Collective dynamics in binary games with memory.

Giacomo Baldo

The University of Leeds

School of Mathematics

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

October 2020

Acknowledgements

My PhD programme was funded by the Leeds Anniversary Research Scholarship (LARS) and the School of Mathematics at the University of Leeds. This research was supervised by Richard Mann, Sandro Azaele and Mauro Mobilia. I acknowledge the use of the High Performance Computing facilities (ARC) at the University of Leeds.

Abstract

Cyclic dynamics are displayed in many social, economic, physical and biological systems governed by negative frequency-dependent selection and delayed feedback. Here, I study the emergence and collapse of cyclic dynamics in an anti-coordination binary game with memory. In this model, individuals are faced with two options and aim to choose the option adopted by a minority in the group. Between adaptation moves, individuals learn about the behaviour of others by attending to and recollecting their choices. I introduce a novel double-fold definition of individual memory, incorporating a rate of observation (how often one collects new information) and duration (how long one remembers information). In the context of bounded rationality, these parameters mirror limitations due to cognitive and environmental constraints.

I show that finite and infinite observation rates generate different collective dynamics. In the limit of an infinite rate, the population exhibits deterministic and thus cyclic dynamics. In contrast, finite rates generate noise, enriching the game dynamics with different equilibria. The time evolution of the frequency of play of the two options in the group is investigated theoretically and computationally, and characterized as a function of the system parameters, particularly in terms of amplitude and period. The onset, decay and robustness of cyclic behaviour is discussed. To make this more concrete, the dynamics at play are illustrated with a specific example from fashion, intended as a sociological subject with collective and individual dynamics and imitation and distinction motifs. Applications to a number of other social, economic and biological contexts are also discussed.

Contents

List of Figures

Chapter 1

Introduction

From bacteria to people, when acting in groups, individuals often coordinate their actions giving rise to collective phenomena. Collective phenomena are patterns in the group behaviour that often stem from simple laws and mechanisms governing the interactions between individuals, such as positive and negative feedback and response threshold (Sumpter, 2005). Fascinatingly, when rules are simple enough, mathematics helps to model and predict the emergence of such patterns. At other times, computer simulations are necessary. Examples of collective phenomena include ants trailing (Wilson, 1962), crowd dynamics (Helbing *et al.*, 2005a; Schadschneider *et al.*, 2010) and the heartbeat (Strogatz, 2004). Multi-agent synchronisation is present across animal groups and biological systems. Rhythmic synchronisation can lock cells, such as pulsing pacemaker cells, and individuals, such as flashing fireflies, in phase. On the contrary, the evolution of turn-taking, such as in traffic congestion models (Helbing *et al.*, 2005b) and sentinel behaviour (Bednekoff, 1997), is an example of anti-phase coordination. Likewise, a double pendulum, which is an inanimate object, can reach anti-phase synchronisation of the two pendulums.

In the study of group behaviour, collective intelligence is the organisation into an emergent functional behaviour that results from individuals' interactions rather than from individual reasoning or global optimisation (Helbing & Johansson, 2013). In particular, the limits imposed on individuals to deliberate and optimise their behaviour are to be taken into account. In social and economic systems, the theory of bounded rationality (Simon, 1955) states that people's behaviour is constrained by the level of access to information and computational capacities compatible with both the individual and the environment. This theory is in opposition to the view that individuals possess perfect rationality entailing access to complete information and abilities to solve optimisation problems that maximise personal utility (Wheeler, 2018). The El Farol Bar problem, which I describe in chapter 3, is a computational model exemplifying how social and fi-

nancial agents under the constraints of bounded rationality can coordinate their actions using trial-and-error rather than deductive thinking.

Within the field of collective intelligence, my research interest lies in understanding the effects that bounded memory and stochastic differences in information have on group coordination. Specifically, I propose a binary choice model with memory in which individuals, between adaptation moves, observe a number of times the behaviour enacted by others in the group. Individuals who collect a finite number of observations are said to collect a finite batch. Likewise, individuals who collect an infinite number of observations are said to collect an infinite batch. Memory entails the ability to recollect a simple average quantity, such as the average number of people with a clean-shave, the average number of times that 'selling' was the minority choice and the average density of public-good molecules dispersed in a medium. The group behaviour will then be described, primarily, with the dynamics of the recurrence of each option in the population over time. It is well known that systems with a finite time delay between the moments that an agent perceives and reacts to information may cause oscillations in the dynamics (Erneux, 2009). In my original contribution to research, instead, I look at how much information needs sharing among individuals to induce periodic oscillations in the dynamics and to what extent these oscillations are sensitive to model assumptions and robust to individual heterogeneity. Overall, I present an extensive description of the population's emergent dynamics using analytical, numerical and simulation results. I show that infinite and finite batches generate different collective dynamics. The population exhibits deterministic dynamics in the limit of infinite batches and stochastic dynamics for finite batches. I show that deterministic learning always yields periodic oscillating dynamics. In contrast, stochasticity affects the game dynamics, with noise contributing to the formation of different equilibria in the population's dynamics.

The specifications of the model are very general, so that it is paradigmatic of a broad range of situations where individuals in large groups are confronted with two options. In social dynamics, the two options can represent opposite fashion choices, such as clean-shaving or not regarding men's facial grooming (Robinson, 1976) (being in the minority is motivated by the negative frequency-dependent attractiveness of beards (Alderman, 2017; Janif *et al.*, 2014)). In financial markets, the two options can represent buying and selling 'one share' (Challet, 2000; Challet *et al.*, 2013). The two options can also represent the choices available, within a bacterial population, to the single bacterium that chooses whether to contribute to the production of a public good or not (Cavaliere & Poyatos, 2013). Previous research has studied a similar problem. For example, the minority game (Challet, 2000; Challet *et al.*, 2013) is a financial market model that examines agents' efficiency at coordinating their actions when choosing between buying and selling 'one share'. However, my research will differ fundamentally when

compared to this game. The minority game explores how agents coordinate their actions using different heuristics, also called predictors, whereas I study the case in which all individual use the same heuristic.

In financial and economics applications, such as in the minority game and supply-demand chains, the price fluctuations are often seen as a source of market instability (Mackey, 1989). Consequently, studies intend to discover mechanisms that reduce such fluctuations. Nonetheless, oscillations are ubiquitous in commodity markets, and their mechanisms are still debated. Researchers are asking whether fluctuations are due to endogenous or exogenous effects and whether agents are rational or not (Gouel, 2012). The model that I present shows that endogenous fluctuations can arise once differences in information among agents are reduced.

Finally, I am interested in linking my model to some specific previous studies in the natural sciences. Multi-cellular systems, in which synchronisation occurs by global coupling of individual oscillators, have been widely studied according to the well-known principle first shown in the Kuramoto model. However, my interest lies in more recent models of cells that are quiescent below a critical cell density and oscillate synchronously above such criticality (De Monte *et al.*, 2007; Taylor *et al.*, 2009). The model that I present exhibits similar features. It shows a transition from a stable steady state to collective limit-cycle oscillations at a critical level of 'information availability'. Here, 'information availability' is controlled by the rate at which individuals sense or collect information about the state of the system.

Overall, in this thesis, I identify a general scenario for the emergence and decay of collective oscillations regulated by memory and an information-driven mechanism. Next is an outline of the content of the thesis.

In chapters 2 and 3, I review some of the methods and definitions used in game theory and evolutionary game theory. As part of this review, I have replicated a selection of results from the literature. I introduce the Prisoner's Dilemma (section 2.1) and spatial versions of the Prisoner's Dilemma without (sec. 2.2) and with memory (sec. 2.3), the definitions of Nash Equilibrium and evolutionarily stable strategy (sec. 2.4) and the Hawk-Dove game followed by an example of replicator equation (sec. 2.5). I conclude chapter 2 reviewing relevant features of a model of the Hawk-Dove game with memory (sec. 2.6). In chapter 3 I present a review of the El Farol bar problem (sec. 3.1) and qualitative features of the canonical minority game (sec. 3.2). The Hawk-Dove game with memory and canonical minority game constitute the basis for developing and discussing a new model in subsequent chapters.

Chapter 4 includes the definition of a new model and its dynamics (sec. 4.2 and 4.3.1), simulation data (sec. 4.3.2), results for the amplitude and period of the oscillations for deterministic dynamics (sec. 4.4), scaling properties (sec. 4.5),

a discussion on the transition between stable state and limit cycles (sec. 4.5.1) and a characterization of the 'space of strategies' (sec. 4.7). Chapter 5 discusses the model under different scenarios, including: heterogeneous memory duration (sec. 5.1.1), synchrony of the processes (sec. 5.2), the presence of an hysteresis (sec. 5.2.1), a scenario without individual memories (sec. 5.1.2), random choice (sec. 5.3), small populations (sec. 5.6.1) and more. Section 5.7 presents a review of how each 'dimension' of the model impacts on the amplitude of the oscillations. Chapter 6 contains a summary of the research and a discussion on its applications and implications.

Chapter 2

Evolutionary outcomes in games

2.1 The Prisoner's Dilemma

In game theory, the Prisoner's Dilemma (PD) models an interaction between two individuals faced with a binary decision. The game's name and its description are due to one of the earliest interpretation of the game given by the mathematician Albert Tucker when illustrating the game to an audience of psychologists (Morgan, 2007). The PD is described as follows. Two people were involved in the same crime but, interrogated by the police separately, could individually testify against the other or remain silent. The two individuals are guilty but there is not adequate evidence to convict them. Each prisoner has two options: to confess the crime or not to confess. If they both do not confess, then they can only be charged with a minor punishment; if they both confess they will be prosecuted but will receive less than the most severe sentence; but if one confesses and the other does not, then the confessor will receive lenient treatment while the latter will get the full sentence (Luce & Raiffa, 1957; Morgan, 2007).

Normally a two-player game is described using a payoff matrix (Nowak, 2006). For the Prisoner's Dilemma the matrix is given as follows (Poundstone, 1992):

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} \tag{2.1}$$

The entries T, R, P, S quantify the payoffs obtained by a focal individual depending on the behaviour of both players. The following conditions define a Prisoner's Dilemma:

$$T > R > P > S. \tag{2.2}$$

The two strategies available are denoted with C for cooperation and D for defection. It is assumed that a prisoner cooperates (with the other prisoner) when

remaining silent and defects when confessing. If both prisoners remain silent (FCC), they receive the payoff R. If a prisoner confesses (D) and the other does not (C), the first receives the highest payoff available T and the latter receives the lowest payoff available S. When both confess (DD), each receives the payoff P. Any self-interested prisoner who aims to maximise his or her reward based solely on payoff matrix 2.1 will choose strategy D because this choice yields the highest payoff regardless of the other prisoner's strategy, given that $T > R$ and $P > S$. Therefore, self-interested individuals end up defecting, obtaining a reward of value P. However, if they both had cooperated, the two individuals would have got a higher payoff, R (Axerold & Dion, 1988). As a result, the Prisoner's Dilemma has become paradigmatic of the 'tragedy of the commons' and challenges faced in establishing cooperative behaviour among purely self-interested individuals. Cooperation, however, can be established and maintained through several mechanisms. In the following sections, I describe two such mechanisms, one based on repetition and reciprocity (section 2.1) and another based on spatial assortment (section 2.2).

Sometimes the payoffs of the PD are given as shown in matrix 2.3. In this payoff matrix, the value b is the benefit acquired from the interaction with a cooperator, and c is the cost associated with cooperative behaviour, with the benefit exceeding the cost, i.e. $b > c$. The benefit b is sometimes referred to as 'donation' and, for this reason, the game has also been named 'donation game' (Hilbe *et al.*, 2013).

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \end{array} \tag{2.3}$$

The Iterated Prisoner's Dilemma

The concept of an Iterated Prisoner's Dilemma (IPD) is introduced to describe the situation in which two players interact repeatedly and each interaction is governed by the PD. The IPD was first studied through two computer tournaments, held by Axelrod in the 1980s, in which various strategies played against each other. Examples of strategy submitted for competition were 'always cooperate' (ALLC) and 'always defect' (ALLD). Both tournaments saw the strategy named Tit-for-Tat (TFT) winning (Imhof *et al.*, 2007). TFT is a strategy in which an individual cooperates in the first interaction and afterwards replicates the other individual's behaviour in the previous turn. As such, TFT's success makes it a candidate strategy for the evolutionary origin of reciprocity (Axelrod & Dion, 1988). Later, Win-stay-lose-shift (WSLS), also known as Pavlov, was introduced by Nowak & Sigmund (1993) as a strategy that outperforms TFT. In the late 1980s and early 1990s, game-theoretic models initiated a new framework based on spatial

modelling, which discovered the spontaneous emergence of cooperation through spatial assortment. As such, TFT has allowed a description of the evolutionary origin of reciprocity (Axelrod & Dion, 1988).

2.2 Spatial models of the Prisoner's Dilemma

Early work on spatial modelling of the Prisoner's Dilemma was performed by Martin Nowak and Robert May (Novak & May, 1992, 1993). They model a population placed on a square lattice, and each individual only interacts with the closest eight neighbours. The dynamics are regulated by pairwise interactions, and each interaction is modelled with a simplified version of the Prisoner's Dilemma. The game's payoffs are described in the matrix 2.4, where b is greater than 1. With this simplification, the authors can describe the behaviour of the system depending on one parameter only.

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix} \end{array} \tag{2.4}$$

The evolution of the prevalence of the traits C and D is simulated computationally. In the first round, each lattice site is randomly assigned either to a cooperator or a defector. Each individual collects nine payoffs in total (one from each interaction). The sum of the payoffs collected in the last round is defined to be the individual's fitness. Each individual copies the strategy of the best performing neighbour (the neighbour with the highest fitness from the previous round) at each subsequent turn. All interactions are synchronous across the lattice and, apart from the random initial distribution of cooperators and defectors, the dynamics of this model is deterministic. I will present a simulation of the model above for a population distributed on a 200 by 200 square grid that evolves according to this scenario for $b = 1.85$. Consistently with the work of Novak & May (1993), I consider the case of periodic boundary conditions that allow avoiding border effects. The authors performed the simulations on a 20 by 20 grid, starting with 90% cooperators and 10% defectors. I suppose that such a high percentage of cooperators is needed to avoid the cooperators' initial extinction. Simulations show that the total number of cooperators can sharply decrease in an initially randomly distributed population, although numbers can recover and reach a static level. On the other hand, in social situations resembling the Prisoner's Dilemma, there is evidence showing that humans 'do not defect in the first round'. Usually, models take this into account by establishing a significant initial fraction of cooperators (McNamara *et al.*, 2004).

Next, I briefly illustrate the model for $b = 1.85$ by comparing the outcome of simulations I run to the original results. In figure 2.1 I present the result of my simulations next to results from the original article. Both graphs show the density of cooperators over 200 time steps.

My result | Novak & May (1993)

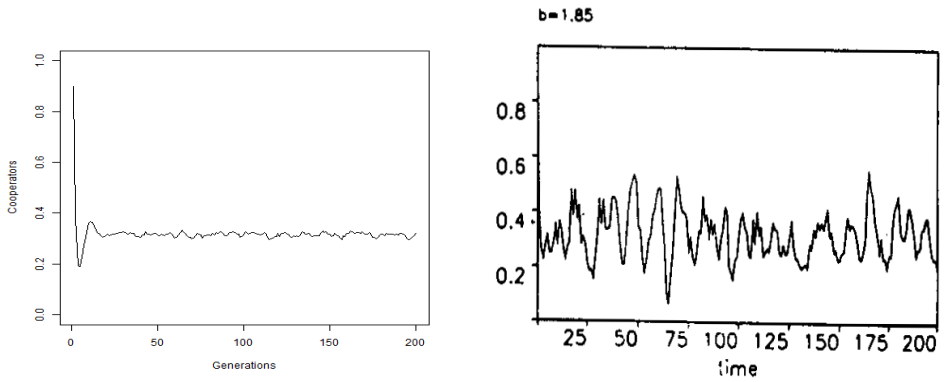

Figure 2.1: Density of cooperators over 200 generations for the simulation of the PD on a square lattice with $b = 1.85$. Initial conditions have 90 % cooperators and 10% defectors. The lattice used is a 200 by 200 grid for the graph on the left and a 20 by 20 grid for the graph on the right, which explains the relative lower variability in the left-hand plot. Both have approximately the same mean after convergence.

Figure 2.2 shows the spatial distribution of the population described in the previous figure for $b = 1.85$ and additionally for $b = 1.45$ for three consecutive time steps. Cooperators are depicted in blue and defectors in red.

Later models have pointed at the biological limits of a deterministic rule and introduced stochastic evolutionary rules. A rule that proved quite popular was inherited by statistical physics and used among the first times by Szabo & Toke (1998). The authors introduced the Fermi-Dirac distribution (2.5) to indicate the probability W that, once player X has picked neighbour Y, he adopts the neighbour Y's strategy, where f_K is the last payoff collected by individual K.

$$W = \frac{1}{1 + \exp[-\beta(f_Y - f_X)]} \tag{2.5}$$

The parameter β, which corresponds to 'inverse temperature' in statistical physics, has the meaning of intensity of selection in evolutionary models and response sensitivity (Galla, 2009) when describing learning dynamics. The value of β describes how social information is used by individuals in order to update their strategy. For $\beta \to \infty$, the system will display strong selection, in the sense that the best

performing strategy will be chosen at each iteration, which corresponds to a deterministic strategy selection. For $\beta \to 0$, the system will display weak selection, in the sense that there will be only a minor correction from a random allocation of strategies and random strategy selection when $\beta = 0$.

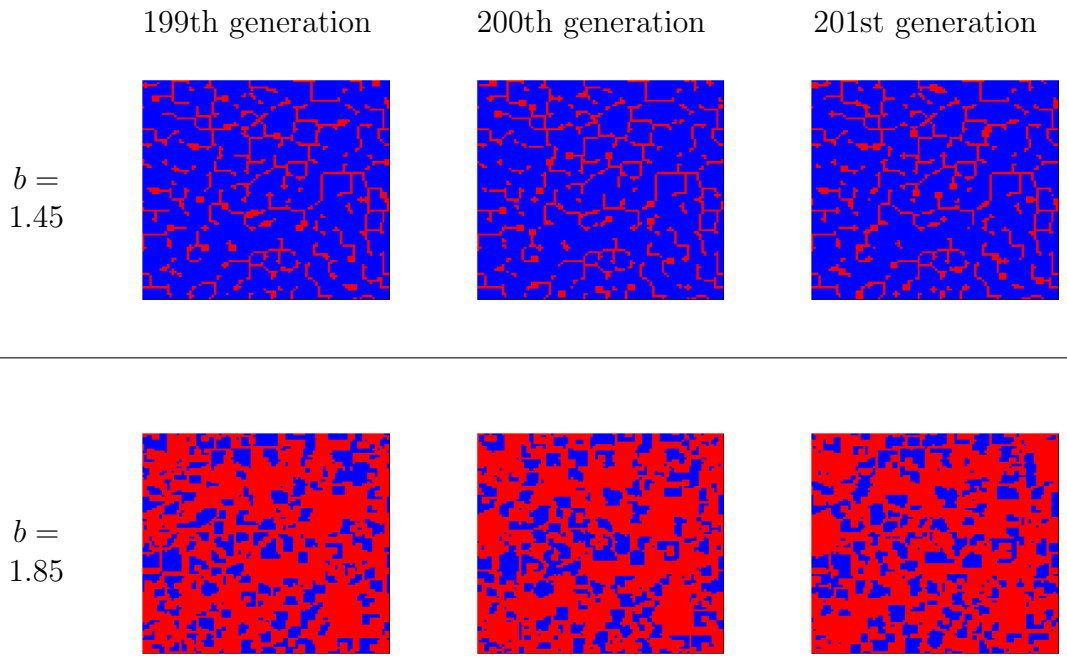

Figure 2.2: Spatial distribution of cooperators (in blue) and defectors (in red) over three consecutive generations for the Prisoner's Dilemma on a square lattice for $b = 1.45$ (majority of cooperators/blue after convergence) and $b = 1.85$ (majority of defectors/red after convergence). Cooperators live on in clusters. This is a replication on a 200 by 200 lattice grid of the model in Novak & May (1993) which was presented on a 20 by 20 lattice grid. The number of neighbours is 8.

In figure 2.3 I present the spatial evolution of cooperation and defection after a simple modification of Novak & May (1993). I model a scenario in which agents choose their strategy with a probabilistic rule based on the fitness of the neighbours and self. Following Ho *et al.* (2007) and Sato & Crutchfield (2003), the probability $W_{\to k}$ to replicate the strategy of neighbour k or self is given in formula 2.6. In this formula, f_k stands for the fitness of the particular agent to be copied, and f_i varies, taking the value of the fitness of all neighbours and self. This scenario differs from Novak & May (1993) only for having introduced a stochastic update rule, and thus, strategy updating remains synchronous across

the population.

$$W_{\to k} = \frac{e^{\beta f_k}}{\sum_{i=1}^{9} e^{\beta f_i}} \tag{2.6}$$

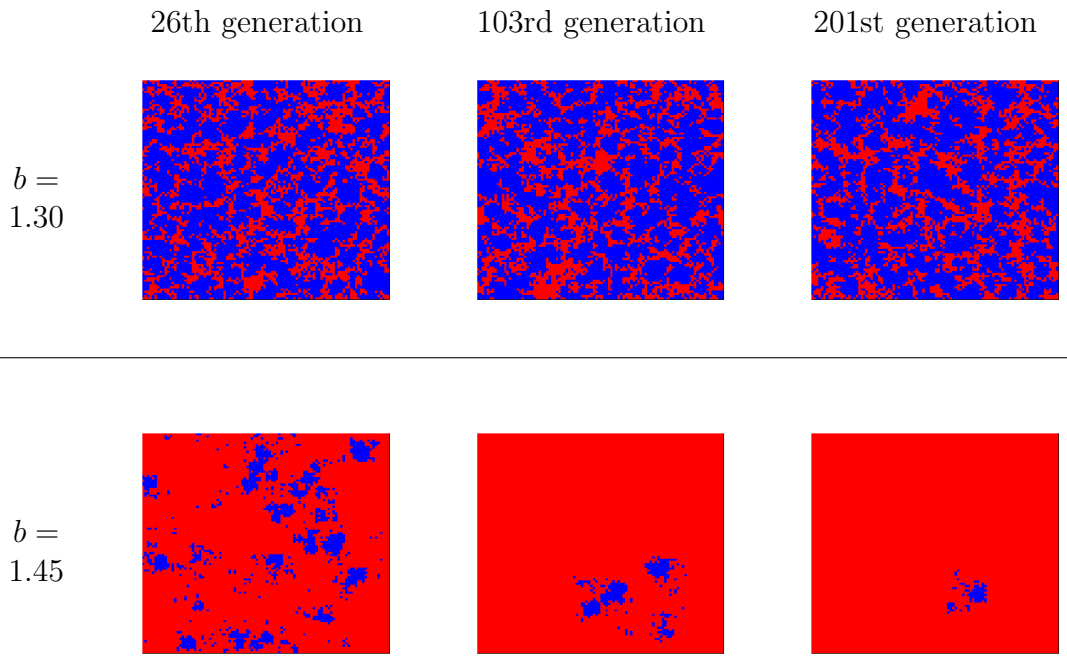

Figure 2.3: Spatial distribution of cooperators (in blue) and defectors (in red) over three generations for the Prisoner's Dilemma on a square lattice with stochastic updating for $b = 1.30$ and $b = 1.45$ and $\beta = 1$.

Comparing figure 2.2 and figure 2.3, one sees that stochastic updating changes the outcome of the game. In both the deterministic and probabilistic scenarios, cooperators live on in clusters, but clusters have a fragmented shape in the probabilistic scenario. To illustrate in more detail how stochasticity affects the game, figure 2.4 compares the evolutionary outcome for a relevant selection of deterministic and probabilistic scenarios.

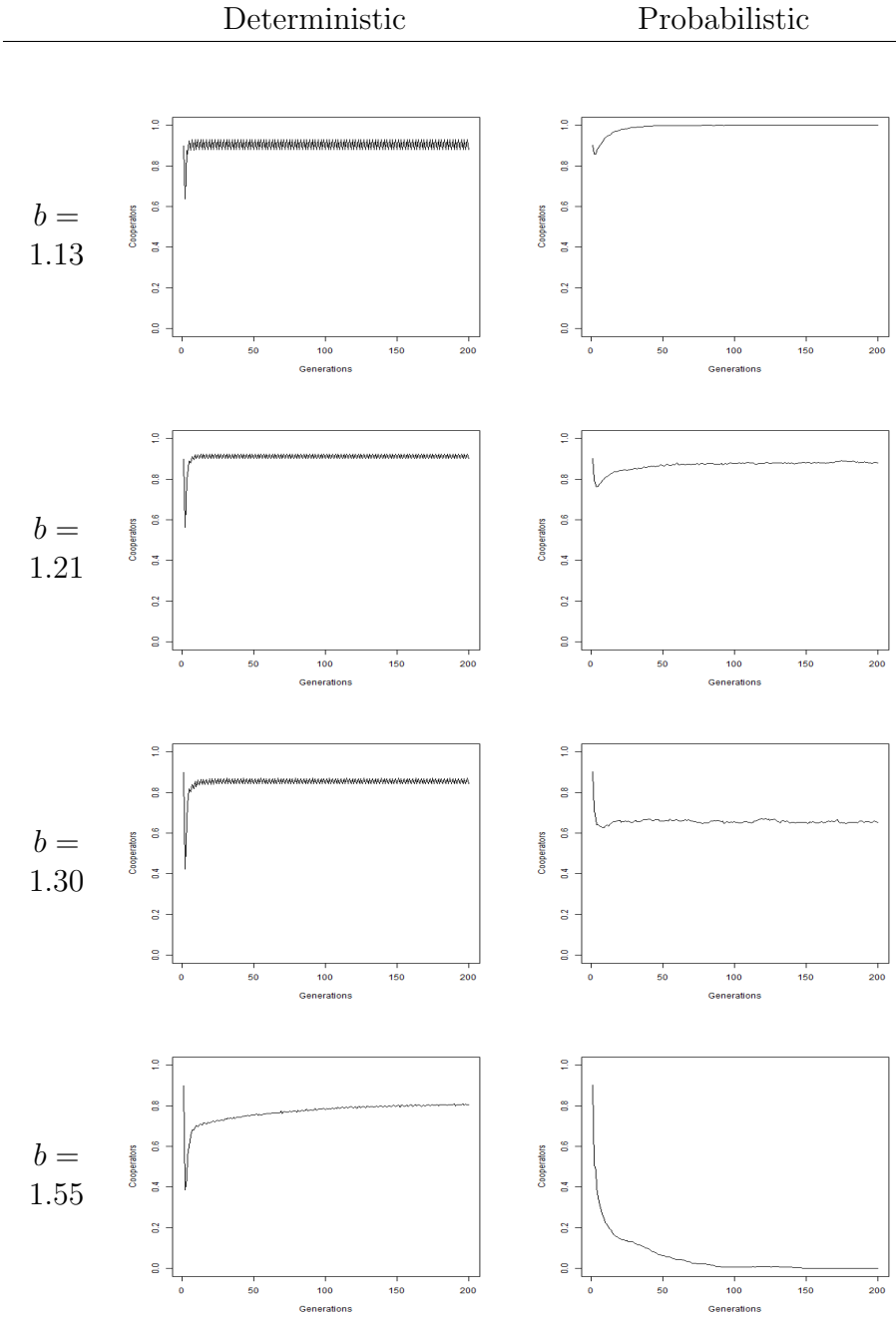

Figure 2.4: Deterministic and probabilistic outcomes of the Prisoners' Dilemma on a lattice. Time evolution over 200 time units of cooperators density with an initial 90% density of cooperators. As a result of the stochastic update rule, cooperation is strengthened for low values of b and becomes strongly negatively affected as b increases.

From figure 2.4 a pattern starts to emerge. The impact of stochasticity on the dynamics is two-sided: after the introduction of a stochastic update rule, the level of cooperation improves for low values of b and is negatively affected as b increases. This result is mediated by clusters' fragmentation, which increases the benefit acquired by defectors from having spatial proximity to cooperators when b is large.

Overall, in this section, I presented models of the Prisoner's Dilemma in a two-dimensional lattice showing that spatial assortment alone can sustain cooperation. In the next section, I review a study showing how spatial assortment coupled with a memory effect maintains cooperation in a one-dimensional lattice.

2.3 A one-dimensional spatial model of the Prisoner's Dilemma

In this section, I review and partially replicate a model (Gelimson *et al.*, 2013) of competition between cooperators and defectors for a Prisoner's Dilemma game in a one-dimensional lattice. I will also point out a specific result for which I disagree with the authors (Gelimson *et al.*, 2013). The dynamics in the original work are simulated both on a one- and two-dimensional lattice, but here I only consider the one-dimensional case. In the one dimensional lattice, each individual has two neighbours only, and, unlike in the two-dimensional lattice, the spatial effect is not strong enough to sustain clusters of cooperators. However, the authors propose that if the current fitness of an individual takes into account past interactions than cooperation can still be sustained in clusters. Specifically, the authors use a fitness collection process that entails accumulating payoffs during an individual's entire lifetime. Consequently, individual fitness is determined by all collected payoffs and not only those collected on the last interaction. For individual i, his fitness f_i is defined as follows, where n is the number of interactions in his lifetime and p_i^l is the payoff collected by individual i in his l-th interaction.

$$f_i = 1 + \frac{1}{n}\sum_{l=1}^{n} p_i^l \tag{2.7}$$

Next, the authors assume that there are a birth and death process. In this model, individuals do not change or update their trait (cooperator or defector) during their lifespan. Instead, strategies replication is obtained with the reproduction of existing individuals that produce offsprings of an identical trait.

In this system's evolution, the population is initialised with cooperators and defectors randomly distributed with equal probability along the one-dimensional lattice, with no empty sites. Each individual interacts with the two adjacent neighbours only, and three processes are unfolding:

1. payoff collection
2. death
3. birth

Death and birth are intertwined. Every time an individual dies, he is immediately replaced by the offspring of one of the two neighbours. For brevity, I name this event a death-birth. In the original article, the authors also incorporate a parameter modelling mobility, but I restrict the analysis to the scenario without mobility for this review. Following the original work notation, individuals engage

with a neighbour, thus collecting a payoff, with rate r. The authors model the payoffs using payoff matrix 2.3 with $b = 1$ obtaining payoff matrix 2.8, where $0 < c < 1$.

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} 1-c & -c \\ 1 & 0 \end{pmatrix} \end{array} \tag{2.8}$$

The rate for an event of death-birth depends on the fitness of individuals. The rate, let say a, for individual j to die and leave an empty site for an offspring of neighbour i is given in equations 2.9 and 2.10. In these equations, f_i indicates the fitness of neighbour i, while $s_k = 0$ indicates that individual k is a cooperator and $s_k = 1$ indicates that individual k is a defector. Importantly, the rates are strictly positive only for neighbours of differing traits and zero otherwise (meaning that there is no death-birth process between neighbours of the same trait). As a result, the death-birth process only happens at the border between clusters of opposing strategies and does not happen within clusters. As another result, individuals located within clusters develop a fitness based on a 'longer' life-span. Since interactions among cooperators yield higher payoffs (i.e. $1 - c$) than interactions among defectors (i.e. 0), this confers cooperators an advantage.

$$a_{\mathrm{C} \to \mathrm{D}}(j) = s_{\mathrm{j}} \frac{1}{2} \sum_{i=1}^{2} f_{\mathrm{i}}(1 - s_{\mathrm{i}}) \tag{2.9}$$

$$a_{\mathrm{D} \to \mathrm{C}}(j) = (1 - s_{\mathrm{j}}) \frac{1}{2} \sum_{i=1}^{2} f_{\mathrm{i}} s_{\mathrm{i}} \tag{2.10}$$

The control parameters are the cost c linked to cooperation and the rate r, $r > 0$, of payoff collection. However, the authors point out that the dynamics' outcome is not dependent on the specific value of r. The results of their simulations show that for $r = 5$, the cost $c \approxeq 0.348$ is a critical value, with cooperators reaching fixation for $c < 0.348$ and defectors reaching fixation otherwise. I simulated the dynamics, and in figure 2.5 I report four realizations that agree with this finding. For $c = 0.1$ cooperators reach extinction (top figure), for $c = 0.5$ defectors reach extinction (bottom figure), for $c = 0.348$ (middle figures) both scenarios have happened and the fixation time is considerably longer. In the simulations reported in figure 2.5, the results hold for $r = 1$. Overall, these findings show that the payoff-accumulation effect enhances cooperation. This effect is mediated by the impact on clusters formation, with the accumulation effect improving the fitness of individuals located within clusters of cooperators. The accumulation effect is significant because there is no effect of spatial reciprocity that can alone sustain cooperation in a one-dimensional lattice, contrary to what was described for a two-dimensional lattice in section 2.2.

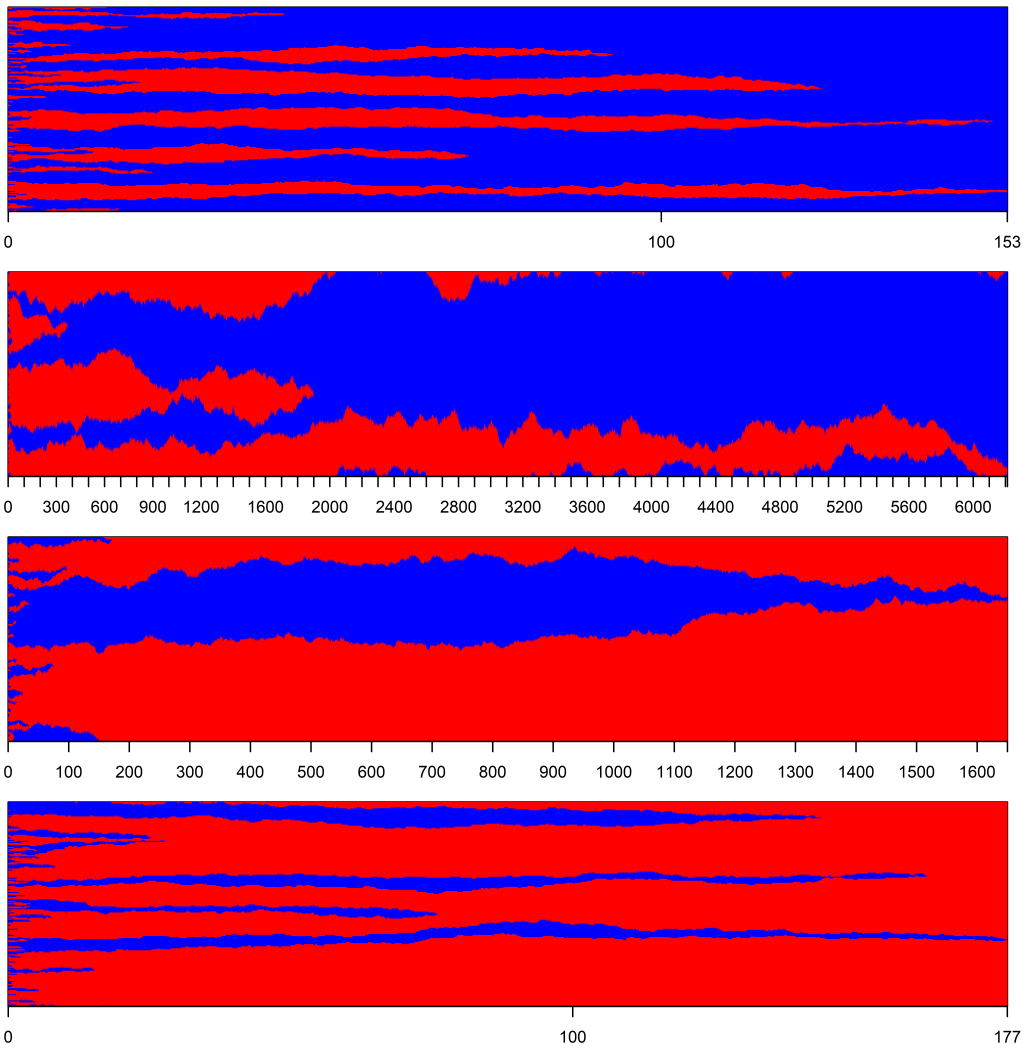

Figure 2.5: Time evolution of cooperators (blue) and defectors (red) on a one-dimensional lattice with 200 individuals. For $c = 0.1$ defectors reach extinction (top figure), for $c = 0.5$ cooperators reach extinction (bottom figure), for $c = 0.348$ both scenarios have happened and the fixation time is considerably longer. The payoff collection rate is set to $r = 1$. Mobility is set to 0.

2.3.1 Predicting the critical value

The aim is now to predict the critical value for the cost c. In this process, I follow the methods developed in Gelimson *et al.* (2013).

Modelling the interface between clusters

I focus on analysing the dynamics at the interface between a cluster of cooperators and a cluster of defectors. It is precisely at the interface that local sites can swap between hosting a cooperator and hosting a defector. It is further assumed that the agents inside the clusters have been in their cluster for a long time so that their accumulated average payoff asymptotically equals a one-shot game payoff. For example, a cooperator that has interacted with cooperators for a long time has accumulated an average payoff equal to $1 - c$ (cf. payoff matrix 2.8). Let us call this his reward. A defector that has interacted with a defector has accumulated an average payoff (his reward) equal to 0. For modelling purposes, only the two individuals at the cluster's interface are no longer assumed to be part of the cluster. Instead, for the two players located at the interface, their reward is calculated as the average of two one-shot games, one with each neighbour. The authors acknowledge that this is a strong assumption that they use to estimate the critical value $c \approxeq 0.348$. Next, I proceed to replicate their calculations to show that they do not predict the critical value.

Figure 2.6 is an illustration of the payoffs (left T shape) and fitness (right T shape) under the strong assumption mentioned. The top rows in figure 2.6 are made of 6 adjacent sites that represent a relevant portion of the lattice where two clusters meet, with the portion of the red cluster of defectors on the right, the portion of the blue cluster of cooperators on the left and two sites in the middle. The transition between cooperation and defection (and vice-versa) can happen only in the middle. The four two-colour combinations below the first row represent the four two-strategy configurations that the two sites at the interface can assume. In addition, the numbers in each cell represent the payoffs (left image) and rewards (right image) obtained under the strong assumption mentioned above. For example, consider the blue-red two-site combination. The blue site represents a cooperator placed between a cooperator and a defector. His reward is the average of the payoffs received, as a cooperator, from the interactions with a cooperator and a defector. This is $0.5(1 - c) + 0.5(-c)$ which equals $\frac{1}{2} - c$ as shown in the relevant site. The corresponding fitness (1+ reward, cf. 2.7) is $\frac{3}{2} - c$ as shown on the relevant site in the right image.

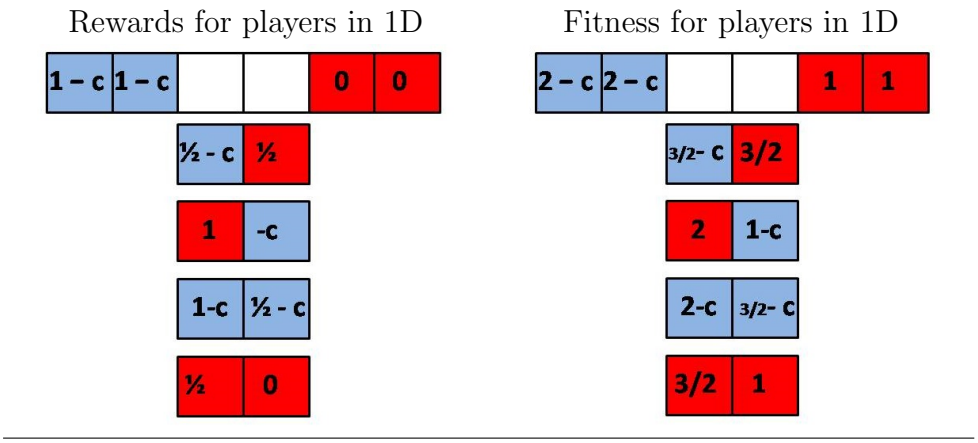

Figure 2.6: Reward and corresponding fitness for six consecutive sites/players in a specific portion of a one-dimensional lattice. Each site is occupied by either a cooperator (blue) or a defector (red); all possible combinations are shown for the two middle sites. For each player, the reward is defined as the average of all payoffs collected (more details in the main text above). The fitness is defined as 1+reward.

Transition rate matrix

Next I can calculate the transition rates between the four two-strategy combinations. The rates for individual j to be taken over by an offspring of individual i (a death-birth) is given in equations 2.9 and 2.10. In these equations, f_i indicates the fitness of neighbour i while $s_{i,j} = 0$ for cooperators and $s_{i,j} = 1$ for defectors. For example the rate $T_{1\to4}$ for blue-red to switch to red-red (from first combination to fourth combination) is given by $\frac{1}{2}\frac{3}{2} + \frac{1}{2}0 = \frac{3}{4}$. Similarly, the rate $T_{4\to1}$ for red-red to switch to blue-red (from fourth combination to first combination) is given by $\frac{1}{2}(2-c) + \frac{1}{2}0 = 1 - \frac{c}{2}$. In this manner, the transition rate matrix describing the continuous Markov Process is given by

$$T = \begin{pmatrix} -\frac{3}{2} + \frac{c}{2} & 0 & \frac{1}{2} & 1 - \frac{c}{2} \\ 0 & -3 + c & 0 & 0 \\ \frac{3}{4} - \frac{c}{2} & \frac{3}{2} - c & -\frac{1}{2} & 0 \\ \frac{3}{4} & \frac{3}{2} & 0 & -1 + \frac{c}{2} \end{pmatrix}$$

where the diagonal elements are placed so that the elements of each column add to 0. The red ink indicates a quantity missing from the original article.

Stationary distribution vector

Following the method proposed in the original article, to find the stationary probability vector, I calculate the vector $\vec{v}$ such that $T \cdot \vec{v} = 0$. From row II, I obtain $v_2 = 0$. One can see that row I is a linear combination of III and IV. As a result:

$$(\frac{3}{4} - \frac{c}{2})v_1 = \frac{1}{2}v_3 \tag{2.11}$$

$$\frac{3}{4}v_1 = (1 - \frac{c}{2})v_4 \tag{2.12}$$

which gives

$$v_3 = (\frac{3}{2} - c)v_1 \tag{2.13}$$

$$v_4 = \frac{\frac{3}{4}}{1 - \frac{c}{2}}v_1 \tag{2.14}$$

The final prevalence of trait C or D is determined by the sign of $v_3 - v_4$ and thus the critical value for c is obtained by setting $v_3 - v_4 = 0$. Calculating $v_3 - v_4$ I obtain $\frac{2c^2-7c+3}{2(2-c)}v_1$. The roots for the numerator are $c_0 = \frac{7\pm5}{4}$.

Since $0 < c < 1$, the solution that one can accept is $c = 1/2$ which does not predict the simulations results correctly. The authors, following from the algebraic error, predicted $c = 0.348$.

