect themselves to important individuals, and disconnect from unimportant ones.

Moreover this adaptive change of neighbourhood must be "local", in the sense that we should model the fact that an individuals will be more likely to interact with other individuals that are close to him, on the social graph.

This defines the social climber problem, how to improve his/her social prominence through local moves in an optimal way.

4 India's cast system or racial segregation in the US and South Africa in the last century, are examples of explicit discrimination of underprivileged groups, that in the course of time has come to be regarded more and more as unacceptable, prompting for explicit measures of *affirmative action* (e.g. quotas for lower casts in India).

5 The expression *social climber* is often used as a derogatory term to identify he who seeks to improve its social position, often with the implication that it does so by being very friendly to people from a higher social class rather than thanks to its capabilities; however we use the term in a neutral way

3.2.2 *Measuring social importance: centrality*

Clearly, this requires that the social climber has some sort of metric to measure social prominence on a social networks. Numerous investigations in social sciences have since the early years of social network analysis, found that the importance of an individual within a social network is related to some quantification of how *central* [102, 103] this agent is. There exist different metrics which measure the centrality of a node (among others: degree, betweenness, closeness, eigenvector centrality see appendix A.1 and [104] for the definition), each one highlighting different facets of this generic concept..

Among the empirical analyses, there is a body of literature showing that centrality explains the role, importance, or pay-offs of the agents constituting the network: in informal structures within organizations, the importance of people is related to their betweenness centrality [105]; students with an higher centrality in the friendship network were found to perform better in education tests [106][6]. From the theoretical side, Ref. [108] shows that in a broad class of games, player's pay-offs increase with their (Bonacich) centrality [109] in the network.

Because of this, if individuals can alter their neighbourhood, the myopic best response strategy is simply to connect to the neighbour who increases their centrality the most.

Interestingly, König *et al.* [107] have shown that when individuals strive to be as central as possible, network evolve to a state where the neighbourhood of any node contains the neighbourhood of the nodes which have a lower degree (*nestedness*). In this kind of networks, the ranking of nodes according to their centrality is the same, regardless of the centrality measure considered [110] and thus the exact measure of centrality is irrelevant.

Remarkably, nested structures have been found in inter-organisational networks of research and development (R&D) alliances [111, 112], in interbank payment networks [113] and in firm competition under oligopolies [114]. This kind of structures will be precisely the ones emerging in the social climbing game. In this respect, our results confer stability to those of Ref. [110] and generalise them in non-trivial ways.

3.3 A MODEL FOR SOCIAL CLIMBING

Let us consider a system composed of N individuals, who are connected through a network which consists of exactly M links. The network is undirected and thus can be specified in terms of a symmetric adjacency matrix G, with elements $g_{ij} = g_{ji} = 1$, if i and j are connected, $g_{ij} = g_{ji} = 0$ otherwise. Agents receive opportunities to use their links in order to get in contact with more "influential" members of the society, in brief to climb the social network.

As we have extensively discussed, rational decision making can be modelled by defining a utility function. To do so, we have to solve the social climber problem, and chose which centrality definition to use. We choose, the simplest

6 A more comprehensive list can be found in Ref. [107].

one, *degree centrality*: the measure of the importance of and individual within a network is given by the number k_i of his/her partners[7]. Degree centrality has also the advantage of being a *a local topological* property of the network. As a measure of the "social capital" of agent i we take the following *local* utility function

$$u_i = \sum_{j,l=1}^{N} g_{ij} g_{jl} + \mu \sum_{l=1}^{N} g_{il} = \sum_{j=1}^{N} g_{ij} k_j + \mu\, k_i, \tag{3.6}$$

that depends both on the centrality k_i of agent i and on the centrality k_j of his/her neighbours, with μ tuning the relative weight between the two terms[8] The efforts of agents to climb the social hierarchy can then be formalised in the maximization of the utility u_i.

We then define the dynamics as follows,

1. At any time, an agent i is picked at random together with one of her neighbours, j. Then, a neighbour l of j is selected at random, $l \neq i$.

2. If l is already connected to i, nothing happens. Otherwise, with probability

$$p_{(i,j)\to(i,l)} = \frac{e^{\beta \Delta u_i}}{1 + e^{\beta \Delta u_i}}, \tag{3.7}$$

the link (i, j) is replaced with (or rewired to) link (i, l), where Δu_i is the corresponding change in i's utility.

The step 1 models random encounters between agents through their network of interactions. In such an encounter, agent i gets to know a friend l of j, as well as his/her importance (the number k_l of l's friends). The probabilistic choice rule in step 2, as we have discussed in section 3.1.1 can be derived from a random utility model where agents maximise a more complex utility function, that accounts for the fact that the social network affects in complex ways the well being of individuals and their unobserved choices in other dimensions. In this view, β plays the role of the relative weight between the observed and the unobserved part of the utility in the the choice of social contacts and it reflects the prevalence of the quest for social status in their decision behaviour[9]. In particular, in the limit $\beta \to \infty$, a move implying a decrease in the utility function is never accepted. This means that the social status is valued so highly by the agents that everything

7 Other measures of centrality can be taken but, as observed in Ref. [110], these rank individuals in the same order in strongly hierarchical networks, that will be stable over time as we shall see later. Conversely, unstructured networks correspond to random rankings with no stable order, with respect to all centrality measures.

8 As will be clear in the following, the second term in (3.6) is irrelevant for the dynamics, but not for the interpretation of the local utility. For example, consider the limit case of a star: while the central node is connected to $N-1$ nodes, all other nodes have only one connection. In this case the first term in (3.6) is equal to $N-1$ for all nodes and only the term proportional to k_i removes this degeneracy. Note that the second term in (3.6) also describes a linear cost $\mu < 0$ to maintain links.

9 For example, Adam may be reluctant to interrupt his relation with Bob, despite his low rank in society, because he is his only friend who shares his interest in Japanese paintings.

else is unimportant. On the contrary, for $\beta = 0$ the probability of accepting a move implying a decrease of the utility function is 1/2, meaning that the social status has negligible importance with respect to the unobserved part of the utility. The general question addressed is then how strong should the parameter β be in order for a social hierarchy to form and be maintained in the long run?

It is worth to remark that if the utility of agent i increases when rewiring the link (i, j) to (i, l), then the utility of agent j decreases, while that of agent l increases. This embodies the fact that the formation of a new link needs the consent of both parties, but their removal can be unilateral. Therefore, we can interpret the rewiring mechanism as a process according to which agent i looks for some social premium (e.g. knowledge of information, professional expertise) that agent l can provide more than agent j. Once agent i secures his/her connection to agent l, agent j essentially represents a redundant, less central source of the same capital, and this is why the rewiring operation happens at his/her expenses. Moreover, the rewiring mechanism described above implies that, in their quest to become central, agents increase the likelihood to be selected by others as new partners.

Notice finally that the number of links is conserved in the dynamics. Hence the density of links is the second important dimension that we shall explore, in order to understand how the structure of social organization depends on it.

3.4 PROPERTIES OF THE DYNAMICS

In the dynamics described above, one can think that the internal state of each agent is define by his/her utility function, and thus it is a function of the topology. The state space of the system can be reduced to the set of states $\omega = (\mathsf{G})$ defined by the adjacency matrix of the system.

The dynamics defined on this space is, as in chapter 2, a Continuous Time Markov Chain defined on a finite state space; a stable state must thus exist.

The constraint that $\sum_{ij} g_{ij} = 2M$ is enforces by the dynamics of the system. It is easy to see, in fact, that the dynamics of the system preserves connected components. Indeed, nodes are never disconnected by the dynamics because, even if they have just one link, this will not be rewired because the neighbour upstream has no second neighbour where to rewire. Therefore, without loss of generality, we restrict attention to the case where $M \geq N - 1$ and the network is composed of a single connected component. Networks composed of disjoint components remain disjoint under the dynamics above, hence the dynamics of different components can be considered independently. Alternative dynamics that do not preserve connectedness – e.g. adding the link (i, l) to a neighbour l of a neighbour j, and removing a link different from (i, j) chosen in any way – would converge to simple structures characterised by cliques of $\sim \sqrt{M}$ nodes in a sea of disconnected nodes. Indeed, it is easy to check that such configurations correspond to absorbing states of the dynamics for all β. On the other hand, as we shall discuss in a moment, it is precisely the rewiring procedure we propose in Section 3.3 that produces non-trivial equilibrium states.

Notice that, since both the number N of nodes and M of edges is conserved during the evolution of the system, the number of fundamental cycles in the graph is also conserved. This follows from the fact that the number of fundamental cycles in a graph is equal to $M - N + K$, where K is the number of connected components (see [8, Ch. 2], appendix **??**). These conservation laws allow us to simplify our discussion to the case where $K = 1$

3.4.1 *Ergodicity of the dynamics*

To proceed further we have to prove that the dynamics of the system is ergodic on the space of adjacency matrices with M edges and one connected component. To do so we use the same strategy used in chapter 2, namely show that given two states ω and ω' it is always possible to reach one form the other with finite number of steps. The states of the system are nothing but the connected graphs with M edges. It is thus sufficient to show that is is possible to pass from one graph to another doing a finite number of rewiring.

Let $\Gamma^C(N, M)$ be the space of connected graphs with N vertices and M edges. In order to prove ergodicity, we have to show that, with a finite number of basic moves, we can reach any connected graph in $\Gamma^C(N, M)$, starting from another arbitrary graph in the same set. Before delving into the technical details, we give a simple intuitive sketch of this proof.

For a finite value of β, the dynamics consists of *reversible* moves as the one depicted in fig. 3.1; such moves can be thought of as a "sliding" of the edge e_{ik} on the path of length one (k, j) from vertex k to vertex j. The key observation to prove the ergodicity by induction is that, since the graph is finite and connected, there always exists a path of minimum length that connects two arbitrarily chosen vertices in the graph. Then, we can proceed in three steps.

1. Let there be two graphs in $\Gamma^C(N, M)$ which differ from each other only by an edge incident on the same vertex, v_k. We first prove that by means of basic moves, we can transform one into the other. To do so, it suffices to slide the edge along the path that connects the other end of the edge, which we know to exist because the graph is connected (Prop. 3.1, Prop. 3.2.

2. Let there be two graphs in $\Gamma^C(N, M)$ that differ by an edge with arbitrary ends. By applying the previous step twice, we show that there exists a finite set of moves that allows us to reach one configuration starting from the other (Prop. 3.2).

3. Finally, let there be two arbitrary graphs in $\Gamma^C(N, M)$. Moving one edge at time, we show by induction that it is possible to reach one graph starting from the other with a finite number of moves. Thus, the ergodicity is proved (Prop. 3.4).

We now proceed with the detailed proofs of egodicity. The uninterested reader can jump directly to section 3.4.2.

DEFINITION 3.1 (c-SWAP). Let us choose a labelling for the space of vertices $V = \{v_1, \dots, v_N\}$ and an induced labelling for the edges $E = \{e_{ij}\}$ where $e_{ij} =$

$e_{ji} = (v_i, v_j)$ denotes the undirected edge between v_i and v_j. Let us define a transformation $\sigma_{ij}^{ik} : \Gamma^C(N, M) \mapsto \Gamma^C(N, M)$, called *corner swaps (c-swaps)*, as following

$$\sigma_{ij}^{ik}(\mathcal{G}) = \mathcal{G}' = (V, E') \tag{3.8}$$

such that

$$E' = \begin{cases} \big(E \setminus \{e_{ik}\}\big) \cup \{e_{ij}\} & \text{if } (e_{kj}, e_{ik} \in E) \wedge (e_{ij} \notin E) \\ E & \text{otherwise} \end{cases}. \tag{3.9}$$

□

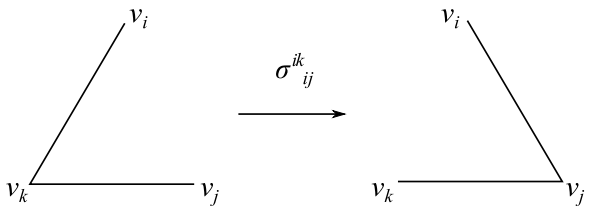


Figure 3.1: The rewiring move (*c*-swap) of the dynamics.

PROPOSITION 3.1. *Let $\mathcal{G} = (V, E)$ and $\mathcal{G}' = (V, E')$ be two graphs in $\Gamma^C(N, M)$ that differ by an edge incident on the same vertex, i.e. $|E| = |E'| = M$, $|E' \cap E| = M - 1$, $E \setminus E' = \{e_{ik}\}$ and $E' \setminus E = \{e_{ij}\}$, and such that the shortest path P from v_k to v_j does not contain neither v_i nor any of its neighbours.*

There exists an integer l and a finite sequence of graphs in $\Gamma^C(N, M)$, $\mathcal{G}^n$ such that:

(i) $\mathcal{G} = \mathcal{G}^0$ and $\mathcal{G}' = \mathcal{G}^l$.

(ii) For all $0 \leq n < l$ there exist adjacent vertices v_{k_n},v_{k_n+1} such that $\mathcal{G}^{n+1} = \sigma_{ik_{n+1}}^{ik_n}(\mathcal{G}^n)$, where $k_0 = k$ and $k_l = j$.

□

Proof. Let l be the length of P.

Let v_{k_1} be the unique neighbour of v_k that lies in P. If we set $\mathcal{G}^1 = \sigma_{ik_1}^{ik}(\mathcal{G})$, the c-swap reduces the distance between v_i and v_j , since the neighbour of v_k that lies in P must have a distance $l - 1$ from v_j. We reiterate the procedure on $\mathcal{G}^1$ and obtain in such a way a sequence of graphs that satisfies property (ii). Now, since at any step the length of P diminishes by 1, after the l-th step, in the graph $\mathcal{G}^l$ v_i and v_j will be neighbours. Thus, since no other edge was changed by applying c-swaps, $\mathcal{G}^l = \mathcal{G}'$ proving property (i). □

PROPOSITION 3.2. *Let $\mathcal{G} = (V, E)$ and $\mathcal{G}' = (V, E')$ be two graphs in $\Gamma^C(N, M)$ which differ by an edge incident on the same vertex, i.e. $|E| = |E'| = M$, $|E' \cap E| = M - 1$, $E \setminus E' = \{e_{ik}\}$ and $E' \setminus E = \{e_{ij}\}$.*

There exists an integer l and a finite sequence of graphs in $\Gamma^C(N, M)$, $\mathcal{G}^n$ such that:

(i) $\mathcal{G} = \mathcal{G}^0$ *and* $\mathcal{G}' = \mathcal{G}^l$

(ii) *For all* $0 \leq n \leq l$ *there exist adjacent vertices* v_{k_n},v_{k_n+1} *such that* $\mathcal{G}^{n+1} = \sigma^{ik_n}_{ik_{n+1}}(\mathcal{G}^n)$.

□

Proof. Let P be the shortest path in $\mathcal{G}$ from v_k to v_j that does not contain (v_k, v_i). There are four possible cases :

(i) P does not contain neither v_i nor any of its neighbours other than v_k. The thesis is proven applying proposition 3.1 directly to P.

(ii) P contains v_i. Let P_1 be the shortest path from v_k to v_i that does not contain the edge (v_k, v_i). let P_2 be the shortest path from v_i to v_j, clearly $P = P_1 \oplus P_2$, where $\oplus$ is the path concatenation. Since by construction there are no neighbours of v_k in P_2 (otherwise P would not contain v_i) we can apply proposition 3.1 and reach $\mathcal{G}'' = (V, (E \setminus \{e_{ki}\}) \cup \{e_{kj}\})$; on the other hand there cannot be neighbours of v_j in P_1 (otherwise there would be a shortest path not containing v_i) and thus applying again proposition 3.1 along P_1 we reach $\mathcal{G}'$ proving the thesis.

