Observational Learning with Finite Memory

by

Kimon Drakopoulos

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

Author . . . .
Department of Electrical Engineering and Computer Science
May 18, 2011

Certified by . . . .
Asuman Ozdaglar
Class of 1943 Associate Professor
Thesis Supervisor

Certified by . . . .
John Tsitsiklis
Clarence J Lebel Professor of Electrical Engineering
Thesis Supervisor

Accepted by . . . .
Professor Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Students

Observational Learning with Finite Memory

by

Kimon Drakopoulos

Submitted to the Department of Electrical Engineering and Computer Science
on May 18, 2011, in partial fulfillment of the
requirements for the degree of
Master of Science in Computer Science and Engineering

Abstract

We study a model of sequential decision making under uncertainty by a population of agents. Each agent prior to making a decision receives a private signal regarding a binary underlying state of the world. Moreover she observes the actions of her last K immediate predecessors. We discriminate between the cases of bounded and unbounded informativeness of private signals.

In contrast to the literature that typically assumes myopic agents who choose the action that maximizes the probability of making the correct decision (the decision that identifies correctly the underlying state), in our model we assume that agents are forward looking, maximizing the discounted sum of the probabilities of a correct decision from all the future agents including theirs. Therefore, an agent when making a decision takes into account the impact that this decision will have on the subsequent agents. We investigate whether in a Perfect Bayesian Equilibrium of this model individual's decisions converge to the correct state of the world, in probability, and we show that this cannot happen for any K and any discount factor if private signals' informativeness is bounded.

As a benchmark, we analyze the design limits associated with this problem, which entail constructing decision profiles that dictate each agent's action as a function of her information set, given by her private signal and the last K decisions. We investigate the case of bounded informativeness of the private signals. We answer the question whether there exists a decision profile that results in agents' actions converging to the correct state of the world, a property that we call *learning*. We first study *almost sure learning* and prove that it is impossible under any decision rule. We then explore learning in probability, where a dichotomy arises. Specifically, if $K = 1$ we show that *learning in probability* is impossible under any decision rule, while for $K > 2$ we design a decision rule that achieves it.

Thesis Supervisor: Asuman Ozdaglar
Title: Class of 1943 Associate Professor

Thesis Supervisor: John Tsitsiklis
Title: Clarence J Lebel Professor of Electrical Engineering

Acknowledgments

First, I would like to express my gratitude to my incredible advisors, Asu Ozdaglar and John Tsitsiklis. First, because they gave me the opportunity to change my life and pursue my dreams in an amazing academic environment. Second, because they made this transition smooth and enjoyable. I have interacted with them for just two years and I can already see the influence they have on me as a researcher but most important as a personality.

There isn't much to say about Asu that she hasn't heard in her everyday life or hasn't read in the acknowledgments section of her other students. Her enthusiasm about research and exploration of problems is the driving force when things don't work out as expected. Her ability to improve everyone surrounding her is incomparable. But the reason that makes me feel most grateful for being her student is condensed in the following story. During my first year and when my research wasn't fruitful, after a meeting she asked me "How are you?" and I started apologizing about not getting results. And then she asked again "I mean how are you?". All the above make her a fantastic advisor, collaborator and person.

Collaborating with John is a wonderful experience that can be described with terms that will be defined later on in this thesis (I am sorry to both for using terms that I haven't yet defined!). In class, when presenting topics related to our problem, John was trying to come up with a good real life example of private signal distributions with Unbounded Likelihood Ratios. Admittedly, it is hard to find a believable one. But here it is. Interacting with him is such an example; you can get arbitrarily strong signals about the underlying truth. Usually, when private signals can be arbitrarily strong learning occurs. Hopefully, this is the case with me.

Moreover I would like to thank Daron Acemoglu for our discussions on the problems tackled in this thesis and for sharing incredible and exciting ideas for future ones.

Those two years at MIT would not be as enjoyable and productive without my new friends, the Greeks and the non-Greeks, the LIDS and the non-LIDS. Thank you all, for the great times that we have shared and for all the support that I have received from you. My life here would be very different without you.