Discussion and conclusion

After replicating the results, I conclude that the assumptions made in the original article to approximate payoff collection at the interface between clusters are too strong. I showed that the analytical method suggested by the authors does not predict the critical cost that separates the scenario in which defection reaches fixation from the scenario in which cooperation reaches fixation. In practice, their method overestimates the resilience of cooperators. It is perhaps not completely incidental that the algebraic error in the computation of matrix T predicts a lower value (i.e. $c = 0.348$) for the critical value c compared to the value obtained with no algebraic error (i.e. $c = 0.5$). In fact, the algebraic error reduced the rate in position $(1, 4)$, namely $T_{4\to1}$. The incorrect value $T_{4\to1} = 1 - c$ could have been derived by assuming that a cooperator in a long-standing cluster has a reward equal to $2 - 2c$ rather than $2 - c$, obtaining the rate $\frac{1}{2}(2 - 2c) + \frac{1}{2}0 = 1 - c$. Compared to the strong assumptions in the original model, this change would model a lower fitness for a cooperator within a long-standing cluster that finds himself next to a defector. This assumption would reflect the reality of the model. I have additionally verified computationally that there is a non-negligible effect on

the rewards of at least three individuals at the interface between clusters, rather than two only. This strongly suggests that the mathematical analysis of the transition rates at the interface between clusters shall take into account the effect on at least three neighbours rather than two. Nevertheless, the above analysis can help understand and model the dynamics at the interface between long-standing clusters. To finish, one limitation of the model reviewed here is that births and deaths only occur at the interface between clusters. This assumption implies that, once one of two strategies in the population has reached fixation, there is no longer an active birth and death process. It seems reasonable to consider that one could extend the birth and death process to all individuals.

2.4 Nash equilibrium and ESS

The Prisoner's Dilemma described earlier is a two-person game in which each player has two pure strategies available, namely cooperation and defection. In game theory, the notion of strategies also includes mixed strategies. A mixed strategy is a probability distribution over the pure strategies available. With two strategies available, a mixed strategy defines the probability of playing each one of the two pure strategies.

A number of concepts have been developed to describe strategies that are 'stable states' of a game. Here, I introduce the notions of Nash equilibrium and evolutionarily stable strategy (ESS) in two-person games.

A Nash equilibrium (NE) defines a set of strategies, one for each player of a game, that has the property that each player's strategy is his best response to the strategy of other players (Holt & Roth, 2004). In a NE, no player can improve his response by unilaterally changing his strategy. A NE can be defined for both pure and mixed strategies. Importantly, a NE can be asymmetric because it is defined as a set of strategies, i.e. one for each player. This asymmetry implies that it can describe a stable state in which different players play different strategies. This scenario can apply to equilibria in economic systems where different actors have different roles in the system. For example, one can consider equilibria for advertising strategies in distribution channels with separate roles for manufacturers and retailers (Jørgensen & Zaccour, 2012). However, in population dynamics, it is often assumed that agents have identical roles (Mesterson-Gibbons, 2001). Consequently, equilibria are symmetric (Nowak, 2006). Next, I present the formal definition of Nash equilibrium for symmetric strategies.

Let $f(u, v)$ denote the payoff for strategy u playing against strategy v in a two-person game. Formally, strategy u is a Nash equilibrium if, for all available v strategies, $v \neq u$:

$$f(u, u) \geq f(v, u). \tag{2.15}$$

In addition, strategy u is a strict Nash equilibrium if, for all available v strategies, $v \neq u$,

$$f(u, u) > f(v, u). \tag{2.16}$$

'Evolutionarily stable strategy' (ESS) is a popular equilibrium notion in evolutionary dynamics. It defines a strategy that, once adopted by all members of a population, resists invasion by any other strategy that is initially rare (Choe, 2019). Formally, strategy u is an ESS if, for all available v strategies, $v \neq u$, either

$$f(u, u) > f(v, u) \tag{2.17}$$

or

$$f(u, u) = f(v, u) \quad \textit{and} \quad f(v, u) > f(v, v). \tag{2.18}$$

Furthermore, strategy u is a weak ESS if, for all available v strategies, $v \neq u$, either

$$f(u,u) > f(v,u) \tag{2.19}$$

or

$$f(u,u) = f(v,u) \quad and \quad f(v,u) \geq f(v,v). \tag{2.20}$$

Thus defined, a weak ESS is stable against invasion by selection but not neutral drift. Overall, strict Nash equilibrium implies ESS, ESS implies weak ESS and weak ESS implies Nash equilibrium. The formal definitions of this section can be found in the book by Nowak (2006).

2.5 The Hawk-Dove game

John Maynard Smith proposed the concept of evolutionarily stable strategy (ESS) and, together with George Price, presented the Hawk-Dove game and the War of Attrition game as the first examples of games modelling animal conflict (Smith & Price, 1973). The Hawk-Dove game is a game with two pure strategies commonly named Hawk and Dove. In its most general definition, the Hawk-Dove is described using payoff matrix 2.21 subject to the constraints $S > P > T > R$.

$$\begin{array}{c} \\ H \\ D \end{array} \begin{array}{c} \begin{array}{cc} H & D \end{array} \\ \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} \tag{2.21}$$

It does differ from the Prisoner's Dilemma. In the Hawk-Dove game, the best strategy adopted by one individual depends on the opponent's strategy: since $S > P$, the best response to Dove is Hawk; since $T > R$, the best response to Hawk is Dove. Smith & Price (1973) introduce the Hawk Dove game to describe the conflict between two animals over the control of a territory. The animals are of the same species but can show different personalities and adopt different strategies. Hawk-types engage in conflicts, whereas Dove-types flee when an opponent wants to engage in a fight. The payoff matrix of the Hawk-Dove game is originally written not as 2.21 but in the less general form 2.22, subject to the constraints $C > V > 0$:

$$\begin{array}{c} \\ H \\ D \end{array} \begin{array}{c} \begin{array}{cc} H & D \end{array} \\ \begin{pmatrix} (V-C)/2 & V \\ 0 & V/2 \end{pmatrix} \end{array} \tag{2.22}$$

In matrix 2.22, C can be interpreted as the cost of engaging in a fight over territory or resources of value V. Let us briefly explore this game. There are four

events possible, namely a Hawk meets a Hawk, a Dove meets a Dove, a Hawk meets a Dove and a Dove meets a Hawk. I denote these events with DD, HH, HD and DH respectively. Suppose that the probability of a focal player to play Hawk is u (his mixed-strategy) and the current probability for him to meet a Hawk in the population is v (on average, this is equivalent to meeting individuals playing Hawk with probability v). I denote the expected payoff for a focal player with $f(u, v)$, i.e. a function of u and v, and obtain

$$f(u,v) = (1-u)(1-v)\frac{V}{2} + u(1-v)V + uv\frac{V-C}{2}. \tag{2.23}$$

By introducing the substitution $\theta = \frac{V}{C}$, equation 2.23 becomes

$$f(u,v) = \frac{1}{2}[V(1-v) + c(\theta - v)u]. \tag{2.24}$$

Equation 2.24 implies that the best response to strategy v is Hawk ($u = 1$) if $v < \theta$ and Dove ($u = 0$) if $v > \theta$. There is no best (or worst) response for $v = \theta$ as all strategies perform identically against it (Mesterton-Gibbons, 2019). Next, I show that the mixed-strategy $u = \theta$ is a Nash equilibrium and ESS. It is easy to verify that, for all available v strategies, $v \neq u$,

$$f(\theta, \theta) - f(v, \theta) = 0 \tag{2.25}$$

and

$$f(\theta, v) - f(v, v) = \frac{1}{2}C(\theta - w)^2 > 0 \tag{2.26}$$

Therefore, since equation 2.25 satisfies condition 2.15 and equation 2.26 satisfies condition 2.18, the mixed-strategy $\theta = \frac{V}{C}$ is a Nash equilibrium and an evolutionarily stable strategy. To sum up, when the payoff matrix is given in form 2.22, the mixed-strategies Nash equilibrium and ESS for the Hawk-Dove game are given by choosing strategy Hawk with probability $\frac{V}{C}$ and strategy Dove with probability $1 - \frac{V}{C}$. In a large well-mixed population, the equilibrium can be replicated with pure strategies. The pure strategies Nash equilibrium corresponds to a fraction $\frac{V}{C}$ of players adopting strategy Hawk and a fraction $1 - \frac{V}{C}$ adopting strategy Dove.

Since its formulation, the Hawk-Dove game has been used in a number of ways to explain, for example, animal conflict (Smith, 1982; Smith & Price, 1973), animal contest with status signalling (Kim, 1995), pedestrian congestions (Helio-vaara *et al.*, 2013), colour polymorphism in birds (Kokko *et al.*, 2014) and cat domestication (Auger & Pontier, 1998).

2.5.1 The replicator equation

Here I introduce the traditional replicator equation (Nowak, 2006) for the specific case of the Hawk-Dove game played in a large well-mixed population. Let ϕ

denote the probability of meeting a Hawk in the population. Next, let us focus on a single individual and call him a focal individual. Using payoff matrix 2.22, the expected payoff for a focal Hawk is given by f_H such that

$$f_H = \phi\frac{V-C}{2} + (1-\phi)V. \tag{2.27}$$

Similarly, for a focal Dove

$$f_D = (1-\phi)\frac{V}{2}. \tag{2.28}$$

The average payoff in the population is given by f such that

$$f = \phi f_H + (1-\phi)f_D. \tag{2.29}$$

By equating fitness with payoff, the replicator equation for Hawks is given by

$$\dot{\phi} = \phi(f_H - f). \tag{2.30}$$

The replicator equation is fundamental in evolutionary dynamics and describes frequency-dependent selection in an infinitely large population. It links the variation in the frequency of a trait to the difference between the fitness of the trait and the average fitness in the population (Nowak, 2006). Equation 2.30 can be expanded and simplified to obtain

$$\dot{\phi} = \phi(1-\phi)\frac{1}{2}(V - \phi C). \tag{2.31}$$

Equation 2.31 has three fixed points, namely $\phi = 0$, $\phi = 1$ and $\phi = \theta = \frac{V}{C}$, where $0 < \frac{V}{C} < 1$ by definition of the Hawk-Dove game. By looking at the sign of $\dot{\phi}$ as a function of θ, one finds that the only stable fixed point is $\phi = \frac{V}{C}$. This result shows that the ESS of the Hawk-Dove game is also the only stable attractor of the replicator equation.

2.6 Hawk-Dove game with memory

2.6.1 The model

Burridge *et al.* (2017) study a well-mixed large population of individuals with memory that play the Hawk-Dove game with mixed-strategies. In the classic well-mixed scenario described above, a population converges to the ESS. However, the authors show that the introduction of memory can create instabilities and give rise to cyclic oscillations. Next, I describe the framework of their study.

The authors model a population of individuals that adopt mixed-strategies. The notation $\phi_i(t)$ is used to denote the probability that individual i plays Hawk at time t. Consequently $1-\phi_i(t)$ is the probability for individual i to play Dove at time t. As the notation suggest, the probability $\phi_i(t)$ is updated over time. The probability of meeting a Hawk in the population at time t is denoted with $\phi(t) = \frac{1}{L}\sum_{i=1}^{L}\phi_i(t)$ for a population of L individuals.

At any given time t, individuals remember for each interaction in the last m time units whether they met a Hawk or a Dove and the payoff collected at that time. Information collected prior to the last m time units is forgotten. Individuals keep track of the proportion of Hawks met and, for this purpose, $\overline{\phi}_i(t)$ is defined as the density of Hawks met by player i in the past m time units. Individuals also keep track of their moving average payoff.

Next, the authors note that when $\phi(t) < \frac{V}{C}$, it is convenient for an individual to switch to Hawk, and Dove otherwise[1] (cf. equation 2.24). Based on this scenario, they further assume that individuals do not immediately switch to the convenient pure strategy but do this in steps as described next.

$$\phi_i(t+\delta t) = \begin{cases} \phi_i(t) & w.p. \quad 1-\delta t \\ \phi_i(t) + \epsilon[1-\phi_i(t)] & w.p. \quad p\delta t \\ \phi_i(t) - \epsilon\phi_i(t) & w.p. (1-p)\delta t \end{cases} \tag{2.32}$$

where p is a function such that

$$p = p(\overline{\phi}_i(t)) = \mathcal{H}\Big[\frac{V}{C} - \overline{\phi}_i(t)\Big]. \tag{2.33}$$

Here $\mathcal{H}$ is the Heavy-side step function such that $\mathcal{H}(x) = 0$ if x is negative, $\mathcal{H}(x) = 1$ if x is positive and $\mathcal{H}(x) = 0.5$ if $x = 0$. The function p acts as a switch that directs the evolution of strategy $\phi_i(t)$. Importantly, each individual updates his strategy using his own estimate of the frequency of Hawks to inform his choice. His estimate is the time-delayed density $\overline{\phi}_i(t)$, rather than the true current average $\phi(t)$. The value ϵ is described as an update rate that regulates how fast players update their strategy. The authors only consider the case of ϵ small. Finally, for the remaining part of this review, according to the original article, I restrict the analysis to the case $V/C = 0.5$.

2.6.2 Model dynamics

Restricting the analysis to small ϵ values, the authors showed that the length m of the memory plays an important role. This role is exemplified in figure 2.7, which

[1] Quantitatively, this switch is modelled in 2.33

I replicated from the original article. Both realisations show the evolution of the average strategy $\phi(t)$ adopted in a population of 1000 individuals for $\epsilon = 0.005$. However, the graph on the left is obtained for memory length $m = 101$ and the graph on the right for $m = 150$. These graphs show that the population reaches the ESS $\phi(t) = \frac{V}{C} = 0.5$ for $m = 101$, while for $m = 150$ the average strategy in the group oscillates around the value $\phi(t) = \frac{V}{C} = 0.5$. Note that although the ESS is defined in terms of V and C, the dynamics of these simulations are independent of payoff collection and dependent on 'memory content' only. This observation will become useful in my research. The simulation data in figure 2.7 is a replication of the original article but differs slightly in the manner time is counted.

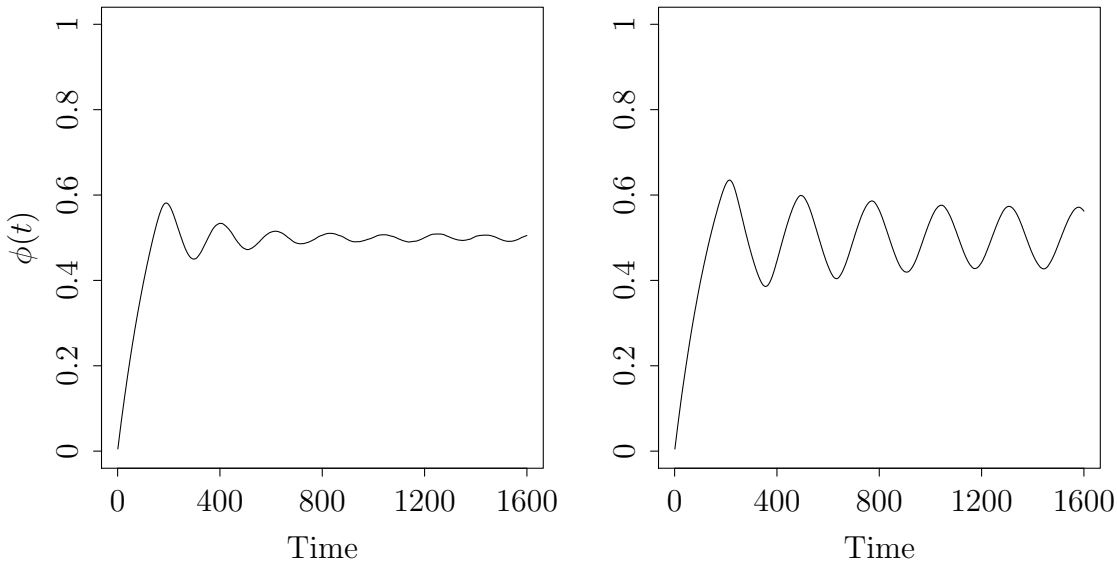

Figure 2.7: Both graphs show the evolution over time of the average strategy $\phi(t)$ for a population of 1000 individuals and $\epsilon = 0.005$. Obtained for individuals having memory length $m = 101$ for the left graph and $m = 150$ for the right.

The authors also showed that for a fixed memory length m, with m large enough, the population can either reach a stable state equilibrium or oscillate depending on the update rate value ϵ. They showed that, for large memory length m, there is a critical value $\epsilon = \epsilon_c$ such that the population reaches equilibrium for $\epsilon < \epsilon_c$ and oscillates for $\epsilon > \epsilon_c$ (in a neighbourhood of ϵ_c). I illustrate this scenario in figure 2.8. In particular, they showed that this critical point is a Hopf bifurcation and provided an analytical expression for ϵ_c as a function of m.

Finally, the authors showed that when a population with a majority of individuals with a long memory oscillate in strategy, these oscillations can be exploited by a minority of players with shorter memory. Only in this last scenario, the authors explicitly modelled the payoff collection process into the dynamics.

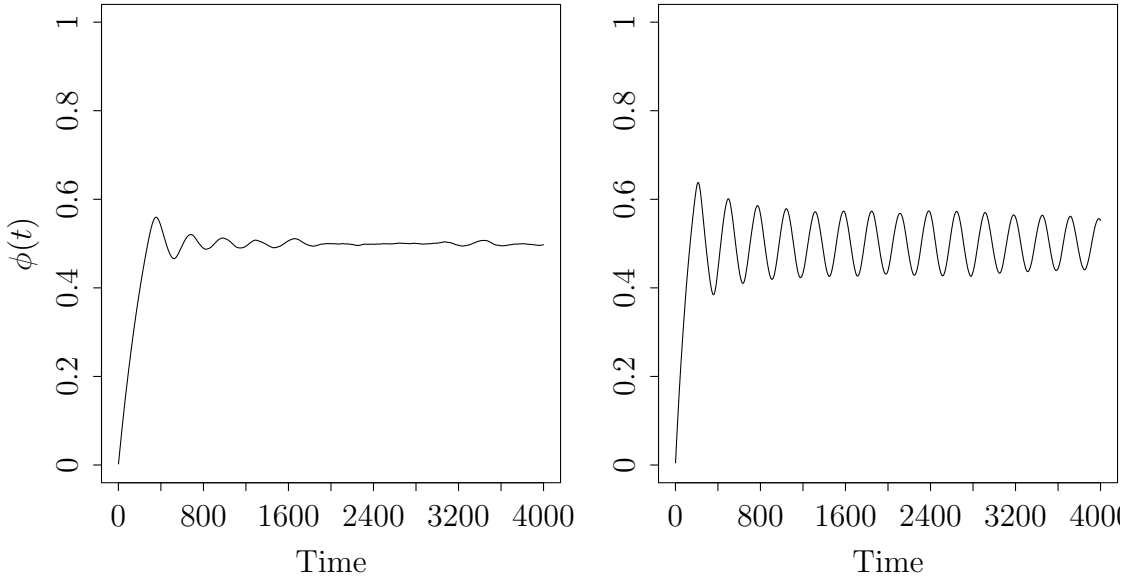

Figure 2.8: Both graphs show the evolution over time of the average strategy $\phi(t)$ for a population of 1000 individuals. Memory length is $m = 150$ for both populations while $\epsilon = 0.0025$ for the left graph and $\epsilon = 0.005$ right graph.

2.6.3 Evolution equation for $\phi(t)$

Burridge *et al.* (2017) sought to linearise the evolution equation for $\phi(t)$ around the fixed point 0.5. Through approximations valid in the limit of small oscillations, population size $L \to \infty$ and memory length $m \to \infty$, the authors derived a delay differential equation for $\psi(t) := \phi(t) - 0.5$ (2.34). This equation is given by

$$\frac{1}{\epsilon}\frac{d\psi(t)}{dt} = -\frac{\sqrt{2m}}{\sqrt{\pi}}(\overline{\psi}_m(t) - \psi(t)) \tag{2.34}$$

where

$$\overline{\psi}_m(t) := \int_{t-m}^{t} \psi(\tau)d\tau. \tag{2.35}$$

By substituting the trial solution $\psi(t) = e^{t(x+iy)}$, the authors calculated an expression for the threshold value ϵ_c. They also identified that the lowest frequency oscillation is at $y = \pi/m$, while highest frequencies are transient (for simulations with $m \to \infty$). In addition, by substituting (2.35) into (2.34) and taking the first derivative of (2.34), I have obtained the following differential equation

$$\dddot{\psi} + \epsilon\ddot{\psi} + \epsilon\frac{\sqrt{2}}{\sqrt{\pi m}}\psi = \epsilon.\frac{\sqrt{2}}{\sqrt{\pi m}}\psi_m \tag{2.36}$$

where m is the delay. Equation 2.36 is a second-order differential equation with delayed positive feedback, a type of equation that does not admit a general analytical solution. Since my research is a generalisation of this model, it is useful to point out that equation 2.36 does not have a general analytical solution.

2.6.4 Open questions

The Hawk-Dove game with memory model described here is interesting because it presents a specific mechanism by which memory can produce coordinated overshooting behaviour at the collective level and suboptimal payoffs for the individuals. Maintaining m large and ϵ small, the authors showed that the amplitude of the oscillations grows as ϵ increases past the critical ϵ_c, or similarly as memory increases past a critical memory length. However, is the onset of oscillations in the Hawk-Dove game with memory reliant on modelling a long memory m and small update rate ϵ? For $\epsilon = 1$ the set of equations in rule 2.32 readily simplifies to represent a model in which the individual probability $\phi_i(t)$ to play Hawk can only take values in the set $\{0, 1\}$, i.e. individuals play pure strategies. Can oscillations in the average group strategy be sustained as ϵ increases towards 1? Are oscillations sustained upon modelling agents that play pure strategy? Is it possible to obtain analytical expressions for amplitude and frequency as a function of ϵ and m? Are the oscillations robust to noise?

In chapter 4 I introduce a model that allows addressing these open questions but first, in chapter 3, I review the role of memory in the canonical minority game, focusing on the areas which are necessary to give context to my research.

Chapter 3

The minority game

The minority game is a mathematical model used to explore how agents in a group coordinate using inductive reasoning. This game was derived as a rigorous mathematical formulation of the 'El Farol bar problem', an earlier model proposed in an article titled 'Inductive reasoning and bounded rationality' (Arthur, 1994), in which results were obtained from simulation. Arthur (1994) was inspired by the theory of bounded rationality developed by Simon (1955) and proposed an alternative to the economic paradigm that considers humans as deductive rational agents able to process an infinite amount of information and solve theoretical complex problems using deductive thinking.

In the following sections, I introduce the 'El Farol bar problem' in more detail and then the minority game, focusing on those features that give context to my research in chapters 4 and 5.

3.1 The 'El Farol bar problem'

The 'El Farol bar problem' is named after and inspired by the 'El Farol' bar in Santa Fe, in the city of New Mexico in the United States. The problem is described as follows. On Thursday night the bar offers Irish music and attendance is enjoyable if there are no more than 60 people present. There are however 100 people interested to go on that night each week. The question posed is 'how should people choose this week?' when the attendance in the past weeks is the only information available. The author does not formulate the problem as a game-theoretical model. He rules out mixed-strategies because he is not interested in the stable mixed-strategies Nash equilibrium (coin tossing) but in the inductive dynamics. This problem becomes interesting in the following scenario: if individuals in Santa Fe build the same expectation about attendance in the current week, their expectation will not happen. For example, if it is reasonable

for everyone to assume that few will go then all would go, and vice-versa. Thus there is not an obvious solution that individuals could adopt. Arthur suggests than one possible solution is allowing agents to formulate a prediction in a variety of ways. He proposes that if agents individually predict next week attendance based on past attendance and if attendance in the past 14 weeks had been

... 44 78 56 15 23 67 84 34 45 76 40 56 22 35

agents could formulate a prediction for the current week attendance using several working hypotheses, named predictors, for example:

- the same as last week [35]
- a mirror image around 50 of last week's [65]
- a (rounded) average of the last four weeks [49]
- the trend in the last 8 weeks, bounded by 0, 100 [29]
- the same as two weeks ago [22] (cycle of period 2)
- the same as 5 weeks ago [76] (cycle of period 5)
- etc.

Each person works with k of such predictors ($k > 0$), ranking them in terms of efficiency. The most accurate predictor, called 'active predictor', is the one employed to decide whether to go to the bar or not. Each week, once the attendance is revealed, individuals update the accuracies of their set of predictors (in a way that is not made explicit by the author) and can change which predictor is their active one. Finally, the efficiency of these sets of strategies is studied through simulations. However, it is not known how many strategies there were overall nor how many each person had. The attendance for one run of simulations is reproduced in figure 3.1 from Arthur (1994). The figure shows that the agents managed to self-organise their predictors with attendance over time fluctuating close to 60 people. Arthur concludes that the predictors are acting like mental models that mutually co-adapt over time. For this process to work, there need to be a variety of different predictors to which he refers as an 'ecology of strategies'. In the minority game, this problem is examined in more detail.

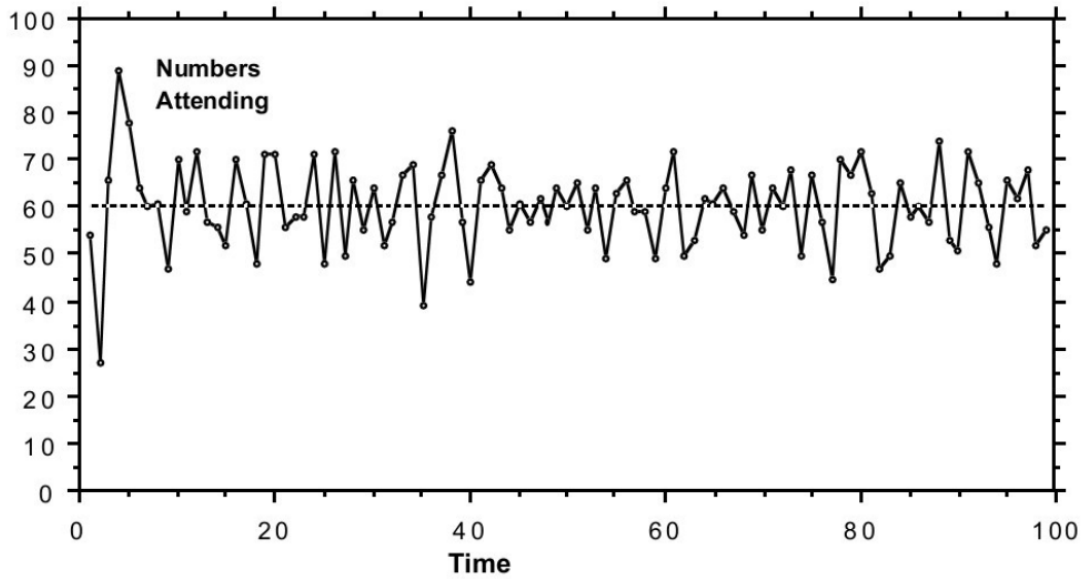

Figure 3.1: Reprinted from Arthur (1994). The number of people going to the bar over 100 weeks.

3.2 The minority game

The minority game (MG), introduced as a toy-model to explore financial markets dynamics (Challet & Zhang, 1997), models a scenario in which individuals in a group choose between two options and there is an advantage to choosing the option being adopted by the minority in the group. Intuitively, the MG is a study of the 'El Farol bar problem' with N people and $N/2$ seats. Next, I present the MG in greater detail.

In the MG, a strategy is a formalisation of the idea of a predictor in the 'El Farol bar problem'. A strategy is a function that maps the set of possible past histories into the set of the two actions available, determining which action is the response to a given history. The history used by agents is however restricted in time to the latest m time steps. For each time step, the information available is also restricted to a 'bit' indicating only the minority side in the group. In computational models of the game, each of the two actions available is denoted with 0 and 1 respectively. Similarly, at each iteration of the game, the minority side is recorded with a 0 or 1 digit. Table 3.1 is an example of a strategy, where an agent chooses an action based on a history of $m = 3$ 'bits'. In table 3.1 there are $2^m = 8$ different possible configurations of the history.

History	Action
000	1
001	0
010	0
011	1
100	1
101	0
110	1
111	0

Table 3.1: Example of a strategy for a memory of 3 bits

In the game, there are N individuals and each individual has S strategies available. Similarly to the El Farol bar problem, each individual ranks his strategies at each iteration of the game. To rank their strategies, individuals update a score for each of their strategies by virtually assessing them. They award one point to a strategy if it would have been successful at predicting the minority, had it been played, and no points otherwise. The score is often referred to as 'virtual score' because it is the result of the accumulation of perceived payoffs. Virtual payoffs differ from real payoffs because individuals overlook the impact that using that strategy would have had at the aggregate level (Challet *et al.*, 2000). The 'computational' description above can be translated into a mathematical formulation as follows. The individual decisions are represented with the set $\{-1, 1\}$ rather than $\{0, 1\}$ obtaining

$$a^{\mu}_{i,s} = 1 \tag{3.1}$$

for one side, and with

$$a^{\mu}_{i,s} = -1 \tag{3.2}$$

for the other side, where μ here represents the history in the last M time steps. $a^{\mu}_{i,s}(t)$ is defined to be the side taken by individual i, at time step t, using strategy s, given history μ. In particular i varies in $\{1, ..., N\}$ and the history μ has m bits of information corresponding to the minority sides in the last m iterations of the game. As a result, the minority side for iteration t can be detected using the opposite of the sign of the following sum:

$$A(t) = \sum_{i=1}^{N} a^{\mu(t)}_{i,s_i(t)}. \tag{3.3}$$

$A(t)$ is the aggregate response at time t while $s_i(t)$ is the 'best strategy' used by individual i at time t. When $\text{sign}(A(t)) = +1$ the minority strategy at time step t is -1, and vice-versa. The dynamics evolve in discrete time steps and thus time t is equivalent to iteration t.

3.2.1 Variance

In the literature, the main quantity of interest to describe the dynamics of the game is the variance of the aggregate quantity $A(t)$. The variance is defined as

$$\sigma^2 = \overline{\langle [A(t) - \langle A(t) \rangle]^2 \rangle}. \tag{3.4}$$

where the angled brackets indicate the time average and the bar indicates the average over many realisations. The variance is studied because it is considered a useful measure of the efficiency of the group at coordinating their actions. This is because the smaller the variance, the larger is the minority group (Savit *et al.*, 1999) and thus the larger the aggregate payoff. Using the notation in 3.1-3.4, one obtains $\langle A(t) \rangle = 0$ and $\sigma^2 = \overline{\langle A(t)^2 \rangle}$. This result is due to symmetry arguments which consider that $A(t)$ fluctuates around the average of the two actions (Challet *et al.*, 2004).

In particular, if we were to assume that each agent i draws action $a_i(t)$ randomly and independently from the set $\{-1, 1\}$ at time t, we would obtain $\langle a_i \rangle = 0$ and $\langle a_i a_j \rangle = \delta_{i,j}$. This implies:

$$\langle A(t) \rangle = \sum_{i=1}^{N} \langle a_i(t) \rangle = 0. \tag{3.5}$$

$$\frac{\sigma^2}{N} = \frac{1}{N}\langle A^2(t) \rangle = \frac{1}{N}\langle \sum_{i=1}^{N} a_i(t) \sum_{i=1}^{N} a_i(t) \rangle = \frac{1}{N}\langle \sum_{i,j=1}^{N} a_i(t) a_j(t) \rangle \tag{3.6a}$$

$$= \frac{1}{N} \sum_{i,j=1}^{N} \langle a_i(t) a_j(t) \rangle = 1 \tag{3.6b}$$

Equations 3.6a-3.6b show that $\frac{\sigma^2}{N} = 1$ for agents behaving randomly. In the minority game literature, the coordination among agents is considered efficient when the variance $\frac{\sigma^2}{N}$ is less than it would be if they were playing randomly and inefficient otherwise. Figure 3.2 depicts how the variance depends on the fundamental quantity $\frac{2^m}{N}$ and shows that it admits a global minimum as a function of $\frac{2^m}{N}$, which has created a lot of interest on this problem since an economic or financial interpretation of this minimum means that a market admits an efficient

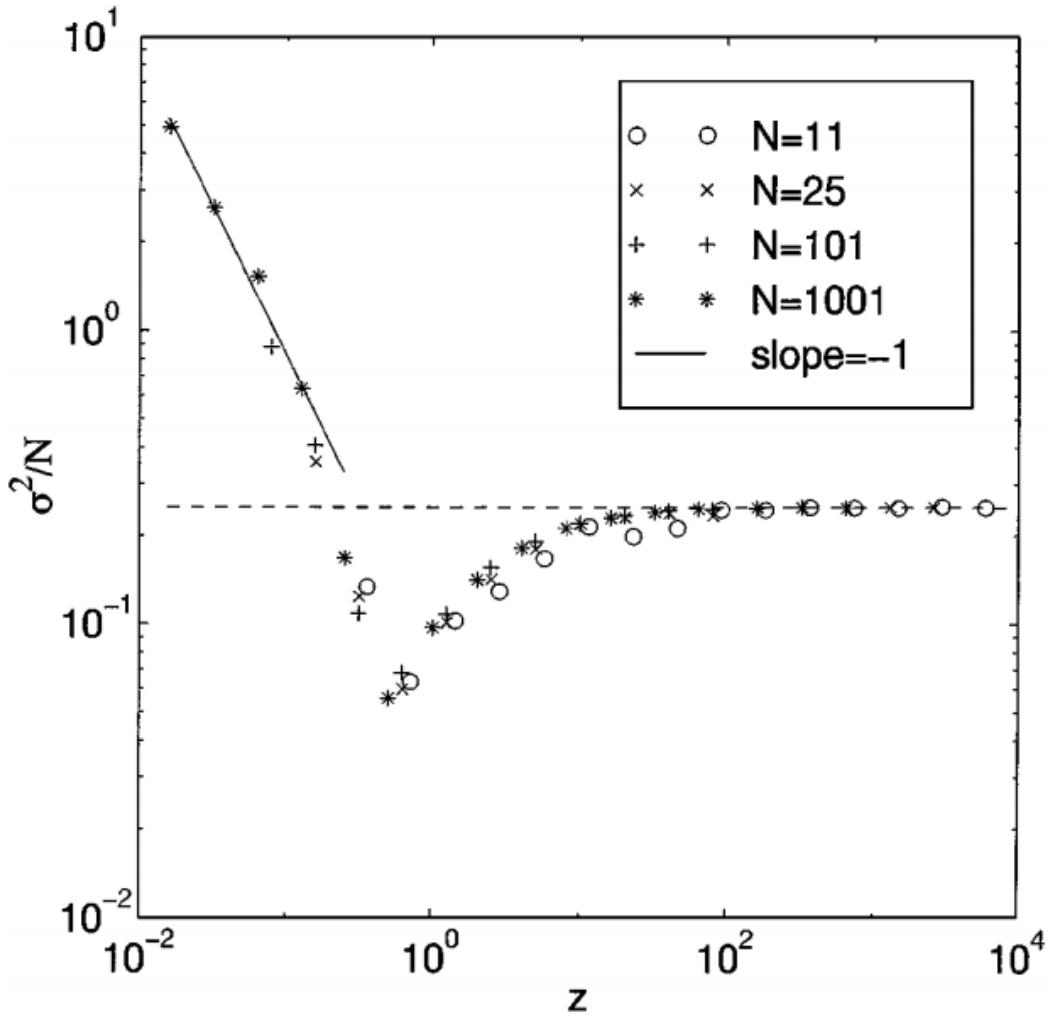

Figure 3.2: Reprinted from Savit *et al.* (1999). σ^2/N as a function of $z = \frac{2^m}{N}$ on a log log scale. Individuals have two strategies available (i.e. $S = 2$).

distribution of allocated resources.
The existence of a minimum has been verified in the literature for $S \geqslant 2$, the case in which agents have two or more strategies. Let us indicate the minimum with $z = z_c$. This value has been shown to be a critical value that separates the region $z < z_c$ from $z > z_c$. Using geometrical arguments it has been proposed that either $2 * 2^m$ (Challet & Zhang (1998)) or 2^m (Savit *et al.* (1999)) are measures for the dimension of the strategy space. In both cases, the conclusion holds that the variance reaches its minimum when the population size is of the same order as the dimension of the strategies space. In the region $z > z_c$, σ^2/N converges to

the value it would take if agents were choosing between the two options randomly. Interpretations for this result are not definitive. Most commonly it is interpreted as the fact that as the (complexity of the) information used by agents increases, agents become increasingly inefficient at coordinating their strategies, or in other words, the information provided becomes too complex such that agents can no longer recognise patterns (Savit *et al.* (1999),Challet *et al.* (2004)). However, as z decreases, due to either an increased number of agents or a reduced amount of information, agents coordinate their strategies more efficiently until reaching a minimum for the variance at $z = z_c$. However, in the region $z < z_c$, σ^2/N increases as z further decreases, reaching an inefficient region where $\sigma^2/N >> 1$. In particular, simulations (Savit *et al.* (1999)) have shown that the aggregate quantity $A(t)$ exhibits periodic fluctuations for small values of σ^2/N (this information will be relevant in my research). On a qualitative level, the emergence of inefficient coordination among individuals is interpreted as a 'crowd effect' (Challet *et al.* (2004),Ghosh *et al.* (2012)).

3.2.2 Probabilistic strategy choice

In the version of the minority game shown above, agents' choices follow a deterministic rule given the virtual scores. However, this formulation limits the analytical tractability of the model. In the article titled 'Thermal Minority Game', Cavagna *et al.* (1999) propose to use the following probabilistic strategy choice rule:

$$Prob\{s_i(t) = s\} = \frac{e^{\Gamma U_{i,s}(t)}}{\Sigma_{s'} e^{\Gamma U_{i,s'}(t)}} \tag{3.7}$$

with $\Gamma > 0$. $U_{i,s'}(t)$ is defined to be, at time t, the virtual score of strategy s' for agent i. Γ has the role of 'inverse temperature', regulating the stochasticity of individuals strategy choices. Similarly, in the Gibbs distribution, temperature regulates the probability that a system is in a certain state depending on the temperature of the system itself (Yeomans, 1992). The form of update rule 3.7 is identical to the form of update rule 2.6 described earlier.

3.2.3 Random history

Initially, the study of the MG assumed that individuals respond to endogenous information, i.e. the history they have themselves created. However, Cavagna (1999) showed that there is no feedback loop of information from the past. In other words, the real history of the aggregate quantity $A(t)$ does not impact the dynamics of $A(t)$ itself. This result is obtained by comparing the dynamics of populations fed with real and random histories. Note, however, that the

information is homogeneous (i.e. the same for all agents) even when randomly generated. The equivalence between the dynamics generated by real and random histories is sometimes termed 'Memory irrelevance' (Cavagna, 1999). This equivalence implies that agents can co-adapt their responses using any string of shared information. In particular, in the MG, coordination among individuals evolves without a capacity to forecast the future based on the understanding of the past (Challet *et al.*, 2013).

3.3 Conclusions

The canonical version of the minority game has been formalised by Challet & Zhang (1997) and stemmed from the previous work of Arthur (1994). In this chapter, I have described some of its features. The canonical version rests firmly on the ideas of deductive thinking and complexity to the extent that the minority game is regarded as a case study on those topics. The authors themselves have called the formulation described in this chapter the standard MG (Challet *et al.*, 2013). In the rest of this thesis, I refer to the combination of MG and TMG as the canonical MG. On the other hand, a simple description of the aim of the game is that "a finite number of players have to choose between two sides; whoever ends up in the minority side is a winner" (Challet *et al.*, 2013). In light of this primary definition, I introduce and study a non-canonical version of the minority game in the next chapter. As a fundamental departure from the canonical model, inspired by the work of Burridge *et al.* (2017) on the Hawk-Dove game, I will propose that agents use one predictor only, which (in the language of the 'El Farol bar problem') entails the use of the mean of 'past attendance' as recorded by the single agents. This predictor is an intuitive and sensible use of memory. It is also used by Burridge *et al.* (2017). This difference is fundamental to the extent that endogenous history is now crucially important. The ensuing phenomenology is rich and arguably presents the potential to explain observational data and real-world phenomena.

Chapter 4

Emergence and collapse of oscillations in binary choice dynamics.

4.1 Introduction

In previous chapters, I reviewed the canonical description of minority games (in chapter 3) and a specific version of the Hawk-Dove game with memory (in section 2.6). In essence, both games model the social scenario in which individuals in a group choose between two options. At the same time, the highest reward goes to the individuals who have selected the option taken up by a minority in the group. Both games also model a learning process that employs memory and online learning. The two games, however, adopt opposing views concerning the granularity of the information available to agents. Specifically, data is finely grained and stochastic in the Hawk-Dove game with memory. Individuals learn through one-to-one random interactions, within a large group, the option taken up by one other individual at a time and, as soon as they have learnt this, update their strategy. On the contrary, in the canonical minority game, individuals have access to coarse-grained information. At the end of each turn, they learn which option was taken up by a minority in the group. This information is available to everyone in the group, neglects any form of stochasticity and originates deterministic dynamics.

My goal is to present a more comprehensive study of the learning dynamics in binary games with memory, bridging the gap between the two aforementioned models. I propose a binary choice model with memory in which individuals, between adaptation moves, observe a number of times the behaviour enacted by others in the group. Individuals who collect a finite number of observations

are said to collect a finite batch. Individuals who collect an infinite number of observations are said to collect an infinite batch. I show that infinite and finite batches generate different collective dynamics. The population exhibits deterministic dynamics in the limit of infinite batches and stochastic dynamics for finite batches. I show that deterministic learning always yields periodic oscillating dynamics. In contrast, stochasticity enriches the game dynamics, with noise contributing to the formation of different equilibria in the population's dynamics.

This chapter specifies the dynamical system, describing the processes, variables, parameters, and initial conditions. I characterise the stochastic dynamics presenting extensive simulation data for finite batches and derive, through a combination of analytical and numerical means, a description of the deterministic dynamics in the case of infinite batches. Comparing the results for finite and infinite batches allows assessing the impact of stochasticity on the system. I also characterise, depending on the parameters, the transition from a stable collective state to oscillations.

In chapter 5, I discuss the effects on the dynamics of i) making memory duration heterogeneous, ii) making individual actions simultaneous and iii) changing the population size. I also discuss to what extent the convergence property of the iterated update rule affects the breakdown of oscillations in the simulations. I conclude the chapter with a review of how each 'dimension' of the model impacts on the amplitude of the oscillations. Chapter 6 contains a summary of the research and a discussion on its applications.