(iii) P does not contain v_i but two of its neighbours, c and f such that $c \neq v_k$, $f \neq v_k$ and $|c, v_k| < |f, v_k|$, where $|\cdot, \cdot|$ represents the graph distance between two vertices. We first note that c and f must be neighbours, otherwise P should include v_i. Then, as in case (ii), by minimality we can write $P = P_1 \oplus (c, f) \oplus P_2$ where P_1 is the shortest path from v_k to c and P_2 is the shortest path from f to v_j. It is easy to see that $Q_2 = (v_i, f) \oplus P_2$ is a shortest path from v_i to v_j: if it were not so, there would exist a path Q_2' from v_i to v_j strictly shorter than Q_2, but in that case $P_1 \oplus (c, v_i) \oplus Q_2'$ would be a shortest path from v_k to v_j containing v_i, in contradiction with our hypotheses. A similar argument holds for Q_1. As before, since, by minimality, there cannot be neighbours of v_k in P_2, it is possible to reach the graph $\mathcal{G}'' = (V, (E \setminus \{e_{ki}\}) \cup \{e_{kj}\})$ by applying proposition 3.1 to Q_2; since by minimality there cannot be neighbours of v_j in P_1, we can apply proposition 3.1 to $\mathcal{G}''$ along Q_2 and reach $\mathcal{G}'$ proving the thesis.

(iv) The shortest path P contains only one neighbour of v_i other than v_k, let us call it m. As before, $P = P_1 \oplus P_2$ where P_1 is the shortest path from v_k to m and P_2 is the shortest path from m to v_j. Since by construction there cannot be other neighbours of i in P_2, we can apply proposition 3.1 to P_2 and reach the graph $\mathcal{G}^* = (V, (E \setminus \{e_{im}\}) \cup \{e_{ij}\}$. On the other hand, by construction there cannot be neighbours of v_i in P_1 other than v_k and thus we can apply proposition 3.1 to P_2 and reach $\mathcal{G}'$ proving the thesis.

□

PROPOSITION 3.3. *Let* $\mathcal{G} = (V, E)$ *and* $\mathcal{G}' = (V, E')$ *be two graphs in* $\Gamma^C(N, M)$ *such that* $|E| = |E'| = M$ *and* $|E \cap E'| = M - 1$. *Let us assume that, in particular,* $E = \{e_{ij}\} \cup (E \cap E')$ *and* $E' = \{e_{hk}\} \cup (E \cap E')$.

Thus there exists an integer l and a finite sequence of graphs in $\Gamma^C(N,M)$, $\mathcal{G}^n$ such that:

(i) $\mathcal{G} = \mathcal{G}^0$ *and* $\mathcal{G}' = \mathcal{G}^l$

(ii) For all $0 \leq n < l$ *there exist adjacent vertices* v_{k_n}, v_{k_n+1} *such that* $\mathcal{G}^{n+1} = \sigma_{ik_{n+1}}^{ik_n}(\mathcal{G}^n)$.

□

Proof. Let us define the graph $\mathcal{G}'' = (V, E'')$ such that $E'' = (E \setminus \{e_{ij}\}) \cup \{e_{ih}\}$. Applying proposition 3.2 first to graphs $\mathcal{G}$ and $\mathcal{G}''$ and then to graph $\mathcal{G}''$ and $\mathcal{G}'$ proves the thesis. □

Definition 3.2 (g-swap). Let $\mathcal{G} = (V, E)$ and $\mathcal{G}' = (V, E')$ be two graphs in $\Gamma^C(N,M)$ which differ at most by an edge, that is such that $|E| = |E'| = M$ and $|E \cap E'| = M - 1$. Let us assume that, in particular, $E = \{e_{ij}\} \cup (E \cap E')$ and $E' = \{e_{hk}\} \cup (E \cap E')$.

We define a *global swap or g-swap* of the edge e_{ij} to the edge e_{hk} a transformation such that:

$$\mathcal{G}' = \Sigma_{ij}^{hk}(\mathcal{G}) \tag{3.10}$$

□

Proposition 3.3 simply states that any global swap can be obtained as the composition of a minimal set of corner swaps between adjacent vertices.

Proposition 3.4. *Let $\mathcal{G} = (V, E)$ and $\mathcal{G} = (V, E')$ be two graphs in $\Gamma^C(N,M)$. There exists an integer d and a sequence of graphs $\mathcal{G}^n(V, E_n)$ in $\Gamma^C(N,M)$ such that:*

(i) $\mathcal{G} = \mathcal{G}^0$ *and* $\mathcal{G}' = \mathcal{G}^d$

(ii) For all $0 \leq n < d$ *there exist four vertices* v_i, v_j, v_h *and* v_k *such that* $\mathcal{G}^{n+1} = \Sigma_{ij}^{hk}(\mathcal{G}^n)$

□

Proof. Let $\mathcal{Z} = (V, Z = E \cap E')$, and let us define $\delta = |Z|$. We proceed by induction on the number δ.

base case If $\delta = M - 1$, the Thesis is trivially true because of Proposition 3.3.

inductive step Let us assume that the Thesis holds for $\delta = M - d$, we want to show that this implies that it also holds for $\delta = M - d - 1$, with $d < M - 1$. Let us assume that $\mathcal{G} = (V, E)$ and $\mathcal{G}' = (V, E')$ are such that $|E' \cap E| = M - d - 1$. Let $e_{ij} \in E \setminus (E \cap E')$ and $e_{hk} \in E' \setminus (E \cap E')$. Moreover, let $E'' = (E \setminus \{e_{ij}\}) \cup \{e_{hk}\}$. By construction, $|E \cap E''| = M - 1$ and $|E' \cap E''| = M - d$. Finally, let $\mathcal{G}'' = (V, E'')$. Since $\mathcal{G}''$ and $\mathcal{G}'$ differ by $M - d$ edges, by inductive assumption there exists a sequence $\mathcal{G}^i$, with $i \in [0, d]$, such that $\mathcal{G}^0 = \mathcal{G}'$ and $\mathcal{G}^d = \mathcal{G}''$, that satisfies the Thesis. Moreover, by Proposition 3.3, there exists a g-swap such that $\mathcal{G} = \Sigma_{hk}^{ij}(\mathcal{G}'')$. Thus, the complete sequence $\mathcal{G}' = \mathcal{G}^0, \mathcal{G}^1, \cdots, \mathcal{G}'' = \mathcal{G}^d, \mathcal{G} = \mathcal{G}^{d+1}$ satisfies the Thesis.

□

Proposition 3.4 and proposition 3.3 state simply that any two connected graphs with the same number of edges can be obtained one from the other applying a finite sequence of c-swaps. Moreover, since the number of edges is finite, then there must be a minimal sequence of c-swaps that connects any two of such graphs. Since, for finite β, all c-swaps are allowed with non-zero probability, this proves the ergodicity.

3.4.2 *Potential function and Hamiltonian*

The fact that the system is ergodic ensures us that there exist a unique stationary distribution. In this case it is possible to obtain the shape of the invariant distribution directly from the definition of utility.

The dynamics of the model, in fact, admits a potential which is just the *global* utility, i.e. the sum of the utilities $U = \sum_i u_i$. Indeed, let us consider the change Δu_x in the utility of the agent x when the rewiring (i, j) into (i, l) occurs. Depending on the position of x in the network, the following changes are obtained:

$$\Delta u_i = k_l - k_j + 1 \tag{3.11a}$$
$$\Delta u_j = 1 - k_i - \mu \tag{3.11b}$$
$$\Delta u_l = k_i - 1 + \mu \tag{3.11c}$$
$$\Delta u_h = -1 \qquad \forall h \in \partial j \setminus \{i, l\} \tag{3.11d}$$
$$\Delta u_g = +1 \qquad \forall g \in \partial l \setminus \{j\} \tag{3.11e}$$
$$\Delta u_x = +0 \qquad \forall x \neq i, j, l, x \notin \partial j \cup \partial l\,, \tag{3.11f}$$

where ∂x is the set of the neighbours of x, before the move.

In the total variation of the utility $\Delta U = \sum_x \Delta u_x$, the term Δu_h appears $k_j - 2$ times, while the term Δu_g appears $k_l - 1$ times, because k_x is the degree of the node x *before* the rewiring. Gathering all the contributions one has:

$$\begin{aligned} \Delta U &= \Delta u_i + \Delta u_j + \Delta u_l + (k_j - 2)\Delta u_h + (k_l - 1)\Delta u_g \\ &= 2(k_l - k_j + 1) = 2\Delta u_i\,. \end{aligned} \tag{3.12}$$

The last point implies that, since the dynamics is ergodic, the system converges to thermal equilibrium with Hamiltonian

$$\mathscr{H} = -U = -\sum_i k_i^2 - \mu \sum_i k_i,$$

and fixed density of links at temperature $2/\beta$[10]. Notice that the second term does not play any role, being $\sum_i k_i$ a fixed quantity in our case. Indeed the dynamics in Eq. (3.7) is equivalent to Metropolis dynamics, and hence it samples the Gibbs distribution $P\{\hat{a}\} \propto e^{\beta U(\hat{a})/2}$, which is known in sociology as the 2-star model.

10 The factor 2 comes from the fact that the variation of the global utility is the double of the variation of the local utility.

Park and Newman [115, 116], have shown that the 2-star model where the density of links is not fixed, exhibits a sharp phase transition. This result suggests that there might be a phase transition also in the model we study in this paper. As a by-product, our discussion also provides a microscopic derivation for the 2-star model[11].

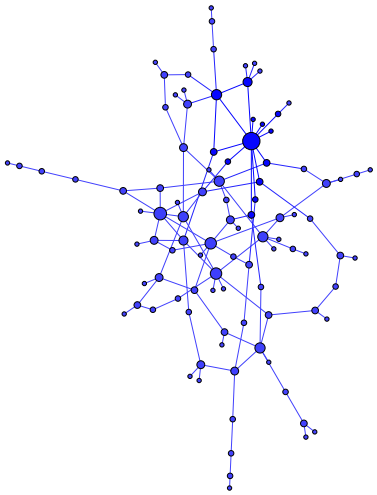
3.5 NUMERICAL SIMULATIONS

In order to investigate the behaviour of the model, we performed extensive numerical simulations sampling the Gibbs distribution $P\{\hat{a}\} \propto e^{\beta U(\hat{a})/2}$ using the Metropolis algorithm based on the rewiring moves introduced in Sect. 3.3. All the results to be presented throughout the rest of this section were obtained, for each value of β, by performing R rewiring proposals per node, and we checked that the value $R = 5 \cdot 10^5$ is large enough to always ensure the attainment of an equilibrium state. Fig. 3.2 shows two typical realizations of the social network for small and large values of β (see caption for more details). Fig. 3.2 suggests that, as anticipated in the previous section, the social climbing model undergoes a transition from hierarchical to random structures. In the following, we will show the presence of a phase transition between these two states.

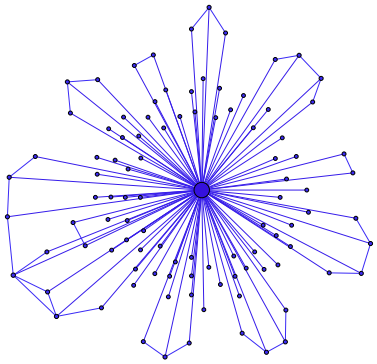
In Fig. 3.3 we show the largest degree of the network $\Phi = \max_i k_i$ as a function of the inverse temperature β for systems with $N = 100$ nodes and $M = 110$, 200, 300, 500 links. As can be seen, in all cases the system actually undergoes a transition, going from a phase where the largest degree Φ is roughly of order $1 - 10$ (depending on the relative size of N and M) to a phase where the largest degree is of order N. These observations qualitatively match the findings of [117], where a prediction for the critical temperature $T_c = 1/\beta_c$ characterizing this phase transition was also derived, from combinatorial arguments, for networks with average degree $\bar{k} = 2M/N < 2$, i.e. for disconnected graphs. The nature of the phase transition depicted in Fig. 3.3 is further investigated in Fig. 3.4. In the left panel we show the full distribution of Φ/N with respect to $\bar{k}$ obtained by binning the results relative to 100 networks, for $\beta = 0.01$ and $N = 500$. For low (high) values of $\bar{k}$ the distribution is sharply concentrated around zero (one) and a steep transition occurs at a critical value of $\bar{k}$, meaning that the average is representative of the distribution of Φ/N. Completely analogous results are found for different values of β. Therefore, in order to characterise the transition more precisely, in the right panel of Fig. 3.4 we show the relation between the average of Φ/N over 100 networks with respect to both $\bar{k}$ and β.

In Fig. 3.5 we analyse the dependence of the critical value of β with respect to the size N of the network, while keeping the average degree fixed, for $\bar{k} = 2.5,\ 5.0$. Qualitatively it is clear that, increasing N, both the transition becomes sharper

11 The case studied in ref. [115, 116] where the number of links is allowed to change, can be recovered in a model where, in addition to rewiring steps discussed above, we also allow for link creation upon random encounters and link obsolescence (i.e. decay). More precisely, consider a model where each agent receives opportunities *i)* to rewire his/her links (as above) at rate ν and *ii)* to form new links (with randomly chosen agents), with rate $\eta/2$. In addition, each link decays with rate 1. Then, in a time interval Δt, the number of links changes by $\Delta M = \eta N \Delta t - M \Delta t$, which means that in the stationary state $\langle M \rangle = \eta N$.



(a)



(b)

Figure 3.2: Snapshot of networks of the social climbing game for $N = 100$, $M = 125$ for $\beta = 0.03$ (a) and $\beta = 0.1$ (b). Size of the nodes is proportional to the degree.

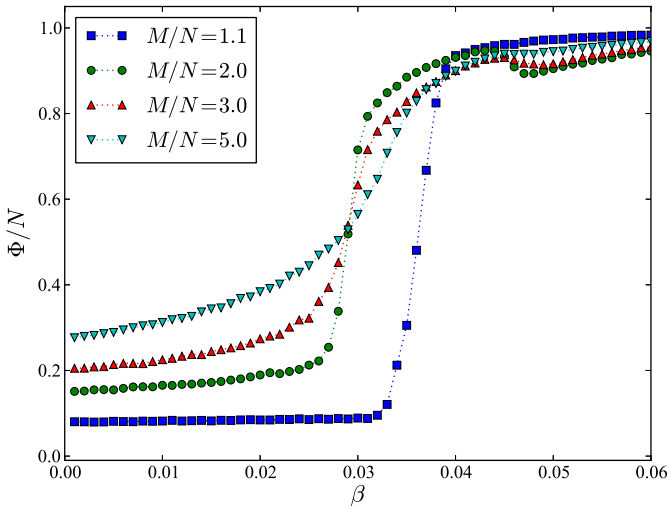
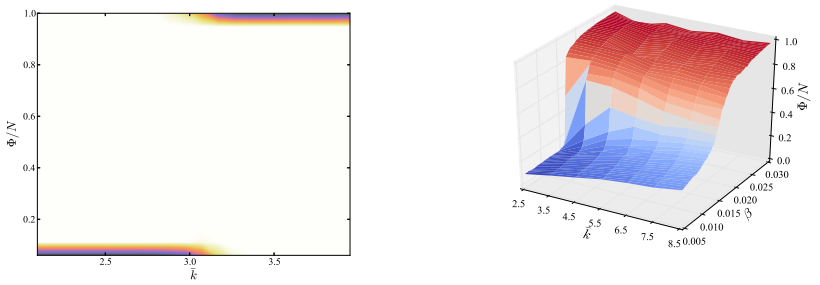
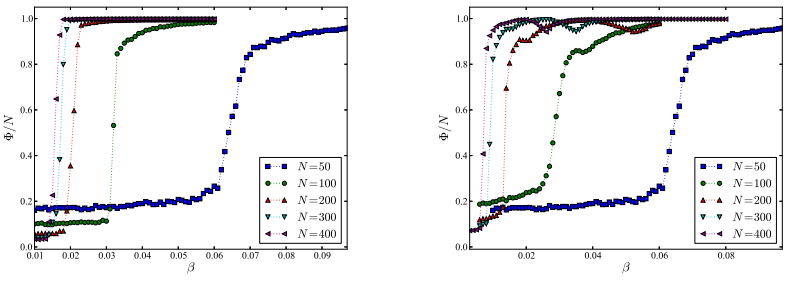


Figure 3.3: Dependence of the largest degree Φ (divided by N) in the social climbing network as a function of the inverse temperature (or intensity of choice parameter) β. The different curves refer to $N = 100$ and $M = 110, 200, 300, 500$. For each value of β the reported values of Φ are obtained by averaging over 100 networks. An abrupt change in is observed in all curves after a threshold value of β, with Φ/N going from low values to values close to one, signalling the emergence of a star, i.e. a link with $\mathcal{O}(N)$ links, in the network.