This thesis though, is dedicated to the most important people in my life, the ones that are back home and support me from abroad; my parents Silia and Kostas and my girlfriend Aimilia. Mama, Mpampa and Aimilaki I owe you everything.

Contents

1 Introduction **13**

1.1 Problem motivation 13

1.2 Contributions 15

1.3 Related literature 16

1.3.1 Statistics/engineering literature 16

1.3.2 Economics literature 17

1.4 Organization 18

2 The model **21**

2.1 Formulation 21

2.1.1 Observation model 21

2.1.2 Bounded and unbounded likelihood ratios 23

2.2 Almost sure learning versus learning in probability 25

2.3 Designed decision rules versus strategic agents 26

3 Designing decision rules for almost sure learning **29**

3.1 Unbounded likelihood ratios: Cover's construction 29

3.2 No almost sure learning for the case of Bounded Likelihood Ratios . . 30

3.3 Discussion and Conclusions 36

4 Designing decision rules for learning in probability **37**

4.1 No learning in probability when K=1 37

4.2 Learning in probability when $K \geq 2$ 49

4.2.1 A first approach 49

4.2.2 Biased coin observation model 50

4.2.3 Cover's decision and Koplowitz's modification 51

4.2.4 Learning in probability when observing two or more of the immediate predecessors 53

4.2.5 Proof . . . 55
4.3 Discussion and Conclusions . . . 60

5 Forward looking agents: Bounded Likelihood Ratios **63**
5.1 Basic definitions . . . 63
5.2 Characterization of Best Responses . . . 68
5.3 No learning in probability for forward looking agents . . . 71
5.4 Discussion and Conclusions . . . 82

6 Forward looking agents: Unbounded Likelihood Ratios **85**
6.1 Characterization and properties of monotone equilibria. . . . 85
6.2 Myopic behaviour as equilibrium for symmetric distributions . . . 90
6.3 Construction of a non-learning equilibrium . . . 95
6.4 Existence of monotone Perfect Bayesian Equilibria . . . 102
6.4.1 Existence of a monotone Perfect Bayesian Equilibrium for the finite horizon altruistic learning game . . . 103
6.4.2 Existence of a monotone Perfect Bayesian Equilibrium for the forward looking game . . . 106
6.5 Discussion and Conclusions . . . 107

7 Conclusions **109**
7.1 Summary . . . 109
7.2 Research directions—Open problems . . . 110

List of Figures

2-1 The observation model. Agents receive an independent private signal drawn from the distribution $\mathbb{F}_\theta$, and observe the last K immediate predecessors' actions. If agent n observes the decision of agent k, we draw an arrow pointing from n to k. 22

2-2 Likelihood ratios for the coin tossing example 24

4-1 The Markov chains that model the decision process for $K = 1$. States correspond to observed actions while transitions correspond to agents' decisions. 38

4-2 Proof sketch for theorem 2. Divide agents into blocks so that the sum, in each block, of transition probabilities from 0 to 1 are constant. If during such a block the sum of transition probabilities from 1 to 0 is small there is positive probability of "getting stuck" at state 1 under $\theta = 0$. Similarly, if it is large there is positive probability of getting stuck at state 0 under $\theta = 1$. 44

4-3 The Markov chains that correspond to the decision process when K=2. 50

4-4 Dividing agents in R-blocks and S-blocks 53

4-5 The decision profile during an S-block. (a) Current state is 11 anywhere in the block (b) The state at the beginning of the block is 00 (c) The current state is 01. (d) The current state is 10. 55

4-6 An illustration of the algorithm. (a) A state automaton that represents the possible histories observed. (b) The decision process for of a signal realization of all $H * H * H * H * H *$.., where $*$ represents any possible private signal, during an S block that begins from 00 and for which the *search phase is initiated.* (c) An R block that begins with 00. (d) The decision process during an S-block where the search phase is initiated in case of signal realization of $H * T * * * *$... 62