4.2 The model

In this section, I introduce a detailed model of a scenario in which everyone within a large group is faced with a choice between two options, say A and B, and aims to adopt the option taken up by a minority in the group. I assume that individuals learn over time about the two options' prevalence but recollect any information for a limited time only. Agents learn through an observation process, which is modelled as a Poisson process with rate r. An observation is an event in which an individual observes the option, namely A or B, chosen by another individual. Individuals remember this information for m time units and after that discards it; m is defined as the memory duration. The product mr is, on average, the number of observations recollected by an individual. I assume that each individual uses the information available to him to estimate which option is taken up by a minority in the group. This estimation is calculated by computing a simple statistic. The statistic at time t for individual i, say $\mu_i(t)$, is the fraction of B players[1] observed in the time interval $[t-m, t]$. I have earlier mentioned that

[1] A B player is an agent that currently adopts option B.

agents choose between option A and B. I further assume that this choice is driven by a mixed-strategy probability, say $\phi_i(t) \in [0,1]$. The mixed-strategy $\phi_i(t)$ is the probability that individual i chooses option B at time t, while $1-\phi_i(t)$ is the probability that he chooses option A. The statistic $\mu_i(t)$ is employed at the times when individual i updates his strategy $\phi_i(t)$. Individuals adapt their strategies to increase their chances of providing the best response to the group's behaviour. An heuristic approach to problem solving guides their adaptation move: when an agent estimates that most individuals have been adopting option B, he responds by incrementally decreasing his mixed-strategy probability (his probability of adopting option B); and vice-versa. The strategy update process is a Poisson process with rate ρ. It is formally defined in equations 4.1a-4.1c in the limit $\Delta t \to 0$.

$$\phi_i(t+\Delta t) = \begin{cases} \phi_i(t) & \text{w.p. } (1-\rho\Delta t) \quad (4.1a) \\ \phi_i(t) + \epsilon[1-\phi_i(t)] & \text{w.p. } \rho\tilde{p}\Delta t \quad (4.1b) \\ \phi_i(t) - \epsilon\phi_i(t) & \text{w.p. } \rho(1-\tilde{p})\Delta t \quad (4.1c) \end{cases}$$

where $\tilde{p}$ is a function of $\mu_i(t)$ that takes values in {0,0.5,1} and defined as

$$\tilde{p}(\mu_i(t)) = \mathcal{H}\Big[0.5 - \mu_i(t)\Big]. \quad (4.2)$$

$\mathcal{H}$ is the Heaviside step function that takes value 0 when the argument is negative (i.e. when the estimated fraction of B players is greater than 0.5), 1 when the argument is positive (i.e. when the estimated fraction of B players is less than 0.5) and 0.5 otherwise. Individuals adapt their strategy incrementally, with the intensity of response ϵ modulating the size of the increment; ϵ can represent a measure of caution in decision making (Fecteau *et al.*, 2007), a learning rate (Burridge *et al.*, 2017) or mirror the concept of decision inertia (Alos-Ferrer C, 2016; Pitz G F, 1968), a tendency to repeat previous choices regardless of their outcome. The constant ρ is the rate at which individuals update their strategy. The unit of measure for ρ is $[T]^{-1}$, whereas $\phi_i(t)$ and ϵ are dimensionless. The memory duration m has unit of measure $[T]$. The rate of observation r has unit of measure $[T]^{-1}$. The product mr is therefore dimensionless. Given a population of size L, the average group strategy at time t is given by $\phi(t) = \frac{1}{L}\sum_{i=1}^{L}\phi_i(t)$.

In equations (4.1a)-(4.1c), individuals adopt mixed strategies for $\epsilon < 1$, while they adopt pure strategies for $\epsilon = 1$. In the latter case, individuals switch between pure strategy A ($\phi_i = 0$) and B ($\phi_i = 1$). Otherwise, for $0 < \epsilon < 1$, individuals display intermediate reactions. In the case of pure strategies, equations (4.1a)-(4.1c) reduce to:

$$\phi_i(t+\Delta t) = \begin{cases} \phi_i(t) & \text{w.p. } (1-\rho\Delta t) \quad (4.3a) \\ 1 & \text{w.p. } \rho\tilde{p}\Delta t \quad (4.3b) \\ 0 & \text{w.p. } \rho(1-\tilde{p})\Delta t \quad (4.3c) \end{cases}$$

We will see that pure and mixed strategies yield similar collective dynamics, even if they necessarily yield different dynamics at the individual level. In this chapter I present the simulation results for finite batches in the case $\rho = 1$ (in section 4.3), $\rho \neq 1$ and $\epsilon = 1$. To lighten the notation throughout the thesis, at times, I will omit ρ from the notation when assumed equal to 1.

Further comments

The model described above is a generalisation of the one proposed by Burridge *et al.* (2017) in the context of Hawk-Dove games with memory. In their study, Burridge *et al.* (2017) conflate the two processes, namely observation and strategy update (online learning), and the rate of observation is thus constrained to be equal to the rate of strategy update. This is also a feature of the canonical minority game. Importantly, a novel feature of my model is the separation of the two processes, observation and strategy update, with the introduction of an observation rate $r \neq 1$ [1]. Except for the results concerning the co-evolution of two memory traits that make use of payoffs, the results by Burridge *et al.* (2017) become a subset of the results that I will describe. The two models are equal when we assume that the two aforementioned processes are paired with r being implicitly equal to 1 and $\rho = 1$. In addition, while Burridge *et al.* (2017) studied their model for $\epsilon \approx 0$, I extend the analysis to the entire interval $[0, 1]$ ($0 \leq \epsilon \leq 1$). Note that pure strategies (obtained for $\epsilon = 1$) represent more closely the best-response in the Hawk-Dove game (cf. 2.24 for a derivation of the result).

Compared to the work of Burridge *et al.* (2017), another novel parameter in my model is the rate of strategy update ρ, which will become helpful in thoroughly characterising the dynamics of the system. In addition, except for section 5.1.3, I do not explicitly model payoff collection. Nevertheless, I will still make qualitative comments on the evolutionary implications of the equilibria reached throughout the thesis.

Lastly, the novel feature, i.e. the separation of the processes of observation and strategy update with the subsequent introduction of an observation rate $r \neq 1$, allows obtaining an exact analytical and numerical description of the dynamics of $\phi(t)$ in the limiting case $r \to \infty$. This scenario models individuals that are equipped with an exact knowledge of the behaviour of the group restricted to the time span of the memory duration m. On the other hand, for finite batches, i.e. r finite, I investigate the dynamics mainly through extensive simulations, which I proceed to describe next.

[1] Note to the reader: the rate r is not explicitly present in 4.1a-4.1c.

4.3 Simulation results for individual and collective dynamics for finite batches, r finite.

4.3.1 The algorithm

The algorithm used for simulations is based on the Gillespie algorithm (Gillespie, 1976, 1977). Specifically, the succession of events is determined according to the Gillespie algorithm; however, the waiting time elapsed between events is neither recorded nor generated. Fundamentally, the algorithm models the two processes mentioned above, observation and strategy update, and records the evolution of the average strategy $\phi(t)$ in the population. It can easily be adapted to record other quantities, such as the evolution of the individual strategies $\phi_i(t)$ and standard deviation of the set of individual strategies. Figure 4.1 contains a diagram representing the workflow of the algorithm. In appendix A, I present an annotated code followed by its output. The output is a plot showing the evolution of $\phi(t)$ together with the standard deviation of the individual strategies. Here I list the set of parameters used to initialise the algorithm:

(i) m, the memory duration;

(ii) r, the observation rate;

(iii) ρ, the strategy update rate;

(iv) ϵ, the intensity of response;

(v) L, the population size;

(vi) the end time of the simulation.

The following variables are updated over time:

(vii) $\phi_i(t)$, the probability that individual i plays option B at time t;

(viii) $\mu_i(t)$, the fraction of B players observed in the group by individual i in the last mr observations;

(ix) $\phi(t)$, the average of all $\phi_i(t)$ at time t.

To summarise, the two possible events are observation and strategy update, having probability $\frac{r}{L(r+\rho)}$ and $\frac{\rho}{L(r+\rho)}$ respectively.

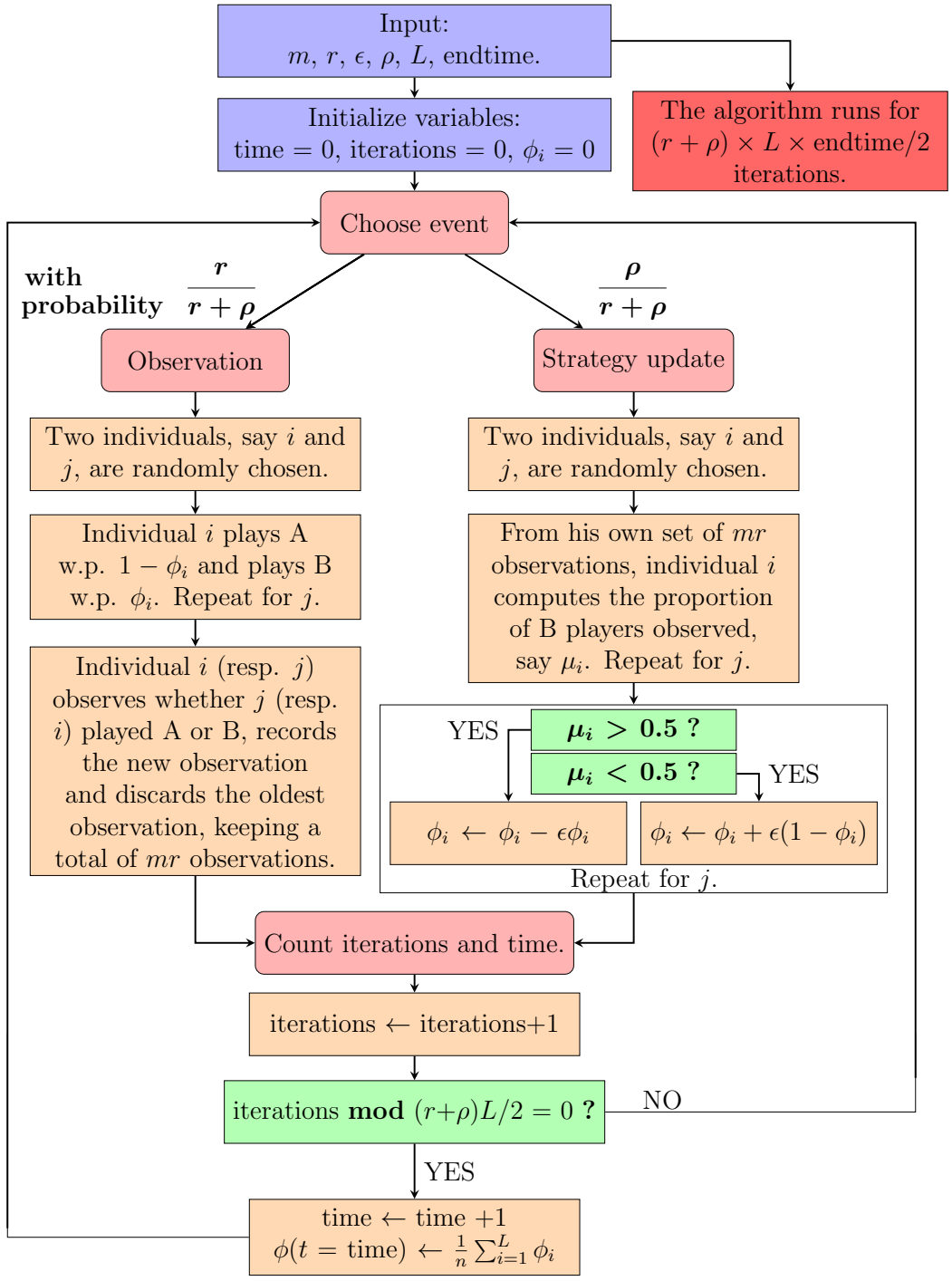

Figure 4.1: Diagram visualising the workflow of the algorithm used in simulations. Note that, at the beginning of simulations, agents have an 'empty memory' and do not discard observations. An agent discards observations only after having first collected mr observations.

Time keeping

In order to increase computational speed, the algorithm does not keep track of time for single events, whether the event is a strategy update or an observation. Instead, it keeps track of how many events occurred. Then, time is estimated. The rationale behind the estimation is described with the following concrete example. Suppose the population has 1000 individuals updating their strategy with rate 1 and 'observing' with rate 10. On average, in this scenario, one expects $1000 \times (10 + 1)$ events in one time unit. Consequently, the algorithm counts time by increments of one time unit every $1000 \times (10 + 1)$ events. Accordingly, simulation data is recorded at intervals of one time unit. In general, this estimation is accurate as long as the expected number of events in one time unit is large.

Memory

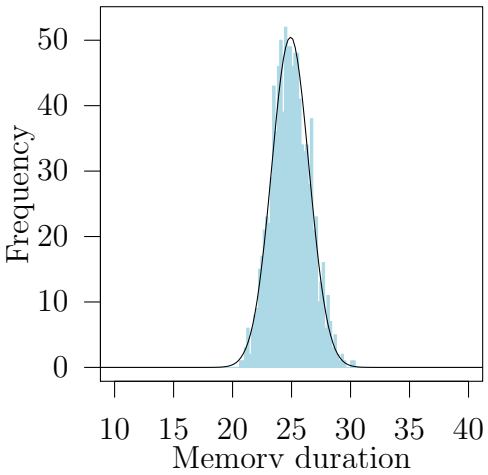

Figure 4.2: Frequency plot of the simulated memory duration of 1000 agents with 250 observations at rate $r = 10$ (blue bars) compared to the normal distribution $\mathrm{N}(\mu = mr/r, \sigma^2 = mr/r^2)$ (black line) with $m = 25$ and $r = 10$. The width of each frequency bar is 0.2.

In the description of the model given in section 4.2, I stated that agents recall the events observed in the last m time units. In simulations, however, the algorithm does not keep track of time. Instead, I assume that each agent recollects exactly the last mr observations. One can show that, as $mr \to \infty$, the probability distribution of the memory duration converges to the normal distribution $N(m, m/r)$. This result follows from the fact that: i) the duration is given by the sum of a large number (mr) of independent waiting times; ii) waiting

times follow an exponential distribution with rate r, having expected value $\frac{1}{r}$ and variance $\frac{1}{r^2}$.

In particular, the expected value of the memory duration is m. The variance decreases as r increases. Figure 4.2 shows how memory duration may distribute in simulations. The figure consists of a frequency plot obtained by simulating the duration of one thousand memories, each made of $mr = 250$ observation at rate $r = 10$. Each memory duration is calculated by adding 249 waiting times generated with an exponential distribution with rate 10. The figure shows that the distribution of memory duration is well approximated by a normal distribution with mean $m = 25$ and variance $m/r = 2.5$.

4.3.2 Results

In figure 2.8 one is reminded with a specific example ($m = 150$, $r = 1$ and $\rho = 1$) that there is a critical value ϵ_c such that for $\epsilon > \epsilon_c$ the group average strategy $\phi(t)$ oscillates around the optimal equilibrium. Considering the same set of parameters and increasing ϵ further, I show that (figure 4.3) the oscillations expand for 'intermediate' values of ϵ and collapse for 'large' values of ϵ. In figure 4.4 I plot the evolution of the strategy $\phi_i(t)$ of a typical individual alongside the group average strategy $\phi(t)$. One can notice that the observation of the aggregate quantity $\phi(t)$ does not help inferring the individual dynamics in the group. In order to characterize the dichotomy between individual and group behaviour, I have first run extensive simulations.

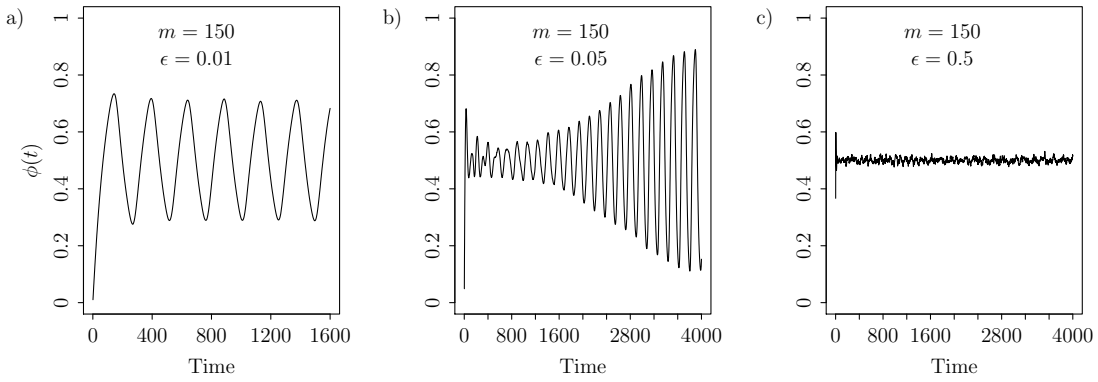

Figure 4.3: Evolution over time of the average strategy $\phi(t)$ for a population of 1000 individuals. Rate of update ϵ varying in $\{0.01, 0.05, 0.5\}$ for fixed memory duration $m = 150$.

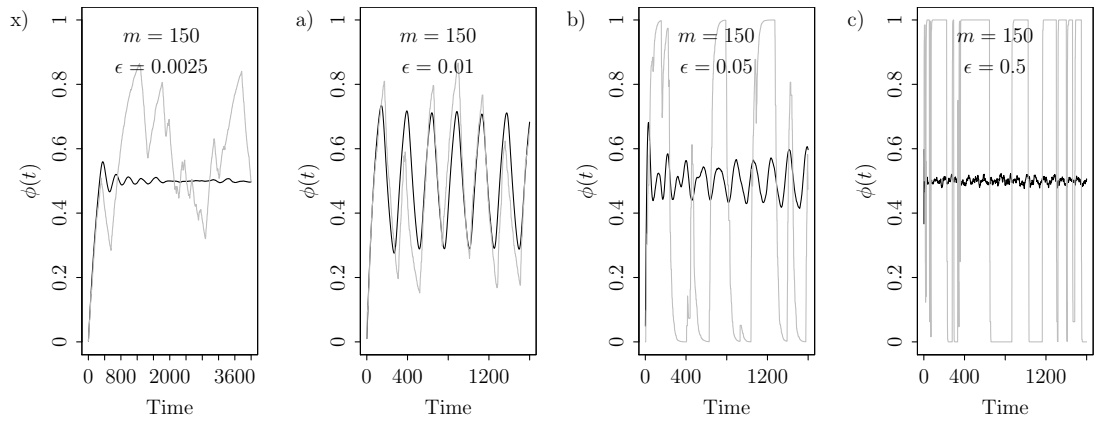

Figure 4.4: Evolution over time of the average strategy $\phi(t)$ for a population of 1000 individuals and ϵ varying in $\{0.025, 0.01, 0.05, 0.5\}$ for fixed memory duration $m = 150$ (black line). Evolution over time of the strategy $\phi_i(t)$ of one typical individual in each population (grey line). Here, images a to c correspond to images a to c in figure 4.3.

Preliminary simulations had shown that the relevant parameters are the product $\epsilon\rho$, memory duration m and rate of observation r. Unless I state otherwise, simulations results are always obtained from large well-mixed populations, where demographic noise can be largely ignored. In practice, a population of 1000 individuals satisfies this constraint. Figure 4.5 is one way to illustrate the range of outcomes both at the individual and collective level for agents employing mixed strategies. In this figure, results refer to populations that are simulated for different sizes of ϵ, while memory duration, $m = 50$, rate of observation, $r = 10$, and rate of update, $\rho = 1$, are fixed. The group average strategy (black line) is the mean of all individual strategies (orange points) at any moment in time. I remind the reader that individual strategies $\phi_i(t)$ take value in the interval $[0, 1]$, thus the mean value for the collective $\phi(t)$ also takes a value between 0 and 1. Dependent on having simulated a large population of 1000 individuals, I can observe approximately three phases for the collective average behaviour (black line): stable equilibria, bounded oscillations and random fluctuations around the equilibrium. Additionally, on the individual level the population shifts from being closely aligned to being split into two groups. This can be deduced from observing the individual distribution of strategies (dotted orange colour) that transition from steady unimodal to moving unimodal to bimodal distribution.

For cyclical periodic oscillations, I measure the amplitude and period of the trajectory of $\phi(t)$. In order to do this, I execute a code that first detects local

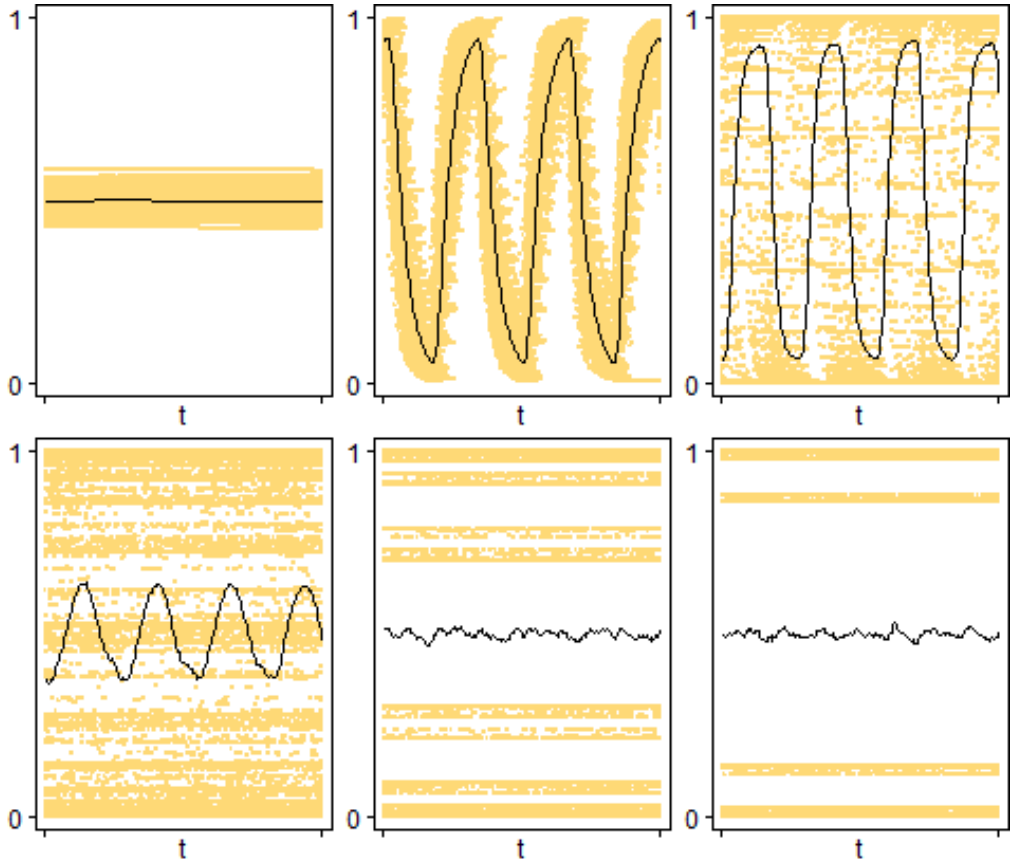

Figure 4.5: Each quadrant shows $\phi(t) = \frac{1}{n}\sum_{i=1}^{n}\phi_i(t)$, i.e. the trajectory of the average collective strategy over time (black line), coupled with $\phi_i(t)$, i.e. the individual strategies of the one thousand individuals in the population (dotted coloured area). Each quadrant shows one of the realisations obtained for six populations, $\epsilon = 0.0001$ (top left), 0.1005 (top middle), 0.3255 (top right), 0.4755 (bottom left), 0.7005 (bottom middle) and 0.8655 (bottom right). For all six realizations, $m = 50$, $r = 10$ and $\rho = 1$.

maxima and minima from simulation data. In figure 4.6 I exemplify with an illustration the quantities A for amplitude and T for period for $\phi(t)$ from simulation data.

Figure 4.7 depicts the amplitude, A, of the oscillations of the average group strategy, $\phi(t)$, generated from simulation data. Each line represents the amplitude generated by populations simulated for m, r and ρ fixed and ϵ varying from 0 towards 1. Populations are simulated for two observation rate values: a less detailed memory on the left graph ($r = 5$) and a more detailed memory on the right graph ($r = 10$). The value of ϵ increases from 0 towards 1 in both

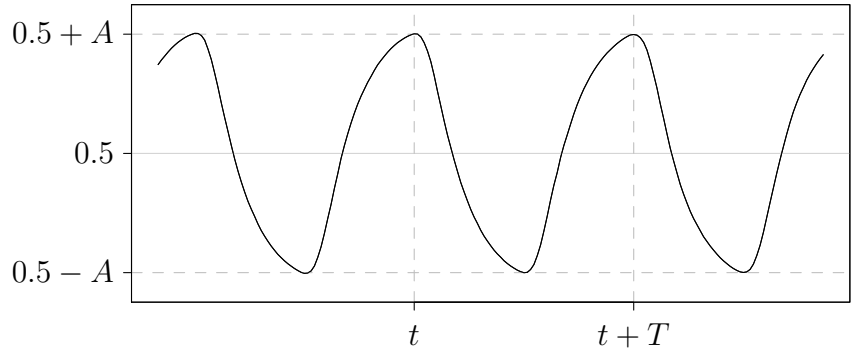

Figure 4.6: An illustration of the period T and amplitude A for the periodic evolution of the average population strategy $\phi(t)$ (black line) for $m = 50$, $\epsilon = 0.1005$ and $r = 50$ over a time of approximately 200 time units. Here, $T \approx 66$ time units and $A \approx 0.43$.

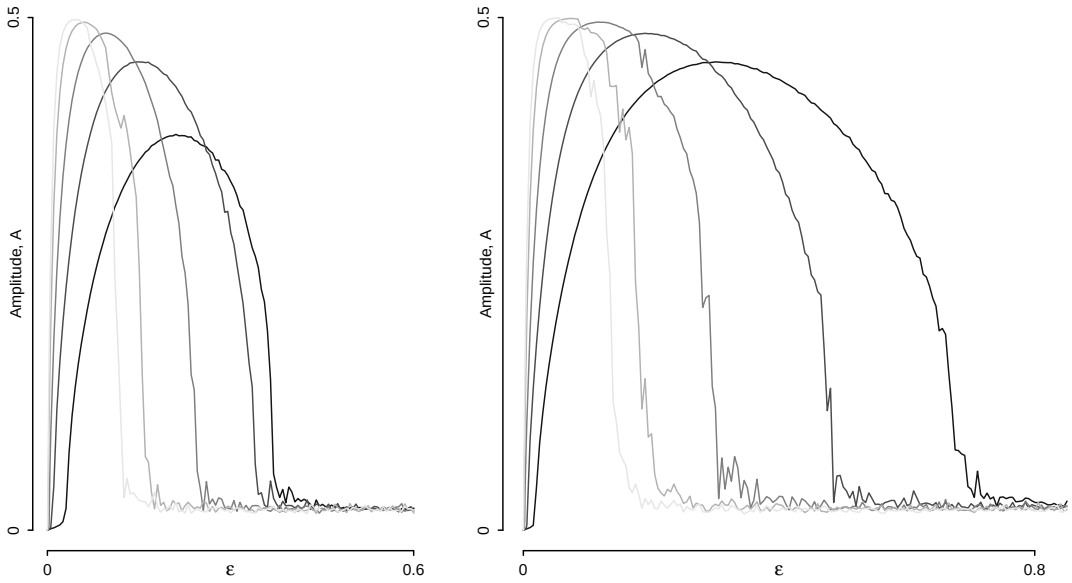

Figure 4.7: The amplitude, A, of the oscillations of the group average strategy $\phi(t)$ generated from simulation data. Each line refers to the amplitude of the oscillations generated collectively by populations defined in terms of fixed m, fixed r and fixed $\rho = 1$ with ϵ varying from 0 towards 1. The observation rate is lower for the left plot ($r = 5$) and higher for the right plot ($r = 10$). Each line shows that, for a given r and m, the dynamics of the group average strategy $\phi(t)$ undergoes a Hopf-bifurcation at a small value of ϵ. The memory duration takes a value in the set $\{25, 50, 100, 200, 300\}$. The colour code is a gradient of greys: a lighter colour refers to a longer memory.

graphs. In each graph, the memory duration of individuals take a value in the set $\{25, 50, 100, 200, 300\}$. Different shades of grey are used to indicate different memory durations. For a lighter colour, individuals have a longer-lasting memory (i.e. the darkest shade refers to $m = 25$ and the lightest to $m = 300$). This figure confirms a recurring pattern that, as mentioned earlier, comprises three phases: a stable steady state followed by bounded oscillations followed by a noisy steady state of dumped fluctuations. Extensive simulations have shown that, given the condition that the product mr is large but finite, this pattern is universal. What else does the simulation data reveal? I will show this in section 4.5. Next, however, I characterize the system dynamics with a combination of analytical and numerical results for $r \to \infty$.

4.4 Deterministic dynamics in the limit $r \to \infty$

In this section, by analytical and numerical means, I study an infinitely large population of individuals equipped with a fully detailed memory of the social environment for the time spanned by their memory duration, corresponding to the limiting case of an infinite batch ($r \to \infty$) with m finite and fixed across the population. Importantly, this allows us to understand better the behaviour of individuals collecting finite batches (r finite) by describing how the latter relates to the former.

The collective trajectory.

Here I assume that $r \to \infty$ and $n \to \infty$, where r is the observation rate and n is the population size. The assumption $r \to \infty$ implies that every agent at time t has recorded the exact time-averaged fraction of B players in the population in the time interval $[t - m, t]$. In our model, each individual samples the social environment, recording the behaviour enacted at different times by other individuals, excluding himself. However, in the limit of infinite population size, the impact of a single individual's strategy is negligible on an agent's memory content. Therefore, in this limiting case, memory content is identical for all individuals. I can rewrite the statistic $\mu_i(t)$ as a time-averaged integral over the time interval $[t - m, t]$ as follows:

$$\begin{aligned} \lim_{n\to\infty} \lim_{r\to\infty} \mu_i(t) &= \lim_{n\to\infty} \frac{1}{m} \int_{t-m}^{t} \frac{1}{n-1} \sum_{j\neq i, j=1}^{n} \phi_j(\tau) d\tau = \\ &= \frac{1}{m} \int_{t-m}^{t} \phi(\tau) d\tau. \end{aligned} \tag{4.4}$$

By recalling the definition in formula (4.2), one has that the individual dynamics is regulated by the switch "$\mu_i(t) > 0.5$?" (i.e. an agent chooses how to update his strategy depending on whether he believes that B players constitutes the majority or not). Therefore, for $r \to \infty$ and $n \to \infty$, individuals form identical expectations (identical memories μ_i) and make the same choices. This implies that agents are synchronous. It follows that individual and group trajectories coincide.

The exact solution

By inspection of the simulation results, I earlier observed that the collective trajectory $\phi(t)$ exhibits periodic oscillations in a specific range of parameters. Figure 4.8 presents 'oscillations' obtained from simulation data. The image is annotated to clarify the argument that is about to follow. The argument is based on a thought experiment.

Assume that agents are oscillating synchronously in a periodic fashion with an arbitrarily fixed amplitude, A, according to equations (4.1b)-(4.1c). Solving

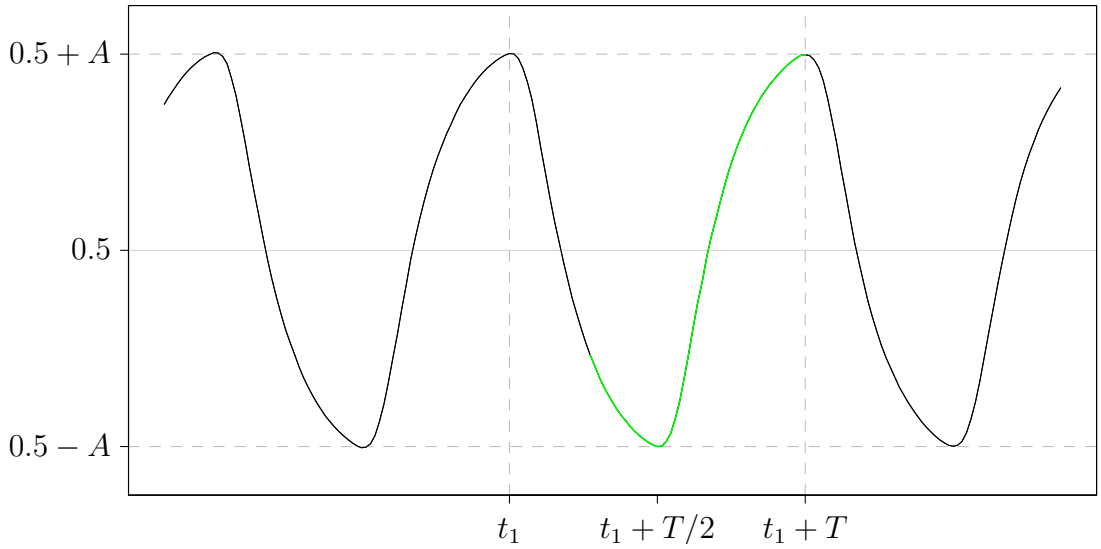

Figure 4.8: Illustration of the assumptions in the thought experiment that helps determining amplitude and period. The portion of the trajectory in green shows the reach into the past for the individual memory of an agent at time $t_1 + T$. Here, $m = 50$, $T \approx 66$ time units and $A \approx 0.43$.

equations (4.1b)-(4.1c) in the limit $\Delta t \to 0$, one obtains an analytical piecewise description of the collective trajectory strictly valid for $n, r \to \infty$ and m

arbitrarily fixed and finite. Equations (4.1b)-(4.1c) can be rewritten, in the limit $\Delta t \to 0$, as follows.

$$\frac{\phi_i(t+\Delta t) - \phi_i(t)}{\Delta t} = \epsilon\rho[1 - \phi_i(t)] \tag{4.5}$$

$$\frac{\phi_i(t+\Delta t) - \phi_i(t)}{\Delta t} = -\epsilon\rho\phi_i(t). \tag{4.6}$$

Using a more concise notation, the same equations can be rewritten as

$$\dot{\phi}_i(t) = \epsilon\rho[1 - \phi_i(t)] \tag{4.7}$$

$$\dot{\phi}_i(t) = -\epsilon\rho\phi_i(t). \tag{4.8}$$

The general-form solution to equation 4.8 is $\phi_i(t) = k\exp(-\epsilon\rho t)$. This general-form solution describes a descending curve and, notably, approximates well the trajectory of the plot in the interval in which it decreases (figure 4.8). Therefore, for the purpose of writing the solution, it is useful to define time with reference to the plot in figure 4.8. Without loss of generality, one can define the initial condition $\phi_i(t_1 + 0) = 0.5 + A$. Consequently, the solution to equation 4.8 is $\phi_i(t_1 + t) = (0.5 + A)\exp(-\epsilon\rho t)$. Similarly, one finds a solution to equation 4.7. Finally, I choose to write the two solutions in the following form, where time is defined in reference to figure 4.8.

$$\phi\left(t_1 + \frac{T}{2} + t\right) = 1 - (0.5 + A)\exp(-\epsilon\rho t) \qquad \text{(when ascending)} \tag{4.9}$$

$$\phi(t_1 + t) = (0.5 + A)\exp(-\epsilon\rho t) \qquad \text{(when descending)} \tag{4.10}$$

Notably, this piece-wise exponential solution approximates well the trajectory observed in figure 4.8, but fails to describe the collective trajectory $\phi(t)$ for all values of $\epsilon\rho$ in the range $[0, 1]$ in the case of finite batches (i.e. r finite).

Amplitude and period.

Now consider that the solutions in 4.9-4.10 are symmetric and $\phi(t)$ reaches a maximum at times t_1 and $t_2 = t_1 + T$, and a minimum at time $t_1 + \frac{T}{2}$, as annotated in figure 4.8. One has that $\phi(t)$ evaluated at $t_1 + \frac{T}{2}$ is $0.5 - A$. Thus, from equation 4.10, it must hold that

$$(0.5 + A)\exp\left(-\epsilon\rho\frac{T}{2}\right) = 0.5 - A \tag{4.11}$$

By expanding equation 4.11 and solving for A, I obtain

$$A = 0.5 \cdot \tanh(\epsilon\rho \cdot T/4). \tag{4.12}$$

The next argument is based on the assumption that agents aim to be in the minority side and update their strategy $\phi_i(t)$ based on their current estimate, $\mu_i(t)$, of the fraction of B players in the group. In the model this corresponds to the fact that agent i chooses to decrease $\phi_i(t)$ when $\mu_i(t) > 0.5$ and increase $\phi_i(t)$ when $\mu_i(t) < 0.5$. This implies that $\mu_i(t) - 0.5$ changes sign exactly at the time when agent i switches between 'increasing $\phi_i(t)$' and 'decreasing $\phi_i(t)$'. From this it follows that, with reference to figure 4.8, μ_i takes value 0.5 at time $t_1 + T$. By using the formula for μ_i given in 4.4, one obtains

$$\mu_i(t_1 + T) = \frac{1}{m} \int_{t_1+T-m}^{t_1+T} \phi(\tau) d\tau = 0.5. \tag{4.13}$$

In figure 4.8, the green segment overriding the black curve shows exactly the region of $\phi(t)$ that is being integrated in expression 4.13. The segment extends backwards for m time units starting at time $t_1 + T$.

Since the ascending and descending parts of the trajectory $\phi(t)$ have an horizontal axes of symmetry at 0.5, the integral calculated over one period gives $\frac{1}{T}\int_{t_1}^{t_1+T} \phi(\tau) d\tau = 0.5$. Equating the last two integrals, one obtains:

$$\frac{1}{T-m} \int_{t_1}^{t_1+T-m} \phi(\tau) d\tau = 0.5 \tag{4.14}$$

Integral 4.14 is calculated between time $t_1 + T - m$ and t_1, i.e it refers to the descending tract of the trajectory. Thus, in equation 4.14, one can replace $\phi(t)$ with the solution given in 4.10 and A with the expression given in 4.12. The resulting equation can be manipulated to obtain the following equation:

$$\Big[1 + \tanh(\epsilon\rho T/4)\Big]\Big[1 - \exp[-\epsilon\rho(T-m)]\Big] = (T-m)\epsilon\rho. \tag{4.15}$$

I was not able to solve this equation globally but was able to find an approximate solution in the limit $\epsilon\rho \to 0$ for m finite and fixed. The derivation of this solution is in section 4.8.1. The approximation is shown in equation 4.16 and shows that $T \to 2m$ as $\epsilon\rho \to 0$.

$$\frac{T(\epsilon\rho, m)}{m} = 2 - \frac{1}{3}m\epsilon\rho + \frac{1}{18}(m\epsilon\rho)^2 + \frac{2}{135}(m\epsilon\rho)^3 + \mathcal{O}((\epsilon\rho)^4) \tag{4.16}$$

In remarkable agreement, separate calculations for finite batches (r finite) show that oscillations arising near the Hopf-bifurcation (cf. equation 4.67) have the property that $T \to 2m$ as $\epsilon \to \epsilon_c$ (assuming $\rho = 1$). Substituting equation 4.16 (for T) into equation 4.12 (for A), one obtains an expression for the amplitude in the limit $\epsilon\rho \to 0$

$$\begin{aligned} A(\epsilon\rho, m) = \frac{1}{2} \cdot \tanh\Big(\frac{1}{2}m\epsilon\rho - \frac{1}{12}(m\epsilon\rho)^2 + \frac{1}{72}(m\epsilon\rho)^3 + \\ \frac{1}{270}(m\epsilon\rho)^4 + \mathcal{O}((\epsilon\rho)^5)\Big). \end{aligned} \tag{4.17}$$

In these approximations T/m and A are functions of the 'fundamental variable' $\epsilon\rho m$ as $\epsilon\rho \to 0$. Now let us consider the substitution $\tau = \frac{T}{m}$. Then equation 4.15 becomes

$$\Big[1 + \tanh(\epsilon\rho m\tau/4)\Big]\Big[1 - \exp[-\epsilon\rho m(\tau - 1)]\Big] = \epsilon\rho m(\tau - 1). \tag{4.18}$$

Again, this means that, if 4.15 admits a solution, it must be of the form $T/m = f(\epsilon\rho m)$. The existence of such solution is confirmed numerically[1] and plotted in figure 4.9. Once the period is determined numerically, the numerical values can be used to compute the amplitude, A, through equation 4.12. The numerical solution confirms that A and T/m are functions of $m\epsilon\rho$, thus effectively reducing the description of the behaviour of the system to just two universal curves. One can also verify that T/m and $m\epsilon\rho$ are dimensionless. Figure 4.9 also point out that T/m takes values in the interval [1,2]. This fact is consistent with my thought experiment in which I assumed that memory lasts between T and $T/2$ time units.

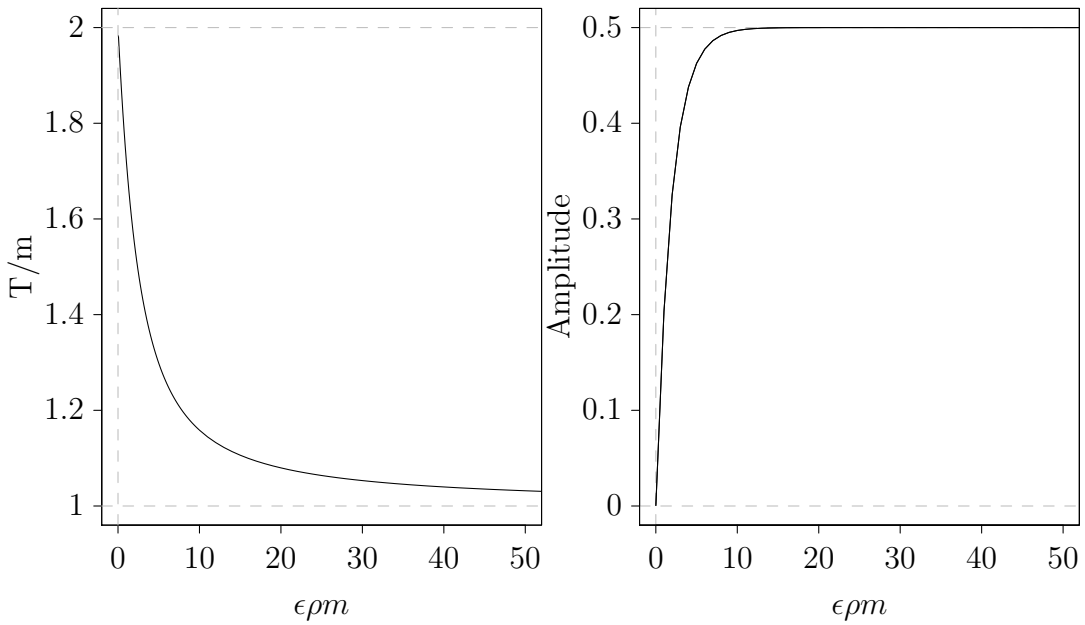

Figure 4.9: Plots of the two universal curves (black lines) that describe the amplitude and period of the oscillations in the limiting case $r \to \infty$. The curve for T/m tends asymptotically to 1 as $\epsilon\rho m \to \infty$. In reality it seems more plausible that $\rho \to \infty$, rather than $m \to \infty$. Grey dashed lines only help the eye.

[1]The equation is solved using the function *roots.all* from the RootSolve package (Karline Soetaert, 2009).

4.5 Scaling properties

At this stage we can ask whether the scaling properties observed in section 4.4 for $r \to \infty$ hold for r finite. In order to check this hypothesis, in figure 4.10 I compare the numerical values of amplitude and period for $r \to \infty$ (black lines) and for $r = 10$ and $r = 5$ obtained from simulation data (red and orange lines). More specifically, figure 4.10 (A) compares the amplitude A as ϵ ranges in $[0, 1]$ for $\rho = 1$ fixed in the following three scenarios:

(i) $r \to \infty$

(ii) $rm = 250$ for the two combinations (r, m) equals $(5, 50)$ and $(10, 25)$

(iii) $rm = 500$ for the two combinations (r, m) equals $(5, 100)$ and $(10, 50)$

Similarly, figure 4.10 (B) compares the period T/m for $r \to \infty$ and $rm = 250$ with ϵ in the range $[0, 1]$. Figure 4.10 (C) compares the period T/m for $r \to \infty$ and $rm = 500$ with ϵ in the range $[0, 1]$. These figures show that $\epsilon\rho m$ and rm are the two fundamental parameters that regulate the system's behaviour even when r is finite. The reader can note that there is a region in figure 4.10 in which the period stops being a smooth line. This effect indicates that the algorithm used to calculate the period, which works by identifying peaks, does not provide a reliable estimate. In turn, this is due to the fact that the periodic cyclic oscillations broke down. However, by performing a spectral analysis of the trajectories, I can confirm the value of the period before the oscillations break down and further identify the main frequency in the entire range displayed. The value of T/m for T calculated through spectral analysis is shown in figure 4.24.