(a) (b)

Figure 3.4: (a) density plot depicting the full distribution of the maximum degree (divided by N) as a function of the average degree obtained by binning the results relative to 100 networks, at inverse temperature $\beta = 0.01$ and number of nodes $N = 500$. High (low) values are darker (lighter). (b) average maximum degree (divided by N) as a function of the average degree and β for $N = 100$; results obtained by averaging over 100 networks. Clearly, for low values of $\bar{k}$ the network is in the disordered phase, while for high values of β it is in the ordered phase.



(a) (b)

Figure 3.5: (a) Dependence between the maximum degree (divided by N) and β, for $\bar{k} = 2.5$ and different values of N. Results are averaged over 100 networks. The transition become sharper and the critical value of β shifts to the left for increasing values of N. (b) as in the left panel for $\bar{k} = 5.0$.

and the critical value of β shifts to the left. In order to understand if the critical value in thermodynamic limit β_c is non-zero, we analyse the finite-size scaling behaviour, assuming $\beta^*(N) = \beta_c + aN^{-b}$, where $\beta^*(N)$ is the critical value at size N and β_c, a and b are free parameters. Since b is expected to be universal (i.e. not dependent on the other parameters, like $\bar{k}$), it is reasonable to choose it by plotting $\beta^*(N)$ against N^{-b} until straight lines are obtained. Both a and β_c are then found by a best fit. The value of $\beta^*(N)$ is obtained by a linear interpolation of the curves in Fig. 3.5 and calculating the value of β such that $\Phi/N = 1/2$. From Fig. 3.6 it can be clearly seen that for $b = 1.25$ the assumed functional form is fully consistent with numerical simulations up the investigated system size. The values we find for β_c are soundly different from zero within 95% confidence intervals provided by best fit. In particular we find for $\bar{k} = 2.5$: $\beta_c = (1.1 \pm 0.2) \cdot 10^{-2}$, while for $\bar{k} = 5.0$: $\beta_c = (2.6 \pm 0.2) \cdot 10^{-3}$.[12]

Following [117], let us define a star as a node whose degree is of order N. Then, it is immediate to figure out that, depending on the ratio M/N, different number of stars might emerge in the network for temperatures lower than T_c. Clearly, in the case $N = 100$, $M = 110$ (i.e. $\bar{k}$ only slightly larger than 2), the appearance of a star ($\Phi \simeq 100$ in this case) below the critical temperature leaves very few links to be distributed amongst the remaining nodes. On the other hand, increasing the number of links provides enough room for the emergence of a larger number of stars. In other words, it is intuitively reasonable to expect a system with an average degree $\bar{k} \simeq 2n$ to produce, for sufficiently low temper-

12 Inspired by [117] we also performed finite-size scaling analysis according to the functional form: $\beta^*(N) = \beta_c + a(M/\log(N))^{-b}$, which also gives values of β_c soundly above zero and consistent with the ones discussed in the main text.

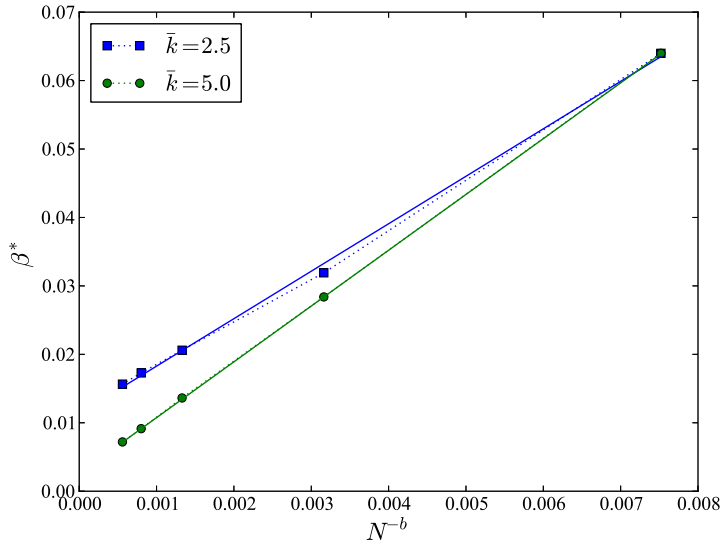


Figure 3.6: Finite-size scaling for determining the critical value in the thermodynamic limit β_c, assuming the functional form $\beta^*(N) = \beta_c + aN^{-1.25}$, where $\beta^*(N)$ is the critical value at size N. By best fitting we find: for $\bar{k} = 2.5$: $\beta_c = (1.1 \pm 0.2) \cdot 10^{-2}$, while for $\bar{k} = 5.0$: $\beta_c = (2.6 \pm 0.2) \cdot 10^{-3}$.

atures, exactly n stars. In order to support such an intuitive line of reasoning, we computed the inverse participation ratios (IPRs) of the degree sequences $\boldsymbol{v} = (k_1, k_2, \dots, k_N)/\sqrt{\sum_{i=1}^{N} k_i^2}$ of several networks with different numbers of nodes and links. Given a normalised vector $\boldsymbol{v}$, its IPR is defined as

$$I(\boldsymbol{v}) = \left(\sum_{i=1}^{N} v_i^4 \right)^{-1} . \tag{3.13}$$

The IPR of a completely localised vector, say $\boldsymbol{v} = (1, 0, \dots, 0)$, is equal to one. On the other hand, the IPR of a fully de-localised vector, whose components are all equal to $v_i \simeq 1/\sqrt{N}$, is of order N. In our case $I(\boldsymbol{v})$ gives an estimate of the number of dominant nodes in the network. In Fig. 3.7 we plot the IPR $I(\boldsymbol{v})$, as functions of β, for $N = 100$ nodes and $M = 110, 200, 300$. As can be seen, the IPR of the sparsest network, i.e. the one with $M = 110$, essentially drops down to one right below its critical temperature. On the other hand, systems with a larger number of links undergo a less trivial evolution: after the initial drop below the critical temperature, the IPR increases and eventually reaches a steady state. In the example shown in Fig. 3.7, the system with $M = 200$ links reaches a steady value $I(\boldsymbol{v}) \simeq 2.12 \pm 0.03$, whereas the system with $M = 300$ reaches $I(\boldsymbol{v}) \simeq 3.28 \pm 0.06$ (where the errors represent the 68% confidence intervals obtained by averaging over 100 networks), and such values clearly show that the maximal number of stars allowed by the relative sizes of N and M has been achieved. Moreover, these observations are consistent with the small temporary decrease of the largest degree Φ which can be observed in Fig. 3.3 for systems with $\bar{k} > 2$ when the inverse temperature is slightly larger than its critical value.

3.5.1 *Correlations and social mobility*

As already explained in Sect. 3.2, one of the goals of the present paper is to model the positive feedback mechanism between the individuals' effort to climb the social hierarchy and the subsequent reinforcement of the social hierarchy itself. Suppose that a given social network reaches its equilibrium state, at a certain inverse temperature β, after t_0 steps of the social climbing dynamics described in Sect. 3.3. Let us denote the corresponding graph's adjacency matrix as $\mathsf{G}(t_0)$. Then, one way of quantitatively describing how mobile or "frozen" a society is would be to assess the level of correlation, according to some proper notion, between $\mathsf{G}(t_0)$ and a following configurations $\mathsf{G}(t)$, where $t = t_0 + \Delta t$ for some positive Δt. We will now measure correlations by means of Kendall's rank correlation coefficient. Given the joint set of all matrix entries in $\mathsf{G}(t_0)$ and $\mathsf{G}(t)$, let us focus, for example, on entries (i, j) and (h, l) in both matrices. Then, if both $g_{ij}(t_0) > g_{hl}(t_0)$ and $g_{ij}(t) > g_{hl}(t)$, or if both $g_{ij}(t_0) < g_{hl}(t_0)$ and $g_{ij}(t) < g_{hl}(t)$, the pairs $(g_{ij}(t_0), g_{hl}(t_0))$ and $(g_{ij}(t), g_{hl}(t))$ are said to be concordant. On the contrary, if $g_{ij}(t_0) \gtrless g_{hl}(t_0)$ and $g_{ij}(t) \lessgtr g_{hl}(t)$ the pairs $(g_{ij}(t_0), g_{hl}(t_0))$ and $(g_{ij}(t), g_{hl}(t))$ are said to be discordant. Of course, since the adjacency matrix entries equal zero

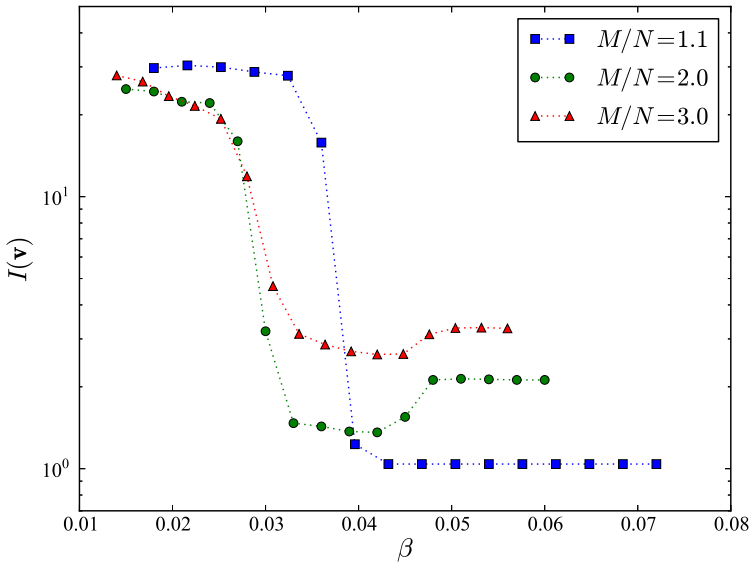


Figure 3.7: Inverse participation ratio of the normalised degree sequence $I(\boldsymbol{v})$ as a function of the inverse temperature β. Different curves refer to networks with $N = 100$ and $M = 110, 200, 300$. Results obtained by averaging over 100 networks. For large enough β one recovers the maximal number of stars.

or one at each time, ties will often happen either at time t_0 or at time t (or at both times). Kendall's correlation coefficient τ reads

$$\tau(\Delta t) = \frac{C - D}{\sqrt{C + D + T_{t_0}}\sqrt{C + D + T_t}}, \tag{3.14}$$

where C (D) is the numbers of concordant (discordant) pairs, whereas T_{t_0} (T_t) denotes the number of time-t_0 (time-t) ties. Pairs where ties happen both at t_0 and t are not taken into account.

In Fig. 3.8 a few examples of Kendall's τ coefficient's time evolution are sketched. All plots refer to networks with $N = 100$ nodes and $M = 300$ links. Here, $\Delta t = t_0 = 5 \cdot 10^6$ elementary Monte Carlo moves, i.e. rewiring proposals. As can be seen, when the social climbing game takes place for temperatures higher than the critical one, Kendall's τ quickly starts to fluctuate around zero, denoting no genuine correlation between configurations distant (in time). This we take as indication of a large social mobility. On the other hand, for temperatures slightly lower than the critical one, Kendall's τ remains significantly larger than zero over several time lags. However, a downward trend is clearly visible in this case, meaning that for temperatures $T \lesssim T_c$ social mobility is recovered after a sufficiently long time. On the contrary, for temperatures significantly lower than the critical one Kendall's τ essentially remains constant and very large (i.e. close to one) over large time lags, hinting at an extremely reduced social mobility, possibly preventing the majority of individuals from climbing the social ladder.

The above considerations on individuals' mobility in the social climbing game can be further clarified and understood more deeply. For these purposes, let us denote as q_i the fraction of agents who, at a given time, have a strictly lower degree than agent i, i.e.

$$q_i = \frac{1}{N} \sum_{j \neq i} \theta(k_i - k_j), \tag{3.15}$$

where $\theta(x) = 0$ for $x \leq 0$ and $\theta(x) = 1$ for $x > 0$. The variable defined in (3.15) clearly represents a suitable definition of the social ranking, hence the social status, of a given individual in the network. Thus, a reasonable measure of the individuals' mobility in the social climbing game is given by the change in the quantity defined above over a certain time lag Δt, i.e. $\Delta q_i(\Delta t) = q_i(t + \Delta t) - q_i(t)$, for $i = 1, \ldots, N$.

In Fig. 3.9 some typical behaviours of the q index defined in (3.15) are shown. All examples refer to networks with $N = 100$ nodes and $M = 1000$ links. In such plots, the average of the change Δq is shown as a function of q, in order to provide information about the typical social mobility over a time lag Δt for an agent whose social ranking at the beginning of such a time lag is quantified by q. As can be seen, depending on the preference for social status, i.e. on the inverse temperature parameter β, very different situations can happen. In a rather disordered society (low values of β) the relation between q and Δq is

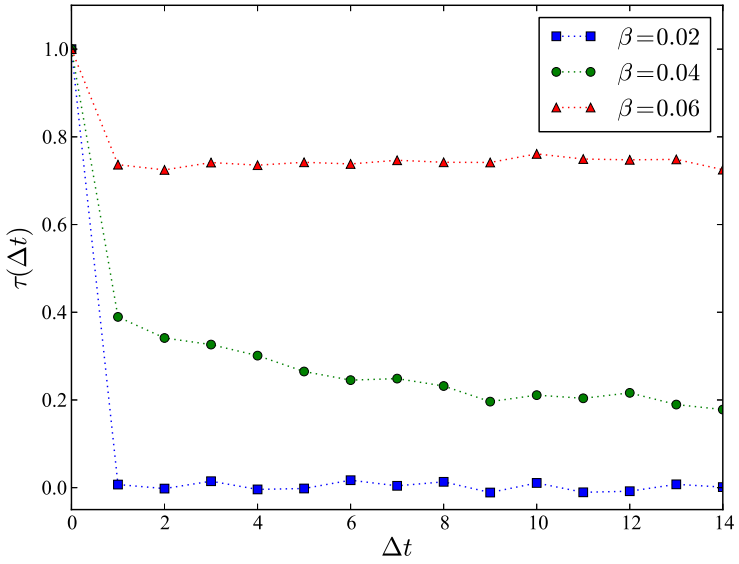


Figure 3.8: Kendall's τ coefficient (see (3.14)) measurements for networks with $N = 100$ and $M = 300$. All measurements are performed between an initial equilibrium configuration $\mathsf{G}(t_0)$ and later configurations $\mathsf{G}(t_0 + n\Delta t)$, with $t_0 = \Delta t = 5 \cdot 10^6$ Monte Carlo steps. The different curves refer to inverse temperatures $\beta = 0.02, 0.04, 0.06$, respectively corresponding to values below, slightly above and well above the critical value for the system under study (see also Fig. 3.3). Results obtained by averaging over 100 networks.