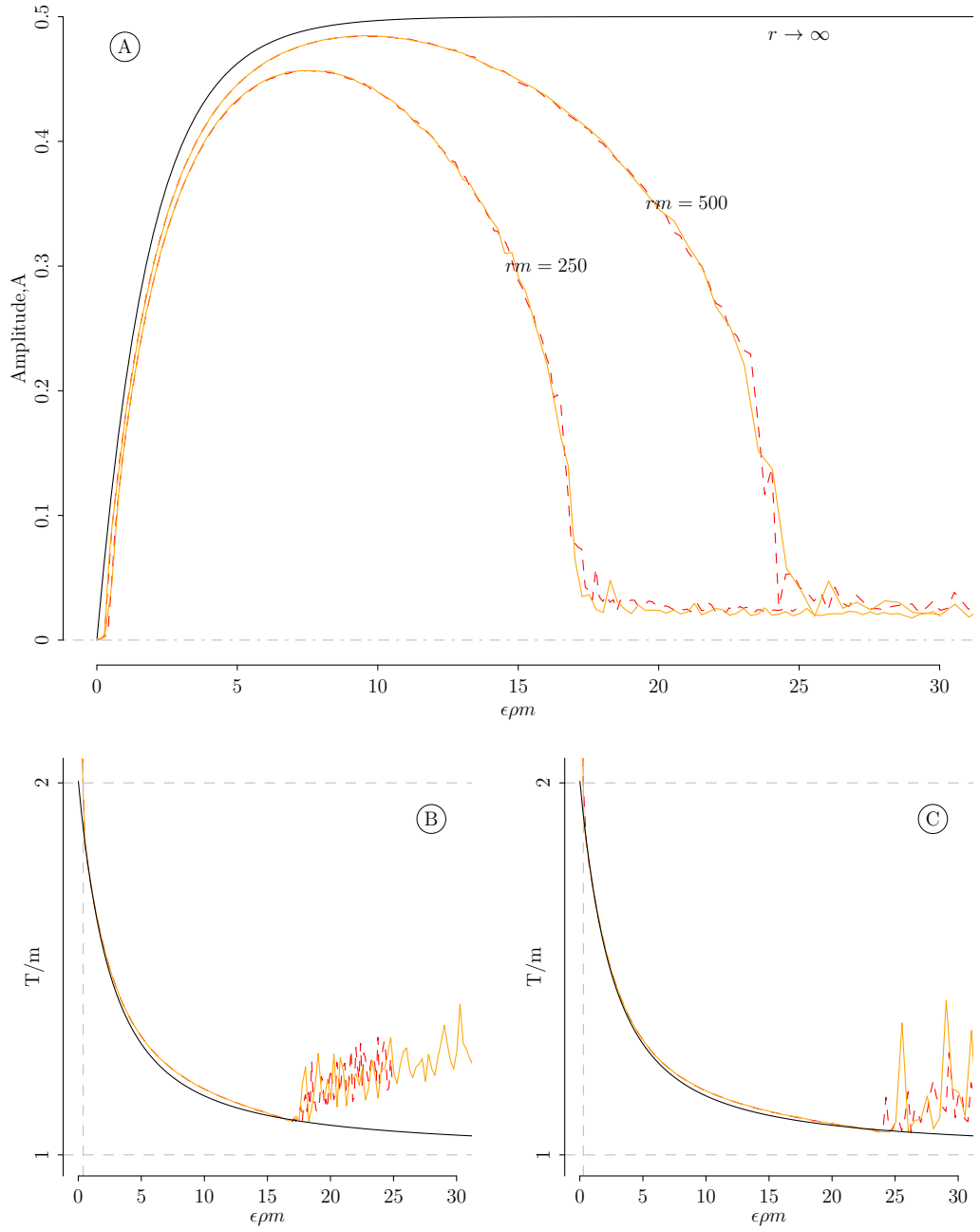

Figure 4.10: (A) A as a function of $m\epsilon\rho$ ($\rho = 1$) from simulation data (red dashed and orange lines) and exact results (black line). Here, the simulations results for $r = 10$ with $m = 25, 50$ (red dashed lines) and $r = 5$ with $m = 50, 100$ (orange lines) are merged from the two panels in fig. 4.7. The amplitude and period for the limiting case $r \to \infty$ are obtained via analytical and numerical means (cf. equations 4.12,4.15). The range for ϵm is chosen to highlight the full range of interesting outcomes. (B) Period T/m for $r \to \infty$ and $rm = 250$. A vertical line is drawn at $\epsilon_c m = 0.3875$. (C) Period T/m for $r \to \infty$ and $rm = 500$. A vertical line is drawn at $\epsilon_c m = 0.275$. In both cases ϵ_c is retrieved from the simulations.

Figure 4.11 shows that the scaling property obtained for $\rho = 1$ fixed and ϵ variable (figure 4.10) is preserved for $\epsilon = 0.01$ fixed and ρ variable.

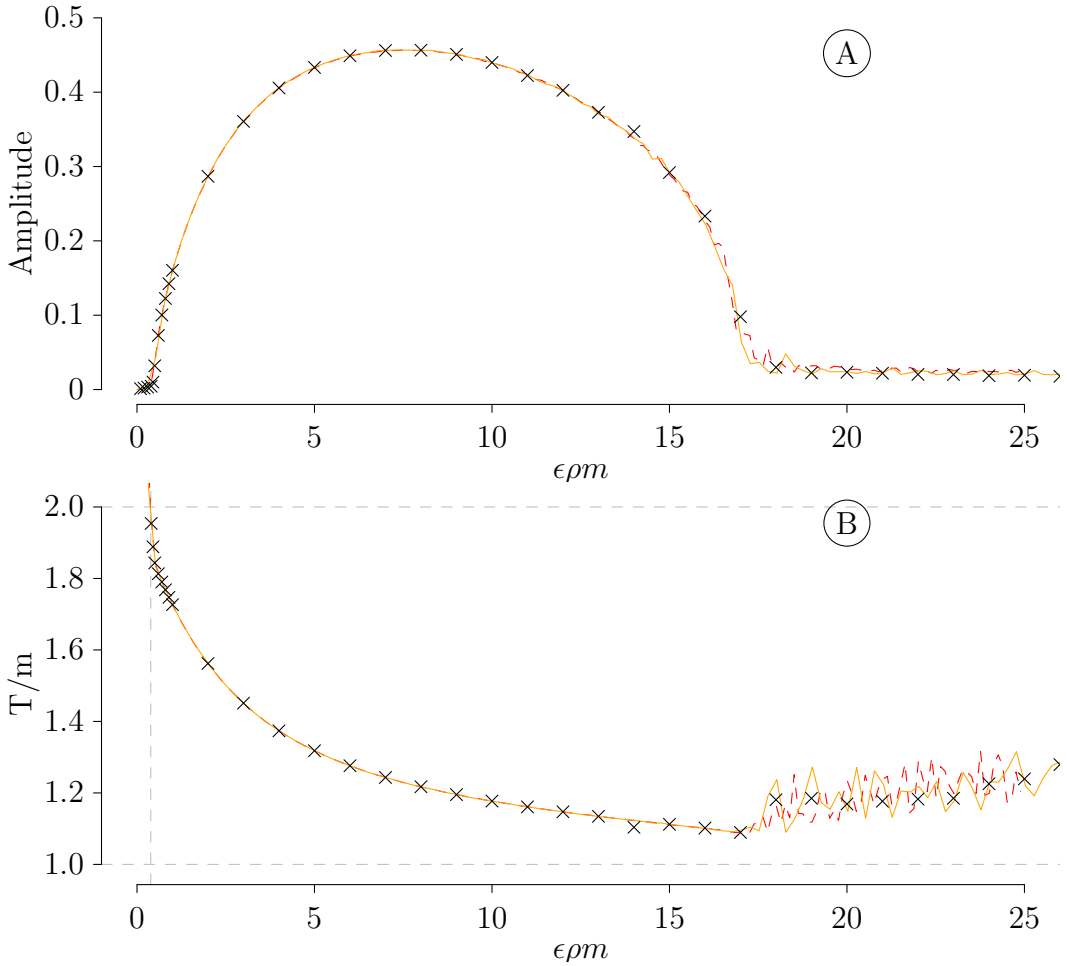

Figure 4.11: The black crosses correspond to the values of the amplitude, in (A), and period, in (B), for simulation results obtained for $mr = 50 \times 5$, $\epsilon = 0.01$ and ρ varying in $[0, 50]$. A closer look at the transition from steady state to regular oscillation for $\phi(t)$ at the bifurcation in this scenario is given in figure 4.23. The red and orange lines correspond to the cases $mr = 250$ in figure 4.10. Here, the dynamics for ϵ fixed and variable fully match.

Figure 4.12 shows that the scaling property valid for $\rho = 1$ fixed and ϵ variable (figure 4.10) is largely preserved for $\epsilon = 1$ fixed and ρ variable.

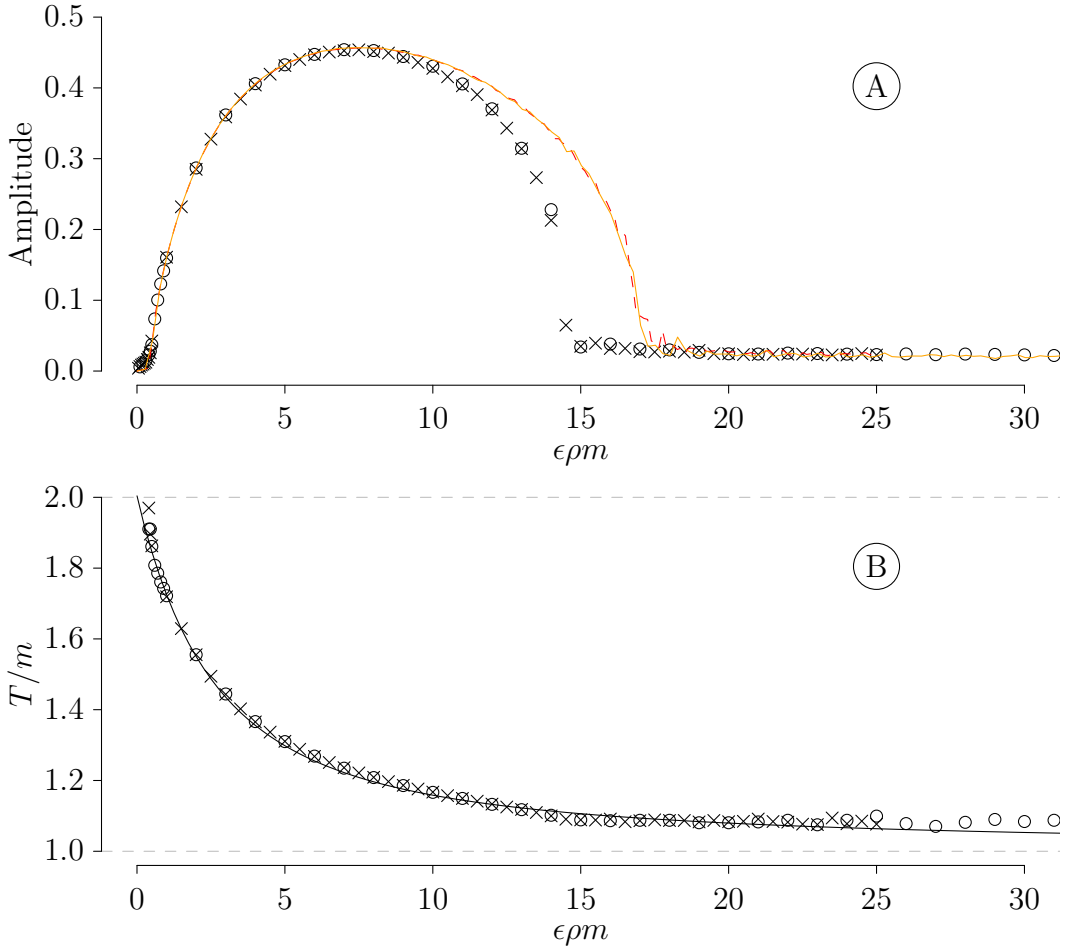

Figure 4.12: (A) Amplitude for $\epsilon = 1$ fixed and ρ in $[0, 1]$ for $mr = 25 \times 10$ (black triangles) and $mr = 50 \times 5$ (black circles) compared to $mr = 250$ with ϵ in $[0, 1]$ and $\rho = 1$ fixed (red and orange lines). (B) Period calculated through spectral analysis for $\epsilon = 1$ fixed and ρ in $[0, 1]$ for $mr = 25 \times 10$ (black triangles) and $mr = 50 \times 5$ (black circles) compared to the prediction for $r \to \infty$ (black line). Overall, the behaviour near the Hopf-bifurcation is preserved. The overall phenomenology is also preserved.

4.5.1 Analytical predictions at the bifurcation

Bifurcation as ϵ increases

Here, I present an analytical derivation of the value of ϵ at the transition from steady states to limit cycles for given memory duration m and observation rate r. The derivation follows the steps proposed by Burridge *et al.* (2017), here adapted to include the parametrization of the rate of observation r. In this section, I present the main results while the calculations with a comment on the derivation and accuracy of the prediction are presented in section 4.8.2. The calculations correctly estimate the critical ϵ_c as well as the period of the oscillations near the bifurcation. They also confirm that the fundamental variables are $\epsilon\rho m$ as well as rm. To highlight the role of ϵ, and without loss of generality, in this section I write ϵ_c in place of $(\epsilon\rho)_c$, thus assuming $\rho = 1$.

In 4.19, I present an expression for the critical ϵ at the bifurcation, namely ϵ_c. This expression is an approximation of the analytical derivation that preserves the scaling properties and still provides a good estimation. For a full expression see 4.70.

$$\epsilon_c \approx \frac{1}{2}\sqrt{\frac{\pi^5}{2rm^3}} + \frac{\pi^3}{2rm^2} + \frac{3}{4}\sqrt{\frac{\pi^7}{2r^3m^5}} + ... \tag{4.19}$$

The accuracy of the analytical prediction 4.19 is shown in figure 4.13 where it is shown to match the bifurcation point for simulation data. In addition, by rewriting equation 4.19, one obtains 4.20. In this expression, one readily sees that $\epsilon_c m$ is a function of rm.

$$\epsilon_c m \approx \frac{1}{2}\sqrt{\frac{\pi^5}{2rm}} + \frac{\pi^3}{2rm} + \frac{3}{4}\sqrt{\frac{\pi^7}{2r^3m^3}} + \frac{12\pi^4 + \pi^6}{24r^2m^2} + ... \tag{4.20}$$

Next I report a truncated version of the solution obtained in 4.67 for the frequency y at the Hopf-bifurcation.

$$y = \frac{\pi}{m} + \frac{\pi^{1.5}}{\sqrt{2r}m^{1.5}} + \frac{\pi^2}{2rm^2} + ... \tag{4.21}$$

Since the frequency is calculated in radiants per time unit, from the frequency I calculate the period, $T = 2\pi/y$, to obtain at leading order that $T \to 2m$ as $\epsilon \to \epsilon_c$. The precision of this relationship is shown in figure 4.10, panels (B) and (C), where $T/m \to 2$ as $\epsilon \to \epsilon_c$ for the simulation data. Studying the dynamics of this system by varying ϵ ($\epsilon\rho$) has shown us that there are three phases: stable equilibrium, oscillations and a breakdown of the oscillations. However, studying the same system as r varies gives us a different insight.

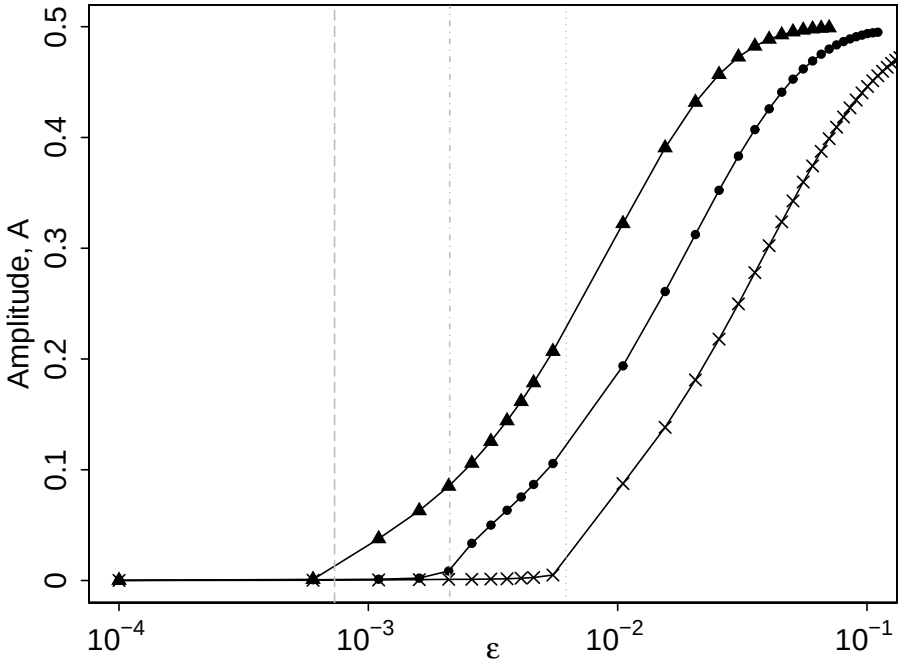

Figure 4.13: The vertical lines show the analytical approximation of ϵ_c for memory duration 200 (dashed), 100 (dot dashed), 50 (dotted), $\rho = 1$ and $r = 10$. The curved lines show the amplitude, A, of the trajectories $\phi(t)$ obtained with the simulations for memory duration 200 (triangles), 100 (circles), 50 (crosses), $\rho = 1$ and $r = 10$ as ϵ varies.

Bifurcation as r increases

By varying r the system displays only one transition, this being from a stable equilibrium towards determinist cycles. In figure 4.14 I show two examples of this type of bifurcation. For low rates of observation (i.e. small r), the collective trajectory $\phi(t)$ reaches a steady state. As r increases past a critical value $\phi(t)$ develops bounded oscillations. This means that, for a fixed memory duration, oscillations emerge at a critical level of information detail. In particular, this transition captures the role of information detail in the coordination of a large group of agents. The critical value r_c at which oscillations arise can be estimated numerically by solving 4.70 for r, with $\epsilon\rho$ and m fixed. In figure 4.14, the vertical dashed lines indicate the numerical value predicted for r_c while the horizontal dashed lines indicates the amplitude A calculated numerically for the case $r \to \infty$ (cf. 4.12 and 4.15).

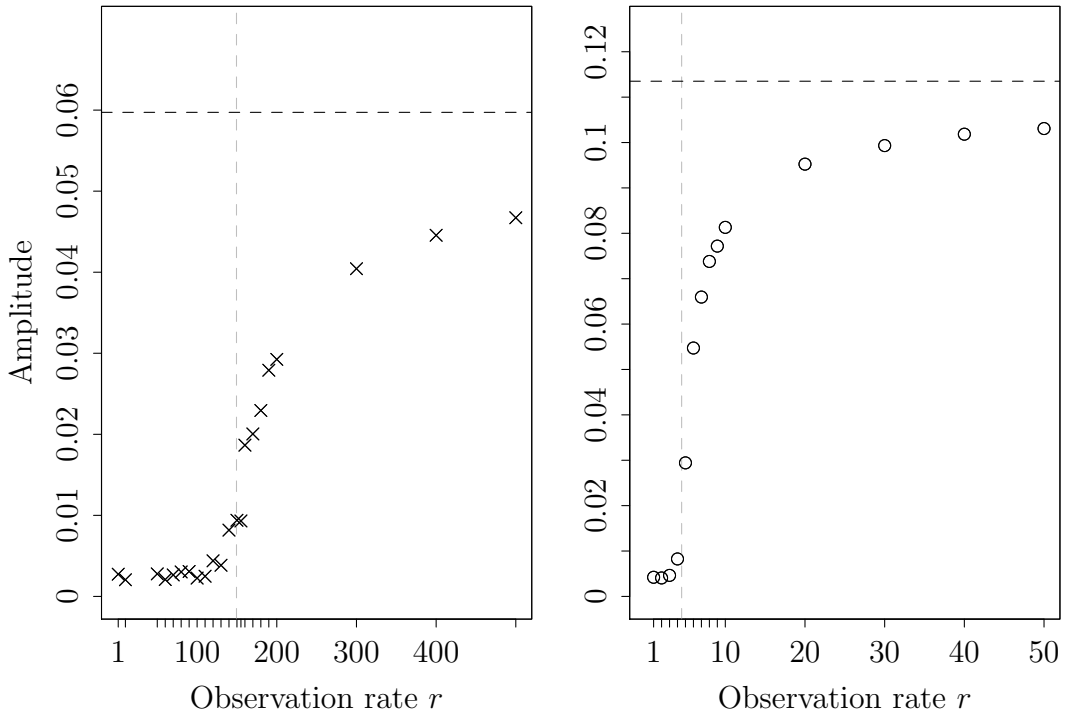

Figure 4.14: Amplitude, A, of the fluctuations of $\phi(t)$ as a function of r from simulation data. Simulation results are obtained for $m = 5$, $\rho = 25$ (crosses) and $m = 50$, $\rho = 5$ (circles) while $\epsilon = 0.002$ in both plots. There is a Hopf-bifurcation at a critical observation rate r_c. The horizontal lines indicate the amplitude obtained numerically for $r \to \infty$. The vertical lines show the prediction for the value of r_c at the bifurcation, obtained numerically by solving 4.70 for r, with $\epsilon\rho$ and m fixed. The prediction for r_c equals 149.44 and 4.52 respectively.

Therefore, as shown in figure 4.14, I can predict the emergence of coordinated oscillations in the population as a function of the rate of sampling and to further estimate the upper bound of the amplitude, reached as the sampling rate tends to infinity. The plots show a slight difference between the upper bound and the actual amplitude measured from the simulation data. This difference is mainly due to simulation data being recorded in time steps of one unit. The simulation data is only recorded at times 1, 2, 3, 4, etc. Therefore, the maximum amplitude is not always captured in the data, especially for small memory duration values. The impact of this discretization on the amplitude measurement is negligible for long memories but non-negligible for short memories. Therefore, for short memories, the measurement can be improved by recording the simulation data at a finer resolution. I suggest that the observation rate acts as an 'inverse temperature of selection' analogous to the role of selection intensity in the 'thermal minority

game' (TMG) (Cavagna *et al.*, 1999). The TMG differs from the MG as it introduces a parameter Γ modelling selection intensity, which is considered equivalent to an 'inverse temperature of selection' (see equation 3.7). In the limiting case $\Gamma \to \infty$, agents choose precisely the best response and the system is deterministic; the limiting case $\Gamma \to 0$ models random choice. In the TMG, a higher 'inverse temperature' produces the 'remarkable' (Challet *et al.*, 2013) result of increasing fluctuations, precisely as the observation rate does in my model. Since, in the minority game, the group's objective is to coordinate their actions to reduce fluctuations, Cavagna *et al.* (1999) concludes that agents perform better when operating with a non-zero degree of individual error, i.e. when Γ is finite. Further work should address similarities and differences between my model and the TMG and investigate whether the model that I studied corresponds to a limiting case of the thermal minority game, namely the limit $z = D/N \to \infty$, where D is the size of the strategy space and N the population size. It is precisely in this limit that Cavagna *et al.* (1999) drew the conclusion that I mentioned.

In addition, having modelled the rate of sampling allows us to suggest a reinterpretation of previous studies, in particular referring to the work of Burridge *et al.* (2015) and Burridge *et al.* (2017). Burridge and co-authors study two-strategy anti-coordination competition models in which they incorporate memory duration but not a rate of observation. The authors show that long memories give agents a 'statistical' advantage against short memory agents (insofar as memory is not too long since this would originate oscillations). The statistical advantage derives from a larger sample but vanishes as the sample becomes too large, thus triggering fluctuations. In particular, a greater memory duration (when restricted to values that do not originate oscillations) provides a statistical advantage since it allows agents to locate the best response and reduces randomness in the system. In this manner, the authors show that memory duration acts equivalently to an 'inverse temperature' (Burridge *et al.*, 2015). In this regard, it would be interesting to investigate whether the statistical advantage in these systems, regulated by the sample size, is primarily linked to the sampling rate rather than memory duration.

4.6 Comments

'Information flow'

Until now, in this chapter, we have seen through analytical, numerical and simulation results that the quantities rm and $\epsilon\rho m$ essentially regulate A and T/m. Here I provide an explanation of why these two quantities are crucial.

First, let us assume that $\epsilon = 1$ is fixed and study the system in terms of the rates r and ρ. On average, r defines the number of samples collected in one unit of time, whereas $1/\rho$ defines the time elapsed between two consecutive strategy updates. Next, consider two populations are having identical values for rm, and also for ρm. This implies that the ratio $rm/\rho m$ is identical for the two populations and equals r/ρ. This ratio is important because it defines the average size of the sample of observations collected by any one of the agents between two consecutive strategy updates. This implies that, on average, agents in the two populations share the sample size rm of the information stored in their memory and the size of the sample collected between consecutive strategy updates. Thus the 'flow of information' remains unchanged when considering 'equivalent' systems.

The roles of r and m

Having modelled memory both in terms of accuracy (rate r per unit of time) and duration (m time units), I find that the role of the memory size is two-faceted. Although rm is itself a control parameter, r and m contribute differently to the dynamics.

A higher level of information detail unequivocally increases coordination among agents. As $r \to \infty$, since all agents use the same learning heuristic, agents form identical expectations. On the contrary, as $r \to 0$, agents collect widely diverse information (i.e. with significant variation and variance) and cannot coordinate their behaviour.

A change in memory duration can also mediate coordinated behaviour. Memory duration, however, produces a double-fold effect. Memory duration can indeed influence a collective phenomenon (oscillations) that requires a degree of coordination among agents. However, a larger memory duration causes an 'earlier' break down of the oscillations. As shown in figure 4.7, a longer memory duration corresponds to a smaller range of values for which the variable ϵ sustains oscillations (for r and ρ fixed).

4.7 Strategy space

In a previous section, figure 4.5 shows how the individual strategies (the ϕ_i values shown as orange dots) are distributed for six different values of ϵ, while $r = 10$, $m = 50$ and $\rho = 1$ are fixed. In particular, the values of the individual strategies are displayed over 200 consecutive time units for populations made of 1000 agents.

Here, figure 4.15 portrays the entire 'strategy space' used by agents. For each of 209 different ϵ values in $[0, 1]$ and r, m, ρ fixed as above, I plot the set of all values assumed by the individual strategy ϕ_i adopted by each of 1000 agents over 200 consecutive time units. For each ϵ value, a total of 200000 points is plotted (1000 individual strategies recorded for 200 successive time units).

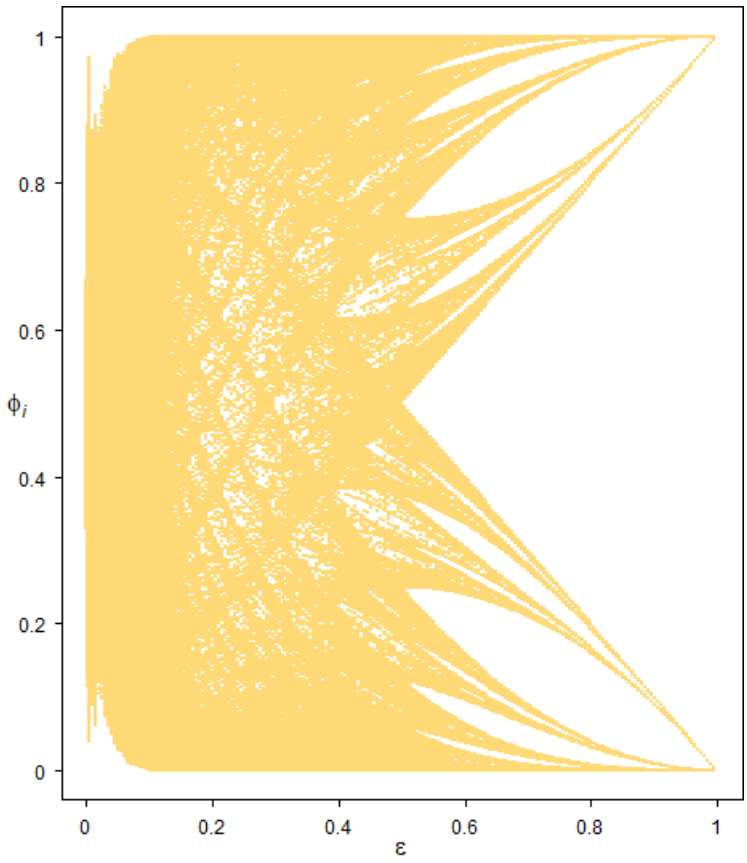

Figure 4.15: For each of 209 values of ϵ on the x-axis, I have plotted, as points, all the ϕ_i values assumed by 1000 agents over 200 time units, resulting in 200000 points for each ϵ value. All 209 simulated populations have parameters $r = 10$, $m = 50$ and $\rho = 1$ fixed.

In figure 4.16, the data from the last figure is displayed differently. On the $y-$axis, the interval $[0, 1]$ is divided into 400 small intervals of width 0.0025 and, for each ϵ value, the 200000 data points are grouped into the appropriate intervals. A vertical line is drawn at $\epsilon = 0.5$ to help the eye see where the strategy space seems to diversify. It is helpful to note that this plot does not alter the visual representation of the strategy space significantly. However, it allows me to analyse the data to show how the frequency of strategies varies across the different intervals. In figure 4.17 I use a colour code to show precisely this.

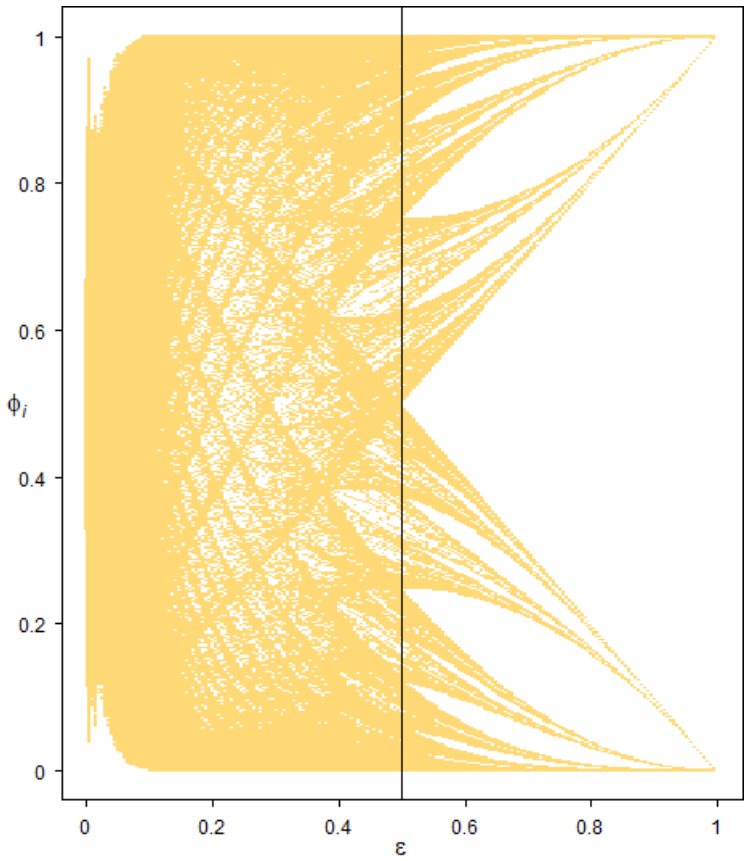

Figure 4.16: This figure represents the same data as in figure 4.15. However, the $[0, 1]$ interval for the y-axis is divided into intervals of duration 0.0025, and the data is grouped accordingly. At the mid-point of each group, a dot is plotted if a strategy is recorded within the group at least once (i.e. if the group has a strictly positive frequency). Only 400 buckets are available to depict 200000 points. Additionally, a vertical line is drawn at $\epsilon = 0.5$ to help the eye see where the strategy space diversify.

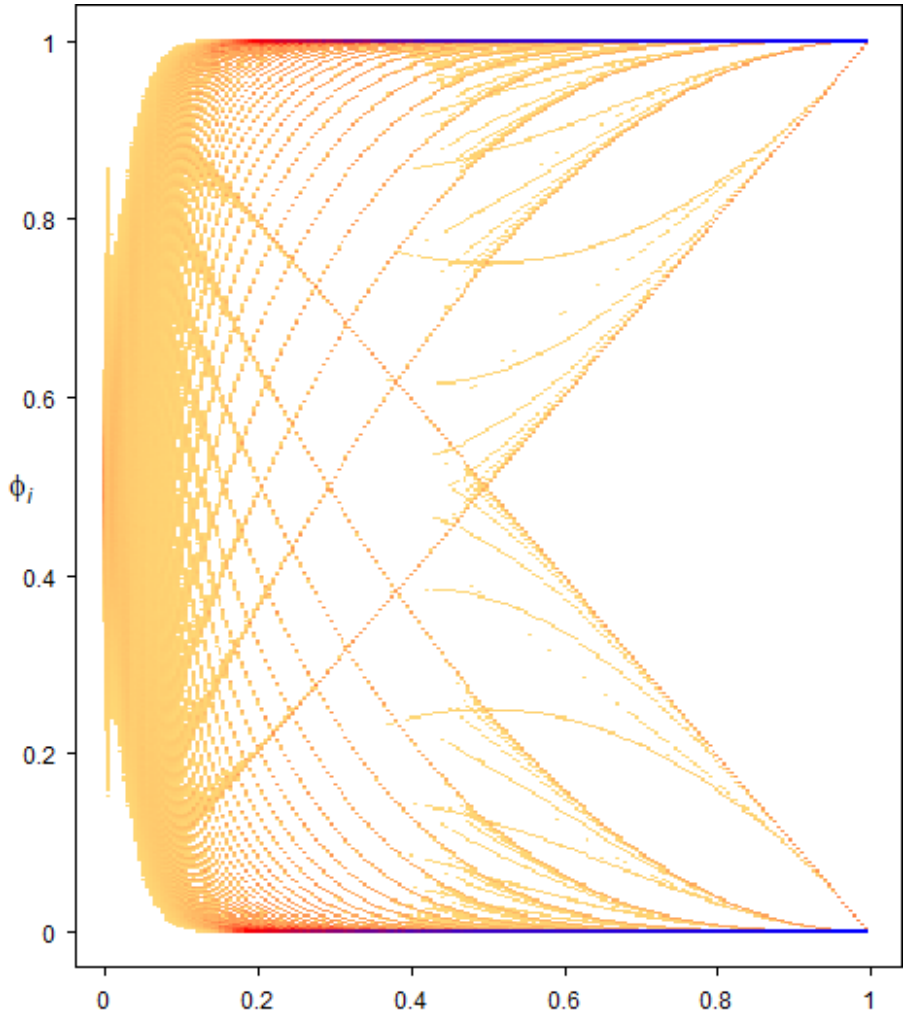

Figure 4.17: As in 4.16, the data is grouped. However, groups are not displayed if the frequency is lower than 101. Since there are 200000 data-points, having fewer than 101 recurrences within a group represents a probability of less than or equal to 0.05%. The colour-coded legend below shows how the probability that an agent is in a given group is linked to the colour it is plotted with.

4.7.1 Characterizing the strategy space

The images above offer an insight into the structure of the strategy space as a function of the variable ϵ. Notably, an evolving strategy travels a path produced by a sequence of steps, with each step governed by one of the two equations 4.1b-4.1c. Here, I characterize the strategy space by considering the set of all possible travel paths. First, I rewrite equations 4.1b-4.1c in a convenient form as recursive equations, obtaining, respectively,

$$x_{n+1} = x_n + \epsilon(1 - x_n) \tag{4.22}$$
$$x_{n+1} = x_n - \epsilon x_n. \tag{4.23}$$

In addition, equations 4.22-4.23 can be rewritten as

$$x_{n+1} = x_n(1 - \epsilon) + \epsilon \tag{4.24}$$
$$x_{n+1} = x_n(1 - \epsilon). \tag{4.25}$$

Using equations 4.24-4.25, I explicitly calculate all possible outcomes of the evolution of the system of recursive equations starting from x_0 and up the third iteration. The set of all trajectories is displayed as a tree diagram in figure 4.18. This tree helps inferring a closed form expression for the general solution

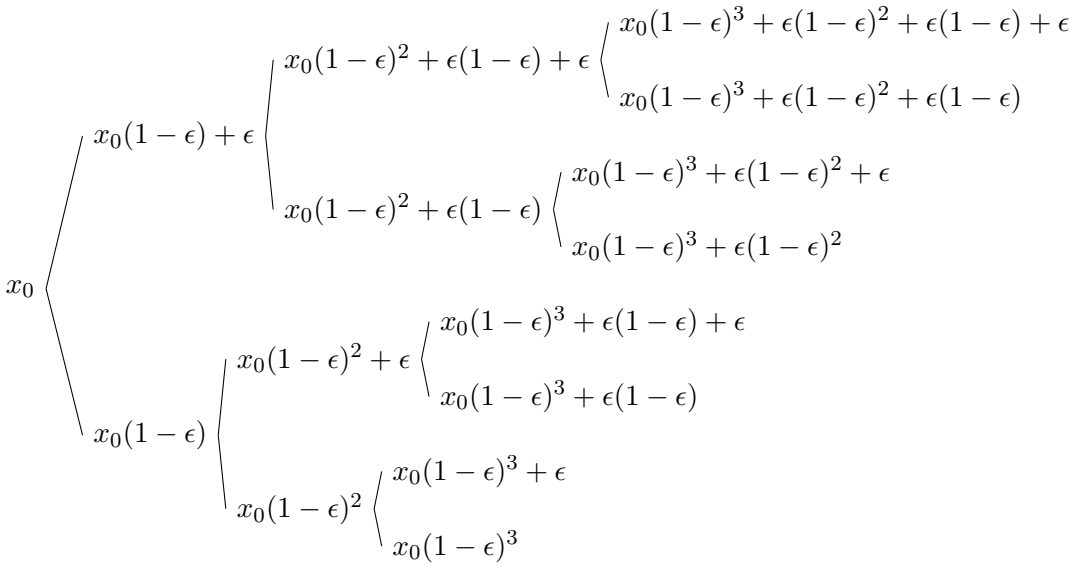

Figure 4.18: All possible outcomes of the evolution of the system of recursive equations starting from x_0 and up the third iteration.

after n steps, based solely upon knowing which iterations obeyed equation 4.24

rather than 4.25. Let define N to be the set of natural numbers from 1 to n, i.e. $\{1, 2, ..., n-1, n\}$. Let define S to be the subset of N such that the iteration from x_{k-1} into x_k is increasing (i.e. it follows 4.24) for $k \in S$ and decreasing for $k \in N \setminus S$. Using this definition, the value of the variable x_n with initial condition x_0 is

$$x_n = x_0(1-\epsilon)^n + \sum_{k \in S} \epsilon(1-\epsilon)^{n-k}. \tag{4.26}$$

From equation 4.26, one can see that the total number of paths can be computed as $\sum_{r=0}^{n} \binom{n}{r}$. For each r, $\binom{n}{r}$ represents how many (different) paths are made of $n-r$ terms, each term having the form $u(1-\epsilon)^i$ with $0 \leq i \leq n$ and $u = x_0$ or $u = \epsilon$. For example, for $r = n$, there is only one path ($\binom{n}{n} = 1$), namely $x_0(1-\epsilon)^n$, which has no terms of the form $\epsilon(1-\epsilon)^i$. Alternatively, one can note that the number of paths is doubled at each iteration and therefore the number of paths for x_n is 2^n. By counting the solutions through the two methods proposed here, one obtains the identity $2^n = \sum_{r=0}^{n} \binom{n}{r}$, i.e. a case of the binomial theorem.

Equation 4.26 also shows that x_0 only appears in the highest-order term, $x_0(1-\epsilon)^n$, which is negligible as n tends to infinity. This fact implies that, for large n, ,x_n is independent of the initial condition x_0.

Considering each recursive equation separately, its evolution admits a closed-form solution in the following form.

$$x_n = 1 - (1 - x_0)(1-\epsilon)^n \tag{4.27}$$
$$x_n = x_0(1-\epsilon)^n \tag{4.28}$$

where x_0 is the initial value and n the number of iterations. Equation 4.28 follows directly from equation 4.25 whereas a proof for equation 4.27 is given next.

Brief proof. Using equation 4.22 recursively, the value of x_n for $n = 1, 2, 3$ can be written as follows.

$$\begin{aligned}
x_1 =& x_0 + \epsilon(1 - x_0) \\
x_2 =& x_0 + \epsilon(1 - x_0) + \epsilon\left(1 - (x_0 + \epsilon(1 - x_0))\right) \\
x_3 =& x_0 + \epsilon(1 - x_0) + \epsilon\left(1 - (x_0 + \epsilon(1 - x_0))\right) + \\
&+ \epsilon\left(1 - (x_0 + \epsilon(1 - x_0) + \epsilon\left(1 - (x_0 + \epsilon(1 - x_0))\right)\right)\right)
\end{aligned}$$

By means of expanding, collecting and rearranging, one obtains

$$x_1 = 1 - (1 - x_0) + \epsilon(1 - x_0) \tag{4.29}$$
$$x_2 = 1 - (1 - x_0) + 2\epsilon(1 - x_0) - \epsilon^2(1 - x_0) \tag{4.30}$$
$$x_3 = 1 - (1 - x_0) + 3\epsilon(1 - x_0) - 3\epsilon^2(1 - x_0) + \epsilon^3(1 - x_0) \tag{4.31}$$

The coefficients of $(1 - x_0)$ are the binomial expansion of $(1 - \epsilon)^n$, for $n = 1, 2, 3$. By generalising this result, equation (4.23) becomes

$$x_n = 1 - (1 - x_0)(1 - \epsilon)^n. \qquad \text{End of proof.} \tag{4.32}$$

The results collected so far allow to better understand the structure of the strategy space. The solution equation 4.26 tells us that all points in the strategy space are reached as a modification of the power series in the variable $(1-\epsilon)$ with coefficient x_0 for the higher order term and ϵ for every other term. Ignoring the trivial case $\epsilon = 0$ and $\epsilon = 1$, equation 4.32 tells us that the power series converges towards 1 as n tends to infinity. The leading terms of this series are the lower order terms, namely ϵ, $\epsilon(1-\epsilon)$, $\epsilon(1-\epsilon)^2$, and so forth. This implies that a point, say p, in the strategy space may be written as

$$p = \epsilon + \epsilon(1 - \epsilon) + \epsilon(1 - \epsilon)^2 + \mathcal{O}\left((1 - \epsilon)^3\right). \tag{4.33}$$

The corresponding point inversion (across the point 0.5) can be written as

$$q = 1 - \left[\epsilon + \epsilon(1 - \epsilon) + \epsilon(1 - \epsilon)^2 + \mathcal{O}\left((1 - \epsilon)^3\right)\right]. \tag{4.34}$$

Following this notation, I list all points obtained as a combination of their leading terms up to the third leading term, for ϵ fixed. There are 16 (2×2^3) such points.

$$\begin{aligned}
p_1 &= \epsilon + \epsilon(1 - \epsilon) + \epsilon(1 - \epsilon)^2 + \mathcal{O}\left((1 - \epsilon)^3\right). \\
p_2 &= \epsilon(1 - \epsilon) + \epsilon(1 - \epsilon)^2 + \mathcal{O}\left((1 - \epsilon)^3\right). \\
p_3 &= \epsilon + \epsilon(1 - \epsilon)^2 + \mathcal{O}\left((1 - \epsilon)^3\right). \\
p_4 &= \epsilon(1 - \epsilon)^2 + \mathcal{O}\left((1 - \epsilon)^3\right). \\
p_5 &= \epsilon + \epsilon(1 - \epsilon) + \mathcal{O}\left((1 - \epsilon)^3\right). \\
p_6 &= \epsilon(1 - \epsilon) + \mathcal{O}\left((1 - \epsilon)^3\right). \\
p_7 &= \epsilon + \mathcal{O}\left((1 - \epsilon)^3\right). \\
p_8 &= 0 + \mathcal{O}\left((1 - \epsilon)^3\right). \\
p_9 &= 1 - \left[\epsilon + \epsilon(1 - \epsilon) + \epsilon(1 - \epsilon)^2 + \mathcal{O}\left((1 - \epsilon)^3\right)\right]. \\
p_{10} &= 1 - \left[\epsilon(1 - \epsilon) + \epsilon(1 - \epsilon)^2 + \mathcal{O}\left((1 - \epsilon)^3\right)\right]. \\
p_{11} &= 1 - \left[\epsilon + \epsilon(1 - \epsilon)^2 + \mathcal{O}\left((1 - \epsilon)^3\right)\right]. \\
p_{12} &= 1 - \left[\epsilon(1 - \epsilon)^2 + \mathcal{O}\left((1 - \epsilon)^3\right)\right]. \\
p_{13} &= 1 - \left[\epsilon + \epsilon(1 - \epsilon) + \mathcal{O}\left((1 - \epsilon)^3\right)\right]. \\
p_{14} &= 1 - \left[\epsilon(1 - \epsilon) + \mathcal{O}\left((1 - \epsilon)^3\right)\right]. \\
p_{15} &= 1 - \left[\epsilon + \mathcal{O}\left((1 - \epsilon)^3\right)\right]. \\
p_{16} &= 1 - \mathcal{O}\left((1 - \epsilon)^3\right).
\end{aligned}$$

I look at these points as the 'limit values' for all realizations of the system of two iterative equations 4.22-4.23, for fixed ϵ as $n \to \infty$. In figure 4.19, the 16 points are plotted as ϵ varies in $[0, 1]$ and ignoring the $\mathcal{O}(\cdot)$ term. Notably, these points generate lines that accurately represent the borders of the strategy space on the right hand side of the plot. In particular, the lines separate the right hand side of the strategy space into 8 separate segments. These results can be repeated for

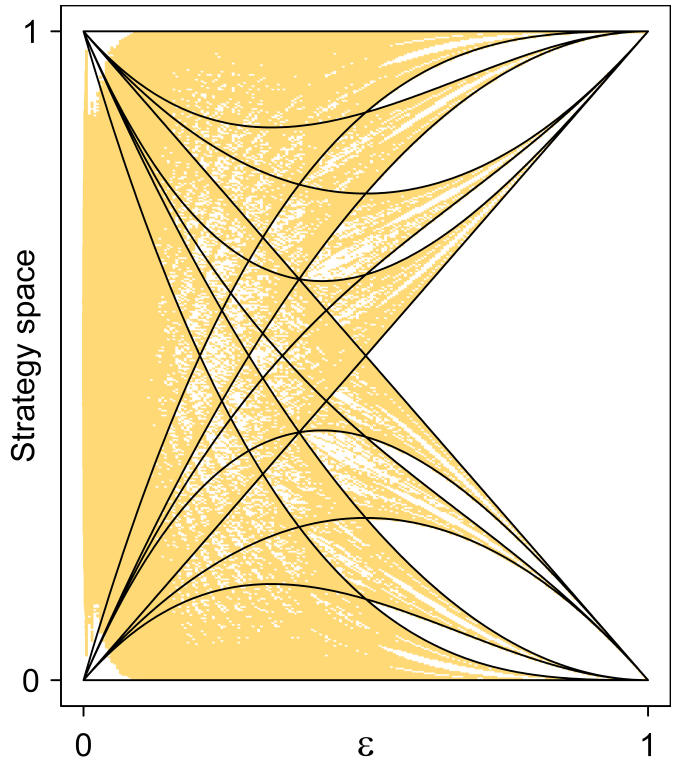

Figure 4.19: Truncated series describe paths in the strategy space predictive of its structure.

any number of 'leading terms'. For example, four terms would generate 32 lines, including the 16 lines already generated, splitting the right hand side of the plot into 16 separate segments.

Lastly, a short argument allows to prove that, for $0.5 < \epsilon \leq 1$, the strategy space is restricted to the area above the line $f(\epsilon) = \epsilon$ and below the line $f(\epsilon) = 1 - \epsilon$. By taking any point $x_n \in [0, 1]$, x_{n+1} can either take value $x_n(1 - \epsilon)$ or $x_n(1-\epsilon)+\epsilon$. Since $x_n \in [0, 1]$, it follows that $x_n(1-\epsilon) \leq 1-\epsilon$ and $x_n(1-\epsilon)+\epsilon \geq \epsilon$. This shows that either $x_{n+1} > \epsilon$ or $x_{n+1} < 1 - \epsilon$.

4.8 Analytical results

4.8.1 Analytical approximation of the period T near $\epsilon = 0$

In this section, I approximate the solution for $T(\epsilon)$ in equation 4.15 (here rewritten in equation 4.35) as a power series expansion valid near $\epsilon = 0$.

$$\Big[1 + \tanh(\epsilon T/4)\Big]\Big[1 - exp[-\epsilon(T-m)]\Big] = (T-m)\epsilon. \tag{4.35}$$

I rewrite equation 4.35

$$f(\epsilon) = (T-m)\epsilon - \Big[1 + \tanh(\epsilon T/4)\Big]\Big[1 - exp[-\epsilon(T-m)]\Big].$$

and aim to find $T(\epsilon)$ for $f(\epsilon) = 0$. I assume $T(\epsilon) = x_0 + x_1\epsilon + x_2\epsilon^2 + x_3\epsilon^3 + x_4\epsilon^4 + \mathcal{O}(\epsilon^5)$ and rewrite $f(\epsilon)$ as

$$f(\epsilon) = (T-m)\epsilon - 1 - \tanh(\epsilon T/4) + exp[-\epsilon(T-m)] + \tanh(\epsilon T/4)exp[-\epsilon(T-m)].$$

I derive the following Taylor expansions for $\epsilon \to 0$.

(i) $(T-m)\epsilon = (x_0 - m)\epsilon + x_1\epsilon^2 + x_2\epsilon^3 + x_3\epsilon^4 + x_4\epsilon^5 + \mathcal{O}(\epsilon^6)$.

(ii) $\tanh(\epsilon T/4) = \frac{x_0}{4}\epsilon + \frac{x_1}{2}\frac{\epsilon^2}{2} + (\frac{3}{2}x_2 - \frac{1}{32}x_0^3)\frac{\epsilon^3}{6} + (6x_3 - \frac{3}{8}x_0^2x_1)\frac{\epsilon^4}{24} + \mathcal{O}(\epsilon^5)$.

(iii) $\exp[-\epsilon(T-m)] = 1 + (m-x_0)\epsilon + [(m-x_0)^2 - 2x_1]\frac{\epsilon^2}{2} + [-6x_1(m-x_0) + (m-x_0)^3 - 6x_2]\frac{\epsilon^3}{6} + [-12x_1(m-x_0)^2 - 24x_2(m-x_0) + (m-x_0)^4 + 12x_1^2 - 24x_3]\frac{\epsilon^4}{24} + \mathcal{O}(\epsilon^5)$.

By substituting the Taylor expansions into the function above, one obtains

$$T(m\epsilon) = 2m - \frac{m^2}{3}\epsilon + \frac{m^3}{18}\epsilon^2 + \frac{2m^4}{135}\epsilon^3 + \mathcal{O}(\epsilon^4). \tag{4.36}$$

Finally, one obtains the following power series expansion for T:

$$\frac{T(\epsilon, m)}{m} = 2 - \frac{1}{3}(m\epsilon) + \frac{1}{18}(m\epsilon)^2 + \frac{2}{135}(m\epsilon)^3 + \mathcal{O}(\epsilon^4). \tag{4.37}$$

Substituting equation 4.36 into equation 4.12 I obtain an expression for A that depends on ϵm as follows:

$$A(m\epsilon) = \frac{1}{2} \cdot tanh\Big(2m\epsilon - \frac{1}{3}(m\epsilon)^2 + \frac{1}{18}(m\epsilon)^3 + \frac{2}{135}(m\epsilon)^4 + \mathcal{O}(\epsilon^5)\Big) \tag{4.38}$$

The expression for $T(\epsilon)$ in equation 4.36 has a region of non-uniformity of order $\mathcal{O}(\frac{1}{m})$. I expect that equation 4.36 approximates the solution for $\epsilon \in [0, \frac{1}{m})$.

4.8.2 Hopf-bifurcation

Burridge *et al.* (2017) have rewritten the set of equations 4.1a-4.1c as one single delay differential equation that specifies the evolution of $\phi(t)$. From the linearization of the equation, the authors obtained a critical value of ϵ at the Hopf-bifurcation, say ϵ_c. An important step in this process consists of approximating the memory content of an individual with, first, a binomial distribution and, second, a normal distribution. This approximation works better for large samples of observations. In this chapter, the size of the memory content of an individual is given by rm. While the work of Burridge *et al.* (2017) corresponds to the case $r = 1$, requiring a long memory duration for the approximation to hold, I study the case $r > 0$, which allows having large samples for a short memory duration m and large observation rate r. Here I present a derivation for ϵ_c valid for any positive observation rate r. Although it is still convenient to assume m large in analytical derivations, eventually the results hold for mr large (i.e. it is the product of r and m that matters, rather than m alone). Overall, in the following calculations, I follow the ideas presented by Burridge *et al.* (2017) although, at times, I follow different analytical derivations of the results. In particular, a powers expansion method is introduced in steps 4.64-4.71. In steps 4.44a-4.53, a novel approach is proposed for the derivation of equation 4.48.

Finding the evolution equation. Here, I am interested in deriving an evolution equation for $\phi(t)$. However, first, it is useful to study the evolution equation of a single individual. From equations (4.1a)-(4.1c), one can derive an expected value for the individual displacement $\Delta\phi_i(t)$ during an infinitesimal time interval Δt. Assuming that $\mu_i(t)$ is a stochastic variable (more on this later), one obtains:

$$\mathbb{E}\left[\phi_i(t+\Delta t) - \phi_i(t)\right] = \epsilon\rho\left(\mathbb{E}\left[\tilde{p}(\mu_i(t))\right] - \phi_i(t)\right)\Delta t. \tag{4.39}$$

Without loss of generality, to simplify the notation, I assume that $\rho = 1$ for the remainder of this section.

In equation 4.39, the quantity $\mu_i(t)$ is stochastic. Its distribution can be approximated as follows. For memories made of large samples, i.e. $mr \to \infty$, the distribution of the set of sampling times in agents' memory at time t approaches a uniform distribution in $[t-m, t]$. Consequently, $\mu_i(t)$ is approximated by the time-averaged average population strategy as follows:

$$\mu_i(t) \to \frac{1}{m}\int_{t-m}^{t} \phi(\tau)\delta\tau \text{ as } mr \to \infty. \tag{4.40}$$

It is now useful to establish the following notation:

$$\overline{\phi}(t) := \frac{1}{m}\int_{t-m}^{t} \phi(\tau)\delta\tau. \tag{4.41}$$

The number, say d, of B players that an agent observes in mr observations will have a probability mass function which, for $\phi(t)$ constant, is binomial and, at time t, follows:

$$\mathbb{P}(d = h) \approx \frac{(m \cdot r)! \cdot (\overline{\phi}(t))^h \cdot (1 - \overline{\phi}(t))^{mr-h}}{h!(mr - h)!}. \tag{4.42}$$

Thus d can be approximated with a binomial distribution with 'probability of success' $\overline{\phi}(t)$, that is $d_t \sim B(m\,r, \overline{\phi}_t)$, where $\overline{\phi}_t$ is a shorthand notation for the moving average $\overline{\phi}(t)$. When mr is large enough, the binomial can be approximated with a normal distribution to obtain $d_t \sim N\Big(\bar{\phi}_t\, m\, r, \bar{\phi}_t\,(1-\bar{\phi}_t)\, m\, r\Big)$. This allows to approximate the expected mean of the memory of a general individual, $\mu_i(t)$, obtaining:

$$\mu_i(t) = \frac{d_t}{m\,r} \sim N\left(\bar{\phi}_t, \frac{\bar{\phi}_t\,(1-\bar{\phi}_t)}{m\,r}\right). \tag{4.43}$$

At this stage, we have the tools to answer the following question: what is the value of $\mathbb{E}\left[\tilde{p}(\mu_i(t))\right]$? To answer this question, it is useful to remind the reader that $\tilde{p}(\mu_i(t))$ can only take value in $\{0,1\}$ with the following probabilities.

$$\tilde{p}(\mu_i(t)) = \begin{cases} 0 & \text{w.p. } P\left(N\left(\bar{\phi}_t, \frac{\bar{\phi}_t\,(1-\bar{\phi}_t)}{m\,r}\right) > 0.5\right). \quad (4.44a) \\ 1 & \text{w.p. } P\left(N\left(\bar{\phi}_t, \frac{\bar{\phi}_t\,(1-\bar{\phi}_t)}{m\,r}\right) < 0.5\right). \quad (4.44b) \end{cases}$$

Assuming that $\phi(t)$ makes small oscillations around 0.5, I define

$$\psi(t) := \phi(t) - 0.5. \tag{4.45}$$

and the corresponding time average

$$\overline{\psi}(t) := \frac{1}{m}\int_{t-m}^{t} \psi(\tau)d\tau. \tag{4.46}$$

In light of the last two definitions, equations 4.44a-4.44b can be rewritten as

$$\tilde{p}(\mu_i(t)) = \begin{cases} 0 & \text{w.p. } P\left(N\left(\bar{\psi}_t, \frac{0.5^2-\bar{\psi}_t^2}{m\,r}\right) > 0\right). \quad (4.47a) \\ 1 & \text{w.p. } P\left(N\left(\bar{\psi}_t, \frac{0.5^2-\bar{\psi}_t^2}{m\,r}\right) < 0\right). \quad (4.47b) \end{cases}$$

From equations 4.47a-4.47b, it follows that

$$\mathbb{E}\left[\tilde{p}(\mu_i(t))\right] = P\left(N\left(\bar{\psi}_t, \frac{0.5^2 - \bar{\psi}_t^2}{m\,r}\right) < 0\right). \tag{4.48}$$

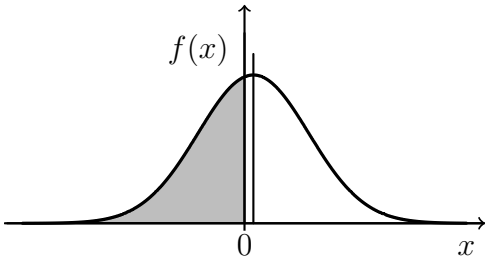

Figure 4.20: Probability distribution function $f(x)$ of a normal random variable having a small positive mean $\bar{\psi}$ and variance $\frac{0.5^2-\bar{\psi}_t^2}{m\,r}$. There is a vertical line at $x = \bar{\psi}$. The shaded area corresponds to the value of $\mathbb{E}\left[\tilde{p}(\mu_i(t))\right]$ according to equation 4.48.

As a helpful visual reference, the shaded area in figure 4.20 corresponds to the value of $\mathbb{E}\left[\tilde{p}(\mu_i(t))\right]$. To evaluate its value, one can perform a Taylor expansion for the normal cumulative distribution function around 0 and calculate its value to first order in $\bar{\psi}_t$. First, I can rewrite the argument of the probability as follows.

$$P\left(N\left(\bar{\psi}_t, \frac{0.5^2-\bar{\psi}_t^2}{m\,r}\right) < 0\right) = P\left(N\left(0, \frac{0.5^2-\bar{\psi}_t^2}{m\,r}\right) < -\bar{\psi}_t\right). \tag{4.49}$$

The value of the former can be computed as the following integral.

$$\int_{-\infty}^{-\bar{\psi}_t} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx, \tag{4.50}$$

with $\sigma^2 = \frac{0.5^2-\bar{\psi}_t^2}{m\,r}$. The $\exp(\cdot)$ function can be written as a Taylor expansion around $x = 0$ to give

$$\int_{-\infty}^{-\bar{\psi}_t} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{-\bar{\psi}_t} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{2^n n! \sigma^{2n}} dx. \tag{4.51}$$

Since $\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = 0.5$, the value of the former becomes

$$\frac{1}{2} + \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=0}^{\infty} (-1)^n \frac{\left(-\bar{\psi}_t\right)^{2n+1}}{2^n n! \sigma^{2n} (2n+1)}. \tag{4.52}$$

Additionally, a Taylor expansion for $\frac{1}{\sigma}$ in $\bar{\psi}_t$ around 0 gives

$$\frac{1}{\sigma} = 2\sqrt{mr} + \mathcal{O}(\bar{\psi}_t{}^2). \tag{4.53}$$

Equations 4.52 and 4.53, taken together, give the following first order approximation, valid for $\bar{\psi}_t \to 0$.

$$\mathbb{E}\left[\tilde{p}(\mu_i(t))\right] \approx \frac{1}{2} - \frac{\sqrt{2mr}}{\sqrt{\pi}}\bar{\psi}_t. \tag{4.54}$$

Alternatively, equation 4.54 can be derived with an 'holistic' approach. Figure 4.20 displays the probability distribution function $f(x)$ of $\mathrm{N}(\bar{\psi}_t, \frac{0.5^2-\bar{\psi}_t^2}{mr})$ and a vertical line at $x = \bar{\psi}_t$. With reference to this figure, the value of $\mathbb{E}\left[\tilde{p}(\mu_i(t))\right]$ corresponds the the grey shaded area underlying the bell curve. Use $\bar{A}$ to denote the value of the area of the small rectangular stripe underlying the bell curve for $x \in [0, \bar{\psi}_t]$. For $\bar{\psi}_t$ infinitesimal, which corresponds to small oscillations, one has $\bar{A} = 0.5 - f(\bar{\psi}_t)\bar{\psi}_t$. The following equality concludes the alternative proof.

$$f(\bar{\psi}_t) = \frac{\sqrt{2mr}}{\sqrt{\pi}}. \tag{4.55}$$

Finally, equation 4.54 helps writing an evolution equation for the group average strategy $\phi(t) = 0.5 + \psi_t$. At this stage, the aim becomes to approximate the value of

$$\frac{\Delta\psi(t)}{\Delta t} = \frac{1}{n}\sum_{i=1}^{n}\frac{\Delta\psi_i(t)}{\Delta t}. \tag{4.56}$$

For $n \to \infty$, the evolution of $\phi(t)$ can be approximated with the following differential equation:

$$\frac{\Delta\phi(t)}{\Delta t} = \epsilon\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{p}(\mu_i(t)) - \phi(t)\right). \tag{4.57}$$

The corresponding evolution equation for ψ_t is

$$\frac{\Delta\psi(t)}{\Delta t} = \epsilon\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{p}(\mu_i(t)) - \psi(t)\right). \tag{4.58}$$

Since individual memories are independent and identically distributed, I approximate the average $\frac{1}{n}\sum_{i=1}^{n}\tilde{p}(\mu_i(t))$ with the expected value of $\tilde{p}(\mu_i(t))$, to obtain the following delay differential equation valid for small oscillations:

$$\frac{\Delta\psi(t)}{\Delta t} = \epsilon\left(\frac{\sqrt{2mr}}{\sqrt{\pi}}\overline{\psi}(t) - \psi(t)\right). \tag{4.59}$$

Finding the Hopf bifurcation. To find a solution to equation 4.59, I substitute $e^{\lambda t}$ as a trial solution to obtain the characteristic equation

$$\lambda^2 + \lambda\epsilon + \frac{\sqrt{2r}}{\sqrt{\pi m}}\epsilon\Big(1 - e^{-\lambda m}\Big). \tag{4.60}$$

By substituting λ with $x + iy$, I obtain

$$x^2 - y^2 + \epsilon x + \frac{\sqrt{2r}}{\sqrt{\pi m}}\epsilon\Big(1 - e^{-mx}\cos(my)\Big) = 0. \tag{4.61}$$

$$2xy + \epsilon y + \frac{\sqrt{2r}}{\sqrt{\pi m}}\epsilon\Big(e^{-mx}\sin(my)\Big) = 0. \tag{4.62}$$

The simulations show a Hopf-Bifurcation arising for $\phi(t)$ at critical values of ϵ, which depend on r and m. To calculate these critical values, I set $x = 0$ in equation (4.62), obtaining

$$\sin(my) = -\frac{\sqrt{\pi}}{\sqrt{2mr}}my \tag{4.63}$$

In order to approximate the solution for the frequency y, I suppose that the frequency is a function of m, $y(m)$, and approximate $my(m)$ with the following power series expansion in the limit $m \to \infty$:

$$my(m) = x_0 + x_1\eta + x_2\eta^2 + x_3\eta^3 + ... \qquad \text{with} \quad \eta = m^\alpha, \alpha < 0. \tag{4.64}$$

By using the substitution 4.64, equation 4.63 becomes

$$\sin(x_0 + x_1\eta + x_2\eta^2 + x_3\eta^3 + ..) = -\frac{\sqrt{\pi}}{\sqrt{2mr}}(x_0 + x_1\eta + x_2\eta^2 + x_3\eta^3 + ...). \tag{4.65}$$

The left hand side of 4.65 can be approximated with a Taylor expansion centred at $\eta = 0$, which gives

$$\sin(x_0) + x_1\cos(x_0)\eta + \Big(x_2\cos(x_0) - \frac{1}{2}{x_1}^2\sin(x_0)\Big)\eta^2 + ... = -\frac{\sqrt{\pi}}{\sqrt{2mr}}(x_0 + x_1\eta + ..). \tag{4.66}$$

The leading term gives $\sin(x_0) = 0$. This has infinite solutions but I look for the slowest oscillation frequency. This is also called the fundamental solution (Jarvis, 2017). Thus one obtains $x_0 = \pi$. Using the last result, the next two leading terms now give $-x_1\eta = -\frac{\sqrt{\pi}}{\sqrt{2mr}}\pi$. This is in a convenient form to see that $\eta = \frac{1}{\sqrt{m}}$

and $x_1 = \frac{\pi^{\frac{3}{2}}}{\sqrt{2r}}$. Continuing iteratively in this manner, eventually the frequency can be written as

$$\begin{aligned} y(m) =& \frac{\pi}{m} + \frac{\pi^{1.5}}{\sqrt{2r}m^{1.5}} + \frac{\pi^2}{2rm^2} + \frac{6\sqrt{2}\pi^{2.5} + \sqrt{2}\pi^{4.5}}{24r^{1.5}m^{2.5}} + \\ & \frac{3\pi^3 + 2\pi^5}{12r^2m^3} + \mathcal{O}(\frac{1}{m^{3.5}}) \quad \text{as} \qquad m \to \infty. \end{aligned} \tag{4.67}$$

Simulations results confirm that this is a very accurate approximation of the frequency at the Hopf-bifurcation: in solution 4.67, at leading order, the period of the oscillations (period= $2\pi/y$) at the Hopf-bifurcation is $2m$ which is confirmed in simulation data shown in figure 4.10, panels (B) and (C). To obtain the critical value for ϵ, one can substitute expression (4.67) into equation (4.61) and solve for $x = 0$. Equation (4.61) becomes

$$y(m)^2 = \frac{\sqrt{2r}}{\sqrt{\pi m}}\epsilon\Big(1 - \cos(my(m))\Big). \tag{4.68}$$

One can expand $cos(\cdot)$ in the RHS, for $m \to \infty$, to obtain

$$\cos(my(m)) = -1 + \frac{\pi^3}{4rm} + \frac{\pi^3.5}{2\sqrt{2}r^1.5m^{1.5}} + \frac{12\pi^4 + \pi^6}{32r^2m^2} + ... \tag{4.69}$$

Finally, the critical value can be written, for $m \to \infty$, as follows:

$$(\epsilon\rho)_c = \frac{\frac{1}{m}\left(\sqrt{\frac{\pi^5}{2rm}} + \frac{\pi^3}{rm} + \frac{3}{2}\sqrt{\frac{\pi^7}{2(rm)^3}} + \frac{12\pi^4+\pi^6}{12(rm)^2} + \frac{15\pi^{4.5}+5\pi^{6.5}}{12\sqrt{2}(rm)^{2.5}}\right) + ..}{2 - \frac{\pi^3}{4rm} - \frac{\pi^{3.5}}{2\sqrt{2}(rm)^{1.5}} - \frac{12\pi^4+\pi^6}{32(rm)^2} - \frac{3\pi^{4.5}+\pi^{6.5}}{6*\sqrt{2}(rm)^{2.5}}..}. \tag{4.70}$$

Although not a formal approximation, a polynomial solution for ϵ_c with $\rho = 1$ can be given by approximating the denominator to 2, in which case one has

$$\begin{aligned} \epsilon_c \quad \approx \quad & \frac{1}{2}\sqrt{\frac{\pi^5}{2rm^3}} + \frac{\pi^3}{2rm^2} + \frac{3}{4}\sqrt{\frac{\pi^7}{2r^3m^5}} + \\ & + \frac{12\pi^4 + \pi^6}{24r^2m^3} + \frac{6\sqrt{2}\pi^{2.5} + \sqrt{2}\pi^{4.5}}{24r^{2.5}m^{3.5}} + \mathcal{O}\left(\sqrt{\frac{1}{m^7}}\right). \end{aligned} \tag{4.71}$$

Comments. The accuracy of the analytical prediction 4.70 is presented in figure 4.13 where it is shown to match the bifurcation point for simulation data. The critical relationship at the Hopf bifurcation in equation 4.70 can be rewritten in terms of the fundamental variables as $(\epsilon\rho m)_c = g(rm)$. The shape of $g(x)$ is plotted in figure 4.21.

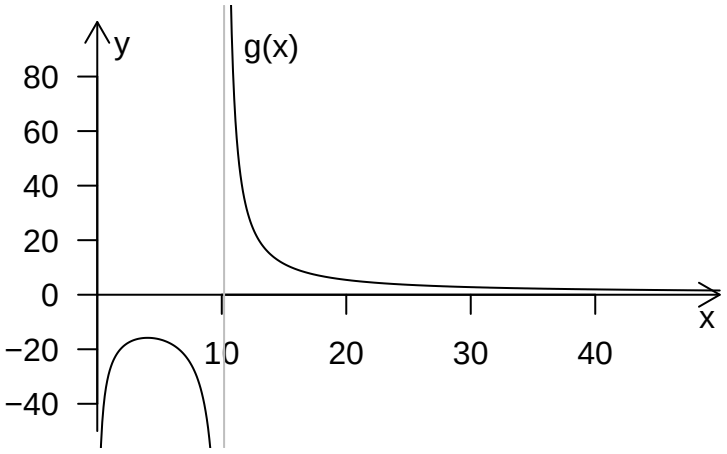

Figure 4.21: Plot of $g(x)$, the function defined such that equation 4.70 can be rewritten as $\epsilon\rho m = g(rm)$.

This plot shows that, for given $\epsilon\rho m > 0$, there exists a unique r that satisfies 4.70. Therefore the critical value of r at the bifurcation, say r_c, can be determined numerically by inverting a truncated version of 4.70. The accuracy of this calculation is presented in figure 4.14 where it is shown to closely predict the bifurcation point for simulation data.

Lastly, consider the denominator of the main fraction in equation 4.70. Setting this denominator to 0 and solving for rm provides the discontinuity point of the function $g(x)$. Numerical solutions show that the discontinuity point is at $x_d \approx 10.1871$. The function $g(x)$ tends to positive ∞ as $rm \to x_d$, with $rm > x_d$. Consequently, as rm approaches x_d, the Hopf bifurcation point $(\epsilon\rho m)_c$ tends to ∞. The discontinuity point also acts as a lower bound such that for $rm < x_d$ the system does not admit a Hopf bifurcation.

An evolutionary interpretation of the magical number seven. The existence of a tipping point, corresponding to only 10 bits of information, provides a reason to motivate the existence and evolution of the magical number seven (Miller, 1956). George Miller's 1956 influential article proposes that humans can hold 7 ± 2 chunks of information in short-term memory on average. Other studies propose that a lower limit of 4 elements applies (Cowan, 2001). In all cases, these limits imply that recollection in short-term human memory does not exceed the critical value of x_d bits. The model studied in this chapter proposes that, when attending to the same information, the use of very small samples prevent a group from incurring the cost of inefficient collective decision making (i.e. fluctuations).

Therefore, the model suggests, the onset of fluctuating group behaviour might provide an evolutionary pressure on the evolution of the size in human working memory.

4.9 Additional material

In this section, I describe additional aspects of the dynamics of the model studied. A brief introduction explains the significance of the four figures that follow.

Memory dynamics.

In section 4.4, the derivation of the exact evolution equation in the limit of infinite batches ($r \to \infty$) is presented. The derivation is based on the self-consistent assumption that '$0.5-\mu_i(t)$'[1] changes sign when $\phi_i(t)$ changes the direction of evolution from ascending to descending as well as vice versa. In this section, I show that this assumption can be observed in simulation data for finite batches ($r = 10$). Figure 4.22 shows how the average memory content, $\mu(t) = \frac{1}{N}\sum_{i=1}^{N}\mu_i(t)$, oscillates around 0.5 as a specific population evolve. In particular, the black dots represent the evolution of $\phi(t) = \frac{1}{N}\sum_{i=1}^{N}\phi_i(t)$ and the blue dots are a good indication of the corresponding dynamics of the average 'proportion of B players observed' in the group.

Regularity near the bifurcation.

The derivation of the Hopf bifurcation, say at $\epsilon = \epsilon_c$, tells us that prior to the onset of oscillations, say for $\epsilon < \epsilon_c$, the system dynamics converges to a stable state. However, the simulation data presented in this chapter consistently shows (cf. figure 4.10) that the period T, which equals $2m$ at the bifurcation, increases past $2m$ prior to the bifurcation. Here, I briefly examine, for $\epsilon \to \epsilon_c$, whether the dynamics displays random fluctuations or regular periodic oscillations. In particular, the presence of micro periodic oscillations prior to the Hopf bifurcation would suggest that the group behaviour is coordinated. In fact, figure 4.23 suggests that regular periodic oscillations are present for $\epsilon\rho$ values prior to the onset of the Hopf-bifurcation. This result suggests that group coordination emerges gradually.

Period calculated using spectral analysis.

In section 4.5 I discuss the significance of figure 4.10. The figure reports the values of the period calculated through a specific algorithm that works by detecting

[1] The variable $\mu_i(t)$ represents the proportion of B players observed by agent i.

maxima and minima of trajectories. Here, the same simulation data is analysed through spectral analysis. The results from the spectral analysis confirm the value of the period before the oscillations break down and further identify the main frequency of oscillation in the entire $\epsilon\rho m$ range displayed in figure 4.10. The spectral analysis is performed using the function *spectrum()* from the package *spectral* in R, in particular using the fast Fourier transform algorithm (FFT). The period is obtained by selecting the frequency with the highest spectral density.

Fluctuations size assessed via standard deviation.

Figure 4.10 reports the value of the amplitude of the oscillations of $\phi(t)$. The amplitude is calculated through a specific algorithm that works by detecting maxima and minima of trajectories. As a result, the precision of this algorithm can be affected by demographic noise. However, the use of the standard deviation to assess the size of the fluctuations is more robust to demographic noise, albeit less informative. Here, in figure 4.25, I use the standard deviation rather than the amplitude to describe the size of the fluctuations of $\phi(t)$. A selection of populations already analysed in this chapter is assessed via standard deviation. The figure confirms the scaling properties of the system. In particular, it confirms that the case $\epsilon = 1$ (pure strategies) has slightly different collective dynamics compared to the other cases studied.

Figures

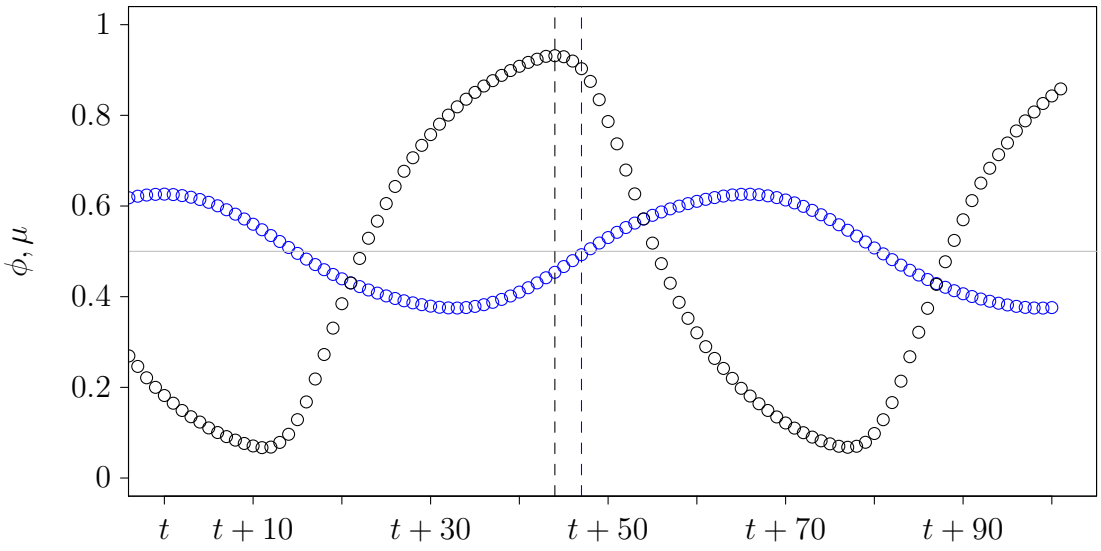

Figure 4.22: A detailed illustration of the collective strategy (black dots) and collective memory (blue dots) for $m = 50$, $\epsilon = 0.1005$, $\rho = 1$ and $r = 10$ over a time of 100 time units. $T \approx 66$ time units and $A \approx 0.43$. The dynamic of a fictional memory $\mu(t)$ is obtained as a moving average of $\phi(t)$ according to expression 4.4. One can note that in this simulation data, due to the granular nature of the memory ($r << \infty$) and stochastic effects, $\mu(t) - 0.5$ changes sign with a small delay with respect to the time in which the collective trajectory reaches a maximum. Here the delay is quantified to be between 3 and 4 time units. The data is plotted at integer units of time.

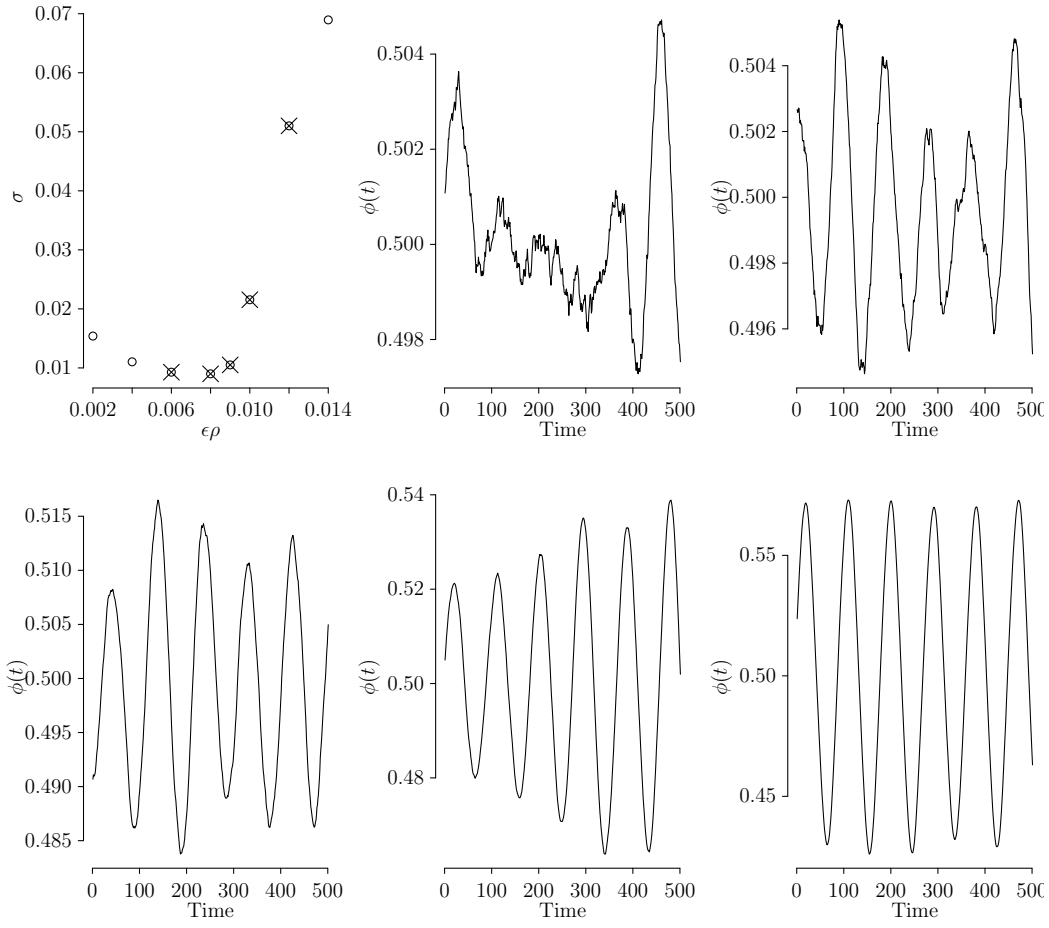

Figure 4.23: An illustration of the oscillations at/near the Hopf-bifurcation. For each $\epsilon\rho$ value denoted with a cross I have plotted the corresponding trajectory for $\phi(t)$ generated after 25000 time units. Regularity arises at the bifurcation where the period is already well-defined and the collective trajectory generates micro-oscillations. The system is simulated for $m = 50$, $r = 10$, $\epsilon = 0.01$ fixed and ρ variable.

.

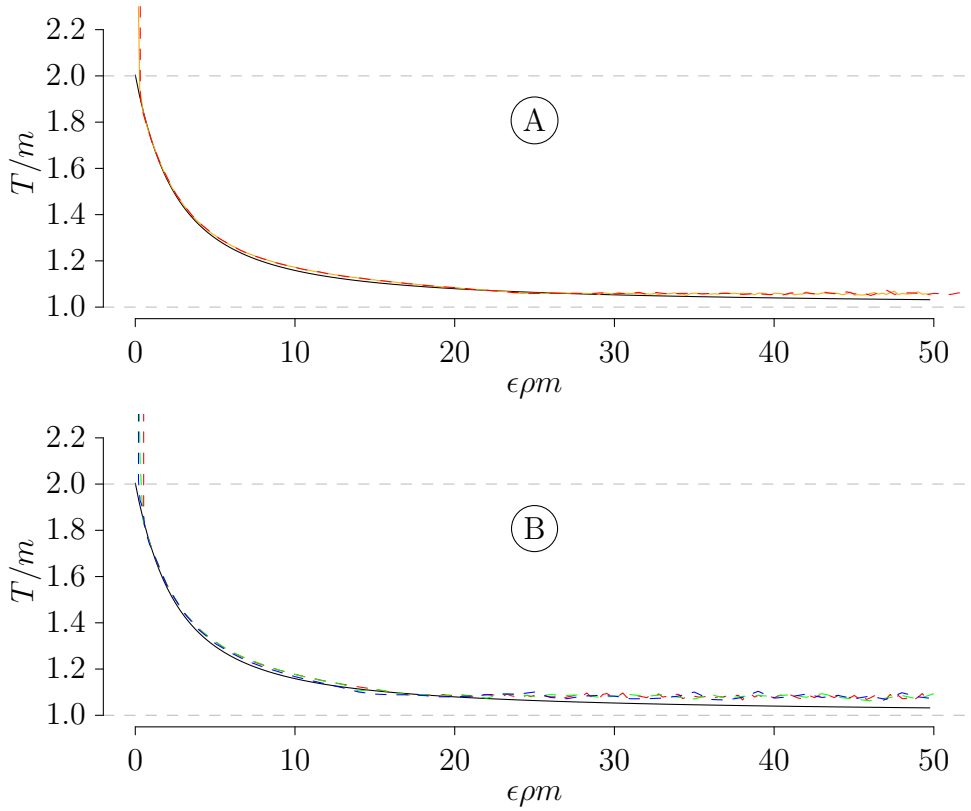

Figure 4.24: (A) The period T for $r \times m$ equals 5×100 (dashed red line) and 10×50 (orange line) with $\rho = 1$ and ϵ variable. The black line refers to the case $r \to \infty$. (B) The period T for $r \times m$ equals 5×50 for: $\rho = 1$ and ϵ variable (dashed red line); ρ variable and $\epsilon = 0.01$ fixed (dashed green line); ρ variable and $\epsilon = 1$ (dashed blue line). For $\epsilon\rho m > 30$, the period T has average value: (A) 52.90 with s.d. 0.168 (orange line) (B) 54.04 with s.d 0.338 (red dashed line), 54.042 with s.d. 0.419 (green dashed line), 54.07 with s.d. 0.392 (blue dashed line).