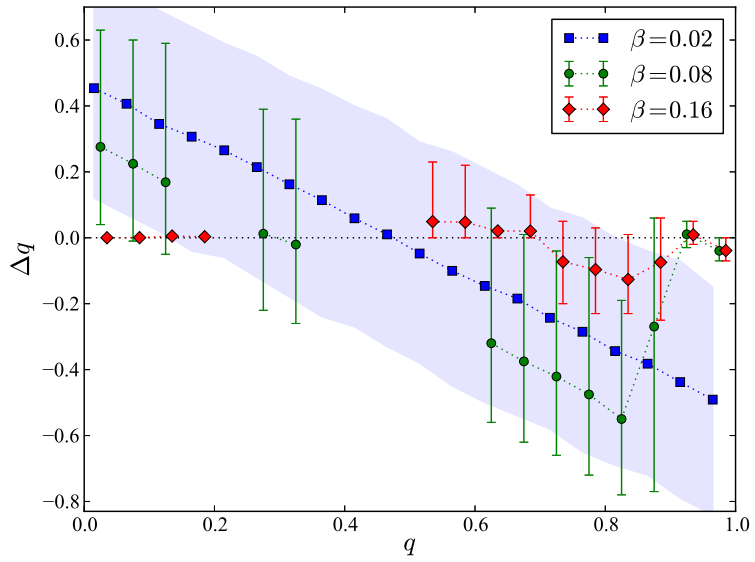


Figure 3.9: Relation between the q index defined in (3.15) and its variation Δq over a given time lag $\Delta t = 10^4$ Monte Carlo steps for a network with $N = 100$ and $M = 1000$. The different curves refer to inverse temperatures $\beta = 0.02$, 0.08, 0.16, respectively corresponding to values below, slightly above and well above the critical value for the system under study. Points refer to the average variation Δq over an equally spaced grid of q values (going from 0 to 1 in steps of 0.05). Results obtained by averaging over 100 networks. Shaded area (for $\beta = 0.02$) and error bars (for $\beta = 0.08, 0.16$) represent the central 68% of the events. Points and error bars relative to different values of β have been shifted to enhance readability.

clearly linear, and does not depend strongly on the time lag size Δt. In particular, it can be seen that, on average, individuals sitting at the bottom of the ranking typically end up higher in the social ladder after some time, whereas individuals sitting atop the hierarchy are prevented from keeping their social status intact for a long time. When the preference for social status crosses its critical value, such a picture starts changing quite dramatically. For values of β slightly larger than the critical value β_c agents with low degrees still have a chance to climb up the social ladder, especially over rather long time lags, whereas the dominant individuals ($q \gtrsim 0.9$) typically get to keep their social ranking. It is worth to remark that for low β the distribution of k_i is not very skewed, so changes in social ranking Δq_i are more frequent. In this sense, our notion of social mobility captures aspects related to social dynamics but it also depends on the stationary distribution of q_i's, i.e. on the degree of inequality. As β increases, the degree distribution acquires skewness, with few individuals having many links and reduced social mobility. For $\beta > \beta_c$ the population separates into two groups, those with k_i of order N and those with very few links, with suppressed mobility across the whole social hierarchy. Also, as can be seen from the right plot in Fig. 3.9, when the critical threshold is crossed the social network becomes "fragmented", as the q index is no longer defined over the whole interval $[0, 1]$. In a strongly ordered society, i.e. β well above its critical value, agents with low degrees are almost completely stuck, and all of the social mobility happens in the top half of the social network, i.e. amongst agents with $q > 0.5$, and this is precisely due to the freezing of the dominant individuals inducing social mobility to disappear completely also amongst nodes with small degrees. These results complement, at a "microscopic" level, those presented in Fig. 3.8.

3.6 CONCLUSIONS

In summary, we have discussed a very simple model for the dynamics of a social network where the agents' quest for high status in the social hierarchy reinforces the latter while reducing social mobility. The model is very stylised and far from a realistic description of social dynamics. Yet, it captures some key ingredients that are enough to reproduce stylised facts known at least since the work of Vilfredo Pareto [100]. Namely, Pareto observed that societies tend to organise in a hierarchical manner, with the emergence of "social elites"[13]. Our model, as well as Refs. [107, 110], provides a formal framework showing that individual incentives for high social status are enough to confer this property to the social network, even in the absence of explicit discrimination of particular groups (e.g. cast system or racial segregation) or preferential biases (e.g. hereditary rules).

This model suggests that the assumption that individual freedom promotes social mobility is a non-trivial one. This is because the structure of a society, while constraining the set of opportunities that are available to individuals, depends on the very incentives of individuals in complex ways.

We have shown, in fact, that the interplay between agents that are free to modify their neighbourhood in order to increase their centrality can produce

13 A similar concept of "power elites" has been discussed in [118].

very "rigid" societies, with extremely low social mobility, characterised by persistent inequalities between *a priori* equal individuals We find, in fact, that the hierarchical state is remarkably stable, with suppressed social mobility in the upper and lower parts of the hierarchy. Furthermore, our model exhibits a negative dependence between mobility and inequality, in the sense that more hierarchically structured (i.e. unequal) societies manifest a lower degree of mobility. It is tempting to relate this to the pervasive empirical observation that more unequal societies tend to have lower inter-generational mobility [119–121]. Our model neglects important dimensions, such as wealth or political power that, however, likely contribute to reinforce our results.

Another important aspect of the model is the fact that social climbing game admits a potential function, thereby allowing us to deploy techniques and concepts of statistical mechanics to understand the behaviour of the system. The understanding of this phenomenon, indeed, hinges on the concept of ergodicity breaking that occurs in strongly interacting systems, when a symmetry – here related to the *a priori* equality among individuals – is spontaneously broken. This phenomenon, well known in statistical physics, is an emergent collective property, and it manifests only when the system is large enough. Remarkably, we find that persistent inequality with low mobility occurs precisely when the quest for "power" – i.e. for occupying the most central or important place in the social hierarchy – becomes a dominant component of what motivates the behaviour of individuals.

In words, thus, our model epitomises an apparent positive feedback between the intensity of the efforts of individuals to "climb" the social hierarchy and the structure of a society: on the one hand, the more a society is hierarchically structured, the easier it is for individuals to understand how to climb it. On the other, the efforts of agents to climb the hierarchy reinforce the social ranking as individuals rewire their links from less to more influential individuals. We discuss this interplay in a highly stylised model of a society, that while being very far from realistic, serves as a proof of concept and allows us to unveil the mechanism responsible for the emergence of a persistent inequality in a transparent manner.

We solved this model mostly through numerical simulations. The mean field approach discussed in Ref. [115, 116] is not applicable in our case, because the density of links that plays the role of an order parameter in Ref. [115, 116] is fixed in our case. Indeed, the phenomenology we find is different from that of Ref. [115, 116] as we do not find evidence of hysteresis phenomena: there is no range of parameters where the disordered and the ordered societies are both stable. We speculate that this might be related to the fact that in the social climbing game there are mechanisms by which a social hierarchy can "nucleate" gradually in an ordered society, by forming social elites that grow over time.

This and other issues can in principle be addressed within more sophisticated statistical mechanics approaches. In this respect, it is worth to mention that it is possible to map the problem into that of an interacting lattice gas that possibly admits for a full and exact statistical mechanics treatment.

Part III

CONCLUSION

4

CONCLUSION

In this thesis we have focused on modelling decision making in complex social environments. The two specific case studies are completely different in their nature: in chapter 2 we described a model for consensus decision making in large animal groups, in chapter 3, instead, we modelled a society of rational self-interested individuals that strive to enhance their importance inside the society in which they live.

In both cases, notwithstanding the deep and substantial differences, we could use a similar theoretical framework to address the decision problem. This was made possible by the high degree of abstraction that network models have. As long as it possible to offer a reasonable definition of *social interaction* and as long as it is reasonable to stylise the behaviour of the individuals interacting in a society in a simple way, a network approach is viable.

Of course, this abstraction comes with a price: as Ising model is not a good model to describe quantitatively the details of magnetism in ferromagnetic metals, it is unlikely that we shall be able to exact quantitative results from network models, possibly measurable in real experiments; however, as for Ising model in Statistical Physics, these network models offer us a qualitative picture of the behaviour of the system that we are modelling, and highlight the possible mechanisms that originate these behaviour.

In some cases—as the one discussed in chapter 2— it is possible to attempt a more realistic microscopic description of the system, at the price,however ,of loosing any hope for a analytical solution. Still, in most cases we still miss a more realistic microscopic description and thus the network description is the best possible one(this is the case discussed in chapter 3).

A general feature of adaptive networks is the fact that there is a positive feedback between the inner dynamics of the agents and the social one: in chapter2, the adaptive change of the neighbourhood reinforce the decision made by the individuals, in chapter 3 the change of the neighbourhood, driven by the desire of centrality, leads the birth of local hierarchies that reinforce the adaptive dynamics; as a result of these positive feedbacks, for some range of the parameters, the system shows abrupt transition between disordered and ordered phases (such behaviours have also been found in other adaptive models [15, 63, 64, 78]).

Such abrupt changes are indeed a general feature observed in biology and sociology which has for long puzzled scientist; early attempts to provide a general, mainly phenomenological, framework to model these phenomena , such as catastrophe theory [122] , essentially failed, and a general microscopic theory for modelling these phenomena is still missing. Adaptive networks may provide a possible reasonable microscopic approach able to justify some of these patterns.

Finally, if on the one hand the model we discussed in this thesis provided a better insight on the two problems we have addressed, they are not perfect and miss some aspects. A particular relevant one, is the emergence of communities.

In chapter 2 states with more than one community appear as metastable states of the dynamics, however these states eventually vanish in the long run. Similarly in the model discussed in 3 the stable state is either completely disordered or completely hierarchic. Understanding how to extend the kind of models we have discussed in this thesis to account for the emergence of stable communities is a modelling challenge in which is worth investing time

Part IV

APPENDIX

A

APPENDIX TO CHAPTER 1

A.1 GRAPHS: SOME CONCEPTS AND NOTATION

The mathematical foundation of network science is graph theory, a branch of combinatorics and topology, with a long tradition [8]. Its origins are traditionally tracked back to the famous solution to the problem of the Seven Bridges of Königsberg that Leonard Euler found in 1735. During its first two centuries of development, graph theory showed strong bonds with other branches of mathematics such as topology, group theory, number theory, and self organised in several branches. Combinatorial Graph Theory— the branch of Graph theory that on counting graphs which given properties— has in particular offered some of the most relevant result for network science. In particular a paper by Paul Erdős and Albert Rényi [9] and (independently by Gilbert[10]) focused on applying the so called *probabilistic method* to graph theory.

This seminal paper started an entire new line of research in graph theory, the random graph theory, where rather that looking to the property of a single graph, one focuses on the average properties of graphs satisfying certain properties. This random graphs can be seen as a product of some random process that generates them and ,as such, they offer an ideal modelling framework adapted to capture the nature of real world networks.

The increase of computational power available to researcher and the development of the Internet, which offer a easily accessible experimental set-up, have boosted the empirical studies of real world networks. Such studies have made,however, clear that Erdős-Rényi (ER) model was not able to capture the nature of the real networks. ER model, in fact, completely ignores the correlation between elements in a network and proved to be unable to reproduce some topological properties of various empirical networks. The mathematical research in random graphs continued in the attempt to build better models; an acceleration in this modelling effort occurred at the end of the 90s when new empirical results came out and several seminal papers appeared. In particular we cite here two popular models, Watts and Stogats *Small World* model [11] and the Barabasi Albert model [12] . In the last fifteen years a plethora of network models has been developed, and network models have been applied to a rich variety of systems.

In this section we shall review briefly some basic concepts of graph theory with the end of setting up the notation and nomenclature that we shall use afterwards.

A.1.1 *Graphs*

Given a set V, *a graph* $\mathcal{G}$ is defined as an ordered pair of disjointed sets $\mathcal{G}(V, E)$ where $E \subseteq V \times V$.

The definition is general but usually V is at least countable. If $|V| \leq \infty$, the graph is said to be *finite*. For finite graphs V it is always possible to map V in the set formed by the first $|V|$ positive integers. If not otherwise specified, V must be considered as that subset.

An element $v \in V$ is called a *vertex* or *a node*; the sets of nodes of a graph shall be written as $V(\mathcal{G})$. A pair $e \in E$ is called an *edge* or a *link*; the set of edges of a graph shall written as $E(\mathcal{G})$. If we consider E as ordered couples we obtain a *directed graphs* , if instead we consider the pairs unordered we obtain an *undirected graph*.

Two vertices v and w for which $(v, w) \in E$ or $(w, v) \in E$ are said to be *adjacent* or *neighbours*. The set $\mathcal{N}_{\mathcal{G}}(v) = \{w \in V(\mathcal{G}) \ : \ (v, w) \in E(\mathcal{G})\}$ is called the *neighbourhood of* v[1]. Given an edge$e_{vw} = (v, w)$ we shall also say that the vertices v and w are incident on the edge e.

A subgraph $\mathcal{G}'(V', E')$ is a graph such that $V' \subseteq V$ and $E' \subseteq E$. We shall write $\mathcal{G}' \subset \mathcal{G}$

.If no pairs of the form (i, i)(self edges) are permitted, the graph is *simple*. If more than one edge is allowed, we refer to it a *multigraph*.

A.1.2 *Isomorphism, graph properties and graph invariance*

Two graphs $\mathcal{G}(V, E)$ and $\mathcal{G}'(V', E')$ are said to be *isomorphic*, and we shall write $\mathcal{G} \sim \mathcal{G}'$ if there exist a function $\varphi : V \mapsto V'$ such that $\forall (i, j) \in V \times V$, if *inE* then $(\varphi(i), \varphi(j)) \in E$'.

A class of graphs that is closed under isomorphisms is called a *graph property* (e. g. planarity , containing a triangle etc..).

A map ψ defined on graphs is a graph invariant if $\psi(\mathcal{G}) = \psi(\mathcal{G}')$ for all $\mathcal{G} \sim \mathcal{G}$ (e. g. the number of vertices, of edges, the average degree etc.)

A.1.3 *Degree of a vertex, average degree*

Let $\mathcal{G}(V, E)$ be a non empty undirected graph. And let $v \in V$ be a node. The quantity $k_v = |\mathcal{N}_{\mathcal{G}}(v)|$ is called the *degree of vertex* v .

The degree of a vertex is not a graph invariant (since it depends on the labelling of the vertices). However we can construct different graph invariants out of them. In particular we can define the minimal degree $k_{\min} = \min_{v \in V}(k(v))$, the maximal degree $k_{\max} =_{v \in V} (k(v))$ and *average degree of* $\mathcal{G}$

$$\bar{k} = \frac{1}{|V|} \sum_{v \in V} k_v. \tag{A.1}$$

1 If the graph is directed this expression defines out neighbourhood. A similar expression with w and v inverted shall define instead the in neighbourhood

Since in an undirected graph each edge connects only two vertices on can easily check that $\sum_{v \in V} k(v) = 2|E|$ and thus

$$\bar{k} = \frac{2|E|}{|V|}. \tag{A.2}$$

For undirected graphs another relevant invariant is the degree sequence:

$$D_{\mathcal{G}} = \{d_i\}_{i=1}^{|V|} \tag{A.3}$$

such that

1. $a_i \leq a_j$ for all $i, j \in \{0, \dots, V\}$
2. $\forall i \in \{0, \dots, |V|\},\ \exists v \in V : a_i = k_v$.

Degree sequence of a graph bears important information about the topology of the graph and has to satisfy the condition of equation A.2 and of Erdős-Gallai theorem [123]:

$$\sum_{1}^{k} a_i \leq k(k+1) + \sum_{k+1}^{|V|} \min(a_j, k). \tag{A.4}$$

A.1.4 *Adjacency Matrix*

The structure of a finite graph may be described in synthetic way using a matrix, called the *adjacency matrix*.

Given an graph $\mathcal{G}(V, E)$ it is possible to define a $V \times V$ matrix, $\mathsf{G} = (g_{ij})$ as follow:

$$g_{ij} = 1, \quad \text{if } (i, j) \in E \tag{A.5}$$

$$g_{ij} = 0, \quad \text{otherwise.} \tag{A.6}$$

If the graph is *undirected* the matrix shall be symmetric with null trace, if the graph is is *directed* it might be that $g_{ij} \neq g_{ji}$. Since G has the same amount of information of E we shall often use the notation $\mathcal{G}(V, \mathsf{G})$ to refer to a graph.