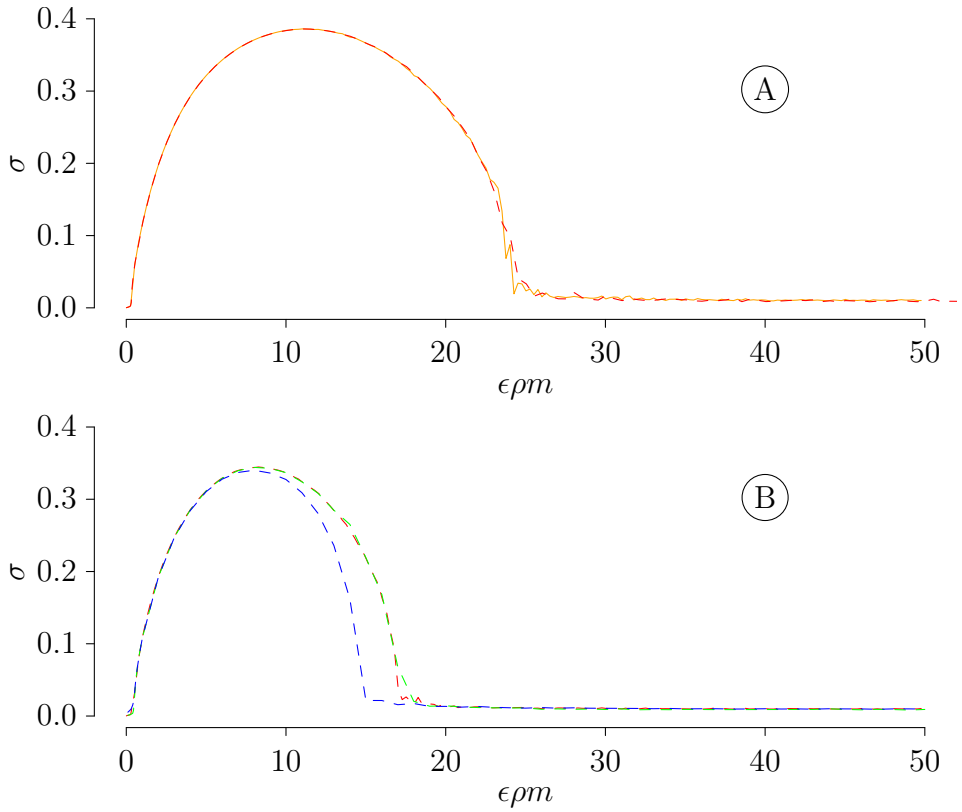

Figure 4.25: (A) S.d. of $\phi(t)$ for $r \times m$ equals 5×100 (dashed red) and 10×50 (orange) with $\rho = 1$ and ϵ variable. (B) S.d. of $\phi(t)$ for $r \times m$ equals 5×50 for: $\rho = 1$ and ϵ variable (red dashed); ρ variable and $\epsilon = 0.01$ (dashed green); ρ variable and $\epsilon = 1$ (dashed blue). In each of the 5 cases, the s.d. for $\epsilon\rho m > 35$ is constant or has a microscopic trend which I have not detected. For all five case, for $\epsilon\rho m > 35$, σ has mean between 0.010 and 0.009 (with standard deviation around the mean between 0.0002 and 0.0004). A least squares linear regression of the 5 different lines shows that each line has a gradient of value $\approx 10^{-5}$ where both negative and positive values have been obtained.

Chapter 5

Further features

5.1 Memory

We have seen that the onset of oscillations is regulated by the extent to which agents develop similar backward-looking expectations based on the information collected. Given this result, it is natural to ask if variations on how memory is modelled and information is used impact on the emergence of collective phenomena. In this section I investigate three interesting scenarios: heterogeneous, fictitious and mutant memories.

5.1.1 Heterogeneous memories

The oscillations are a collective phenomenon that results from a degree of coordination between agents. This coordination depends on the level of homogeneity of the information owned by different individuals. In the model, I have assumed that individuals have identical memory capacity in terms of the observation rate and duration. Within an evolutionary setting, however, it is reasonable to admit differences in the phenotypical expression of individual memory capacity. Therefore, I am interested to know to what extent the collective behaviour so far described is maintained when individuals exhibit heterogeneous memory capacity. Note that in the El Farol bar problem and the minority game, heterogeneous strategies are studied to find the combination of strategies that reduces fluctuations. Similarly, Burridge *et al.* (2017) are interested in the selection dynamics, between competing agents having different memory durations, that drive the system to a stable equilibrium. Whereas these studies remain interesting, here I want to analyse heterogeneous memory to test the robustness of the oscillation, not the convergence to optimal group coordination nor the ability of the group to withstand invasions from an intruder with a different memory capacity. Here, I have simulated the evolution of populations with memory duration varied across individuals, while

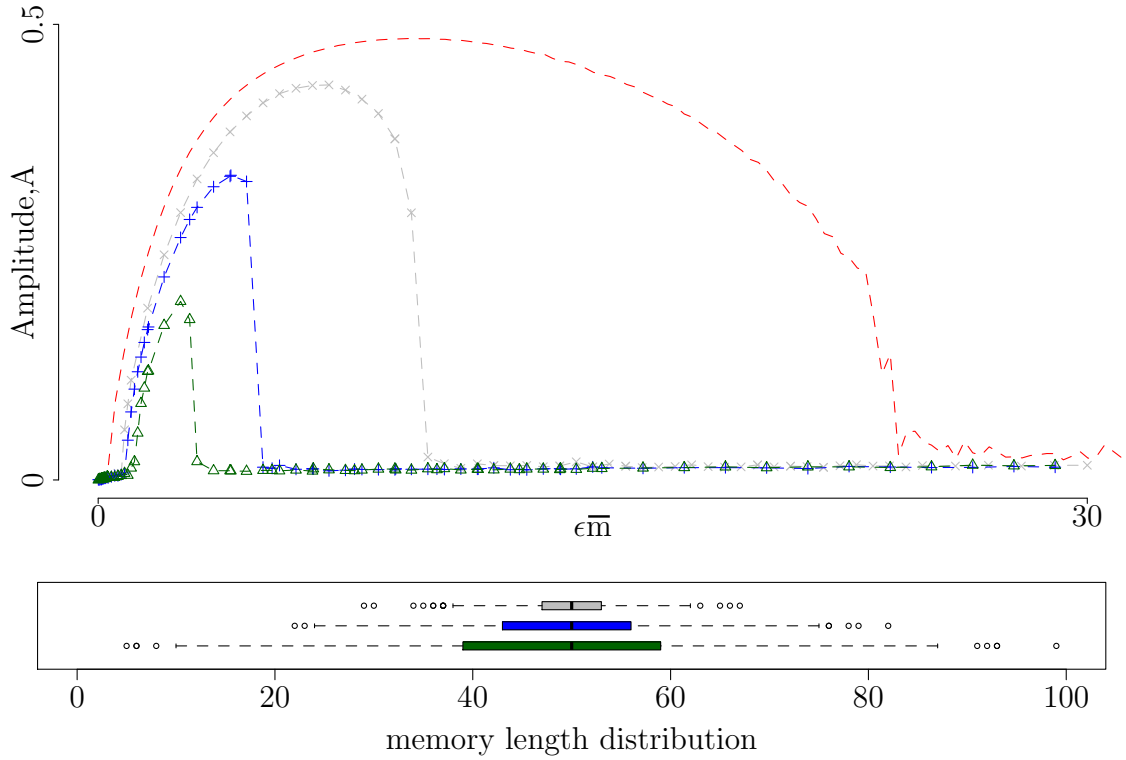

Figure 5.1: The Cartesian quadrant shows the amplitude of the oscillations for different realizations obtained for population of mean memory duration 50 and rate of sampling 10. The red dashed line is for a homogeneous population, while the other dashed lines are for heterogeneous populations for which memory duration is distributed around the mean $m = 50$. More specifically, the whiskers-plot show the distribution of individual memories duration.

the observation rate is fixed. In figure 5.1 I show the results. This figure compares the amplitude of the oscillations for homogeneous individuals as previously shown in figure 4.10 and for populations with memory duration normally distributed across individuals. The chosen normal distributions $N(mean, var)$ for the duration are the following: $N(50, 5)$, $N(50, 10)$ and $N(50, 15)$. The specific set of durations is generated only once according to each of the three distribution by creating a set of 1000 draws from a normal distribution (note that in computational models population size is 1000). Each draw is rounded to the nearest integer number. For each of the three phenotypical variations of the memory, a population is simulated for several different values of the parameter ϵ varying in $[0, 1]$ with $r = 10$ and $\rho = 1$ fixed. For each realization, the amplitude of the collective trajectory is plotted with a geometric symbol, either a cross (variance 5), a rotation of the same cross (variance 10) and a triangle (variance 15). The rectangular frame below the main graph contains box-plots showing the distribution of the three sets of memory durations generated following a normal distribution.

Figure 5.2 shows in detail a section of figure 5.1, specifically indicating that the Hopf-bifurcation, point at which oscillations arise, moves to the right as memories become phenotypically more diverse. This result is consistent with the previous result stating that the Hopf-bifurcation point moves to the right for less detailed memories (i.e. for lower r, cf. equation 4.70).

Figure 5.3 suggests that the system obeys the scaling properties according to 'new' fundamental quantities $\overline{m}r$ and $\overline{m}\epsilon$, with $\overline{m}$ denoting the mean memory duration in a given population. In particular, it shows different realizations for populations with $\overline{m}r = 50 \times 10$ (grey dash line with crosses) and $\overline{m}r = 25 \times 20$ (black circles). The latter set of memories is obtained by halving the duration of the previous set of memories. The durations' distribution of the previous set is approximately $N(50, 5)$ and thus for the new set is approximately $N(25, 5/4)$. As a result, one can verify that the variance of the set made of the sizes $(m \times r)$ of all memories in a given population is equal to 500 for both sets of memories. This suggests that in this scenario there is an additional fundamental variable, namely the variance of the set of the sizes of all memories in a given population. The parameter $\rho = 1$ is fixed for the investigation of heterogeneous memory duration. Crosses and circles are plotted corresponding to the ϵ values of the population simulated.

To conclude, in this section I introduced a specific form of variability of the capacity of individuals to remember information. The brief analysis shows that periodic fluctuations can arise even under heterogeneous memory capacity, providing evidence that the phenomenology studied in the model withstand biological and evolutionary scrutiny.

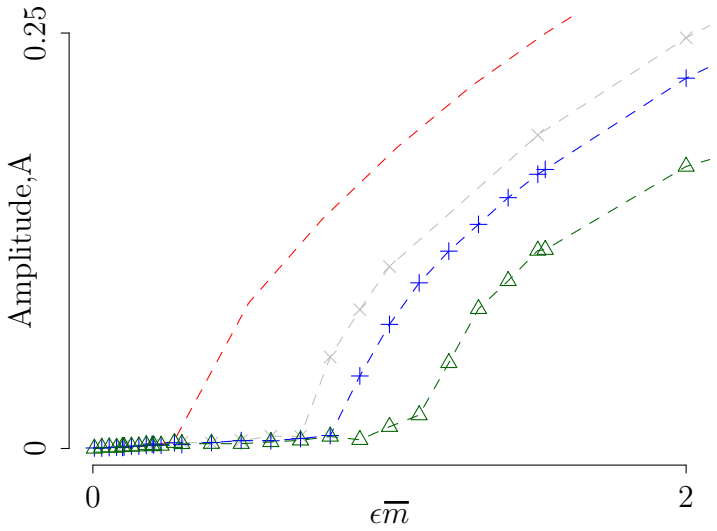

Figure 5.2: A close look at the bifurcation for heterogeneous memory capacity.

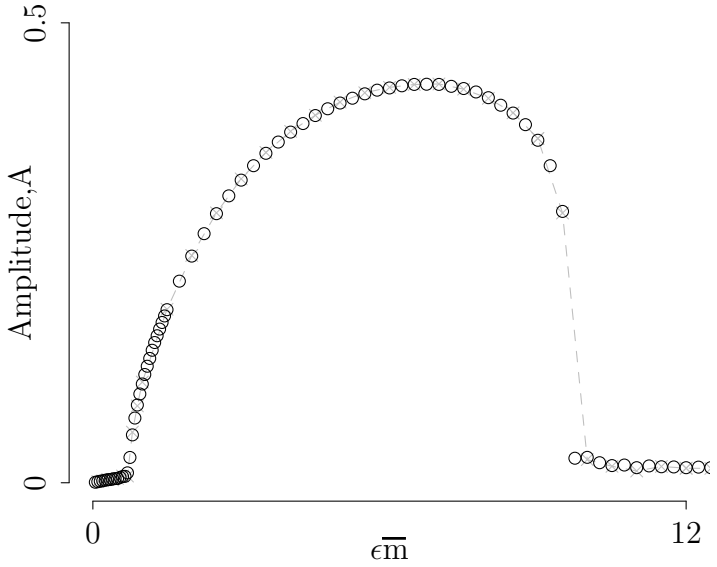

Figure 5.3: The amplitude as a function of $\overline{m}\epsilon$ scales according to the fundamental quantities $\overline{m}r$ and $\overline{m}\epsilon$ under the assumption of heterogeneous memory duration (m) normally distributed. Here, $\overline{m} = 50$ and $r = 10$ (grey dashed line with crosses) and $\overline{m} = 25$ and $r = 20$ (black circles).

5.1.2 Fictitious memory

There is a difference between the theoretic mean-field formulation of equations 4.41-4.42 and the operative individual memories coded in simulations.

On one hand, the operative approach uses the realistic assumption that agents have individual memories, which are personal recordings of the history of the system. These recordings are updated over time. In this updating process, individual memories are subject to biased sampling. For example, an individual may observe an unusually high number of A players which does not reflect the state of the population. These observations could sway the average memory content of that individual until, after several more observations, the earlier observations are eventually replaced in the memory of the individual.

On the other hand, the theoretical mean-field approach of equations 4.41-4.42 entails the fact that a biased sample does not persist in individual memory. In other words, although the content of memories is subject to variance, bias is not carried on. The theoretical formulation corresponds more practically to agents that make decisions based on memories sampled anew every time that a decision is made. Nonetheless, the analytical estimate of the Hopf-bifurcation still predict correctly the 'experimental' Hopf-bifurcation observed in the simulation data for populations with individual memories (see figure 4.13). This suggests that the theoretical mean-field approximation of individual memories is valid in a neighbourhood of the bifurcation observed for simulation data. However, this approximation may not hold globally for all $\epsilon\rho > 0$. In particular, such a theoretical mean-field approximation may not characterise the decay of oscillations as well as it does their emergence.

To investigate this last point I simulated a population of individuals that do not accumulate biased information in their memory. On the contrary, these individuals sample the history of the strategies used in the population anew each time they perform a strategy update. In the code, this feature is obtained by maintaining a large set of recorded strings of the history of the population, which are updated as if belonging to fictitious agents. When an agent needs to use their memory, he resorts to using the memory of a randomly picked fictitious agent.

The simulation results reported in figure 5.4 compare two sets of simulations obtained for fictitious memories (black crosses joined by a dashed-dotted line and black circles) and two sets of simulation reproduced from chapter 4 (red and orange lines). These simulations reveal that the size of the oscillations is invariant near the bifurcation and is affected (although only slightly) when ϵm is 'large'. This indicates a limitation on the use of the above mentioned theoretical approximation of memory to fully characterise the behaviour of a population with agents holding individual memory. Such an approximation will not characterise the decay of oscillations as well as it does their emergence.

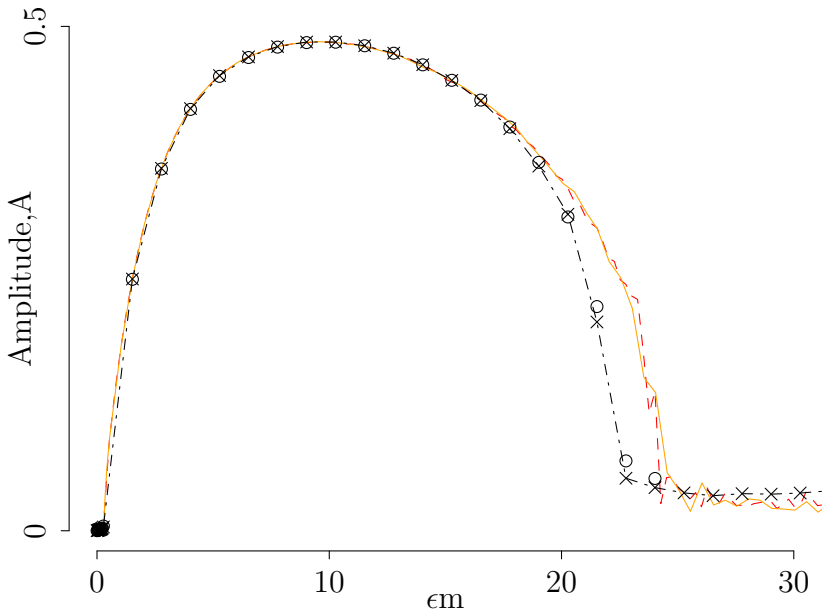

Figure 5.4: The black crosses joined by a dotted-dashed line and the black circles show the amplitude of the oscillations of $\phi(t)$ for populations having parameters, respectively, $mr = 50 \times 10$ and 25×20 with $\rho = 1$ and ϵ varying. In these populations, individuals employ 'collective' memories, i.e at each strategy update, individuals employ at random one memory from a set of 1000 memories each belonging to a fictitious agent. This simulation data is compared to data for processes with $rm = 500$ and $\rho = 1$ (red and orange lines), which are results reproduced from chapter 4.

Overall, this brief study tells us that fictitious memory, rather than individual memory, corresponds more closely to the theoretical mean-field analysis presented in this thesis. This suggests that further analytical mean-field quantitative explorations of the decay of the oscillations may more closely describe a scenario of agents with fictitious memory.

However, having decoupled agents from their memories allows focusing on the evolution of memories as 'separate entities'. Here, in particular, I suggest a description of the evolution of individual memories as a binary De Bruijn graph. This method is introduced by Challet & Marsili (2000) in their study of the minority game, enabling them to study in details the dynamics of the histories in the game.

In the model studied in chapters 4 and 5, individual memory can be encoded

by a binary sequence of length mr which consists of mr ordered elements where each element belongs to the set $\{0, 1\}$. After an observation, the new individual memory is obtained by adding 0 or 1 to the right and erasing the oldest observation on the left of the memory sequence. Thus, for a given memory, there are two possible updates. This updating rule defines the De Bruijn graph of order mr (Challet & Marsili, 2000).

The two main differences between this model and the minority game consist of the number of histories and strategies that define the game. The minority game has one history that encodes the last m minority groups and 2^m strategies available, whereas the model studied in this thesis has N histories, one for each agent, encoding the last mr observations, and one strategy. Further research is needed to explore this method and find out whether one can establish the transition probabilities between individual histories in a De Bruijn graph. It may be convenient to first address the scenario involving only pure strategies (i..e. $\epsilon = 1$).

5.1.3 Modelling an invader

As part of the investigation on the role of memory duration into the emergence of fluctuations, Burridge *et al.* (2017) model the competition between long and short memories. Upon setting up selection dynamics operating between two specific memory durations, they show how a population evolves close to a Hopf-bifurcation point by maintaining a balance between the proportion of agents with short and long memories thus preventing the emergence of oscillations. Following on from the idea that under evolutionary pressure a population self-organises, in this section I model a rare mutant that competes with the host population. In particular, I take a preliminary view at the benefits and costs that an invader, as a rare mutant, may initially face.

Here, the aspect of the individual that is considered subject to variation and evolution is the memory duration. However, I do not model the variation process but directly introduce an invader. In practice, I assume that memory duration is fixed for a large population of 1000 individuals and introduce 20 individuals having a different memory duration. The host population has memory duration $m = 50$, while the mutant population is simulated for a broad range of durations, namely $m_{inv} \in \{1, 2, ..., 99\}$. Note that the interaction between $m = 50$ and each mutant m_{inv} is simulated and analysed separately. The remaining parameters are fixed for both the host and mutant populations as follows: $\epsilon = 0.03$, $r = 10$ and $\rho = 1$. Figure 5.5 presents the results. Each graph depicts the evolution of the average strategy $\phi(t)$ of the host population averaged over 1000 agents (black line) and the strategy of the invading population averaged over 20 agents (red line). Figure 5.5 displays the simulations outcome for invaders having memory duration

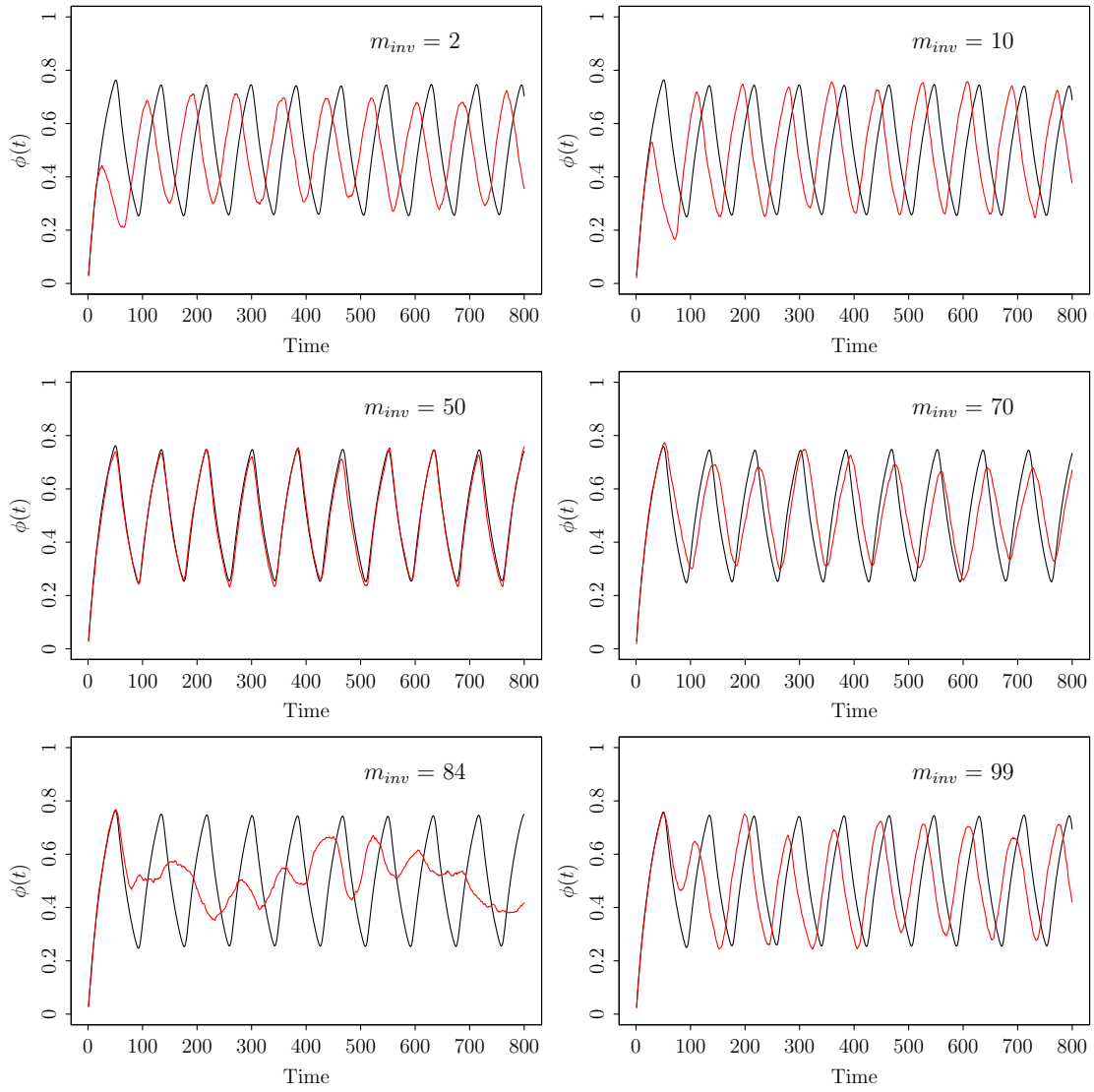

Figure 5.5: Strategies evolution over time of the host population averaged over 1000 agents (black line) and strategies evolution of the invading population averaged over 20 agents (red line). The host population has parameters $m = 50$, $r = 10$, $\epsilon = 0.03$ and $\rho = 10$. The invading population has the same parameters apart from the memory duration m. Here I display the simulation results for m in the set $\{2, 10, 50, 70, 84, 99\}$.

m_{inv} in the set $\{2, 10, 50, 70, 84, 99\}$. Dependent on the host population having $m = 50$, the mutant agents are entrained to the host population when $m_{inv} = 50$ and approximately reach anti-phase synchronisation for $m_{inv} = 2$. Having

simulated the evolution of the strategy for invading agents, it is sensible to ask which mutant memory provides the greatest advantage in terms of evolutionary fitness. This question can be answered by introducing payoff collection dynamics. I choose to model the pay-off yielding interactions using payoff matrix 2.22 with $V = 1$ and $C = 2$, thus effectively maintaining the ESS $\phi(t) = 0.5$ (see section 2.5). In figure 5.6, I plot the expected payoff for an invader as a function of its

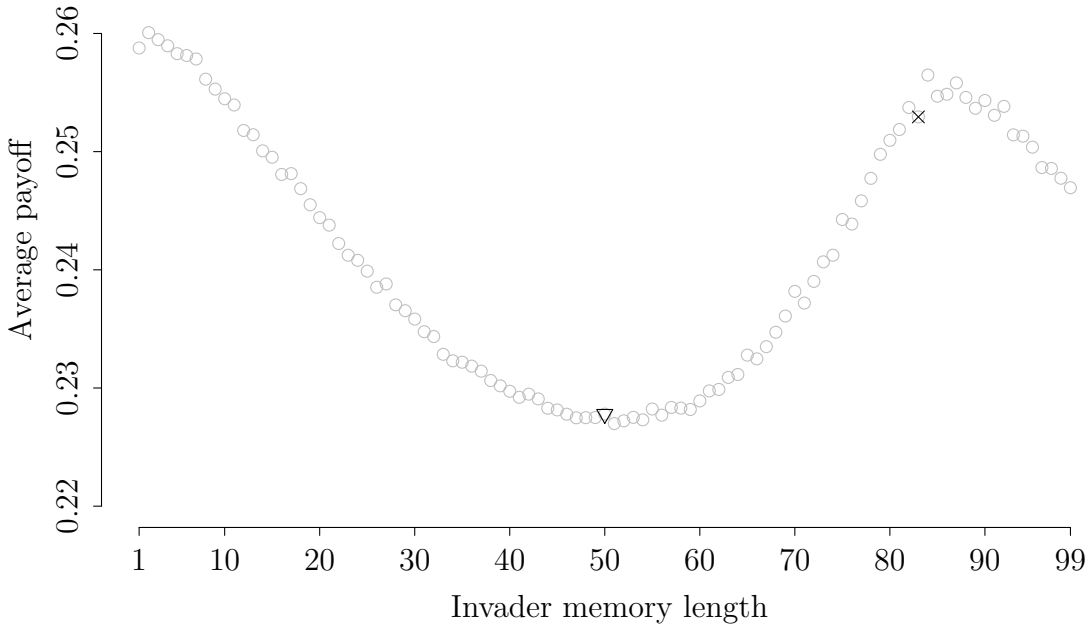

Figure 5.6: Average payoff for mutant agents (y-axis) as a function of the mutant's memory m_{inv} (x-axis) A triangle is drawn at $m_{inv} = 50$, which matches the memory duration of the hosting population. A black cross is at $m_{inv} = 83$, matching the period of the hosting population, calculated directly from the simulations, measuring 83.25.

memory duration m_{inv}. The expected payoff is calculated using equation 2.24 where u is the average strategy of an invader (the red line in figure 5.5) and v the probability to meet a Hawk in the host population (the black line in figure 5.5). I assume that individuals collect a payoff once per unit of time and that the expected payoff is obtained by averaging the payoff over time for 1200 time units. In this manner, the payoff is only an expectations. Nonetheless, figure 5.6 shows quite clearly that a short memory is an advantage for an invader (global maximum payoff). The payoff also shows a local maximum just after the value of the period generated by the oscillation in strategy of the host population. In fact, the period

computed from the simulations for the host population is 83.25 ± 0.1. Further to this, the synchronisation of the strategy with the host population gives the worst outcome and deviating from the memory duration of the host population always provides an advantage.

This brief analysis allows to speculate further questions for future research: (i) will the invader eventually take over, or reach coexistence? will coexistence depend on the specific payoff matrix? (ii) what might happen if lots of agents were allowed to evolve their own memories simultaneously? (iii) What might be the trade-offs in the co-evolution of memory duration (m) and observation rate (r) while incorporating a cost specific to memory size (mr)? To explore these open questions, one could set up computer simulations similarly to Arthur (1994) in the El Farol bar problem and to the computer tournaments for the Iterated Prisoner's Dilemma (Axerold & Dion, 1988; Axerold R, 1981), or alternatively employing a genetic algorithm (Guttal & Couzin, 2010; Zafeiris & Vicsek, 2013).

5.2 Synchronous processes

In chapter 4, the model dynamics for r finite are simulated using an agent-based computational model. Although the rates of observation and strategy update are identical for all players, agents are picked at different times to perform the two processes. I refer to this scenario as having asynchronous processes. On the contrary, in this section, I ask what happens if agents perform each event simultaneously. I refer to the last scenario as having 'synchronous processes'. In particular, an event can be thought of as a collective event. For example, all agents update their strategy simultaneously. Figure 5.8 presents a visualisation of the workflow of the algorithm for synchronous processes for $\rho = 1$. Without loss of generality, the algorithm can be described as follows. Given an observation rate $r = c\rho$ and a strategy update rate $\rho > 1$, with $c > 1$, agents perform an observation event c consecutive times, each time simultaneously. Next, they simultaneously update their strategy once. Finally, time is estimated as follows: a time unit corresponds to $r + \rho$ (collective) events.

In figure 5.7, I compare the simulation results from chapter 4 obtained using a modified Gillespie algorithm for asynchronous processes to new simulation data obtained for populations in which observation and strategy update are synchronous. The figure shows a plot of the amplitude of the oscillations obtained for synchronous processes (grey crosses connected with a dashed line to help the eye), asynchronous process (red and orange lines) and for the theoretical prediction in the limiting case $r \to \infty$ (black line).

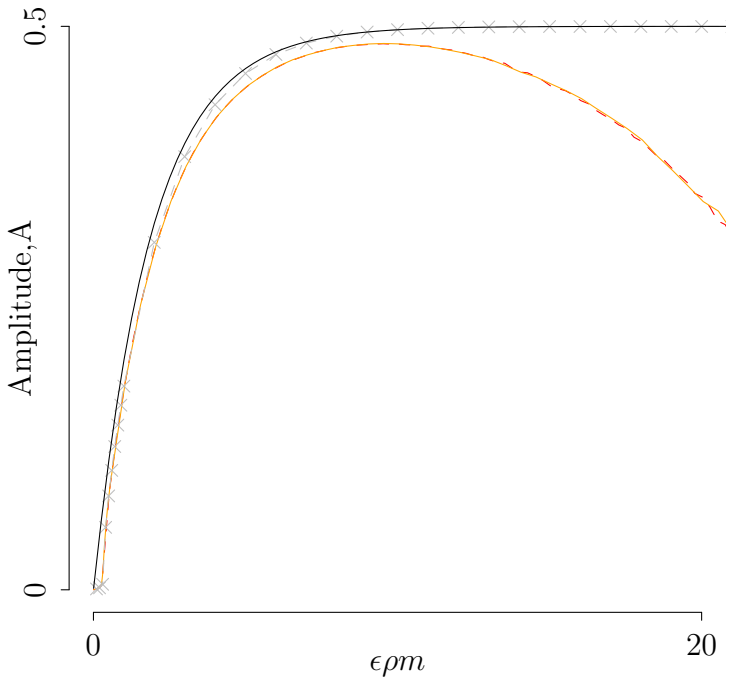

Figure 5.7: The grey crosses, joined by a dashed line, indicate the amplitude of the trajectory $\phi(t)$ obtained for a population of individuals that update their strategies simultaneously, for $mr = 50 \times 10$ with $\rho = 1$ and ϵ variable. This is compared to results reproduced from chapter 4 of populations simulated for asynchronous process with $mr = 50 \times 10$ and $mr = 25 \times 20$ (red and orange lines respectively) and for the limiting case of a deterministic system in which $r \to \infty$ (black line).

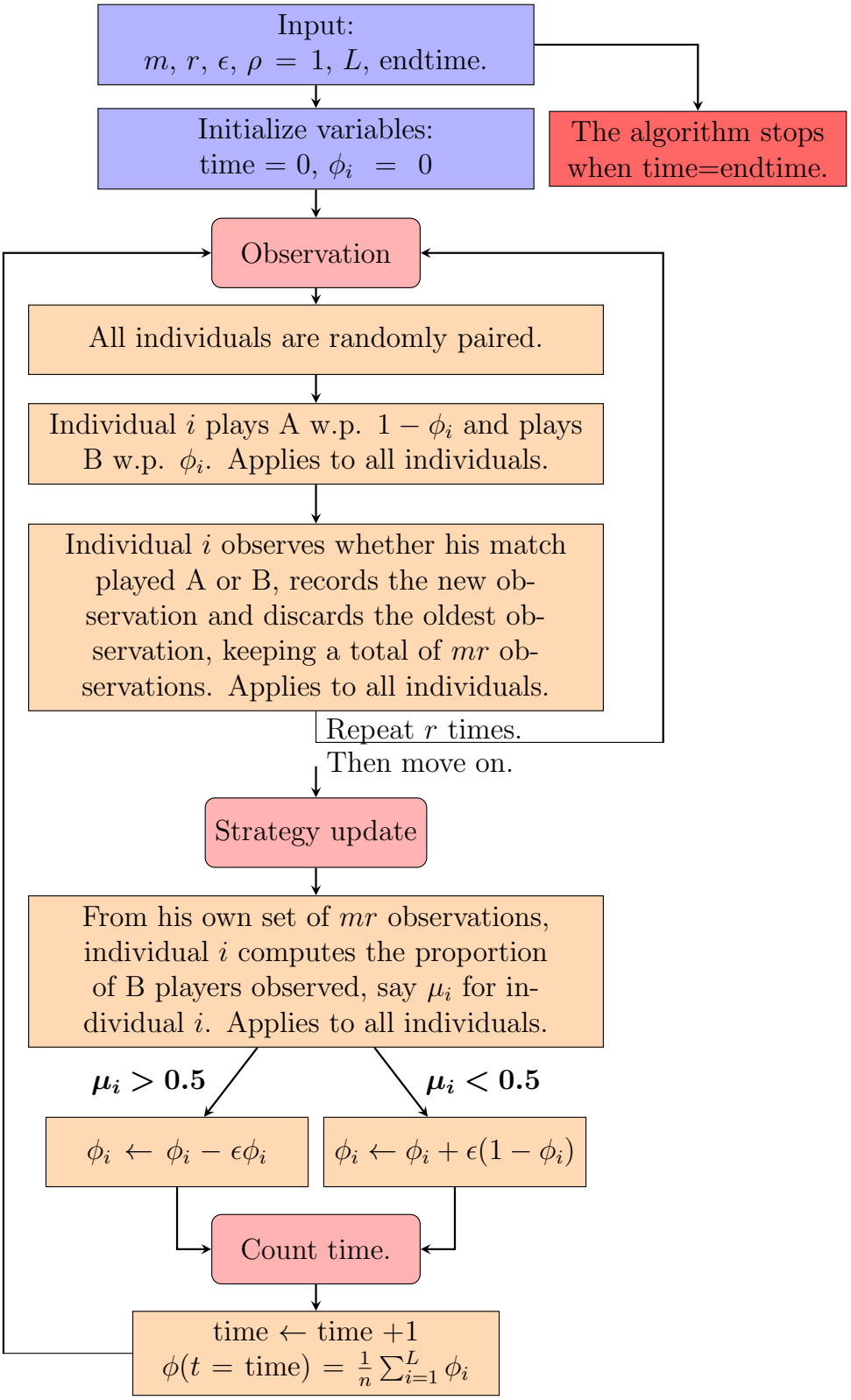

Figure 5.8: Diagram visualising the workflow of the algorithm for synchronous processes used in simulations. In this scenario, each step of the algorithm involving an individual action is performed by all agents simultaneously.

Figure 5.9 presents a closer look at the previous figure near the Hopf-bifurcation point. The two figures, taken together, show that synchronous processes preserve the Hopf-bifurcation (fig. 5.9) and enhance synchrony among agents in a large neighbourhood on the right of the Hopf-bifurcation (fig. 5.7).

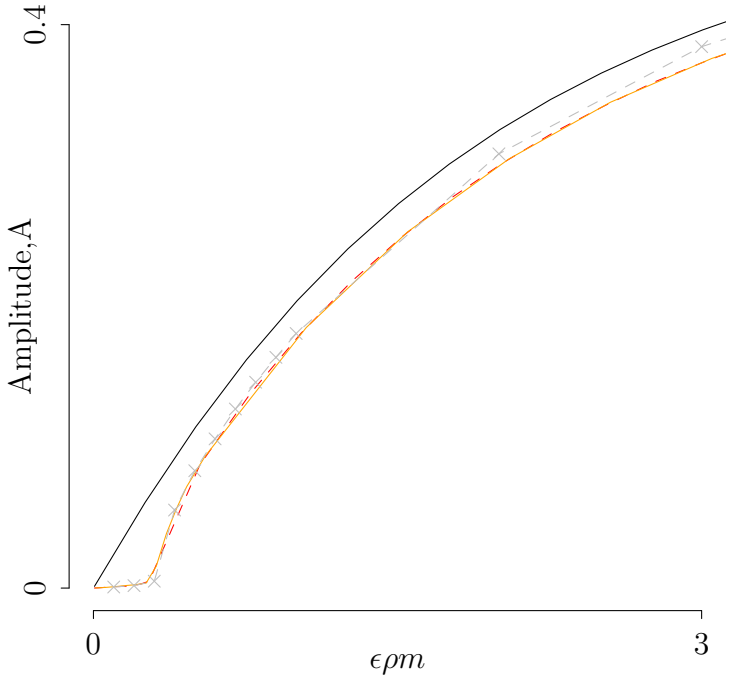

Figure 5.9: The Hopf-bifurcation does not depend on synchrony or lack thereof of the processes involved. The grey crosses, joined by a dashed line, indicate the amplitude of the oscillations of the trajectory $\phi(t)$ obtained for synchronous processes, for $mr = 50 \times 10$ with $\rho = 1$ and ϵ variable. The red and orange lines respectively refer to populations simulated for asynchronous processes with $mr = 50 \times 10$ and $mr = 25 \times 20$ and the black line refers to numerical results for the limiting case $r \to \infty$.

Lastly, figure 5.10 compares two sets of simulations for synchronous processes, respectively for $mr = 50 \times 10$ and $mr = 25 \times 20$, with ϵ variable and $\rho = 1$. This figure indicates that, for synchronous processes, oscillations break down with a sharp transition (in this case at $\epsilon\rho m \approx 30$). These results imply that asynchronous and synchronous processes provide different and gradual approximations to the dynamics studied for $r \to \infty$. Finally, these results point out that the dynamics near the Hopf-bifurcation are strongly determined by the statistical properties of the 'information flow' and not influenced by the synchrony of the processes, or lack thereof.

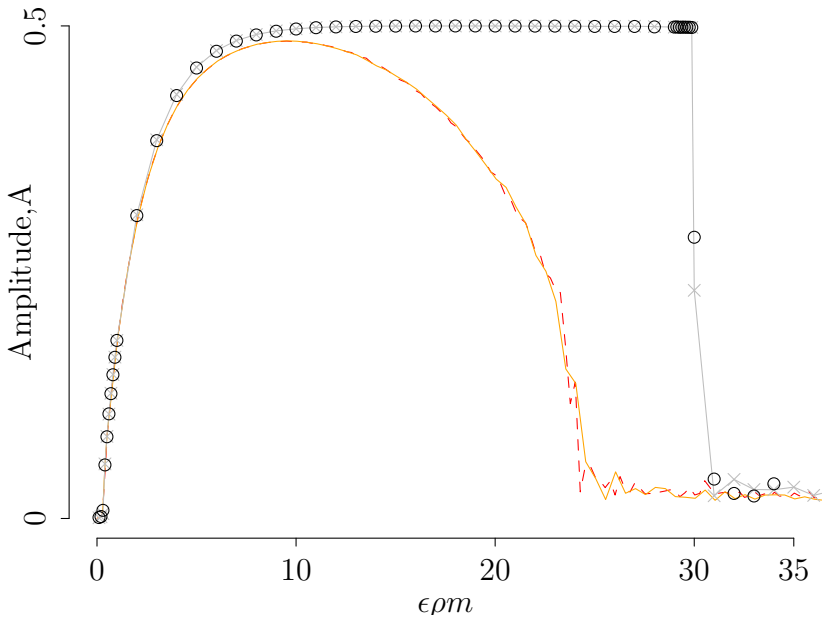

Figure 5.10: The grey crosses, joined by a dashed line, and the black circles show the amplitude of the oscillations of $\phi(t)$ of population simulated with synchronous processes for $mr = 50 \times 10$ with $\rho = 1$ and $mr = 25 \times 20$ with $\rho = 2$, with ϵ variable. These results obeys the scaling acording to $\epsilon\rho m$ and rm as predicted. The breakdown of oscillation happens for values of ϵ between 0.5975 and 0.6, i.e at $\epsilon\rho m \approx 30$.

5.2.1 An hysteresis region

Further investigating the features of the dynamics with synchronous processes uncovers the presence of a thin hysteresis region. The hysteresis region is uncovered by comparing the results from the last section with simulation data obtained using a novel approach. In all previous simulations, the dynamics for different values of ϵ are always tested by completely restarting the simulation for each value of ϵ. Indeed, this holds true for all system parameters, namely m, r, ρ and ϵ. In this section, however, I present the simulation results obtained for one single simulation with ϵ updated en-route while all other parameters are fixed. Specifically, this entails that, after $\phi(t)$ had reached equilibrium for a given ϵ value, the value of ϵ was increased. To verify that the system had reached equilibrium, the simulations were run for 20000 time units for each ϵ value. Since the memory duration, m, equals 50 and the period, T, tends to m as ϵ increases, the trajectory of $\phi(t)$ performed nearly 4000 periodic oscillations before incrementing the value

of the parameter ϵ.

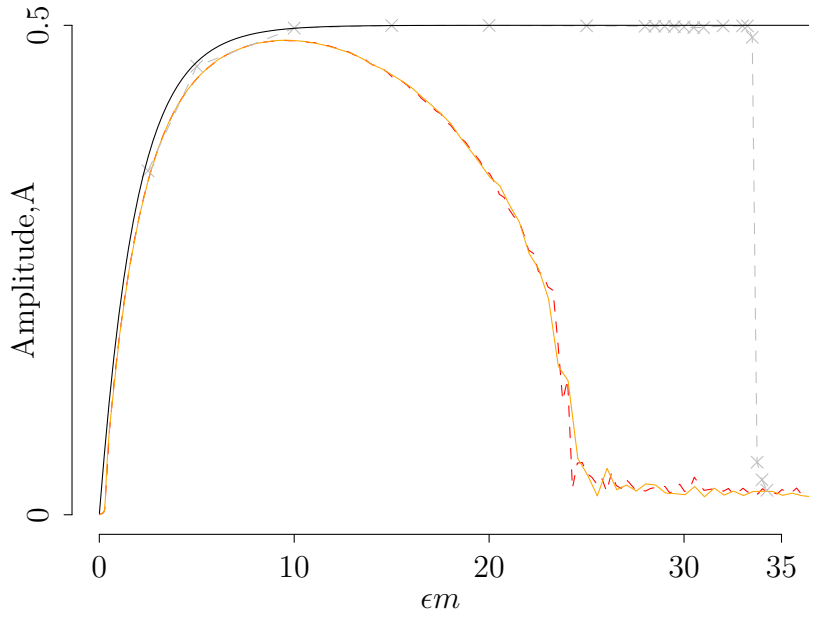

Figure 5.11: The grey crosses indicate the amplitude of $\phi(t)$ obtained with one simulation for $mr = 50 \times 10$ and $\rho = 1$ with ϵ progressively updated 'en-route'. Overall, in this figure, ϵ is consecutively increased following this list of values: 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.56, 0.57, 0.58, 0.59, 0.60, 0.61, 0.62, 0.64, 0.66, 0.665, 0.67, 0.675, 0.68, 0.685. The grey crosses showing the amplitude of the asymptotic trajectories for $m = 50$ and $r = 10$ are compared with the black line showing the numerical values for the limiting case $r \to \infty$ and $m = 50$ and the orange and red lines showing the amplitude for asynchronous populations for $mr = 50 \times 10$ and 100×5.