It is straightforward, by definition, to see that:

$$k_i = \sum_j g_{ij} \tag{A.7}$$

and thus

$$\sum_{ij} g_{ij} = 2|E|. \tag{A.8}$$

In particular the average degree reeds:

$$\bar{k} = \frac{\sum_{ij} g_{ij}}{|V|} \tag{A.9}$$

A.1.5 *Paths, distances connected components*

Let $V = \{v_1, \dots, v_n\}$ and $E = \{(v_1, v_2), \dots, (v_{n-1}, v_n)\}$. The graph $P_{v_1,v_n}(V, E)$ is called a *path*. The length of a $l(P)$ of a path is $|E|$ and, by definition is equal $|V| - 1$.

Let $\mathcal{G}(V, E)$ be a graph and $v_1, v_2 \in V$. If there exists a path P_{v_i,v_2} which is also a subgraph of $\mathcal{G}$, we say that v_1 and v_2 are *connected (or linked)*. In this case we call P_{v_i,v_j} a path from v_i to v_j on the graph $\mathcal{G}$.

Let $\mathcal{P}_{v_i,v_j}(\mathcal{G}) = \{P_{v_i,v_j} \ : \ P_{v_i,v_j} \in \mathcal{G}\}$be the set of all the paths from v_i to v_j. We can define a distance between two vertices in a graph $d(v_i, v_j)$ as follow:

$$d(v_i, v_j) = \begin{cases} \min_{P_{v_i,v_j} \in \mathcal{P}_{v_i,v_j}(\mathcal{G})} l(P_{v_i,v_j}) & \text{if} \mathcal{P}_{v_i,v_j} \neq \varnothing \\ +\infty & \text{otherwise.} \end{cases} \tag{A.10}$$

This distance allows us to define an equivalence relation on the nodes of a graph as follow. $v_k \sim v_j$ if $d(v_i, v_j) < +\infty$. This relation of equivalence induces a partition on V_1, V_K in classes of equivalence. By definition there cannot be any edge between vertices in different classes. The maximal subgraphs $K_i(V_i, E_i)$ of $\mathcal{G}$ are called *connected components*. Clearly $\bigcup_i E_i = E$. In particular a graph which has only one connected components is called *connected*.

A.1.6 *Cycles and trees*

A path $P_{v_i v_j}$ such that $v_i = v_j$ is called a *cycle*; we shall used the notation C_{v_i} to refer to it. Let $T(V, E)$ be a connected graph. If, for all $v_i \in V$ it is not possible to find a cycle C_{v_i} such that $C_{v_i} \subset T$ the graph is called a *tree*. It is easy to check that if for a tree |E|=|V|-1.

If $\mathcal{G}(V, E)$ is a connected graph, there exists at least a subgraph T which is a tree and such that $V(T) = V$; this subgraph is called *a spanning tree*[124].

In particular it given if a graph (not necessarily connected) $\mathcal{G}(V, E)$ is such that $|E| > |V| - 1$ the the number of cycle has a lower bound which is given by the so called *cyclomatic number* $\gamma = |E| - |V| + K$ where K is the number of connected components [8]. Intuitively one can see it as follow. Let K_i be one K connected components of the graph; its spanning tree has $V(K_i) - 1$ edges; thus any additional edge in K_i will generate at least a cycle. There are $E(K_i) - V(K_i) + 1$ of such edges and thus summing over i one gets the formula for the cyclomatic number

A.1.7 *Random graphs, degree distribution and giant component*

Networks in real word are dynamic entities; due to this dynamics most of the properties are not fixed once and for all, but varies and fluctuate. That is why one is interested to the rather to the *typical* properties of a set of networks that on the the properties of one of them. In general from the analysis of a system we shall obtain a set of graphs Ω and a probability measure defined on Ω: from this ingredients we can extract the different properties to which we are interested.

To illustrate some of the quantities we are interested in we shall use the famous Erdős-Rényi model as a example. In the original form discussed by Erdős-Rényi [9, 125] Ω was the space $\Gamma(N, M)$ of graphs with N vertices and M edges; the probability measure was the uniform measure $\pi(\omega) = \frac{1}{|G(N,M)|} = \frac{1}{\binom{N(N-1)/2}{M}}$. This probability space is usually denoted $G(M, N)$.

In the version given by Gilbert the space Ω is given by $\Gamma(N)$ that is the set of graphs with N vertices and the probability measure is constructed assuming that the probability that an edge occurs between two arbitrary vertices is p. It is straightforward to see that the probability measure for a graph $\omega \in \Omega$ with m edges is $\pi(\omega) = p^m(1-p)^{\binom{N}{2}-m}$. This probability space is denoted $G(N, p)$.

Since calculations are easier in Gillert model we shall continue focusing on that model.

It is easy to get some quantities. By example the average number of edges:

$$\langle m \rangle = \sum_m \binom{\frac{N(N-1)}{2}}{m} m p^m (1-p)^{\binom{N}{2}-m} = p\frac{N}{2} \tag{A.11}$$

and thus

$$\langle k \rangle = \frac{2\langle m \rangle}{N} = (N-1)p \tag{A.12}$$

Degree distribution

One of the most important quantity that it is possible to calculate in this framework is the degree distribution which is the probability that a node in a graph in $G(N, p)$ has degree k.

Let $N_k(\omega)$ be the number of vertices of ω with degree k. Clearly $N_k(\omega)$ is a stochastic variable and we can define the quantity.

$$p_k = \frac{\langle N_k \rangle}{N} = \frac{1}{N} \sum_{\omega \in G(N,p} N_k(\omega).\pi(\omega) \tag{A.13}$$

The quantity p_k is the degree distribution of the system.

In the case of $G(N, p)$ we can calculate it easily. In fact, since by construction, the probability that a link between two nodes is present is p, and since the maximal degree is limited by $N-1$ we get:

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}. \tag{A.14}$$

One immediately retrieves that $\langle k \rangle = (N-1)p$ but one can also calculate additional moments easily.

Since often we deal with networks with a large number of vertices, we can be interested to take the limit $N \to \infty$ and with $(N-1)p \to c < +\infty$. It is easy to do it in this case and we get:

$$p_k = \frac{c^k \mathrm{e}^{-c}}{k!}. \tag{A.15}$$

Giant component

When $p = 0$ the typical graph has no edges and thus all nodes are isolated. On the contrary when $p = 1$ the typical graph shall be fully connected an thus there shall be a single connected component of size N which is called *Giant component*.

It is convenient to define the size S of the connected component and its densitys. The quantity $u = 1 - s$ measures the probability that a vertex in a typical graph is not in the giant component. If a vertex is not in the giant component this means that it must not be connected through any other vertex j to the giant component: and thus either i is not connected to j or i is connected to j but $j$is not in the giant component. The probability of not being connected to the giant component through a vertex j is thus $(1 - p - pu)$. Thus we can write down a consistency equation for u summing over al vertices and obtain

$$u = (1 - p - pu)^{N-1}. \tag{A.16}$$

In the large N limit assuming that c is constant and writing $s = 1 - u$ we obtain a consistency equation for s

$$s = 1 - e^{-cs}. \tag{A.17}$$

It can be proven that when $c \geq 1$ this equation admits a solution with $s > 0$ which actually corresponds to the relative size of the giant component of the typical graph.

It can also be proven that the there is only one giant component and that all small components are tree like in structure.

A.1.8 *Centrality Indices*

A centrality index can be loosely defined as a real valued function $C(v)$ on the vertices of a graph, such that if $\Phi : \mathcal{G} \to \mathcal{G}'$ is a graph isomorphism such that, for any $v \in V(\mathcal{G})$ $C(v) = C(\Phi(v))$. All centrality indices fall in this category, however not all the function defined as above can be meaningfully interpreted

There exists several centrality indices each of which highlight some a different property of the graph.

Here we define the indices cited in chapter 3.

Degree centrality

The degree centrality $C_D(v)$ is simply given by the degree k_v of vertex v. This quantity measures the local influence of a node in a graph.

Closeness centrality

Closeness centrality measures how central a node is with respect to the metrics induced by shortest paths. Let $d(v, w)$ be the graph distance as defined in (A.10). If $\mathcal{G}$ is a connected undirected graph, the closeness centrality is define as

$$C_D(v) = N \left(\sum_{w \in V(\mathcal{G})} d(v, w) \right)^{-1}. \tag{A.18}$$

Closeness is a *geometric* measure; and can measure on how fast a signal may spread from node v to the rest of the graph. If the graph is not connected closeness of isolated vertices is 0.

Betweenness centrality

Let $\mathcal{G}$ be a connected graph, For all $w, z \in V(\mathcal{G})$ let $\sigma_{wz}(v)$ be the number of shortest path between w and z that pass by v. and let Σ_{wz} the the total number of shortest path from w to z. The betweenness centrality is defined as:

$$C_B(v) = \sum_{w,z \in V(\mathcal{G}), w \neq z \neq c} \frac{\sigma_{wz}(v)}{\Sigma_{wz}} \tag{A.19}$$

This index measures the structural importance of a vertex for information spread in a graph, i. e. how much the can control information spread between vertices

Eigenvector centrality (Bonacich centrality)

Eigenvector centrality for node $C_E(v)$ is defined be such that, for all $v, w \in V(\mathcal{G})$

$$C_E(v) = \frac{1}{\lambda_1} \sum_w g_{vw} C_E(v) \tag{A.20}$$

where λ_1 is the largest eigenvalue of the adjacency matrix.

With respect t degree centrality, which essentially weights all the neighbours equally, eigenvector centrality weights all the neighbours by their relevance and thus bears some global information about the entire network

B

APPENDIX TO CHAPTER 2

B.1 PROPERTIES OF LAMBERT'S W FUNCTION

Lambert's function is defined by

$$W(x)e^{W(x)} = x. \tag{B.1}$$

This function has two real branches, plotted in figure B.1 . Using the same notation as in [126]) we have the principal branch $W_0(x)$ is defined in $[e^{-1}, +\infty]$ and takes values between in $[1, \infty]$, the second branch W_{-1} which is defined in$[e^{-1}, 0]$ and takes values in $[-\infty, 1]$.

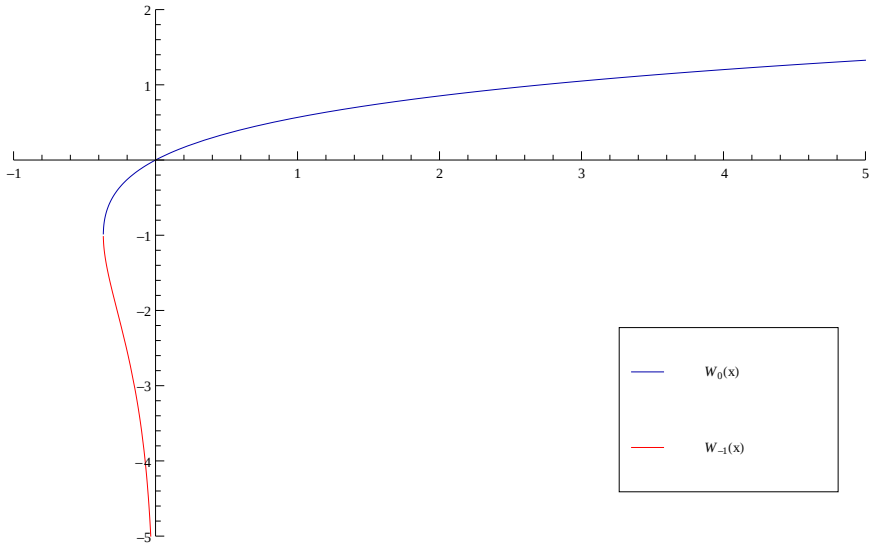


Figure B.1: Plot of Lamber W function

The principal branch can be expanded in a convergent power series around $z = 0$:

$$W_0 = \sum_{n=1}^{\infty} \frac{(-n)^{(n-1)}}{n!} x^n \tag{B.2}$$

which can be extended to define an holomorphic function on the complex plain with a branch cut in $[-\infty, e^{-1}]$..

Its derivative satisfies:

$$\frac{d}{dx}W(x) = \frac{e^{-W(x)}}{1 + W(x)} \underset{x\neq 0}{=} \frac{W(x)}{x(1 + W(x))}. \tag{B.3}$$

As for all inverse function, its numerical evaluation can be done with arbitrary precision using an iterative method; in particular Halley method proves to be quite effective [126]. The evolution rule is:

$$w_{j+1} = w_j - \frac{w_j e^{w_j} - z}{e^{w_j}(w_j+1) - \frac{(w_j+2)(w_j e^{w_j} - z)}{2w_j+2}}, \tag{B.4}$$

where, as a starting point of the iterative process, the first terms of the power expansion (or, for the secondary branch, the first terms of asymptotic expansion) are taken.

B.2 NUMERICAL SOLUTION OF FIRST ORDER CONDITION . CASE $n^1 = 1 - n^0$

In this section we give the details on how to reduce the FOC to a 1 dimensional equation in the case when $n^1 = 1 - n^0$

When we consider only systems where only one direction is preferred, some simplification can be made.

To ease the notation, as in the previous section, we shall write $z = \frac{2\eta}{\lambda}$, $x = n_1$ and $y_i - 1 = n_i$ for $i \in \{2, .., q\}$.

In this case equation (2.83) take the simplified form:

$$\begin{cases} x e^{-x} = z\frac{n^0}{Q} + z e^h \frac{1-n^0}{(e^h-1)e^x+Q} \\ y_i e^{-y_i} = zP = z\frac{n^0}{Q} + z\frac{1-n^0}{(e^h-1)e^x+Q} \quad \text{q-1 times} \end{cases} \tag{B.5}$$

with the conditions

$$x + \sum_{i=0}^{q-1} y_i = z \tag{B.6}$$

$$Q = e^x + \sum_{i=0}^{q-1} e_i^y \tag{B.7}$$

and

$$P = \frac{n^0}{Q} + W = \frac{n^0}{Q} + \frac{1-n^0}{(e^h-1)e^x+Q}. \tag{B.8}$$

Solving equation (B.8) with respect e^x , calling $\Gamma = PQ = n^0 + WQ$ we get:

$$\frac{e^x}{Q} = \frac{1-\Gamma}{\Gamma - n^0}\frac{1}{e^h-1} \tag{B.9}$$

and thus, plugging it in the first equation of (B.5), we obtain:

$$x = \frac{z}{e^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-e^h)n^0 + e^h\Gamma\right] \tag{B.10}$$

Plugging the previous result in (B.9), we get Q. On the other hand form equation 1 in (B.5) we obtain

$$\frac{1-n^0}{(e^h-1)e^x+Q} = P - \frac{n^0}{Q} \tag{B.11}$$

plugging everything in equation (B.5) we obtain the following system of equations written in term of Γ :

$$\begin{cases} x\mathrm{e}^{-x} = \frac{z}{\mathrm{e}^h-1}\left[(1-\mathrm{e}^h)n^0 + \mathrm{e}^h\Gamma\right]\frac{1-\Gamma}{\Gamma-n^0}\mathrm{e}^{-\frac{z}{\mathrm{e}^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]} \\ y\mathrm{e}^{-y} = \frac{z}{\mathrm{e}^h-1}\Gamma\frac{1-\Gamma}{\Gamma-n^0}\mathrm{e}^{-\frac{z}{\mathrm{e}^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]} \qquad \text{q-1 times} \end{cases} \tag{B.12}$$

From previous equation (B.12) we can infer the structure of the solutions.
Both equations have the same shape:

$$x\mathrm{e}^{-x} = c \tag{B.13}$$

where c is a constant to be determined auto-consistently.

Both equations have the same shape:

$$x\mathrm{e}^{-x} = c \tag{B.14}$$

where c is a constant to be determined auto-consistently.

If c is negative the equation has only one negative (thus unphysical) solution, if c is greater than e^{-1} it has no solution otherwise it admits solution which can be expressed in terms of Lambert W functions [126]:

$$x_- = -\mathrm{W}_0(-c) \tag{B.15}$$

and

$$x_+ = -\mathrm{W}_{-1}(-c) \tag{B.16}$$

where $\mathrm{W}_0, \mathrm{W}_{-1}$ represents the two real branches of Lambert W function (using the notation of [126]). It is trivial to check $x_- < 1$ whereas $x_+ > 1$ and that for small values of c, $x_+ > z$ and thus is to be discarded.

In our case c is a complicate function of Γ but, as above, once Γ (and z) is fixed we know that x and y can take only two values $x_\pm$ and $y_\pm$ as defined above and thus we can label all the solutions using two integers α which counts the number of x in + state (which of course is either 0 or 1) and L_+ which counts the number of y in + state.

If we define, for notational ease:

$$B(\Gamma) = \frac{z}{\mathrm{e}^h-1}\Gamma\frac{1-\Gamma}{\Gamma-n^0}\mathrm{e}^{-\frac{z}{\mathrm{e}^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]} \tag{B.17}$$

and

$$A(h,\Gamma) = \left[(1-\mathrm{e}^h)n^0 + \mathrm{e}^h\Gamma\right] \tag{B.18}$$

it is easy to check that have the following hierarchy in the admissible range $\Gamma \in [n^0, 1]$:

$$y_-(A(0,\Gamma)B(\Gamma)) \le x_-(A(h,\Gamma)B(\Gamma)) \le 1 \le x_+(A(h,\Gamma)B(\Gamma)) \le y_+(A(0,\Gamma)B(\Gamma)) \tag{B.19}$$

The normalization equation (B.6) will then become, given integers α and L_+:

$$\begin{aligned}
&\alpha x_+\left(\frac{z}{\mathrm{e}^h-1}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]\frac{1-\Gamma}{\Gamma-n^0}\mathrm{e}^{-\frac{z}{\mathrm{e}^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]}\right)\\
&+(1-\alpha)x_-\left(\frac{z}{\mathrm{e}^h-1}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]\frac{1-\Gamma}{\Gamma-n^0}\mathrm{e}^{-\frac{z}{\mathrm{e}^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]}\right)+\\
&L_+y_+\left(\frac{z}{\mathrm{e}^h-1}\Gamma\frac{1-\Gamma}{\Gamma-n^0}\mathrm{e}^{-\frac{z}{\mathrm{e}^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]}\right)\\
&+(q-1-L_+)y_+\left(\frac{z}{\mathrm{e}^h-1}\Gamma\frac{1-\Gamma}{\Gamma-n^0}\mathrm{e}^{-\frac{z}{\mathrm{e}^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]}\right)=z.
\end{aligned} \tag{B.20}$$

Its solutions will give the Γ auto-consistently.

Since the previous substitution is valid only when (B.10) is assumed we can rewrite the equation as:

$$\boxed{\begin{aligned}
&L_+y_+\left(\frac{z}{\mathrm{e}^h-1}\Gamma\frac{1-\Gamma}{\Gamma-n^0}\mathrm{e}^{-\frac{z}{\mathrm{e}^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]}\right)\\
&+(q-1-L_+)y_+\left(\frac{z}{\mathrm{e}^h-1}\Gamma\frac{1-\Gamma}{\Gamma-n^0}\mathrm{e}^{-\frac{z}{\mathrm{e}^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]}\right)\\
&=z-\frac{z}{\mathrm{e}^h-1}\frac{1-\Gamma}{\Gamma-n^0}\left[(1-\mathrm{e}^h)n^0+\mathrm{e}^h\Gamma\right]
\end{aligned}} \tag{B.21}$$

B.2.1 *Stability*

Here we detailed how we checked the stability of the numerical solution.

When $n^\alpha = 1-n^0$ all the values $n_a^\alpha = 0$ for $\alpha > 1$ because of the normalization constraints; the Lagrangian function, thus, becomes (here $z = 2\frac{\eta}{\lambda}$):

$$\begin{aligned}
\mathrm{L}(\boldsymbol{n}; z, \boldsymbol{h}) = (\beta_1 - 1)\left(n^1 + \sum_a n_a^1\right) - (\zeta_0 - 1)(n^0 - \sum_a n_a^0) - h_1 n_1^1\\
+ \sum_a n_a^0 \log(n_a^0) + n_a^1 \log(n_a^1) - \frac{z}{2}(n_a)^2
\end{aligned} \tag{B.22}$$

To check the stability and the nature of these stationary points we have to check the Hessian L of the Lagrangian restricted to the tangent space $\mathscr{T}$ to the constraints manifold in the stationary point.

The stationary point will be a (local minimum) if and only if:

$$\boldsymbol{y}^T \mathsf{L} \boldsymbol{y} > 0 \text{ for any } \boldsymbol{y} \in \mathscr{T} \tag{B.23}$$

In our case the constraint are linear; therefore the tangent space $\mathscr{T}$ is a $2q-2$ dimensional space and it can be easily seen to be spanned by the orthonormal base:

$$(\boldsymbol{e}_i)_j = \frac{1}{\sqrt{2}}(\delta_{ij} - 2\delta_{2q+1+j \pmod 2, j}) \tag{B.24}$$

where $i \in \{1, .., 2q\}$ and $j \in \{1, .., 2q+2\}$. The projection operator then is given by the matrix

$$\mathsf{M}_{ij} = (\boldsymbol{e}_i)_j \tag{B.25}$$

Any vector $\boldsymbol{y}$ of $\mathscr{T}$ can be expressed by a general vector $\boldsymbol{v}$ of $\mathbb{R}^{2q}$ as $\boldsymbol{y} = \mathsf{M}\boldsymbol{v}$. Equation (B.23) can be then expressed as:

$$\boldsymbol{v}^T \mathsf{M}^T \mathsf{L} \mathsf{M} \boldsymbol{v} > 0 \tag{B.26}$$

Thus in order to check the stability of a stationary point in the constrained problem we can simply apply the usual Hessian criteria to the "effective Hessian $\mathsf{H}_{\text{eff}} = \mathsf{M}^T \mathsf{L} \mathsf{M}$, either through numerical diagonalisation or using Sylvester's Criterion.

This method seems to be numerically more efficient than the usual bordered Hessian Criterion [127, p.155] in particular since we can take advantage of numerical diagonalisation libraries.

B.3 PERTURBATIVE EXPANTION OF THE FREE ENERGY AROUND $z = 0$

In this section we give the details of the perturbative expantion around $z = 0$.

The free energy of the system may be written as:

$$\mathrm{F}(\boldsymbol{n}; z, \boldsymbol{h}) = \mathrm{f}(\boldsymbol{n}(z); \boldsymbol{h}) - z\mathrm{g}(\boldsymbol{n}(z)) \tag{B.27}$$

where f, g are bounded functions of $\boldsymbol{n}(z)$.

For finite z we cannot say anything about the minimum, but from the structure of the free energy we can get the structure of the minimum in the limiting case when $z \to 0$ and $z \to \infty$.

When $z = 0$ the the minimum is obtained by minimizing f. It is easy to check that this solution correspond to the zeroth order of the perturbative expansion of the solutions of (2.83).

B.3.1 *Zeroth order*

Let us assume that $n_a^\alpha(z)$ is analytic around $z = 0$ (i.e.$n_a^\alpha(z) = \sum_k \bar{n}_{a,k}^\alpha z^k$ for $k > 0$). The zeroth order value can be obtained directly substituting z with 0 in eq. (2.76). And, with some algebra we get:

$$\bar{n}_{a,0}^0 = \frac{n^0}{q} \tag{B.28}$$

and

$$\bar{n}_{a,0}^\alpha = \frac{n^\alpha \mathrm{e}^{h_\alpha \delta_{a\alpha}}}{\mathrm{e}^{h_\alpha} - 1 + q} \tag{B.29}$$

thus leading to

$$n_{a,0} = n_a(0) = \frac{n^0}{q} + \sum_\alpha \frac{n^\alpha \mathrm{e}^{h_\alpha \delta_{a\alpha}}}{\mathrm{e}^{h_\alpha} - 1 + q} = \frac{(e_a^h - 1) n^a}{\mathrm{e}_a^h + q - 1} + K \tag{B.30}$$

where $K = \frac{n^0}{q} + \sum_\alpha \frac{n^\alpha}{\mathrm{e}^{h_\alpha} + q - 1}$ and is equal for all

B.3.2 *First order*

The first order of that expansion can be obtained easily deriving eq. (2.83) with respect to z then setting $(z = 0)$.

We have in particular:

$$Q'(z)\Big|_{z=0} = \sum_a n_a(z)\mathrm{e}^{zn_a(z)} + o(z)\Big|_{z=0} = 1 \tag{B.31}$$

and

$$\begin{aligned} W'(z))\Big|_{z=0} &= -\sum_\alpha \frac{n^\alpha}{\left[(\mathrm{e}^{h_\alpha}-1)\mathrm{e}^{zn_a(z)}+Q(z)\right]^2}\left[(\mathrm{e}^{h_\alpha}-1)(n_a(z)+o(z))\mathrm{e}^{zn_a(z)})+Q'(z)\right]\Big|_{z=0} \\ &= -\sum_\alpha \frac{n^\alpha}{\left[(\mathrm{e}^{h_\alpha}-1)\mathrm{e}^{z\bar{n}_{a,0}}+q\right]^2}\left[(\mathrm{e}^{h_\alpha}-1)\bar{n}_{a,0}+1\right] \end{aligned} \tag{B.32}$$

The full equation thus reads:

$$\begin{aligned} \bar{n}_{a,1} = n'_a(z)\Big|_{z=0} &= (n_a(z))^2\mathrm{e}^{-zn_a(z)} + o(z) - \frac{n^0}{Q(z)^2}Q'(z) + W'(z) \\ &-\frac{(e^{h_a}-1)n^a}{\left[(\mathrm{e}^{h_a}-1)\mathrm{e}^{zn_a(z)}+Q(z)\right]^2}\left[(\mathrm{e}^{h_a}-1)(n_a(z)+o(z))+Q'(z)\right]\Big|_{z=0} \\ =&\bar{n}^2_{a,0} - \frac{n^0}{q^2} - \frac{(e^{h_a}-1)n^a}{\left[(\mathrm{e}^{h_a}-1)+q)\right]^2}\left[(\mathrm{e}^{h_a}-1)(\bar{n}_{a,0})+1\right] \\ &-\sum_\alpha \frac{n^\alpha}{\left[(\mathrm{e}^{h_\alpha}-1)\mathrm{e}^{z\bar{n}_{a,0}}+q\right]^2}\left[(\mathrm{e}^{h_\alpha}-1)\bar{n}_{a,0}+1\right] \end{aligned} \tag{B.33}$$

For large n^0, as it is natural to expect, this solution corresponds to the sparse network configuration. This is the only solution of eq. (2.83) at $z = 0$. Since the FOC are analytic for $z \to 0$ at least in a neighbourhood of$z = 0$ we can state that the solution obtained perturbatively is the only minimum of the free energy and that it corresponds to a sparse network configuration (indeed for $z = 0$ the network cannot grow at all).

B.4 ASYMPTOTIC EXPANTION OF THE FREE ENERGY AROUND $z + \infty$

In this section we carry on the details of asymptotic expansion around $z = +\infty$

In the case of $z \to \infty$ case the minimum must correspond to a maximum of f($\boldsymbol{n}$). In this case we have q degenerate maxima $\boldsymbol{n}_i^{\mathrm{eq}}$ corresponding to $n_i = 1$ and $n_b = 0\ \forall b \neq i$.

In order to gain some insight on which of these solutions is the true minimum of the free energy for large but finite z we have to make and asymptotic expansion around $z = \infty$.

We proceed by steps expanding $n_a(z)$ in equation (2.83) for large z around the solution at $z = \infty$ $n_\alpha = 1$, $n_a = 0$ for $a \neq \alpha$.

If we assume that $n_a(z)$ has the following form around ∞:

$$n_a(z) = \delta_{a\alpha} + \frac{1}{z}\hat{n}_a(z) + \mathrm{e}^{-z}\bar{n}_a(z) \tag{B.34}$$

where $\hat{n}_a(z)$ and $\bar{n}_a(z)$ are meromorphic around $z = \infty$.

Since the right hand member of equation (2.83) is $o(\mathrm{e}^{-z})$ whereas, for $a \neq \alpha$ the left hand member of the equation is meromorphic, we immediately obtain that $\hat{n}_a(z) = 0$.

Obtaining the non analytic contribution requires, instead, a harder work. We shall proceed by steps. We assume, in the following, that $\bar{n}_a(z) = \sum_{k\geq 0} \bar{n}_a(k) z^{-k}$.

B.4.1 *Left hand member of equation* (2.83)

Case $a = \alpha$

$$\begin{aligned}\bar{n}_\alpha(z)\mathrm{e}^{-z\bar{n}_\alpha(z)} &\simeq \left[1+\mathrm{e}^{-z}\left(\bar{n}_\alpha(0)+\bar{n}_\alpha(1)z^{-1}\right)\right]\mathrm{e}^{-z}\left[1-z\mathrm{e}^{z}\left(\bar{n}_\alpha(0)+\bar{n}_\alpha(1)z^{-1}+\bar{n}_\alpha(2)z^{-2}\right)\right] = \\ &\mathrm{e}^{-z}+\mathrm{e}^{-2z}\left[-z\bar{n}_\alpha(0)+\bar{n}_\alpha(0)-\bar{n}_\alpha(1)+\left(\bar{n}_\alpha(1)-\bar{n}_\alpha(2)\right)z^{-1}\right]\end{aligned} \tag{B.35}$$

Case $a \neq \alpha$

$$\begin{aligned}\bar{n}_a(z)\mathrm{e}^{-z\bar{n}_a(z)} &\simeq \mathrm{e}^{-z}\left[\bar{n}_a(0)+\bar{n}_a(1)z^{-1}\right]\left[1-z\mathrm{e}^{z}\left(\bar{n}_\alpha(0)+\bar{n}_\alpha(1)z^{-1}+\bar{n}_\alpha(2)z^{-2}\right)\right] = \\ &\mathrm{e}^{-z}\left(\bar{n}_a(0)+\bar{n}_a(1)z^{-1}\right)+\mathrm{e}^{-2z}\left[z\left(\bar{n}_a(0)\right)^2+2\bar{n}_a(0)\bar{n}_a(1)+\left(\left(\bar{n}_a(1)\right)^2+\bar{n}_a(0)\bar{n}_a(2)\right)z^{-1}\right]\end{aligned} \tag{B.36}$$

B.4.2 *Right hand member of equation* (2.83)

Expansion of Q

$$\begin{aligned}Q = \sum_a \mathrm{e}^{z\bar{n}_a(z)} &\simeq \mathrm{e}^{z}\left(1+\mathrm{e}^{-z}\left(z\bar{n}_\alpha(0)+\bar{n}_\alpha(1)+\bar{n}_\alpha(2)z^{-2}\right)+\mathrm{e}^{z}\left(q-1\right)\right)\mathrm{e}^{-z} \\ &= \mathrm{e}^{z}\left[1+\mathrm{e}^{-z}\left(z\bar{n}_\alpha(0)+\bar{n}_\alpha(1)+q-1+\bar{n}_\alpha(2)z^{-2}\right)\right]\end{aligned} \tag{B.37}$$

Expansion of $n^0 Q^{-1}$

$$n^0 Q^{-1} \simeq n^0 \mathrm{e}^{-z}\left[1-\mathrm{e}^{-z}\left(z\bar{n}_\alpha(0)+\bar{n}_\alpha(1)+q-1+\bar{n}_\alpha(2)z^{-1}\right)\right] \tag{B.38}$$

Expansion of $((\mathrm{e}^{h_\alpha}-1)\mathrm{e}^{z\bar{n}_\alpha}+Q)^{-1}$

$$\frac{1}{\left(\mathrm{e}^{h_\alpha}-1\right)\mathrm{e}^{z\bar{n}_\alpha}+Q} \simeq \mathrm{e}^{-h_\alpha}\mathrm{e}^{-z}\left[1-\mathrm{e}^{-z}\left(z\bar{n}_\alpha(0)+\bar{n}_\alpha(1)+\mathrm{e}^{-h_\alpha}\left(q-1\right)+\bar{n}_\alpha(2)z^{-1}\right)\right] \tag{B.39}$$