In figure 5.11 I present the simulation data. The grey crosses, joined with a dashed line, indicate the amplitude obtained from one single simulation for synchronous processes with ϵ initialised to $\epsilon = 0.05$ (first grey cross from the left). Next, the algorithm is updated to $\epsilon = 0.1$ (second grey cross from the left) for the next cross displayed, and so on. The grey crosses showing the amplitude of the asymptotic trajectories are compared to the black line showing the theoretical values for the limiting case $r \to \infty$ and the orange and red lines showing the amplitude for asynchronous populations. The latter are results reproduced from chapter 4. This simulation data suggests that the system displays a hysteresis region because it admits two different equilibria in the region $30 < \epsilon\rho m < 35$ depending on the route used to reach this region. One equilibrium has just been

shown, while the other equilibrium, consisting of damped fluctuations, is shown in figure 5.10. To summarise, I have shown that, for synchronous processes, there is a hysteresis region that separates coordinated oscillations from noisy fluctuations.

This result reminded me of a result by Matthews & Strogatz (1990). The authors show that the order parameter of a system of linearly coupled oscillators undergoes several transitions. In particular, the system admits a thin hysteresis region through which the order parameters leaves a locked state corresponding to oscillators rotating around the origin with constant frequency and amplitude. The specific parameter that drives the transition is the width of the initial distribution of the frequencies of the oscillators. The system thus transitions through a hysteresis as the individual oscillators become more diverse. The period then follows a route to chaos. This thin hysteresis region parallels the behaviour observed in my model which also leaves a stable equilibrium through a hysteresis region. In particular, the coordinated state of group oscillations for $\phi(t)$ mirrors the locked state displayed by the coupled oscillators.

5.3 Variance and noise in fluctuations

So far I have studied and extensively characterised a model of binary choice with memory. I showed that oscillations (of the probability that agents choose between the two options) are present in differing scenarios involving a range of ways in which individuals use information. Here, I introduce a population in which individuals have no memory and use no information. I use such a population as a benchmark to study the efficiency of the system.

In the canonical minority game (CMG), the literature focuses on describing the range of condition that minimize the fluctuations of the number of individuals choosing each of the two strategies over time. In order to determine the efficiency of a population, the CMG literature measures the variance of the fluctuations. As explained in chapter 3, a system is considered efficient when the variance is lower than the one produced from random choice (cfr. figure 3.2 and formula 3.4). In relation to this, there are several findings in the CMG, for example: (i) the system converges to the efficiency of random choice when $z = \frac{2^m}{N}$ tends to ∞ (cfr. figure 3.2) ; (ii) the variance is minimized as the process of switching between predictors is slowed down or the 'temperature' in the system increased (cf. Challet *et al.* (2013) pages 36,38,52). Similarly, in this section I compare the standard deviation of the fluctuations for the model I studied in chapter 4 to the same model modified with the introduction of random strategy-updating and no memory. In figure 5.12 I compare the standard deviation σ of the trajectory $\phi(t)$ for $mr = 50 \times 10$ in a population with individual memories (black line) and the standard deviation for a system with random updating and no memory (blue

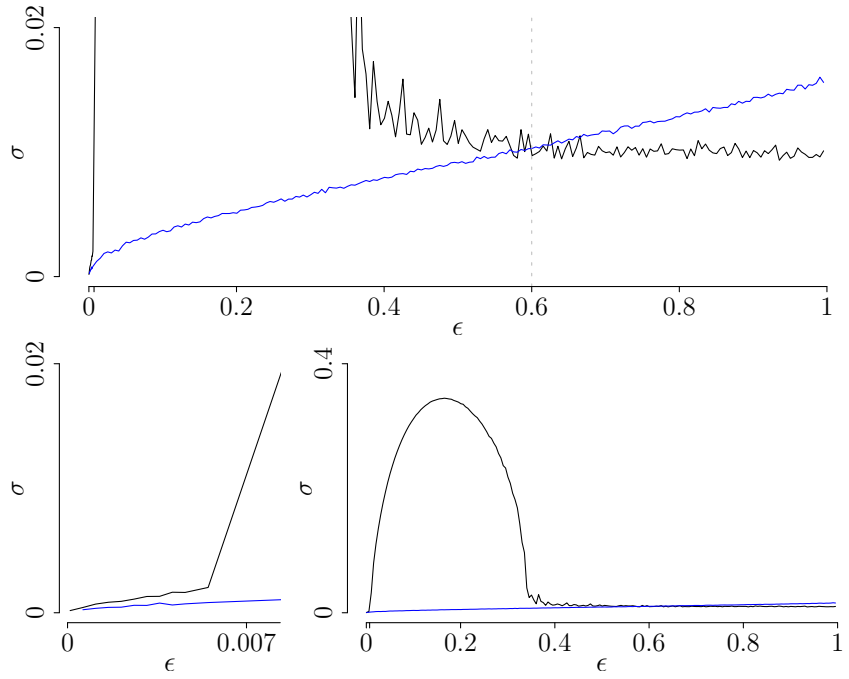

Figure 5.12: σ is the standard deviation of the average strategy $\phi(t)$ over time and a good measure of the size of the fluctuations of $\phi(t)$. In all three panels, I compare the value of σ for two scenarios: (i) a population with memory (black line) and (ii) one made of individuals without memory that update their strategy randomly (blue line). The top panel focuses on values such that the two scenarios are more closely comparable ($\sigma < 0.22$); the smallest panel shows that for $\epsilon \to 0$ random choice is more efficient; the last panel shows the full range of the data. For random updating, the data shows that $\sigma = 0.01559$ when $\epsilon = 0.9955$. The corresponding analytical prediction is $\sigma = 0.0158$ when $\epsilon = 1$. The population with memory has parameters $mr = 50 \times 10$ and $\rho = 1$.

line). It holds that $\sigma = 0$ if and only if $\phi(t)$ is constant over time. For random choice, when $\epsilon = 1$ one can calculate the standard deviation of $\phi(t)$ as follows.

First, for simplicity, I define $\Phi_i(t) = \phi_i(t) - 0.5$. In this manner, $\Phi(t) := \frac{1}{N}\sum_{i=1}^{N}\Phi_i(t) = \phi(t) - 0.5$ and $\Phi(t)$ exhibits fluctuations around 0. Since $\phi_i(t)$ takes random values in the set $\{0, 1\}$, $\Phi_i(t)$ takes random values in the set $\{-0.5, 0.5\}$. In addition, $\mathrm{Var}[\Phi(t)] = \mathrm{Var}[\phi(t)]$.
Therefore, by defining $\sigma^2 := \mathrm{Var}[\phi(t)]$ one obtains:

$$\sigma^2 = \mathrm{Var}\left[\Phi(t)\right] = \mathrm{E}\left[\Phi(t)^2\right] = \mathrm{E}\left[\left(\frac{1}{N}\sum_{i=1}^{N}\Phi_i(t)\right)\left(\frac{1}{N}\sum_{i=1}^{N}\Phi_i(t)\right)\right] \tag{5.1}$$

Since agent randomly updates their strategies, the strategies $\Phi_i(t)$ and $\Phi_j(t)$ are independent for $i \neq j$. Therefore $\mathrm{E}[\Phi_i(t)\Phi_j(t)] = 0$. As a result,

$$\sigma^2 = \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{E}[\Phi_i(t)^2] = \frac{N}{N^2}\mathrm{E}[\Phi_i(t)^2] \tag{5.2}$$

Finally, for a population of 1000 individuals, one obtains

$$\sigma = \sqrt{\frac{0.25}{1000}} = 0.0158. \tag{5.3}$$

The simulation data reported in figure 5.12 confirms this prediction with $\sigma = 0.01559$ for $\epsilon = 0.9955$. Figure 5.12 shows that, for ϵ larger than approximately 0.6, individuals are more efficient using memory rather than choosing randomly. This means that individuals process the information meaningfully. This behaviour resembles an anti-phase coordination effect in which individuals effectively learn to take turns. In section 4.5.1 I suggest that my model is equivalent to a limiting case of the thermal minority game (TMG) (Cavagna *et al.*, 1999), specifically the role of the rate of sampling in my model is equivalent to the role of selection intensity in the TMG. One notable result in the TMG states that the game admits an optimal intensity of selection (Γ^*) such that it minimises fluctuations in the system below the level corresponding to random fluctuations. Compared to the optimal selection intensity, a lower intensity of selection ($\Gamma \to 0$) would drive the system towards random fluctuations, while a greater intensity of selection ($\Gamma \to \infty$) would drive the system towards large cyclic fluctuations. My model predicts the same result. Here, the parameters' space $\epsilon > 0.6$, with $m = 50$, $r = 10$ and $\rho = 1$ is a specific scenario in which the system outperforms random choice. However, I also suggest that (with all parameters constant except r) when $r \to 0$ the system converges to a random choice scenario, whereas when $r \to \infty$ the system converges to coordinated oscillations. To conclude, note that the standard deviation of the average strategy in the group grows linearly for approximately $\epsilon > 0.1$ and non-linearly otherwise. A preliminary analysis of the simulation data suggests that this is due to the changing 'interaction' that agents have with the boundaries (i.e. the boundaries 0 and 1 imposed on the evolution of the individual strategies ϕ_i) as ϵ increases.

5.4 Group synchrony

In chapter 4 and 5 I state that synchrony among agents is necessary for the system to exhibit coordinated oscillations. This is not surprising given that $\phi(t) = \sum_{i=1}^{1000}\phi_i(t)$. However, we do not yet have a measure of such synchrony.

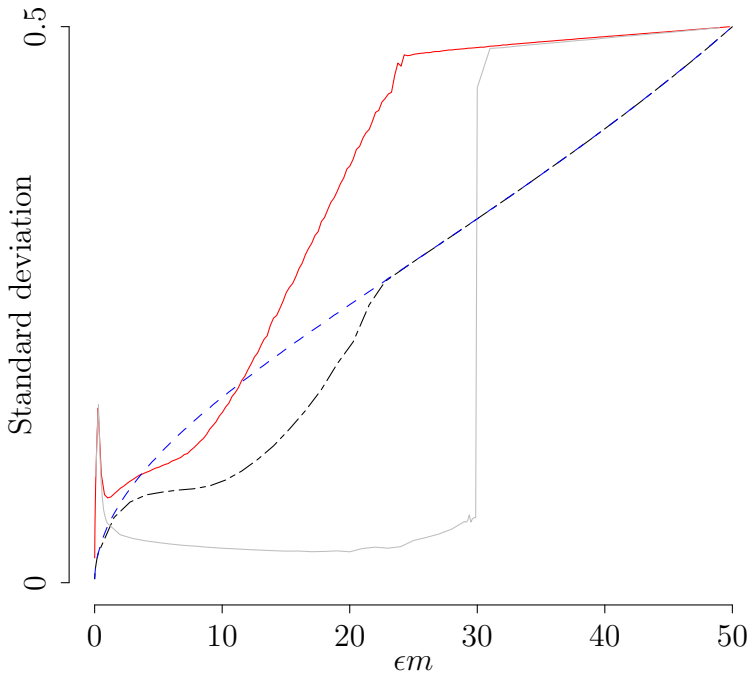

Figure 5.13: The graph shows the group standard deviation for: $mr = 50 \times 10$, $\rho = 1$ and ϵ variable for asynchronous processes (red line) and for synchronous processes (grey line); $mr = 50 \times 10$, $\rho = 1$ and ϵ variable for asynchronous processes and fictitious memory (black dotted-dashed line); $\rho = 1$ and ϵ variable for asynchronous processes and no memory, i.e random updating (blue dashed line). The group s.d. is defined as the average over time of the s.d. of the set of strategies $\{\phi_i(t)\}_{i\in\{1,...,1000\}}$.

In this section I provide a measure of the degree of synchrony among agents and comment on the findings. Here, in order to quantify synchrony in a population, I propose the use of the standard deviation of the set of individual strategies $\phi_i(t)$ in the group (which is different from the standard deviation of the trajectory that give a good estimate of the size of the fluctuations of $\phi(t)$). This measure is calculated as the average (over a period of time of several times the duration of the period of the oscillations) standard deviation of the sample $\{\phi_i(t)\}_{i\in\{1,...,1000\}}$. This is an indication of how much apart agents are or, in other words, of the spread of the strategies in the group at any one moment in time. Specifically, in figure 5.13 I illustrate the value of the average standard deviation for four different systems: $mr = 50 \times 10$, $\rho = 1$ and ϵ variable for asynchronous processes (red line, corresponds to simulation data already reported in figure 4.7 with the same colour) and for synchronous processes (grey line, corresponds to simulation data

already reported in figure 5.7 with the same colour); $mr = 50 \times 10$, $\rho = 1$ and ϵ variable for asynchronous processes with fictitious memory (black dashed-dotted line, corresponds to simulation data already reported in figure 5.4 with the same colour); $\rho = 1$ and ϵ variable for asynchronous processes with no memory at all, i.e random updating (blue dashed line, corresponds to simulation data reported in figure 5.12 with the same colour). In the last instance, the values of ϵ on the $x-$axis are multiplied by 50 for a direct comparability of the standard deviation with the other results with respect to ϵ.

From the figure, comparing the case of individual memories (red line) and fictitious memories (black line) one infer that the spread of strategies can differ substantially even though the collective behaviour, as evidenced in figure 5.4, is remarkably similar in terms of amplitude and period of the fluctuations. On the other hand, comparing the case of fictitious memories (black dot-dashed line) and no memories (blue dashed line) one can infer that the spread of strategies is similar even though in one case agents use information and in the other do not. In the case of a system with individual memories the group spread has a local maximum at the Hopf-bifurcation. Finally, similarly to the behaviour described in the previous section, for processes without memory and random updates (blue dashed line) the group standard deviation grows linearly when ϵm is roughly greater than 10 and non-linearly otherwise.

The results from this section show that, when comparing two populations, the spread of individual strategies can differ substantially even if the collective average strategy is similar. At the same time, even when the spread of individual strategies is similar, the use of information and the collective average strategy may be considerably different.

5.5 Modelling discrete time steps.

In the limit of an infinite population size and $r \to \infty$, the model studied in chapter 4 is described by a pair of simple ODEs that evolve in continuous time. In that scenario, the amplitude and period are uniquely determined by the fundamental quantities $\epsilon\rho$ and m (cf. 4.12 and 4.15). Therefore, one can argue that the relationship between ϵ and ρ is inversely proportional, i.e. the dynamics are preserved as long as $\epsilon\rho$ is fixed. However, in chapter 4 and 5, I have extensively simulated the model using a Gillespie algorithm (Gillespie, 1977). According to this algorithm, the simulations operate by discretization of the time steps, picking random numbers for the waiting times (Bolker, 2008). Given these considerations, here I ask to what extent the discretization of the time steps induces a sensitivity on the value of ρ and ϵ for the evolution of $\phi(t)$. I show that dynamics evolving in discrete time are not invariant under a constant value of the product $\epsilon\rho$, i.e.

the value of each factor matters. In particular, I include working that quantifies under what conditions the evolution of $\phi(t)$ is preserved when ρ or ϵ are varied. Besides, the working presented in this section is also used to derive closed form solutions for the evolution of the recursive functions 4.27-4.28.

Calculations and discussion

First, I assume without loss of generality that $\tilde{p} = 1$ is satisfied in equations 4.1a-4.1c and thus the evolution of ϕ_i follows equation 4.1b. Working on equation 4.1c would give equivalent results. If one truncates the evolution of ϕ_i to one step, one obtains

$$\phi_i(t + \delta t) = \phi_i(t) + \epsilon[1 - \phi_i(t)] \quad \text{with rate} \quad \rho \tag{5.4}$$

Next, using equation 5.4 recursively, I calculate the value of $\phi_i(t+1)$ for different values of ρ. For deterministic events, the following holds true:

$$\begin{aligned}
\phi_i(t+1) =& \phi_i(t) + \epsilon[1 - \phi_i(t)] \\
& \text{when} \quad \rho = 1 \\
\phi_i(t+1) =& \phi_i(t) + \epsilon[1 - \phi_i(t)] + \epsilon\Big[1 - [\phi_i(t) + \epsilon[1 - \phi_i(t)]]\Big] \\
& \text{when} \quad \rho = 2 \\
\phi_i(t+1) =& \phi_i(t) + \epsilon[1 - \phi_i(t)] + \epsilon\Big[1 - [\phi_i(t) + \epsilon[1 - \phi_i(t)]]\Big] + \\
& \epsilon\Big[\phi_i(t) + \epsilon[1 - \phi_i(t)] + \epsilon\Big[1 - [\phi_i(t) + \epsilon[1 - \phi_i(t)]]\Big]\Big] \\
& \text{when} \quad \rho = 3
\end{aligned}$$

By means of expanding, collecting and rearranging, I obtain

$$\begin{aligned}
\phi_i(t+1) =& 1 - (1 - \phi_i(t)) + \epsilon(1 - \phi_i(t)) && (5.5) \\
& \text{when} \quad \rho = 1 \\
\phi_i(t+1) =& 1 - (1 - \phi_i(t)) + 2\epsilon(1 - \phi_i(t)) - \epsilon^2(1 - \phi_i(t)) && (5.6) \\
& \text{when} \quad \rho = 2 \\
\phi_i(t+1) =& 1 - (1 - \phi_i(t)) + 3\epsilon(1 - \phi_i(t)) - 3\epsilon^2(1 - \phi_i(t)) + \epsilon^3(1 - \phi_i(t)) && (5.7) \\
& \text{when} \quad \rho = 3
\end{aligned}$$

I note that the coefficients of the term $1 - \phi_i(t)$ are the binomial expansion of $(1 - \epsilon)^\rho$. As a result, equation (5.4) becomes

$$\phi_i(t + 1) = 1 - (1 - \phi_i(t))(1 - \epsilon)^\rho \tag{5.8}$$

Now one can study how to preserve the dynamics when changing ϵ, say from ϵ_1 to ϵ_2, by changing ρ, say from ρ_{ϵ_1} to ρ_{ϵ_2}. In order to do this, I compare the two processes ensuing from the use of ρ_1 together with ϵ_{ρ_1} and ρ_2 together with ϵ_{ρ_2}. Here, with the wording 'preserving the dynamics' I mean that $\phi_i(t+1) - \phi_i(t)$ is preserved. If the condition

$$(1-\epsilon_1)^{\rho_{\epsilon_1}} = (1-\epsilon_2)^{\rho_{\epsilon_2}} \tag{5.9}$$

is satisfied, than the two processes make the same change in one unit of time. From this condition, we obtain the relationships

$$\epsilon_2 = 1 - (1-\epsilon_1)^{\rho_{\epsilon_1}/\rho_{\epsilon_2}} \tag{5.10}$$

and, equivalently,

$$\frac{\rho_{\epsilon_1}}{\rho_{\epsilon_2}} = \frac{\log(1-\epsilon_2)}{\log(1-\epsilon_1)} \tag{5.11}$$

From equation 5.11, in the limiting case $\epsilon_1, \epsilon_2 \to 0$, I recover the following

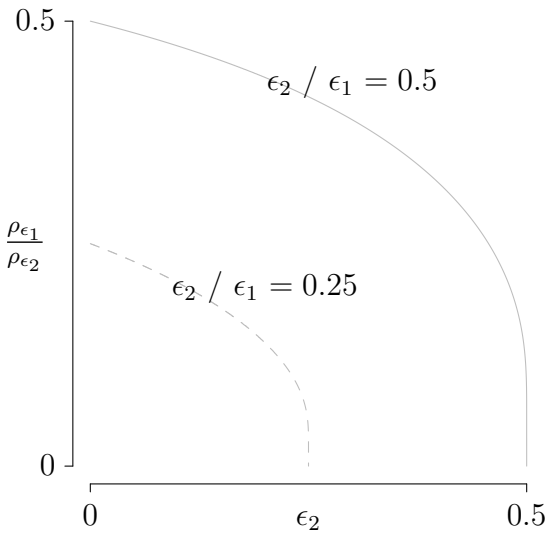

Figure 5.14: An illustration of the 'trade-off' between ϵ and ρ which preserves the dynamics. The ratio $\rho_{\epsilon_1}/\rho_{\epsilon_2}$ is plotted for the case $\epsilon_1 = 2\epsilon_2$ (continuous line) and $\epsilon_1 = 4\epsilon_2$ (dashed line). The value of ϵ_1 is constrained in order for the corresponding value of ϵ_2 to be in the interval $[0, 1]$. In more detail, this illustration answers the following question: in order to preserve the evolution of $\phi_i(t)$, how does ρ vary when ϵ is doubled (or quadrupled)?

equality:

$$\rho_{\epsilon_1}\epsilon_1 = \rho_{\epsilon_2}\epsilon_2 \tag{5.12}$$

This equality recovers the condition which preserves the dynamics in equations 4.12 and 4.15. Figure 5.14 confirms graphically that condition 5.12 holds true in the limiting case mentioned. In this figure, the ratio $\frac{\rho_{\epsilon_1}}{\rho_{\epsilon_2}}$ is plotted on the y-axis for different values of ϵ_2 while keeping fixed the ratio $\epsilon_2/\epsilon_1 = 0.25$ for the dashed line and $\epsilon_2/\epsilon_1 = 0.5$ for the continuous line. One would obtain identical results from studying equation (4.1c). Although these results point out that the trade-off between ϵ and ρ is not inversely-proportional for discrete time dynamics, the trade-off remains inversely proportional for several results of the agent-based Gillespie simulations. Several simulation results consistently show that for the product $\epsilon\rho$ the value of the single factors does not matter. This is supported by the simulation data showing that rescaling according to the fundamental quantity $\epsilon\rho m$ is valid for ϵ in $[0, 1]$ and m fixed when varying ρ (see figure 4.11). However, other simulation data shows that for $\epsilon = 1$ fixed (i.e. pure strategies) and ρ variable, the discretization may have a minor influence on the outcome of the simulations.

5.6 Changing the number of individuals

5.6.1 Small population size

So far, the entirety of theoretical, numerical and simulation results presented relies on the assumption of modelling a large population. The population size tends to infinity in analytical results and is set to one thousand in all performed simulations. However, many biological and social systems evolve under demographic noise and fluctuating population size. Therefore, it is logical to explore to what extent a small population size would affect the collective behaviour of the model studied. In particular, I am interested to know whether small populations merely break down the observable features or, alternatively, originate new features. Before moving on, I point out two relatable results from the literature. First, figure 3.2 shows that the most defining feature of the minority game is derived in the continuum between small and large population size. This feature is regulated by the size of the population (relative to the dimension of the strategy space), with small and large populations yielding differing behaviours. Second, Helbing *et al.* (2005b) study turn-taking equilibria with periodic dynamics for 2-person and 4-person binary anti-coordination games. This study is interesting because the equilibria that holds for 2-person games quickly deteriorates for larger populations. Inspired by these reflections, I decided to explore the existence of new properties for small populations in my model. Upon exploring several scenarios, a new behaviour is uncovered and depicted in figure 5.15. This figure shows that small populations of mixed-strategy agents have the capacity to self-organise into stable groups that play pure strategies. This behaviour contrasts with the one shown in figure 4.5 (top left graph). In the latter, each individual fluctuates closely around the equilibrium. In the former, individual strategies evolve into either: (i) two separating groups of opposing strategies if the number of agents n is even (left hand plot, $n = 10$) or (ii) into two groups plus an undecided agent if the number of agents is odd (right hand plot, $n = 5$). Therefore, while in large groups agents throw a dice at each turn, in small groups agents freeze into pure strategies which entails a different ability to exploit the information collected. To illustrate the transition between the two scenarios, I have plotted figure 5.16. In this figure I plot the asymptotic distribution of strategies as a function of different population sizes. The plots show a transition from a bi-modal (pure strategies) towards a uni-modal distribution of strategies (mixed-strategies) as the number of agents increases.

In this new setting, I discover that the scaling properties uncovered for large population are preserved for small populations. In particular the transition between uni-modal and bi-modal is regulated by the fundamental quantities rm and ϵm ($\rho = 1$ fixed). This is illustrated in figure 5.17. In each plot, the left

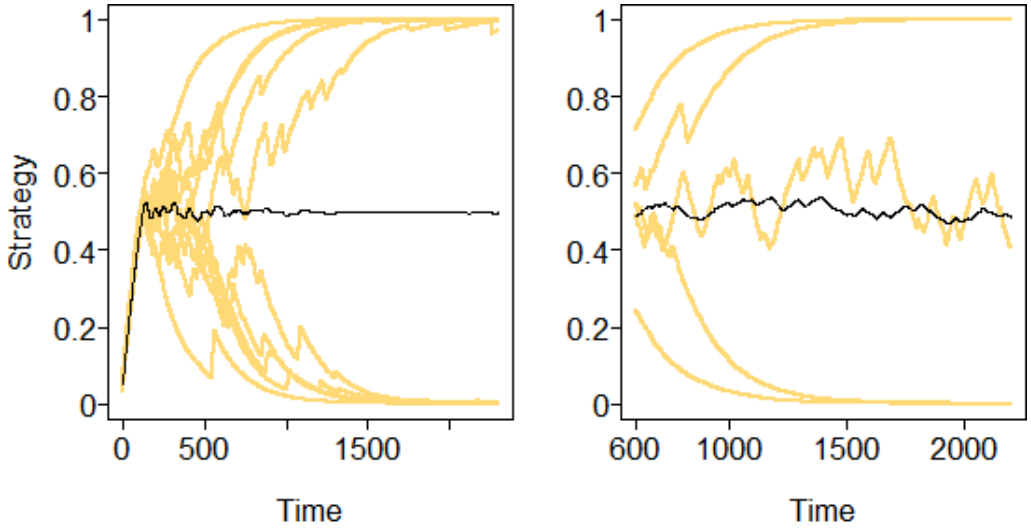

Figure 5.15: Evolution over time of individual (orange lines) and group average (black lines) strategies. Left plot: realization of a population of 10 agents using mixed-strategies which spontaneously converges to a pure-strategies equilibrium. Right plot: realizations of a population of 5 agents using mixed strategies which spontaneously converges to 4 pure-strategy agents and 1 mixed-strategy agent. Parameters used $m = 50$, $r = 10$, $\epsilon = 0.0026$ and $\rho = 1$.

y-axis shows the percentage of agents using a pure strategy, while the x-axis shows the number of agents. The value of rm is 250 and 500 respectively, while $\epsilon m = 0.1$ and $\rho = 1$ are fixed. On the same plots, the right vertical axes shows a theoretical approximation of the probability that an agent collects a misleading sample of the strategies played in the group. In this case, a sample is considered to be misleading when it forecasts the incorrect minority in the population. The probability is calculated by approximating memory with a binomial distribution. The probability of a misleading sample is calculated as follows. Let consider a population of $N = 10$ individuals that are evenly split into two groups of 5. Next, consider a focal individual from one group, say group A. This individual can observe 4 individuals in his own group, group A, and 5 individuals in the other group, group B. If the focal individual randomly samples this population, there would be a probability $p = 4/9$ that he samples an individual in his group (A) and $p = 5/9$ that he does otherwise. Since, the aim of the focal agent is to correctly identify the minority group in the population (without counting himself in), a sample is misleading when it predicts a majority of A individuals. The probability of collecting a misleading sample increases as the population size increases and decreases as the sample size increases. Now, let $X(n, p)$ be binomially distributed with n trials and p chances of success in each trial, where $n = rm$ is

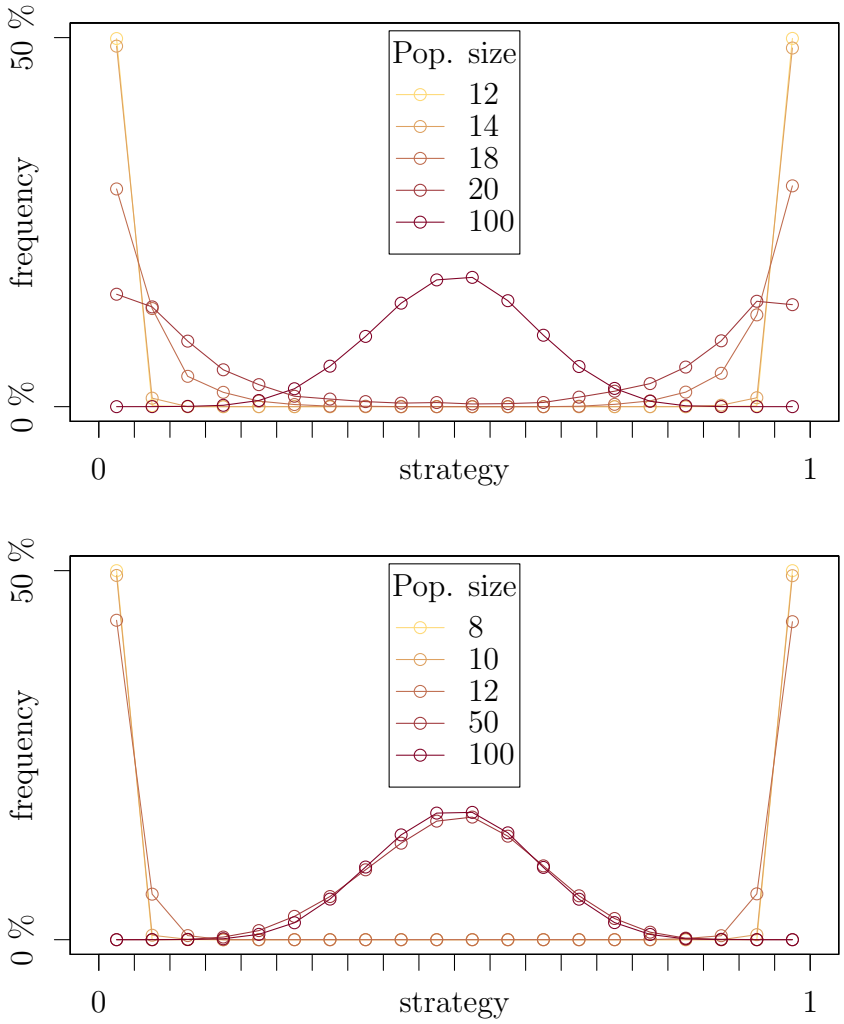

Figure 5.16: Strategies frequency measured over 10000 time units as a function of population size. Parameters: $r = 20$ (top) and $r = 10$ (bottom); $m = 50$, $\rho = 1$ and $\epsilon = 0.002$ fixed. For a small population size the group self-organises by adopting pure strategies whereas it assumes a mixed-strategy equilibrium depicted as a unimodal distribution for large population sizes.

the size of the memory string of an agent and $p = (N/2)/(N-1)$ is the probability that a focal individual samples a B-group member. Let $F(k) = Pr(X \leq k)$ be the cumulative distribution function for $X(rm, (N/2)/(N-1))$. Using this notation, one obtains that the probability of collecting a misleading sample is approximately $F(rm/2-1)$. This quantity is calculated numerically and plotted as a grey line in figure 5.17 with a corresponding y-axes on the right side of the

plots. One can see that this probability is a good indicator of the transition between the uni-modal and bi-modal distribution and captures well the population size at which a group fails to split into two pure strategies groups.

5.6.2 Comments

In chapters 4 and 5, the model assumes that individuals estimate the prevalence of B players by collecting samples of the behaviour of others through observations. A corollary of the mean-field model for large populations is that, in the limiting case of large samples (i.e. $r, m \to \infty$), individuals build expectations (of the proportion of B players) that are convergent to the same value. However, this is not true for small populations.

In small groups, individuals form diverging estimates of the proportion of B players in the group. Specifically, individuals' estimates differ depending on whether they themselves are A, B or mixed-strategies players. For example, in small groups with an even population size N, the observed proportion of B players at equilibrium is $\frac{N/2}{N-1}$ for an A player and $\frac{N/2-1}{N-1}$ for a B player. However, both proportions tend to 0.5 as N tends to ∞. Therefore, for large populations, the two proportions become indistinguishable and the mean-field approximation is valid.

Overall, I would expect that the small-group equilibrium described in section 5.6.1 (for $N < 20$, $mr \geq 500$) depends on the values of N and mr and holds generally for $N << mr$.

Further research is required to fully characterise small-group dynamics, for example, as a function of N, $\epsilon\rho m$, rm. It is useful to note that the computational cost of simulations increases with rm: when the individual memory duration m increases, the amount of data to be held in RAM storage increases; when the observation rate r increases, in addition to RAM storage, the number of iterations also increases.

Overall, this brief study suggests that individuals, although using the same learning strategy, coordinate differently within small groups compared to large groups. In particular, small group size does not cause demographic noise. Small groups can 'naturally' converge to pure-strategy efficient stable equilibria.

From an evolutionary perspective, one might consider a cost associated with memory size and information storage and recollection. In this scenario, individuals may have an incentive to incur the cost of a detailed perception in small groups but not large ones. This consideration leads to a research question. Would it be possible that individuals have an incentive to incur the cost of a detailed perception in small groups but not large ones? Could a transition from small scale communities to large scale communities, such as from rural areas to cities, disrupt the use of established effective learning strategies?

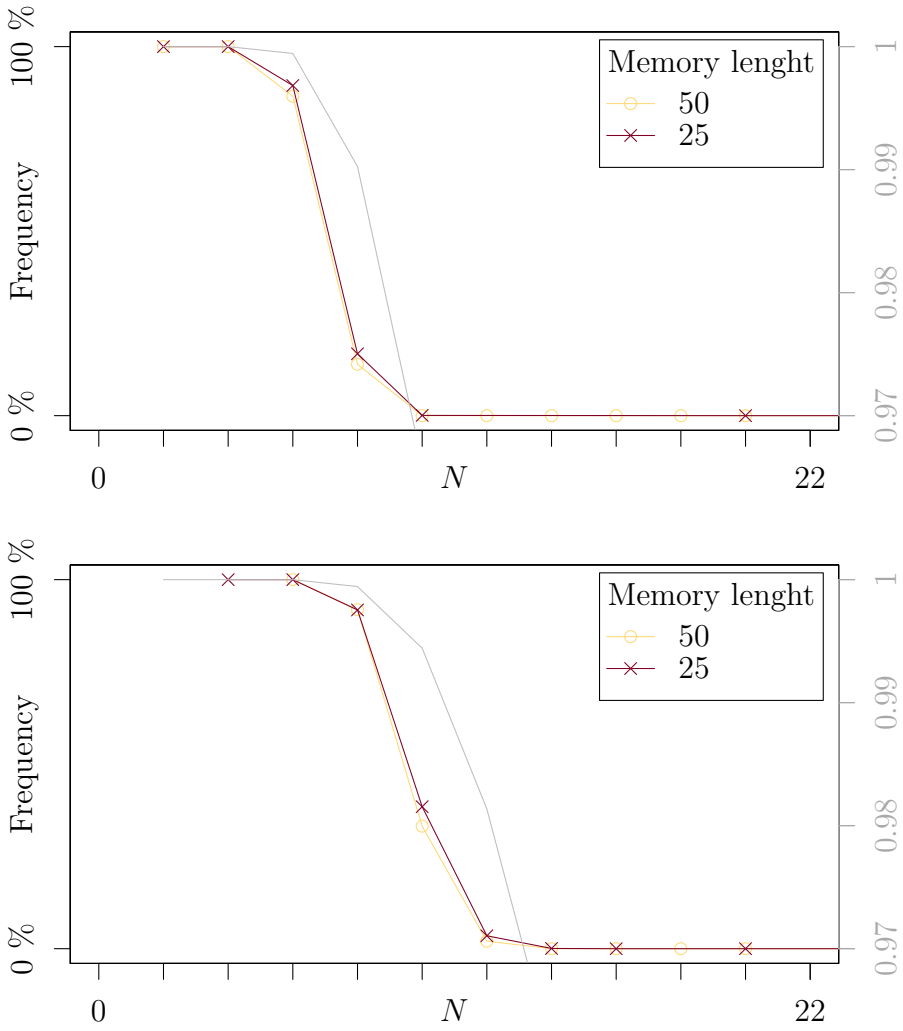

Figure 5.17: Top: $mr = 25 \times 10$ (yellow) and $mr = 50 \times 5$ (red); bottom: $mr = 25 \times 20$ (yellow) and $mr = 50 \times 10$ (red). Fixed $\epsilon m = 0.1$ and $\rho = 1$. Left side y-axes: frequency of individuals holding a pure strategy, i.e. $\phi_i(t) = 0$ or $\phi_i(t) = 1$ with a precision of 0.0005, calculated from simulations over 10000 time units. Right side y-axes: theoretical probability that an individuals does not make a mistake due to biased sampling (grey line). This quantity is derived using a binomial approximation for the memory content of individuals, assuming they are equally distributed between the two pure strategies. N represents the number of agents.

5.7 Onset and collapse of oscillations: a review.

Amplitude and period in deterministic dynamics. In the limiting case $r \to \infty$, the dynamics of $\phi(t)$ is deterministic and can be reduced to the combined study of two ODEs. In this scenario, the rescaled period (T/m) and amplitude (A) are uniquely determined as a function of $\epsilon\rho m$. The next two tables report the known formulae for amplitude and period in the deterministic scenario. An upward arrow indicates that the variable is positively correlated to the size of the amplitude. A downward arrow indicates that the variable is negatively correlated to the size of the amplitude. The two arrows together indicate that the variable is either positively or negatively correlated to the amplitude depending on the range of values considered.

Variable	Effect on A	Comments
$\epsilon\rho T$	$\uparrow$	The exact solution for the amplitude is $$A(\epsilon\rho T) = 0.5 \tanh\left(\frac{\epsilon\rho T}{4}\right).$$ The function $\tanh(\cdot)$ is positive, monotone and increasing from $[0, \infty)$ to $[0, 1)$.
$\epsilon\rho m\tau$	$\uparrow$	The exact solution for the amplitude is $$A(\epsilon\rho m\tau) = 0.5 \tanh\left(\frac{\epsilon\rho m\tau}{4}\right).$$ with $\tau = T/m = f(\epsilon\rho m)$. Numerical solutions show that $f(x)$ is a decreasing monotone smooth function from $(0, \infty)$ to $(1, 2)$.
$\epsilon\rho m$	$\uparrow$	The exact solution for the amplitude is $$A(\epsilon\rho m) = 0.5 \tanh\left(\frac{\epsilon\rho m f(\epsilon\rho m)}{4}\right).$$ Numerical solutions show that $g(x) = xf(x)$ is an increasing monotone smooth function from $(0, \infty)$ to $(0, \infty)$.

Table 5.1: Impact on the amplitude A of the parameters of the model for deterministic dynamics.

Variable	Effect on T	Comments
$\epsilon\rho$	$\downarrow$	The exact solution for the period is $$T = mf(\epsilon\rho m).$$ Numerical solutions show that $f(x)$ is a decreasing monotone smooth function from $(0, \infty)$ to $(1, 2)$.
m	$\uparrow$	Numerical solutions show that $T(m) = mf(\epsilon\rho m)$ is an increasing monotone smooth function from $(0, \infty)$ to $(0, \infty)$.

Table 5.2: Impact on the period T of the parameters of the model for deterministic dynamics.

Amplitude and period in stochastic dynamics, r finite. The dynamics of the population in the case of finite batches, i.e. r finite, can be understood as the dynamics of the population in the case of infinite batches (deterministic scenario) with added noise. It turns out that such noise can affect the dynamics significantly, leading to damped oscillations and noise-sustained equilibria. Simulation results for finite batches suggest that the period, T, of the oscillations in $\phi(t)$ obeys the formula given for the period in deterministic dynamics[1]. Consequently, in the case of finite batches, the amplitude but not the frequency of the fluctuations is affected by the noise induced by finite sampling.

In the next table, I review the ways in which several variables influence the amplitude of the oscillations in $\phi(t)$ for finite batches. Some of the effects discussed here are equivalent to those already observed for deterministic dynamics. Therefore it is important to differentiate between the new contributions due to stochasticity and those intrinsic to the model (with and without noise). For example, the amplitude increases as m increases for deterministic dynamics; however, m has a double-fold effect on the amplitude in dynamics with noise.

Importantly, in table 5.3, the comparison between the parameters ρ and m helps to identify which parameters contribute to the collapse of the oscillations and which do not. The comparison reveals that the value of T/m is correlated to the collapse, while the value of the period T is not. In fact, figures 4.10, 4.11 and 4.12 consistently show that oscillations decay as the ratio T/m nears the value 1. This implies that the ratio between period and memory duration regulates the breakdown of the oscillations, albeit not providing an exact prediction for a breakdown value.

[1]The formula, $T(m, \epsilon\rho) = mf(\epsilon\rho m)$, is discussed in the tables above.

Variable	Effect on A	Description
Observation rate, r	$\uparrow$	As r increases from 0 to ∞, the system transitions (possibly through different routes depending on the other system parameters) from being random to deterministic. For given ϵ, ρ and m, there is a unique critical value r_c where the system exhibits a Hopf bifurcation.
Agents simultaneity	$\uparrow$	Simultaneity defers the decay of oscillations to larger values of $\epsilon\rho m$ but does not impact on the Hopf bifurcation.
Update rate, ρ	$\uparrow\downarrow$	As it increases, ρ triggers a Hopf bifurcation, say at ρ_c, followed by bounded oscillations until it dumps oscillations as it increases further. As ρ increases past the Hopf bifurcation, the period largely follows the rule valid without noise described above: $T = mf(\epsilon\rho m)$. In particular, as ρ increases: i) the period T decreases, i.e. oscillations become faster; ii) the ratio T/m decreases.
Intensity of response, ϵ	$\uparrow\downarrow$	An identical argument applies to ϵ as to ρ. In addition, ϵ causes the discretization of the 'strategy space', affecting the trajectory of the dynamics. However, the comparison of the simulation results for different values of ϵ indicate that the discretization effect on the dynamics is minor and only qualitative.
Memory duration, m	$\uparrow\downarrow$	As it increases, m triggers a Hopf bifurcation, say at m_c, followed by bounded oscillations until it dumps oscillations as it increases further. As m increases past the Hopf bifurcation, the period largely follows the rule valid without noise described above: $T = mf(\epsilon\rho m)$. In particular, as m increases: i) the period T increases, i.e. oscillations become slower; ii) the ratio T/m decreases.
Heterogeneous memory duration	$\downarrow$	Intra-group variations on memory duration increase the variability of the information intake. The resulting effect is similar to reducing r, which defers the Hopf bifurcation to higher values of $\epsilon\rho m$ and dumps oscillations everywhere else. Notably, oscillations close to the Hopf bifurcation are more robust to this form of noise than those far from it.
Small populations	n/a	Small-population effects can lead to a new pure-strategy stable equilibrium in the game. Whereas in large populations the simulated behaviour of the group is well approximated by a theoretical mean-field description, this approximation fails in small populations.

Table 5.3: Impact on the amplitude A of the parameters of the model for stochastic dynamics.