Expansion of $((\mathrm{e}^{h_a}-1)\mathrm{e}^{z\bar{n}_a}+Q)^{-1}$

$$\frac{1}{\left(\mathrm{e}^{h_a}-1\right)\mathrm{e}^{z\bar{n}_a}+Q} \simeq \mathrm{e}^{-z}\left[1-\mathrm{e}^{-z}\left(z\bar{n}_\alpha(0)+\bar{n}_\alpha(1)+\mathrm{e}^{h_a}+\left(q-2\right)+\bar{n}_\alpha(2)\,z^{-1}\right)\right] \tag{B.40}$$

Expansion of W

$$W=\sum_a \frac{n^a}{(\mathrm{e}^{h_a}-1)\mathrm{e}^{z\bar{n}_a}+Q} \simeq (n^\alpha \mathrm{e}^{-h_\alpha}+\sum_{a\neq\alpha} n^a)\mathrm{e}^{-z}\left[1-\mathrm{e}^{-z}\left(z\bar{n}_\alpha(0)+\bar{n}_\alpha(1)+\bar{n}_\alpha(2)\,z^{-1}\right)\right]$$
$$+\mathrm{e}^{-2z}\left[n^\alpha \mathrm{e}^{-2h_\alpha}(q-1)+\sum_{a\neq\alpha} n^a\left(\mathrm{e}^{h_a}+q-2\right)\right] \tag{B.41}$$

B.4.3 *Full Equation*

Case $a\neq\alpha$

In this case it is sufficient to keep items up to order e^{-z} to obtain non trivial results. The FOC becomes then

$$\mathrm{e}^{-z}\left[\bar{n}_a(0)+\bar{n}_a(1)\,z^{-1}\right]+o(\mathrm{e}^{-2z})=\mathrm{e}^{-z}\left[n^0+n^\alpha \mathrm{e}^{-h_\alpha}+\sum_{a\neq\alpha} n^a+(\mathrm{e}^{h_a}-1)n^a\right]+o(\mathrm{e}^{-2z}) \tag{B.42}$$

and thus equating order by order, and using normalization constraints, we obtain:

$$\begin{cases} n_a(0)=\left[1+n^a(\mathrm{e}^{h_a}-1)-n^\alpha(1-\mathrm{e}^{-h_\alpha})\right] \\ n_a(1)=0 \end{cases} \tag{B.43}$$

Case $a=\alpha$

In this case, at order e^{-z} the equation is trivially satisfied because of normalization constraints; we need thus to keep also the terms of order e^{-2z}.

$$\begin{aligned} \mathrm{e}^{-z}+\mathrm{e}^{-2z}&\left[-z\bar{n}_\alpha(0)+\bar{n}_\alpha(0)-\bar{n}_\alpha(1)+\left(\bar{n}_\alpha(1)-\bar{n}_\alpha(2)\right)z^{-1}\right]= \\ &\mathrm{e}^{-z}\left[n^0+n^\alpha \mathrm{e}^{-h_\alpha}+\left(\mathrm{e}^{h_\alpha}-1\right)\mathrm{e}^{-h_\alpha}n^\alpha+\sum_{a\neq\alpha} n^\alpha\right] \\ &-\mathrm{e}^{-2z}n^0\left(z\bar{n}_\alpha(0)+\bar{n}_\alpha(1)+q-1+\bar{n}_\alpha(2)\,z^{-1}\right) \\ &-\mathrm{e}^{-2z}\left(n^\alpha \mathrm{e}^{h_\alpha}+\sum_{a\neq\alpha} n^a\right)\left(z\bar{n}_\alpha(0)+\bar{n}_\alpha(1)+\bar{n}_\alpha(2)\,z^{-1}\right) \\ &-\mathrm{e}^{-2z}\left(\mathrm{e}^{h_\alpha}-1\right)n^\alpha \mathrm{e}^{-h_\alpha}\left(z\bar{n}_\alpha(0)+\bar{n}_\alpha(1)+\mathrm{e}^{-h_\alpha}\left(q-1\right)+\bar{n}_\alpha(2)\,z^{-1}\right) \\ &-\mathrm{e}^{-2z}\left[n^\alpha \mathrm{e}^{-2h_\alpha}\left(q-1\right)+\sum_a \mathrm{e}^{h_a}+\left(q-2\right)\sum_a n^a\right] \end{aligned}$$

(B.44)

The solution can be obtain equating order by order and one gets:

$$\begin{cases} n_\alpha(0) = -\left\{(q-1)\left[1 + n^\alpha(\mathrm{e}^{-h_\alpha} - 1)\right] + n^0 + n^\alpha - 1 + \sum_{i\neq\alpha} \mathrm{e}^{h_i} n^i\right\} \\ n_\alpha(1) = 0 \end{cases} \tag{B.45}$$

It can be easily checked that these solutions are consistent with the normalization constraint order by order.

B.4.4 *Asymptotic expansion of* n_a^b

We can plug the previous equation in equation (2.76) and obtain, after a little algebra, the asymptotic expansion for the n_a^b variables. We have to distinguish between seven cases. when

Case $b = 0$

When $a \neq \alpha$ we have

$$n_a^0 = \mathrm{e}^{-z} n^0 + o(z^{-1}\mathrm{e}^{-z}) \tag{B.46}$$

in the other case

$$n_\alpha^0 = n^0 - \mathrm{e}^{-z} n^0 (q-1) + o(z^{-1}\mathrm{e}^{-z}). \tag{B.47}$$

Case $b = \alpha$

here we have to distinguish between two cases. When $a = \alpha$ we have:

$$n_\alpha^\alpha = n^\alpha - \mathrm{e}^{-z} n^\alpha \mathrm{e}^{-h_\alpha}(q-1) + o(z^{-1}\mathrm{e}^{-z}) \tag{B.48}$$

when $a \neq \alpha$ we have instead:

$$n_a^\alpha = \mathrm{e}^{-z} n^\alpha \mathrm{e}^{-h_\alpha} + o(z^{-1}\mathrm{e}^{-z}). \tag{B.49}$$

Case $b \neq \alpha, 0$

In this case we have to distinguish among three possibilities. When $a = \alpha$ we have:

$$n_\alpha^b = n^b - \mathrm{e}^{-z} n^b (\mathrm{e}^{h_b} + q - 2) + o(z^{-1}\mathrm{e}^{-z}) \tag{B.50}$$

when $a \neq \alpha$ and $a \neq b$ we have:

$$n_a^b = \mathrm{e}^{-z} n^b + o(z^{-1}\mathrm{e}^{-z}) \tag{B.51}$$

and finally when $a \neq \alpha$ and $a = b$ we obtain:

$$n_b^b = \mathrm{e}^{-z} n^b \mathrm{e}^{h_b} + o(z^{-1}\mathrm{e}^{-z}). \tag{B.52}$$

B.4.5 *Asymptotic expansion of the Free Energy*

We can plug all the equation previously written in the expression of the free energy and, after some algebra obtain the following asymptotic formula for the free energy for $z \to \infty$

$$\mathrm{F} \simeq -\frac{z}{2} + n^0 \log(n^0) + \sum_{i \neq 0} n^i \log(n^i) - n^\alpha h_\alpha + z\mathrm{e}^{-z}\Big(- n^\alpha h_\alpha \mathrm{e}^{h_\alpha} + \sum_i n^i h_i \mathrm{e}^{h_i}\Big) + o(\mathrm{e}^{-z}) \tag{B.53}$$

As the sub-leading term (the constant term) is enough to determine which is the global minimum. In fact the solution with the maximum $n^\alpha h_\alpha$ prevail. In the case in which two or more direction have the same value of $h_i n^i$ the solution with the bigger value of h_i will prevail.

BIBLIOGRAPHY

1. Holland, J. H. *Hidden Order: How Adaptation Builds Complexity* ISBN: 9780201442304 (Basic Books, 1995).
2. Miller, J. H. & Page, S. E. *Complex Adaptive Systems* ISBN: 9780691127026 (2007).
3. Lansing, J. S. Complex Adaptive Systems. *Annual Review of Anthropology* **32,** 183–204. ISSN: 0084-6570 (Oct. 2003).
4. Holland, J. H. Studying Complex Adaptive Systems. *Journal of Systems Science and Complexity* **19,** 1–8. ISSN: 1009-6124 (Mar. 2006).
5. Holland, J. & Miller, J. Artificial adaptive agents in economic theory. *The American Economic Review.* <http://www.jstor.org/stable/10.2307/2006886> (1991).
6. McLuhan, M. *The Gutenberg Galaxy: The Making of Typographic Man* 293. ISBN: 0802060412 (University of Toronto Press, 1962).
7. Newman, M. *Networks* 720. ISBN: 9780199206650. doi:10.1093/acprof:oso/9780199206650.001.0001. <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199206650.001.0001/acprof-9780199206650> (Oxford University Press, Mar. 2010).
8. Bollobás, B. *Modern Graph Theory* corrected. ISBN: 038-7-98488-7 (Springer, Heidelberg, 1998).
9. Erdős, P. & Rényi, A. On random Graph I. *Publicationes Mathematicae Debrecen,* 290–297 (1959).
10. Gilbert, E. N. Random Graphs. *The Annals of Mathematical Statistics* **30,** 1141–1144. ISSN: 0003-4851 (Dec. 1959).
11. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393,** 440–2. ISSN: 0028-0836 (June 1998).
12. Barabási, A. Emergence of Scaling in Random Networks. *Science* **286,** 509–512. ISSN: 00368075 (Oct. 1999).
13. *Adaptive Networks* (eds Gross, T. & Sayama, H.) ISBN: 978-3-642-01283-9. doi:10.1007/978-3-642-01284-6. <http://www.springerlink.com/index/10.1007/978-3-642-01284-6> (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009).
14. Liggett, T. M. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes* xIi + 335. ISBN: 978-3-642-08529-1. doi:10.1007/978-3-662-03990-8. <http://link.springer.com/10.1007/978-3-662-03990-8> (Springer Berlin Heidelberg, Berlin, Heidelberg, 1999).
15. Gross, T., D'Lima, C. & Blasius, B. Epidemic Dynamics on an Adaptive Network. *Physical Review Letters* **96,** 208701. ISSN: 0031-9007 (May 2006).

16. Holme, P. & Newman, M. Nonequilibrium phase transition in the coevolution of networks and opinions. *Physical Review E* **74,** 056108. ISSN: 1539-3755 (Nov. 2006).

17. Gil, S. & Zanette, D. H. Coevolution of agents and networks: Opinion spreading and community disconnection. *Physics Letters A* **356,** 89–94. ISSN: 03759601 (July 2006).

18. Zanette, D. H. & Gil, S. Opinion spreading and agent segregation on evolving networks. *Physica D: Nonlinear Phenomena* **224,** 156–165. ISSN: 01672789 (Dec. 2006).

19. Vazquez, F., Eguíluz, V. & Miguel, M. Generic Absorbing Transition in Coevolution Dynamics. *Physical Review Letters* **100,** 108702. ISSN: 0031-9007 (Mar. 2008).

20. Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. *Reviews of Modern Physics* **81,** 591–646. ISSN: 0034-6861 (May 2009).

21. Clifford, P. & Sudbury, A. A model for spatial conflict. *Biometrika* **60,** 581–588. ISSN: 0006-3444 (1973).

22. Holley, R. A. & Liggett, T. M. Ergodic Theorems for Weakly Interacting Infinite Systems and the Voter Model. *The Annals of Probability* **3,** 643–663. ISSN: 0091-1798 (Aug. 1975).

23. Frachebourg, L. & Krapivsky, P. Exact results for kinetics of catalytic reactions. *Physical Review E* **53,** R3009–R3012. ISSN: 1063-651X (Apr. 1996).

24. Demirel, G., Vazquez, F., Böhme, G. & Gross, T. Moment-closure approximations for discrete adaptive networks. *Physica D: Nonlinear Phenomena.* ISSN: 01672789. doi:10.1016/j.physd.2013.07.003. <http://linkinghub.elsevier.com/retrieve/pii/S0167278913002017> (July 2013).

25. Sumpter, D. J. T. *Collective Animal Behavior* ISBN: 9780691148434. <http://books.google.it/books?id=nB67IvU1qlAC> (Princeton University Press, 2010).

26. Krause, J. & Ruxton, G. D. *Living in groups* xiv, 210 p. ISBN: 9780198508182 (Oxford University Press, 2002).

27. Krause, J., Ruxton, G. D. & Krause, S. Swarm intelligence in animals and humans. *Trends in ecology & evolution* **25,** 28–34. ISSN: 0169-5347 (Jan. 2010).

28. Surowiecki, J. *The Wisdom of Crowds* ISBN: 0385721706. <http://www.worldcat.org/isbn/0385721706> (Anchor, 2005).

29. Bonabeau, E., Dorigo, M. & Theraulaz, G. Swarm intelligence: from natural to artificial systems. <http://jasss.soc.surrey.ac.uk/4/1/reviews/kluegl.html> (1999).

30. Couzin, I. Collective minds. *Nature* **445,** 715. ISSN: 1476-4687 (Feb. 2007).

31. Hofstadter, D. R. *Gödel, Escher, Bach: an eternal golden braid* ISBN: 978-0-465-02656-2 (Basic Books, 1979).

32. Conradt, L. Models in animal collective decision-making: information uncertainty and conflicting preferences. *Interface Focus* **2,** 226–240. ISSN: 2042-8898 (Dec. 2011).

33. Conradt, L. & Roper, T. J. Consensus decision making in animals. *Trends in ecology & evolution* **20,** 449–56. ISSN: 0169-5347 (Aug. 2005).

34. Reynolds, C. W. *Flocks, herds and schools: A distributed behavioral model* in *Proceedings of the 14th annual conference on Computer graphics and interactive techniques - SIGGRAPH '87* (ACM Press, New York, New York, USA, 1987), 25–34. ISBN: 0897912276. doi:10.1145/37401.37406. <http://portal.acm.org/citation.cfm?doid=37401.37406>.

35. Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I. & Shochet, O. Novel Type of Phase Transition in a System of Self-Driven Particles. *Physical Review Letters* **75,** 1226–1229. ISSN: 0031-9007 (Aug. 1995).

36. Dingle, H. & Drake, V. A. What Is Migration? *BioScience* **57,** 113. ISSN: 0006-3568 (2007).

37. Dingle, H. Migration: the biology of life on the move. <http://library.wur.nl/WebQuery/clc/922917> (1996).

38. Southwood, T. Habitat, the templet for ecological strategies? *Journal of Animal Ecology.* <http://www.jstor.org/stable/10.2307/3817> (1977).

39. Southwood, T. Migration of terrestrial arthropods in relation to habitat. *Biological reviews.* <http://onlinelibrary.wiley.com/doi/10.1111/j.1469-185X.1962.tb01609.x/abstract> (1962).

40. Guttal, V., Romanczuk, P., Simpson, S. J., Sword, G. A. & Couzin, I. D. Cannibalism can drive the evolution of behavioural phase polyphenism in locusts. *Ecology letters* **15,** 1158–66. ISSN: 1461-0248 (Oct. 2012).

41. MacKenzie, B. R. & Myers, R. a. The development of the northern European fishery for north Atlantic bluefin tuna Thunnus thynnus during 1900–1950. *Fisheries Research* **87,** 229–239. ISSN: 01657836 (Nov. 2007).

42. Petitgas, P., Secor, D. H., McQuinn, I., Huse, G. & Lo, N. Stock collapses and their recovery: mechanisms that establish and maintain life-cycle closure in space and time. *ICES Journal of Marine Science* **67,** 1841–1848. ISSN: 1054-3139 (June 2010).

43. Brown, C. & Laland, K. N. Social learning in fishes: a review. *Fish and Fisheries* **4,** 280–288. ISSN: 1467-2960 (Sept. 2003).

44. Huse, G., Fernö, a. & Holst, J. Establishment of new wintering areas in herring co-occurs with peaks in the 'first time/repeat spawner' ratio. *Marine Ecology Progress Series* **409,** 189–198. ISSN: 0171-8630 (June 2010).

45. Nøttestad, L. & Tangen, S. Øand Sundby. *Possible mechanisms and explanations for the drastic decline and disappearance of Atlantic bluefin tuna in the Norwegian fisheries since the early 1960s: What went wrong and what ca we do?* in *Submitted to 2008 Bluefin Tuna Symposium; Summary in Collect. Vol. Sci. Pap. ICCAT, 63: 231-232* (2008). <http://www.iccat.int/Documents/Meetings/Docs/BFT%5C_SYMP/pdf/BFT%5C_SYMP%5C_018.pdf>.

46. Mather, Frank J., I., Mason, John M., J. & Jones, A. C. *Historical document: life history and fisheries of Atlantic bluefin tuna* tech. rep. (US Dept. of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Miami, 1995), 165.

47. Doumenge, F. L'histoire des pêches thonières. *Collective volume of scientific papers, ICCAT* **50,** 753–803 (1998).

48. Cury, P., Anneville, O., Bard, F., Fonteneau, A. & Roy, C. Obstinate North Atlantic bluefin tuna (Thunnus thynnus thynnus): an evolutionary perspective to consider spawning migration. *Collective volume of scientific papers, ICCAT* **50,** 239–247 (1998).

49. Fonteneau, A. & Fromentin, J.-M. in *Domestication of the bluefin tuna Thunnus thynnus thynnus* (eds Bridges, C., García, A. & Gordin, H.) 73–76 (Zaragoza : CIHEAM, 2003). <http://om.ciheam.org/om/pdf/c60/03600096.pdf>.

50. Fromentin, J.-M. & Powers, J. E. Atlantic bluefin tuna: population dynamics, ecology, fisheries and management. *Fish and Fisheries* **6,** 281–306. ISSN: 1467-2960 (Dec. 2005).

51. ICCAT. *Report of the 2012 Atlantic Blufin Tuna stock assesment session* tech. rep. (ICCAT, Madrid, 2012), 124. <http://www.iccat.int/Documents/Meetings/SCRS2012/SCI-003%5C_ENG.pdf>.

52. Collette, B. *et al. 2011. Thunnus thynnus.* 2013. <http://www.iucnredlist.org/details/21860/0> (visited on 09/04/2013).

53. Block, B. A. *et al.* Migratory movements, depth preferences, and thermal biology of Atlantic bluefin tuna. *Science (New York, N.Y.)* **293,** 1310–4. ISSN: 0036-8075 (Aug. 2001).

54. MacKenzie, B. R. & Mariani, P. Spawning of bluefin tuna in the Black Sea: historical evidence, environmental constraints and population plasticity. *PloS one* **7,** e39998. ISSN: 1932-6203 (Jan. 2012).

55. Block, B. A. *et al.* Electronic tagging and population structure of Atlantic bluefin tuna. *Nature* **434,** 1121–7. ISSN: 1476-4687 (Apr. 2005).

56. Fromentin, J.-M. Lessons from the past: investigating historical data from bluefin tuna fisheries. *Fish and Fisheries* **10,** 197–216. ISSN: 14672960 (June 2009).

57. Vicsek, T. & Zafeiris, A. Collective motion. *Physics Reports,* 71–140. ISSN: 03701573 (Mar. 2012).

58. Partridge, B. L., Johansson, J. & Kalish, J. The structure of schools of giant bluefin tuna in Cape Cod Bay. *Environmental Biology of Fishes* **9,** 253–262. ISSN: 0378-1909 (1983).

59. Berdahl, A., Torney, C. J., Ioannou, C. C., Faria, J. J. & Couzin, I. D. Emergent sensing of complex environments by mobile animal groups. *Science (New York, N.Y.)* **339,** 574–6. ISSN: 1095-9203 (Feb. 2013).

60. Ward, A. J. W., Herbert-Read, J. E., Sumpter, D. J. T. & Krause, J. Fast and accurate decisions through collective vigilance in fish shoals. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 2312–5. ISSN: 1091-6490 (Feb. 2011).

61. Conradt, L. Collective behaviour: When it pays to share decisions. *Nature* **471,** 40–1. ISSN: 1476-4687 (Mar. 2011).

62. Couzin, I., Krause, J. & James, R. Collective memory and spatial sorting in animal groups. *Journal of theoretical ...* <http://www.sciencedirect.com/science/article/pii/S0022519302930651> (2002).

63. Couzin, I. D. *et al.* Uninformed individuals promote democratic consensus in animal groups. *Science (New York, N.Y.)* **334,** 1578–80. ISSN: 1095-9203 (Dec. 2011).

64. Gross, T. & Blasius, B. Adaptive coevolutionary networks: a review. *Journal of the Royal Society, Interface / the Royal Society* **5,** 259–71. ISSN: 1742-5689 (Mar. 2008).

65. Couzin, I. D., Krause, J., Franks, N. R. & Levin, S. a. Effective leadership and decision-making in animal groups on the move. *Nature* **433,** 513–6. ISSN: 1476-4687 (Feb. 2005).

66. Guttal, V. & Couzin, I. D. Social interactions, information use, and the evolution of collective migration. *Proceedings of the National Academy of Sciences of the United States of America* **107,** 16172–7. ISSN: 1091-6490 (Sept. 2010).

67. Guttal, V. & Couzin, I. D. Leadership, collective motion and the evolution of migratory strategies. *Communicative & integrative biology* **4,** 294–8. ISSN: 1942-0889 (May 2011).

68. Mikhailov, A. S. & Calenbuhr, V. *From Cells to Societies* ISBN: 978-3-642-07574-2. doi:10.1007/978-3-662-05062-0. <http://link.springer.com/10.1007/978-3-662-05062-0> (Springer Berlin Heidelberg, Berlin, Heidelberg, 2002).

69. Viscido, S., Parrish, J. & Grünbaum, D. Individual behavior and emergent properties of fish schools: a comparison of observation and theory. *Marine Ecology Progress Series* **273,** 239–249. ISSN: 0171-8630 (2004).

70. Tegeder, R. W. & Krause, J. Density Dependence and Numerosity in Fright Stimulated Aggregation Behaviour of Shoaling Fish. *Philosophical Transactions of the Royal Society B: Biological Sciences* **350,** 381–390. ISSN: 0962-8436 (Dec. 1995).

71. Partridge, B. L. & Pitcher, T. J. The sensory basis of fish schools: Relative roles of lateral line and vision. *Journal of Comparative Physiology ? A* **135,** 315–325. ISSN: 0340-7594 (1980).

72. Huth, A. & Wissel, C. The simulation of the movement of fish schools. *Journal of theoretical biology.* <http://www.sciencedirect.com/science/article/pii/S0022519305806812> (1992).

73. Huth, A. & Wissel, C. The simulation of fish schools in comparison with experimental data. *Ecological modelling.* <http://www.sciencedirect.com/science/article/pii/0304380094900132> (1994).

74. Mogilner, A., Edelstein-Keshet, L., Bent, L. & Spiros, A. Mutual interactions, potentials, and individual distance in a social aggregation. *Journal of mathematical biology* **47,** 353–89. ISSN: 0303-6812 (Oct. 2003).

75. Romey, W. L. Individual differences make a difference in the trajectories of simulated schools of fish. *Ecological Modelling* **92,** 65–77. ISSN: 03043800 (1996).

76. Katz, Y., Tunstrøm, K., Ioannou, C. C., Huepe, C. & Couzin, I. D. Inferring the structure and dynamics of interactions in schooling fish. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 18720–5. ISSN: 1091-6490 (Nov. 2011).

77. Goldberg, D. E. & Deb, K. in *Foundations of Genetic Algorithms* (ed Spratz, B. M.) 69–93 (Morgan Kaufmann Publishers, Inc., San Mateo, CA, 1991). ISBN: 1558601708.

78. Huepe, C., Zschaler, G., Do, A.-L. & Gross, T. Adaptive-network models of swarm dynamics. *New Journal of Physics* **13,** 073022. ISSN: 1367-2630 (July 2011).

79. Bazazi, S. *et al.* Collective motion and cannibalism in locust migratory bands. *Current biology* **18,** 735–9. ISSN: 0960-9822 (May 2008).

80. Buhl, J., Sumpter, D., Couzin, I. & Hale, J. From disorder to order in marching locusts. *Science.* <http://www.sciencemag.org/content/312/5778/1402.short> (2006).

81. Cavagna, A. *et al.* Scale-free correlations in starling flocks. *Proceedings of the National Academy of Sciences of the United States of America* **107,** 11865–70. ISSN: 1091-6490 (June 2010).

82. Ehrhardt, G., Marsili, M. & Vega-Redondo, F. Phenomenological models of socioeconomic network dynamics. *Physical Review E* **74,** 36106. ISSN: 1539-3755 (Sept. 2006).

83. Marsili, M., Vega-Redondo, F. & Ehrhardt, G. Networks emerging in a volatile world. *EUI Working Paper ECO.* ISSN: 1725-6704 (2008).

84. De Luca, G., Mariani, P., MacKenzie, B. R. & Marsili, M. Fishing out collective memory of migratory schools. arXiv: 1305.6975. <http://arxiv.org/abs/1305.6975> (May 2013).

85. Miller, N., Garnier, S., Hartnett, A. T. & Couzin, I. D. Both information and social cohesion determine collective decisions in animal groups. *Proceedings of the National Academy of Sciences of the United States of America* **110.** ISSN: 1091-6490. doi:10.1073/pnas.1217513110. <http://www.ncbi.nlm.nih.gov/pubmed/23440218> (Feb. 2013).

86. Herbert-Read, J. E. *et al.* The role of individuality in collective group movement. *Proceedings. Biological sciences / The Royal Society* **280,** 20122564. ISSN: 1471-2954 (Feb. 2013).

87. Norris, J. R. *Markov Chains* 254. ISBN: 9780521633963 (Cambridge University Press, Cambridge, 1998).

88. Woess, W. *Denumerable Markov Chains* xii + 351. ISBN: 9783037190715 (European Mathematical Society, 2009).

89. Isella, L. *et al.* What's in a crowd? Analysis of face-to-face behavioral networks. *Journal of theoretical biology.* ISSN: 1095-8541. doi:10.1016/j.jtbi.2010.11.033. <http://www.ncbi.nlm.nih.gov/pubmed/21130777> (Dec. 2010).

90. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* **81,** 2340–2361. ISSN: 0022-3654 (Dec. 1977).

91. Benaim, M. & Weibull, J. W. Deterministic Approximation of Stochastic Evolution in Games. *Econometrica* **71,** 873–903. ISSN: 0012-9682 (May 2003).

92. ICES. *Report of the herring assessment working group for the area south of 62 deg N* tech. rep. (ICES, 2011), 745.

93. ICES. *ICES Advices. Book 6.* tech. rep. (ICES, Copenhagen, 2012).

94. Lieberson, S. *A Matter of Taste: How Names, Fashions, and Culture Change* xvi+334. ISBN: 97803000173871 (Yale Univeristy Press, New Haven & London, 2000).

95. Bernoulli, D. Exposition of a New Theory on the Measurement of Risk (Specimen Theoriae Nova de Mensura Sortis, 1738). *Econometrica* **22,** 23. ISSN: 00129682 (Jan. 1954).

96. Shoham, Y. & Leyton-Brown, K. *Multiagent Systems. Algorithmic, Game-Theoretic, and Logical Foundations* 504. ISBN: 9780521899437 (Cambridge University Press, 2008).

97. Von Neumann, J. & Morgenstern, O. Theory of Games and Economic Behavior. **2,** 625. ISSN: 00280836 (1944).

98. McFadden, D. L. in *Structural Analysis of discrete data with econometric application* (ed Manski, C. F.) (MIT Press, 1980).

99. Challet, D., Matteo, M. & Yi-Cheng, Z. *Minority Games* (Oxford University Press, Oxford, 2005).

100. Pareto, V. *Oeuvres complètes,* (Droz, Geneva).

101. Sen, A. *Development as Freedom,* (Oxford University Press, 1999).

102. Bavelas, A. A mathematical model for group structures. *Human Organization* **7,** 16–30 (1948).

103. Wasserman, S. & Faust, K. *Social Network Analysis: Methods and Applications* (Cambridge University Press, 1994).

104. Freeman, L. C. Centrality in Social Networks Conceptual Clarification. *Social Networks* **1,** 215–239 (1978).

105. Brass, D. J. Being in the Right Place: A Structural Analysis of Individual Influence in an Organization. *Administrative Science Quarterly* **29,** 518–539 (1984).

106. Calvó-Armengol, A., Patacchini, E. & Zenou, Y. Peer Effects and Social Networks in Education. *Review of Economic Studies* **76,** 1239 (2009).

107. König, M. D., Tessone, C. J. & Zenou, Y. *Nestedness in Networks: A Theoretical Model and Some Applications* 2009.

108. Ballester, C., Calvó-Armengol, A. & Zenou, Y. Who's Who in Networks. Wanted: The Key Player. *Econometrica* **74,** 1403–1417 (2006).

109. Bonacich, P. Power and Centrality: A Family of Measures. *American Journal of Sociology* **92,** 1170 (1987).

110. König, M. D. & Tessone, C. J. Network evolution based on centrality. *Physical Review E* **84,** 56108. ISSN: 1539-3755 (Nov. 2011).

111. Uzzi, B. The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review* **61,** 674–698 (1996).

112. Saavedra, S., Reed-Tsochas, F. & Uzzi, B. A simple model of bipartite cooperation for ecological and organizational networks. *Nature* **457,** 463–466. ISSN: 0028-0836 (2008).

113. Soramäki, K., Bech, M. L., Arnold, J., Glass, R. J. & Beyeler, W. E. The topology of interbank payment flows. *Physica A: Statistical Mechanics and its Applications* **379,** 317–333 (2007).

114. Goyal, S. & Joshi, S. Networks of Collaboration in Oligopoly. *Games and Economic Behavior* **43,** 57–85 (2003).

115. Park, J. & Newman, M. E. J. The statistical mechanics of networks. *Physical Review E* **70,** 66117 (2004).

116. Park, J. & Newman, M. E. J. Solution of the two-star model of a network. *Physical Review E* **70,** 66146 (2004).

117. Palla, G., Derényi, I., Farkas, I. & Vicsek, T. Statistical mechanics of topological phase transitions in networks. *Physical Review E* **69,** 46117 (2004).

118. Mills, W. C. *The Power Elite* (Oxford University Press, Oxford, 1956).

119. Wilkinson, R. & Pickett, K. *The Spirit Level: Why Equality is Better for Everyone* (Penguin, 2010).

120. Björklund, A. & Jäntti, M. Intergenerational Income Mobility in Sweden Compared to the United States. *American Economic Review* **87,** 1009–1018 (1997).

121. Andrews, D. & Leigh, A. More Inequality, less social mobility. *Applied Economics Letters* **16,** 1489–1492 (2009).

122. Arnold, V. I. *Catastrophe Theory* ISBN: 978-3-540-54811-9. <http://www.springerlink.com/index/10.1007/978-3-642-58124-3> (Springer Berlin Heidelberg, Berlin, Heidelberg, 1992).

123. Choudum, S. A simple proof of the Erdos-Gallai theorem on graph sequences. *Bulletin of the Australian Mathematical ....* <http://journals.cambridge.org/abstract%5C_S0004972700002872> (1986).

124. Diestel, R. *Graph Theory* ISBN: 978-3540261827. <http://www.springer.com/new+%5C&+forthcoming+titles+(default)/book/978-3-642-14278-9> (Springer-Verlag, Heidelberg, New York, 2005).

125. Erdős, P. & Rényi, A. On the evolution of random graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Sciences,* 17–61 (1960).

126. Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J. & Knuth, D. E. On the Lambert W function. *Advances in Computational Mathematics* **5,** 329–359. ISSN: 1019-7168 (Dec. 1996).

127. Magnus, J. & Neudecker, H. *Matrix Differential Calculus with application in statistics and econometrics* 3rd ed. **4,** 501. ISBN: 0471986321 (John Wiley, Nov. 2007).