To conclude this section, I provide an explanation of the role of T/m in causing the decay of oscillations. To do this, consider that the time-averaged integral of $\phi(t)$ over one period always equals 0.5, i.e.

$$\frac{1}{T}\int_{t-T}^{t}\phi(\tau)d\tau = 0.5 \quad \text{for any } t. \tag{5.13}$$

Consider that, in the limit $r \to \infty$, the proportion of B players in an individual's memory at time t is given by

$$\mu_i(t) = \frac{1}{m}\int_{t-m}^{t}\phi(\tau)d\tau \quad \text{for any } t. \tag{5.14}$$

In the case r finite, equation 5.14 is valid on average. Similarly, it would be valid if one included an appropriate noise term with expected value 0 on the RHS. In the limit $r \to \infty$, the study of the dynamics (cf. figure 4.8) tells us the necessary and sufficient condition such that, at time t, an individual changes the direction of strategy update (increase or decrease $\phi_i(t)$). This necessary and sufficient condition states that

$$\mu_i(t) = \frac{1}{m}\int_{t-m}^{t}\phi(\tau)d\tau = 0.5. \tag{5.15}$$

Next, consider that, as $T/m \to 1$, the individual memory duration converges to the duration of the period, T. Equations 5.13-5.14, taken together, tell us that, in the limit $T \to m$, $\mu_i(t)$ converges to 0.5 (for any t). This also means that $\mu_i(t)$ evolves within a small neighbourhood of 0.5. Condition 5.15 implies that, at $\mu_i(t) = 0.5$, individual i changes the direction of evolution of his strategy $\phi_i(t)$. However, the fact that $\mu_i(t)$ evolves in a small neighbourhood of 0.5 implies that $\mu_i(t)$ is always close to a critical point. In particular, the introduction of a source of noise would increase the likelihood that $\mu_i(t)$ traverses the value 0.5 (the critical point) by random drift.

In the case of finite batches, noise is generated by the stochastic distribution of the duration of memories, as well as by the random distribution of observation times. In fact, figure 4.2 tells us that memory duration is normally distributed with mean m and variance m/r (for large mr). These sources of stochasticity imply that, while $\mu_i(t)$ is close to 0.5, it is also subject to random drift. Consequently, agents are more likely to randomly switch the direction of strategy update. As a result, to sum up, group coordination becomes vulnerable to noise as T/m approaches 1.

In addition, this argument also shows that, as r increases, coordinated oscillations break down at a larger value of T/m. This is due to the fact that, as r increases, observation samples become more detailed and the distribution of

memory durations acquires a lower variance. This concludes my exploration of the role of T/m on the collapse of coordinated oscillations. Overall, this discussion presents a scenario for a spontaneous (occurring without apparent cause) break down of oscillations as the population reaches a tipping point characterised in terms of the value of the period of the oscillations themselves. An exact quantitative prediction of the tipping point requires further research.

Chapter 6

Conclusion

This thesis has progressed from a review of agent-based game-theoretical evolutionary models to studying a scenario for the emergence of collective behaviour in a binary game with memory incorporating difference in information among individuals.

At the beginning of this research project, my research interest lay in investigating the evolutionary origin of differences in learning and response in group behaviour. Chapter 2 contains preliminary results in which I review a 2-D spatial model of the Prisoners' Dilemma showing that the introduction of a stochastic learning rule changes the outcome of the game (section 2.2). For example, I proved that stochastic learning can both strengthen and hinder the evolution of cooperation compared to a deterministic learning rule (figure 2.4). In light of these early results, I introduced a simple selection dynamics operating on different levels of response sensitivity (modelled with the parameter β in formula 2.6). In particular, following from Guttal & Couzin (2010) and Zafeiris & Vicsek (2013), I hypothesised that the evolution of individual sensitivity might depend on the individual position occupied within the clusters that naturally emerge on the lattice and a cost function for the response sensitivity. I preliminarily tested a range of scenarios through simulations not reported in this thesis, and they did not present emerging patterns of interest for the evolution of the response sensitivity. Consequently, I shifted my interest to the study of learning dynamics in well-mixed populations, where the role of learning is un-coupled from the effects of spatial structures. For the Prisoners' Dilemma, however, this topic has already been extensively discussed in differing scenarios (Axelrod & Hamilton, 1981; Galla, 2009; Hauert & Schuster, 1997; Hilbe *et al.*, 2017; Imhof *et al.*, 2005; McNamara *et al.*, 2004; Nowak & Sigmund, 1992; Stewart & Plotkin, 2016).

My research focused on the study of information use in the context of group coordination in an anti-coordination game. A coordination game models a situation in which agents benefit from choosing the same action, such as in the

Stug Hunt game. An anti-coordination game models a situation in which agents benefit from choosing different actions. Examples of anti-coordination games are the minority game, the Hawk-Dove game and congestion games. In sections 2.6 and 3.2 I review a model of the Hawk-Dove game with memory and the canonical minority game, offering an overview of the effects of memory on the learning dynamics in an unstructured population of individuals faced with a binary choice. These models incorporate memory and online learning (Saad, 2009) and provide the basis to investigate the effects of different learning strategies.

In chapter 4, following on from reviewing these models, I define a two-strategy anti-coordination game with memory. Memory has a two-fold description, comprising duration and observation rate. In the game, individuals choose between two options, which correspond to two pure strategies in the language of game theory. Each individual holds a preference for each option, formally referred to as a mixed-strategy probability quantifying the chances to choose each option. The objective of each individual is to choose the option adopted by the minority in the group. Individuals respond to the empirical frequency of play observed (fictitious play (Fudenberg *et al.*, 1998)). Agents observe (through random statistical sampling) the strategies enacted by other agents in the population. Each agent remembers his observations for a fixed time only, and memory does not feature gradual forgetfulness or decay. The observation rate regulates the level of detail of the information collected. In the limit of an infinite rate, the dynamics of the prevalence of the two options over time is deterministic. With a finite rate of observation, the dynamics are stochastic, in which case the rate regulates the degree of stochasticity in the system. This is an innovative approach compared to conventional approaches to learning in evolutionary game theory where stochastic choice is modelled (a posteriori) through a 'temperature of selection' (Cavagna *et al.*, 1999; Galla, 2009; Traulsen *et al.*, 2008). Here, instead, the observation rate is the parameter that models the degree of uncertainty in the decision-making process. Within the field of bounded rationality, this rate can represent the amount of detail on the information available due to constraints on the agent or the environment (Kozyreva & Hertwig, 2019). The model accounts for sampling (e.g. observing or sensing), acquired without engaging in a contest, of the behaviour enacted by other members of the group. The learning experiences are not conflated with the process of accumulation of gains and losses, and therefore the frequency of encoding information in the memory does not correspond to the rates of playing. This definition is based on the separation of the time scales of the play and information collection processes and motivated by neuroscientific theories of model-based reinforcement learning, accounting for algorithms in which individuals collect information about their environment without directly incurring reward or punishment (Lee *et al.*, 2012; Tolman, 1948). Therefore, the model that I propose differs from the Hawk-Dove game with memory proposed

by Burridge *et al.* (2017) in this primary aspect of separation of the two processes and their timescales. On the level of the analysis of the model, this new definition has fruitfully allowed a systematic description of the game, uncovering among other findings: (i) the fundamental variables; (ii) a universal classification of the dynamics based on the fundamental variables; (iii) a deterministic description of the dynamics in the limiting case of an infinite rate of sampling; (iv) a dis-entanglement of the effects of memory size, discerning between the role of duration and observation rate. As outlined in section 4.5.1, given this definition of memory, I suggest that (v) the observation rate acts as an 'inverse temperature of selection' analogous to the role of response sensitivity in the 'thermal minority game' (Cavagna *et al.*, 1999) and following this hypothesis I have proposed further research possibilities, in particular suggesting that (i) my model 'maps' to a limiting case of 'thermal minority game' and (ii) how it may shed light on specific features of other anti-coordination games (Burridge *et al.*, 2015, 2017). In total, the model depends on four quantities corresponding to memory duration (m), observation rate (r), update rate (ρ) and intensity of response (ϵ). All results hold for pure strategies ($\epsilon = 1$) and mixed strategies ($0 < \epsilon < 1$), albeit with some qualitative differences. The thesis presents a detailed discussion on the onset of collective oscillations as well as their decay in a binary anti-coordination game with memory.

In light of this analysis, I propose that the model has the potential to provide a framework to describe mechanisms in sociological and social phenomena such as fashion. As a sociological subject, fashion is related to collective and personal identity dynamics and social distinction and imitation mechanisms (Aspers & Godart, 2013; Berger & Heath, 2007). The term fashion describes prevailing aesthetic styles at particular points in time and inherently refers to temporary cyclical phenomena (Sproles, 1981). It can be used to comprise more than clothing styles (Aspers & Godart, 2013; Kawamura, 2020). Male facial hair choices, although driven by a range of factors (Dixson *et al.*, 2017), are considered a matter of fashion (Dixson *et al.*, 2017; Janif *et al.*, 2014) and appear to stem from a combination of features that my model effectively captures.

Making a binary distinction between having a clean-shave and a beard, data collected from the Illustrated London News from 1842 to 1972 shows how beards choices are subject to cyclic patterns over time (Robinson, 1976). Other studies point out that beard dynamics are subject to negative-frequency selection and consistent with mechanisms operating in fashion trends (Alderman, 2017; Janif *et al.*, 2014). At the same time, studies have hypothesised that the importance of beard choices is amplified in settings with large numbers of people and high anonymity (Dixson *et al.*, 2017) such as crowded urban areas (Dixson *et al.*, 2017; Scott *et al.*, 2014). Such settings may provide 'opportunities to discern relationships between facial traits and behaviour by exposing individuals to large

numbers of unfamiliar faces, revealing patterns too subtle to detect with smaller samples' (Scott *et al.*, 2014).

In my model, for $\epsilon = 1$, individuals use a simple heuristic which operates as a negative frequency-selection rule on a binary choice: an individual keeps a beard as long as it corresponds to the minority style in the group and shaves otherwise. The model is particularly flexible as it predicts that oscillations arise both when individual switch between clear preferences (pure strategies for $\epsilon = 1$) and when they tune their preference more gradually (for $0 < \epsilon < 1$) and in a continuous spectrum of preferences (for $\epsilon \to 0$). However, the analysis points out that the spectrum of preferences, previously called space of strategies in this work, only reduces notably for $\epsilon > 0.5$. The model also parametrises individual memory reaching over an arbitrary duration of m time units into the past. I showed that cyclic patterns are present even in the case of heterogeneous memory duration in the population. As mentioned, beard style dynamics are captured in data collected in urban settings where individuals are exposed to large samples of the styles adopted by people mostly unknown. Similarly, the model analysed in this thesis corresponds to the limiting case of a large population in which behaviour is anonymous instead of tag-based. Finally, large urban groups are usually provided with opportunities to efficiently exchange visual social information through posters, television, and the Internet (Scott *et al.*, 2014). While this itself may influence individual preferences (Batres & Perrett, 2014), it may also provide different individuals with homogeneous visual information, whereby enhancing the synchronisation of their choices.

In addition to beard style dynamics, an analogous mechanism may be at play across different scenarios. It is closely related to other fashion subjects, such as, for example, first names dynamics (Kessler *et al.*, 2012). The framework may also describe phenomena observed in supply-demand chains and bacterial communities.

In supply-demand chains, output and price fluctuations are well documented. It is still debated whether the cause of such fluctuations is endogenous or exogenous and whether agents are rational or resort to rule-of-thumb expectations which are understood as a rational solution to the trade-off between costly and cheap predictions (Brock & Hommes, 1997; Gouel, 2012). In this context, simple backward-looking expectations often 'generate systemic errors with strong cyclic patterns' (Gouel, 2012). A similar effect is described in my model, which explains the emergence of cyclic patterns from backward-looking expectations. My study proposes that oscillations arise endogenously as the information collected is homogeneous across the population. Oscillations are also sustained for a broader range of parameters when the processes in the system are synchronous. Therefore this study underscores the possible side-effects of an ever more connected society in which actors are highly interconnected and with the possibility to respond

synchronously to the same information while located at greater distances.

Finally, the model may capture the emergence of synchronisation in unicellular populations. Although the model is not fine-tuned to describe a particular observed behaviour, I suggest that it models an essential mechanism at play in some populations. Bacteria, for example, use simple rules-of-thumb to regulate their actions when choosing to release a public good (Cavaliere & Poyatos, 2013) and some unicellular organisms also undergo transitions to coordinated activity with increasing cell density or nutrient availability (Gregor *et al.*, 2010). For these organisms, the emergence of coordination is always mediated through communication by chemical signalling via the extracellular solution De Monte *et al.* (2007); Garcia-Ojalvo *et al.* (2004); Taylor *et al.* (2009). When modelling the collective synchronisation in biological populations, researchers often refer to coupled oscillators models with non-linear coupling such as Ariaratnam & Strogatz (2001); Kuramoto (1975); Winfree (1967) and linear coupling such as in Matthews & Strogatz (1990). These studies describe systems of individual oscillators that synchronise by entrainment of their phases. Other studies (De Monte *et al.*, 2007; Gregor *et al.*, 2010; Taylor *et al.*, 2009), however, pointed out that some organisms present a collective transition to synchrony in which organisms are quiescent/random at low information density and display synchronise as they share more and more information, either as a result of increasing the population size or as a result of increasing the transport rate in the external solution. Thus, unicellular populations are usually coupled through shared information available and encoded as chemicals dispersed in the medium and have been shown to synchronise as the information becomes uniform in the medium. My model provides an example of this type of transition in which oscillatory behaviour becomes entrained as individuals share more of the same information. On another note, the emergence of cyclic oscillations can provide stability and predictability to biological systems (Alon, 2019). In accordance with this fact, figure 5.3 shows that the onset of limit cycles corresponds to an increase in the synchrony in the population.

To conclude, in this thesis, I have studied a model for the emergence of collective behaviour in a general anti-coordination game and tested how fluctuations emerge and disappear depending on the system parameters in a wide range of alternative scenarios. This study adds to previous research describing how oscillations develop when negative-frequency selection is coupled to delay. I further formalise a specific mechanism in which information is explicitly modelled and becomes the leading parameter regulating the onset of collective behaviour. Further research should fruitfully extend this model to other games and tune its features to specific scenarios.

Appendix A

Annotated code

A.1 Code for simulation of the stochastic model

```
# R code. As tested in RStudio Version 1.2.5033.
# The output is a plot in the "Plots" window on Rstudio.

### PARAMETERS ###
L = 1000          # Population size: unit of measure 1. Must be even.
m = 25            # Memory duration: unit of measure [T].
r = 20            # Observation rate: unit of measure 1/[T].
rho = 1           # Update rate: unit of measure 1/[T].
epsilon = 0.05    # Update step: unit of measure 1.
TT =  150         # Final time: unit of measure [T].
                  # m,r,rho,TT must be positive integers.

### USEFUL AUXILIARY OBJECTS (matrices, vectors, constants) ###
rp = r + rho      # Sum of the two rates.
```

```
factor = L/2*rp   # Number of events in one unit of time (on average).
                  # An event always involves two agents,
                  # thus the factor "/2".
I = TT*factor     # Number of iterations in TT time units (on average).
players = seq(1:L)# Players are numbered from 1 to L.
strat = c(0, 1)   # Binary encoding of the two options available:
                  # 0 means playing A, 1 means playing B.
phi = rep(0,L)    # Vector of each player individual strategy.
                  # Everyone's strategy is set to 0 at the beginning.
mu = rep(0,L)     # Vector of each player..
                  # ..'proportion of B players met'.
                  # Everyone's proportion is set to 0 at the beginning.
nos = m*r         # Memory size: number of observations recorded.
nosmu = nos - 1   # Costant used to "shift" the memory content.
memory = matrix(NA,L,nos)
                  # Object definition (empty matrix): size L*m*r.
                  # Row 'i' of 'memory' contains the last m*r ..
                  # .. obervations of individual i.
memory[,1]=0      # Memory is initilized at 0 for all agents.
Phi_graph = rep(0,TT)# Vector recording the group average strategy..
                  # ..at consecutive integer time units.
st_dv = rep(0,TT) # Vector recording the s.d. of the strategies
                  # at consecutive integer time units.

### ITERATIONS ###
for(ti in 1:I){
# "ti" is an auxiliary variable that keeps count of the iterations.
```

```
        ### OBSERVATION ###
if(sample(1:rp,1)>rho){
# The "if clause" chooses between the two processes:
# observation and strategy update.
# Here it picks observation.
# To make sense of the ‘if clause’, remember that:
# rp=r+rho.
Two_players = sample(players, 2, replace = FALSE)
# Pick 2 out of L players.
player1 = two_players[1]
# Establish which one is player 1 for this iteration.
player2 = two_players[2]
# Establish which one is player 2 for this iteration.

memory[player1,1:nosmu] = memory[player1,2:nos]
# Player 1 shifts the content of his memory by one position.
memory[player2,1:nosmu] = memory[player2,2:nos]
# Player 2 shifts the content of his memory by one position.
# The shift clears the one "bit" of memory which will be updated next.

zfunction = function(i){sample(strat, 1, prob = c(1-phi[i], phi[i]) )}
# Define function:
# choose action A with prob 1-phi[i] or B with prob phi[i].
z1 = zfunction(player1)
# Player 1 chooses his action.
z2 = zfunction(player2)
# Player 2 chooses his action.
```

```
memory[player1,nos] = z2
# The action of player 2 is observed by player 1.
memory[player2,nos] = z1
# The action of player 1 is observed by player 2.
}
else{
        ### STRATEGY UPDATE ###
two_players = sample(players, 2, replace = FALSE)
# Pick 2 out of L players.
player1 = two_players[1]
# Auxiliary variable for player 1 in this iteration.
player2 = two_players[2]
# Auxiliary variable for player 2 in this iteration.
mu[player1] = mean(memory[player1,], na.rm=T)
# Compute the proportion of B players observed..
# ..by player 1 in the last r*m observations.
mu[player2] = mean(memory[player2,], na.rm=T)
# Compute the proportion of B players observed..
# ..by player 2 in the last r*m observations.
phi[player1] = phi[player1]
              + as.double((0.5 > mu[player1]))*epsilon*(1-phi[player1])
              - as.double((0.5 < mu[player1]))*epsilon*phi[player1]
# Increase phi[player1] if mu[player1] < 0.5. Run this as one line.
phi[player2] = phi[player2]
              + as.double((0.5 > mu[player2]))*epsilon*(1-phi[player2])
              - as.double((0.5 < mu[player2]))*epsilon*phi[player2]
# decrease phi[player1] if mu[player2] > 0.5. Run this as one line.
```

```
}

### COUNTING TIME ###
# Remember the fact that..
# ..factor = L/2*(r+rho) = average number of iterations in 1 time unit.
# ti is the variable counting the iterations.
if(ti %% factor == 0){
# %% means "mode", e.g. 1000 %% 500 = 2
Phi_graph[ti/factor] = mean(phi)
# record "average group strategy" at time "ti/factor"
st_dv[ti/factor] = sd(phi) }
# record "standard deviation among players" at time "ti/factor"
}
### ITERATIONS HAVE ENDED ###

# OUTPUT: a plot that shows the average group strategy (black line)..
# .. +- 1*standard deviation (grey area)
Tm = TT/2
at_v = round(c(1,Tm,TT))
labels_v = round(c(1,Tm,TT))
# Define x-axis labels
plot(Phi_graph, ylim = c(0,1), ylab = expression(phi),
xaxt="n" ,xlab = "Time", type = "l")
# Create a plot with labels
polygon(c(seq(1:TT),rev(seq(1:TT))),c(Phi_graph+st_dv,
rev(Phi_graph-st_dv)),ylim = c(0,1),col="lightgrey", border = "NA")
# Add a filled in polygon
```

```
lines(Phi_graph, ylim = c(0,1))
# Plot the black line over the shaded area
axis(1, at= at_v  , labels=labels_v   )
# Add x-axis and its labels
```

A.2 Code for synchronous processes

```
### PARAMETERS ###


L = 1000          # Population size: unit of measure 1. Must be even.
m = 25            # Memory duration: unit of measure [T].
r = 20            # Observation rate: unit of measure 1/[T].
rho = 1           # Update rate: unit of measure 1/[T].
epsilon = 0.05    # Update step: unit of measure 1.
TT =  150         # Final time: unit of measure [T].
                  # m,r,TT must be positive integers.
                  # Code is written for rho=1. Do not change rho.


### USEFUL AUXILIARY OBJECTS (matrices, vectors, constants) ###
rp = r + rho      # Sum of the two rates.
factor = L/2*rp   # Number of events in one unit of time (on average).
                  # An event always involves two agents,
                  # thus the factor "/2".
players = seq(1:L)# Players are numbered from 1 to L.
strat = c(0, 1)   # Binary encoding of the two options available:
                  # 0 means playing A, 1 means playing B.
phi = rep(0,L)    # Vector of each player individual strategy.
                  # Everyone's strategy is set to 0 at the beginning.
mu = rep(0,L)     # Vector of each player..
                  # ..'proportion of B players met'.
                  # Everyone's proportion is set to 0 at the beginning.
nos = m*r         # Memory size: number of observations recorded.
nosmu = nos - 1   # Costant used to "shift" the memory content.
```

```
memory = matrix(NA,L,nos)
                  # Object definition (empty matrix): size L*m*r.
                  # Row ‘i’ of ‘memory’ contains the last m*r ..
                  # .. obervations of individual i.
memory[,1]=0      # Memory is initilized at 0 for all agents.
Phi_graph = rep(0,TT)# Vector recording the group average strategy..
                  # ..at consecutive integer time units.
st_dv = rep(0,TT) # Vector recording the s.d. of the strategies
# at consecutive integer time units.
n = L/2           # Used below to split the group in two.
n1 = n+1          # Used below to split the group in two.

### ITERATIONS START ###
for(ti in 1:TT){
# "ti" is an auxiliary variable that keeps count of the iterations

for(obs in 1:r ){
        ### OBSERVATIONS ###
zfunction = function(i){sample(strat, 1, prob = c(1-phi[i], phi[i]))}
# Define function: choose action A with prob. 1-phi[i] or..
# .. B with prob. phi[i]. Argument i is the i-th player.
z = sapply(players, zfunction)
# Each agent picks a strategy.
all_players = sample(players, L, replace = FALSE)
# Pick L out of L players in random order.
players1 = all_players[1:n]
# Vector of L/2 players by index.
```

```
players2 = all_players[n1:L]
# Vector of the remaining L/2 players.
memory[players,1:nosmu] = memory[players, 2:nos]
# Shift memory content by one.
memory[players1,nos] = z[players2]
# Agents in players1 observe the strategies enacted by agents in players2.
memory[players2,nos] = z[players1]
# Agents in players2 observe the strategies enacted by agents in players1.
}


        ### STRATEGY UPDATE ###
### After the obsevations, each agent updates mu ###
muf = function(i){mean(memory[players[i],], na.rm=T)}
mu[players1] = sapply(players1, muf)
mu[players2] = sapply(players2, muf)
# Compute the proportion of observed B players for each agents.
phif <- function(i){phi[players[i]] +
as.double((0.5 > mu[players[i]]))*epsilon*(1-phi[players[i]])-
as.double((0.5 < mu[players[i]]))*epsilon*phi[players[i]]}
# Function that calculates the strategy update. Use with mr even.
## phif <- function(i){phi[players[i]]*(1 - epsilon) +
## as.double((0.5 > mu[players[i]]))*epsilon }
## Simpler phif function. Use with mr odd. Runs faster.
phi = sapply(players,phif)
# Strategy update

Phi_graph[ti] = mean(phi)
```

```
# Record "average group strategy" at time "ti"
st_dv[ti] = sd(phi)
# Record "standard deviation among players" at time "ti"
}
### ITERATIONS HAVE ENDED ###

# OUTPUT: a plot that shows the average group strategy (black line)..
# .. +- 1*standard deviation (grey area)
Tm = TT/2
at_v = round(c(1,Tm,TT))
labels_v = round(c(1,Tm,TT))
# Define x-axis labels.
plot(Phi_graph, ylim = c(0,1), ylab = expression(phi),
 xaxt="n" ,xlab = "Time", type = "l")
# Create a plot with labels.
polygon(c(seq(1:TT),rev(seq(1:TT))),c(Phi_graph+st_dv,
rev(Phi_graph-st_dv)),ylim = c(0,1),col="lightgrey", border = "NA")
# Add a filled in polygon.
lines(Phi_graph, ylim = c(0,1))
# Plot the black line over the shaded area.
axis(1, at= at_v  , labels=labels_v   )
# Add x-axis and its labels
```

A.3 Code output

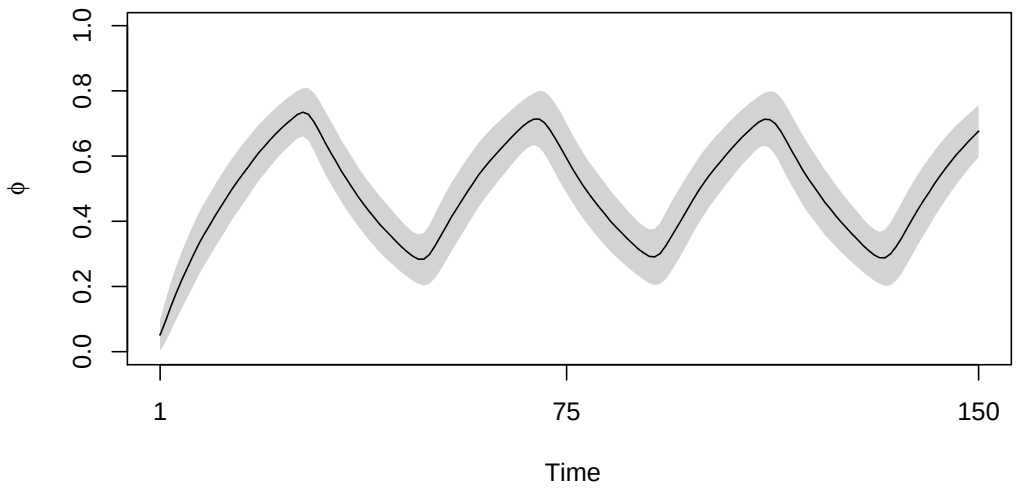

Figure A.1: Output obtained from running the code for the original model in section A.1. Parameters: $m = 25$, $r = 20$, $\rho = 1$, $\epsilon = 0.05$, population size $= 1000$.

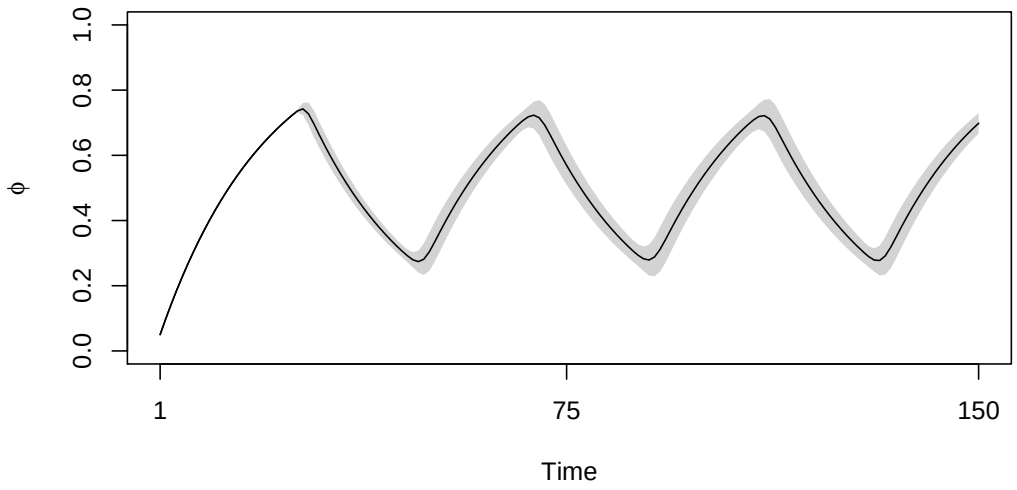

Figure A.2: Output obtained from running the code for synchronous processes in section A.2. Parameters: $m = 25$, $r = 20$, $\rho = 1$, $\epsilon = 0.05$, population size $= 1000$.

References

Alderman, L. (2017). Are men with beard more desirable? *New York Times*. 2, 117

Alon, U. (2019). *An introduction to systems biology: design principles of biological circuits*. CRC press. 119

Alos-Ferrer C, L.J., Hugelschafer S (2016). Inertia and decision making. *Frontiers in Psychology*, **7**. 38

Ariaratnam, J.T. & Strogatz, S.H. (2001). Phase diagram for the winfree model of coupled nonlinear oscillators. *Physical Review Letters*, **86**, 4278. 119

Arthur, W.B. (1994). Inductive reasoning and bounded rationality. *The American economic review*, **84**, 406–411. viii, 28, 29, 30, 35, 91

Aspers, P. & Godart, F. (2013). Sociology of fashion: Order and change. *Annual Review of Sociology*, **39**, 171–192. 117

Auger, P. & Pontier, D. (1998). Fast game theory coupled to slow population dynamics: the case of domestic cat populations. *Mathematical biosciences*, **148**. 22

Axelrod, R. & Dion, D. (1988). The further evolution of cooperation. *Science*, **242**, 1385–1390. 6, 7

Axelrod, R. & Hamilton, W.D. (1981). The evolution of cooperation. *science*, **211**, 1390–1396. 115

Axerold, R. & Dion, D. (1988). The further evolution of cooperation. *Science*, **242**, 1385 – 1389. 6, 91

Axerold R, H.W. (1981). The evolution of cooperation. *Science*, **211**, 1390–1396. 91

Batres, C. & Perrett, D.I. (2014). The influence of the digital divide on face preferences in el salvador: People without internet access prefer more feminine men, more masculine women, and women with higher adiposity. *PLoS One*, **9**, e100966. 118

Bednekoff, P.A. (1997). Mutualism among safe, selfish sentinels: a dynamic game. *The American Naturalist*, **150**, 373–392. 1

Berger, J. & Heath, C. (2007). Where consumers diverge from others: Identity signaling and product domains. *Journal of Consumer Research*, **34**, 121–134. 117

Bolker, B.M. (2008). *Ecological models and data in R*. Princeton University Press. 101

Brock, W.A. & Hommes, C.H. (1997). A rational route to randomness. *Econometrica: Journal of the Econometric Society*, 1059–1095. 118

Burridge, J., Gao, Y. & Mao, Y. (2015). Forgetfulness can help you win games. *Physical Review E*, **92**, 032119. 59, 117

Burridge, J., Gao, Y. & Mao, Y. (2017). Delayed response in the hawk dove game. *The European Physical Journal B*, **90**, 13. 23, 26, 35, 38, 39, 56, 59, 69, 82, 88, 117

Cavagna, A. (1999). Irrelevance of memory in the minority game. *Physical Review E*, **59**, R3783. 34, 35

Cavagna, A., Garrahan, J.P., Giardina, I. & Sherrington, D. (1999). Thermal model for adaptive competition in a market. *Physical Review Letters*, **83**, 4429. 34, 59, 99, 116, 117

Cavaliere, M. & Poyatos, J.F. (2013). Plasticity facilitates sustainable growth in the commons. *Journal of The Royal Society Interface*, **10**, 20121006. 2, 119

Challet, D. (2000). *Modelling markets dynamics: Minority games and beyond*. Ph.D. thesis. 2

Challet, D. & Marsili, M. (2000). Relevance of memory in minority games. *Physical Review E*, **62**, 1862. 87, 88

Challet, D. & Zhang, Y.C. (1997). Emergence of cooperation and organization in an evolutionary game. *Physica A: Statistical Mechanics and its Applications*, **246**, 407–418. 30, 35

Challet, D. & Zhang, Y.C. (1998). On the minority game: Analytical and numerical studies. *Physica A: Statistical Mechanics and its applications*, **256**, 514–532. 33

Challet, D., Marsili, M. & Zhang, Y.C. (2000). Modeling market mechanism with minority game. *Physica A: Statistical Mechanics and its Applications*, **276**, 284–315. 31

Challet, D., Marsili, M. & Zhang, Y.C. (2004). *Minority games: interacting agents in financial markets*. Oxford University Press. 32, 34

Challet, D., Marsili, M., Zhang, Y.C. *et al.* (2013). Minority games: interacting agents in financial markets. *OUP Catalogue*. 2, 35, 59, 97

Choe, J.C. (2019). *Encyclopedia of Animal Behavior*. Academic Press. 20

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, **24**, 87–114. 75

De Monte, S., d'Ovidio, F., Danø, S. & Sørensen, P.G. (2007). Dynamical quorum sensing: Population density encoded in cellular dynamics. *Proceedings of the National Academy of Sciences*, **104**, 18377–18381. 3, 119

Dixson, B.J., Rantala, M.J., Melo, E.F. & Brooks, R.C. (2017). Beards and the big city: displays of masculinity may be amplified under crowded conditions. *Evolution and Human Behavior*, **38**, 259–264. 117

Erneux, T. (2009). *Applied delay differential equations*, vol. 3. Springer Science & Business Media. 2

Fecteau, S., Pascual-Leone, A., Zald, D.H., Liguori, P., Théoret, H., Boggio, P.S. & Fregni, F. (2007). Activation of prefrontal cortex by transcranial direct current stimulation reduces appetite for risk during ambiguous decision making. *Journal of Neuroscience*, **27**, 6212–6218. 38

Fudenberg, D., Drew, F., Levine, D.K. & Levine, D.K. (1998). *The theory of learning in games*, vol. 2. MIT press. 116

Galla, T. (2009). Intrinsic noise in game dynamical learning. *Physical review letters*, **103**, 198702. 8, 115, 116

Garcia-Ojalvo, J., Elowitz, M.B. & Strogatz, S.H. (2004). Modeling a synthetic multicellular clock: repressilators coupled by quorum sensing. *Proceedings of the National Academy of Sciences*, **101**, 10955–10960. 119

Gelimson, A., Cremer, J. & Frey, E. (2013). Mobility, fitness collection, and the breakdown of cooperation. *Physical Review E*, **87**, 042711. 13, 16

Ghosh, A., De Martino, D., Chatterjee, A., Marsili, M. & Chakrabarti, B.K. (2012). Phase transitions in crowd dynamics of resource allocation. *Physical Review E*, **85**, 021116. 34

Gillespie, D.T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, **22**, 403–434. 40

Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, **81**, 2340–2361. 40, 101

Gouel, C. (2012). Agricultural price instability: a survey of competing explanations and remedies. *Journal of economic surveys*, **26**, 129–156. 3, 118

Gregor, T., Fujimoto, K., Masaki, N. & Sawai, S. (2010). The onset of collective behavior in social amoebae. *Science*, **328**, 1021–1025. 119

Guttal, V. & Couzin, I.D. (2010). Social interactions, information use, and the evolution of collective migration. *Proceedings of the national academy of sciences*, **107**, 16172–16177. 91, 115

Hauert, C. & Schuster, H.G. (1997). Effects of increasing the number of players and memory size in the iterated prisoner's dilemma: a numerical approach. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **264**, 513–519. 115

Helbing, D. & Johansson, A. (2013). Pedestrian, crowd, and evacuation dynamics. *arXiv preprint arXiv:1309.1609*. 1

Helbing, D., Buzna, L., Johansson, A. & Werner, T. (2005a). Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. *Transportation science*, **39**, 1–24. 1

Helbing, D., Schönhof, M., Stark, H.U. & Hołyst, J.A. (2005b). How individuals learn to take turns: Emergence of alternating cooperation in a congestion game and the prisoner's dilemma. *Advances in Complex Systems*, **8**, 87–116. 1, 105

Heliovaara, S., Ehtamo, H., Helbing, D. & Korhonen, T. (2013). Patient and impatient pedestrians in a spatial game for egress congestion. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **87**, 012802. 22

Hilbe, C., Nowak, M.A. & Sigmund, K. (2013). Evolution of extortion in iterated prisoner's dilemma games. *Proceedings of the National Academy of Sciences*, **110**, 6913–6918. 6

Hilbe, C., Martinez-Vaquero, L.A., Chatterjee, K. & Nowak, M.A. (2017). Memory-n strategies of direct reciprocity. *Proceedings of the National Academy of Sciences*, **114**, 4715–4720. 115

Ho, T.H., Camerer, C.F. & Chong, J.K. (2007). Self-tuning experience weighted attraction learning in games. *Journal of economic theory*, **133**, 177–198. 9

Holt, C.A. & Roth, A.E. (2004). The nash equilibrium: A perspective. *Proceedings of the National Academy of Sciences*, **101**, 3999–4002. 20

Imhof, L.A., Fudenberg, D. & Nowak, M.A. (2005). Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences*, **102**, 10797–10800. 115

Imhof, L.A., Fudenberg, D. & Nowak, M.A. (2007). Tit-for-tat or win-stay, lose-shift? *Journal of theoretical biology*, **247**, 574–580. 6

Janif, Z.J., Brooks, R.C. & Dixson, B.J. (2014). Negative frequency-dependent preferences and variation in male facial hair. *Biology letters*, **10**, 20130958. 2, 117

Jarvis, M. (2017). Waves and normal modes. 73

Jørgensen, S. & Zaccour, G. (2012). *Differential games in marketing*, vol. 15. Springer Science & Business Media. 20

Karline Soetaert, P.M.H. (2009). *A Practical Guide to Ecological Modelling. Using R as a Simulation Platform*. Springer Link. 51

Kawamura, Y. (2020). *Doing research in fashion and dress: An introduction to qualitative methods*. Bloomsbury Publishing. 117

Kessler, D.A., Maruvka, Y.E., Ouren, J. & Shnerb, N.M. (2012). You name it–how memory and delay govern first name dynamics. *PloS one*, **7**, e38790. 118

Kim, Y.G. (1995). Status signaling games in animal contests. *Journal of Theoretical Biology*, **176**, 221–231. 22

Kokko, H., Griffith, S. & Pryke, S. (2014). The hawk-dove game in a sexually reproducing species explains a colorful polymorphism of an endangered bird. *Proceedings of the Royal Society B*. 22

Kozyreva, A. & Hertwig, R. (2019). The interpretation of uncertainty in ecological rationality. *Synthese*, 1–31. 116

Kuramoto, Y. (1975). Self-entrainment of a population of coupled non-linear oscillators. In *International symposium on mathematical problems in theoretical physics*, 420–422, Springer. 119

Lee, D., Seo, H. & Jung, M.W. (2012). Neural basis of reinforcement learning and decision making. *Annual review of neuroscience*, **35**, 287–308. 116

Luce, R.D. & Raiffa, H. (1957). Games and decisions: Introduction and critical survey. 5

Mackey, M.C. (1989). Commodity price fluctuations: price dependent delays and nonlinearities as explanatory factors. *Journal of economic theory*, **48**, 497–509. 3

Matthews, P.C. & Strogatz, S.H. (1990). Phase diagram for the collective behavior of limit-cycle oscillators. *Physical review letters*, **65**, 1701. 97, 119

McNamara, J.M., Barta, Z. & Houston, A.I. (2004). Variation in behaviour promotes cooperation in the prisoner's dilemma game. *Nature*, **428**, 745–748. 7, 115

Mesterson-Gibbons, M. (2001). *An introduction to game theoretic modelling: Second edition*. AMS. 20

Mesterton-Gibbons, M. (2019). *An introduction to game-theoretic modelling*, vol. 37. American Mathematical Soc. 22

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, **63**, 81. 75

Morgan, M.S. (2007). *The curious case of the prisoner's dilemma: model situation? Exemplary narrative?*. Duke University Press. 5

Novak, M.A. & May, R.M. (1992). Evolutionary games and spatial chaos. *Nature*, **359**, 826–829. 7

Novak, M.A. & May, R.M. (1993). The spatial dilemmas of evolution. *International Journal of Bifurcation and Chaos*, **03**, 35–78. vii, 7, 8, 9

Nowak, M. & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, **364**, 56–58. 6

Nowak, M.A. (2006). *Evolutionary dynamics: exploring the equations of life*. Harvard university press. 5, 20, 21, 22, 23

Nowak, M.A. & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, **355**, 250–253. 115

Pitz G F, R.H. (1968). Payoff effects in sequential decision-making. *Journal of Experimental Psychology*, **77**, 249–257. 38

Poundstone, W. (1992). Prisoner's dilemma. 5

Robinson, D.E. (1976). Fashions in shaving and trimming of the beard: The men of the illustrated london news, 1842-1972. *American Journal of Sociology*, **81**, 1133–1141. 2, 117

Saad, D. (2009). *On-line learning in neural networks*, vol. 17. Cambridge University Press. 116

Sato, Y. & Crutchfield, J.P. (2003). Coupled replicator equations for the dynamics of learning in multiagent systems. *Physical Review E*, **67**, 015206. 9

Savit, R., Manuca, R. & Riolo, R. (1999). Adaptive competition, market efficiency, and phase transitions. *Physical Review Letters*, **82**, 2203. viii, 32, 33, 34

Schadschneider, A., Chowdhury, D. & Nishinari, K. (2010). *Stochastic transport in complex systems: from molecules to vehicles*. Elsevier. 1

Scott, I.M., Clark, A.P., Josephson, S.C., Boyette, A.H., Cuthill, I.C., Fried, R.L., Gibson, M.A., Hewlett, B.S., Jamieson, M., Jankowiak, W. *et al.* (2014). Human preferences for sexually dimorphic faces may be evolutionarily novel. *Proceedings of the National Academy of Sciences*, **111**, 14388–14393. 117, 118

Simon, H.A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, **69**, 99–118. 1, 28

Smith, J.M. (1982). *Evolution and the Theory of Games*. Cambridge university press. 22

Smith, J.M. & Price, G.R. (1973). The logic of animal conflict. *Nature*, **246**, 15–18. 21, 22

Sproles, G.B. (1981). Analyzing fashion life cycles—principles and perspectives. *Journal of marketing*, **45**, 116–124. 117

Stewart, A.J. & Plotkin, J.B. (2016). Small groups and long memories promote cooperation. *Scientific reports*, **6**, 26889. 115

Strogatz, S. (2004). *Sync: The emerging science of spontaneous order*. Penguin UK. 1

Sumpter, D.J. (2005). The principles of collective animal behaviour. *Philosophical transactions of the royal society B: Biological Sciences*, **361**, 5–22. 1

Szabo, G. & Toke, C. (1998). Evolutionary prisoner's dilemma game on square lattice. *Physical Review E.*, **58**, 69 – 73. 8

Taylor, A.F., Tinsley, M.R., Wang, F., Huang, Z. & Showalter, K. (2009). Dynamical quorum sensing and synchronization in large populations of chemical oscillators. *Science*, **323**, 614–617. 3, 119

Tolman, E.C. (1948). Cognitive maps in rats and men. *Psychological review*, **55**, 189. 116

Traulsen, A., Shoresh, N. & Nowak, M.A. (2008). Analytical results for individual and group selection of any intensity. *Bulletin of mathematical biology*, **70**, 1410. 116

Wheeler, G. (2018). Bounded rationality. 1

Wilson, E.O. (1962). Chemical communication among workers of the fire ant solenopsis saevissima (fr. smith) 1. the organization of mass-foraging. *Animal behaviour*, **10**, 134–147. 1

Winfree, A.T. (1967). Biological rhythms and the behavior of populations of coupled oscillators. *Journal of theoretical biology*, **16**, 15–42. 119

Yeomans, J.M. (1992). *Statistical mechanics of phase transitions*. Clarendon Press. 34

Zafeiris, A. & Vicsek, T. (2013). Group performance is maximized by hierarchical competence distribution. *Nature communications*, **4**, 1–8. 91, 115