6-1 In our construction, most agents behave myopically, except for those that are multiples of M. Those behave counter-myopically; they are choosing zero when they believe that the state of the world is one. If M is large enough, we prove that this strategy profile is a Perfect Bayesian Equilibrium of the forward looking game for $\delta \in (1/2, 1)$. Indeed, each of the myopic agents, by acting myopically, drives most future agents (the myopic ones) towards her favored decision and some (the counter-myopic ones) towards the opposite decision, since according to the monotone strategy profiles properties agents are more likely to copy than to switch the observed actions. As long as counter-myopic agents are placed in the tandem sparsely enough, myopic decisions remain optimal. Similarly, for counter-myopic agents it is optimal to choose non-monotonically since the subsequent agents expect them to do so. Clearly, such a construction is possible only when the continuation payoff can be greater than the immediate payoff, which happens only if $\delta > 1/2$. 96

List of Tables

4.1 A decision profile that achieves learning in probability. Note that this is a randomized decision . Our analysis in the previous chapters assumes deterministic decision s but can be extended to randomized by standard arguments that would just make notation harder. . . . 56

7.1 Summary of results. ✓ denotes positive results while ✗ denotes negative results. The references next to each result cite the paper who obtained it. [DOT] stands for this thesis. Theorem 1 of this thesis establishes that if the private signal structure is characterized by Bounded Likelihood Ratios then there does not exist a decision rule that achieves almost sure learning. Therefore there cannot exist an equilibrium strategy profile that achieves almost sure learning for the forward looking game. Theorems 2 and 3 of this thesis establish that if the private signal structure is characterized by Bounded Likelihood Ratios then there does not exist a decision rule that achieves learning in probability if $K = 1$ but there exists one that achieves learning in probability when $K > 1$. (The first result for the special case of coin tossing has been established by Koplowitz in [11]). If the private signal structure is characterized by Unbounded Likelihood Ratios then Cover in [7] has constructed a decision rule that achieves almost sure learning and consequently learning in probability. For the case of Unbounded Likelihood Ratios and forward looking agents we were able to show that there exist equilibria that learn and others that do not. Therefore, one cannot give specific answers for this case. 112

Chapter 1

Introduction

1.1 Problem motivation

Imagine a situation where each of a large number of entities has a noisy signal about an unknown underlying state of the world. There are many scenarios that fit in this framework. One example is a set of sensors each of which takes a measurement from the environment related to an unknown parameter. This unknown state of the world might also concern the unknown quality of a product, the applicability of a therapy, the suitability of a political party for the welfare of the citizens of a country. If the private signals, i.e., the private information that each entity receives is unbiased, their combination - *aggregation* - would be sufficient to "learn" the true underlying state of the world. On the other hand, because of communication or memory constraints, central processing of individuals' private information is usually not possible. Typically agents' information is summarized in a finite valued statistic which is then observed by other agents to refine their own belief about the unknown state. This thesis investigates what type of communication behaviors and information structures accommodate such *information aggregation*.

Such considerations can be modeled as sequential learning problems: there is an unknown state of the world that can take one of two possible values and agents act sequentially, making a binary decision on the basis of their private information and observation of some of the previous agents. These problems have been studied both in the statistics/engineering and economic literatures. The statistics literature focuses on designing decentralized decision profiles under which information aggregation takes place, delineating the impact of different communication structures. The economics literature, on the other hand, considers strategic agents and investigates how information aggregation may fail along the (perfect) Bayesian equilibrium of a dynamic game due to information externalities. Almost all the economics literature assumes that agents are *myopic*,that i.e., they choose the action that maximizes the quality of their decision. Under widely applicable assumptions on the private signal informativity, information aggregation fails when agents are myopic because of the creation of herds; agents copy the observed decisions irrespective of their private signal.

The myopic assumption is a good benchmark but does not capture the behavior of agents in several occasions. Consider the case where the unknown state of the world is indeed the quality of a new product. Individuals receive noisy private information about the unknown quality by testing it in the store where it is demonstrated, but also have access to review sites where others have already expressed their opinion about whether the unknown quality is good or bad. It is reasonable to assume that individuals do not go over the whole history of reviews but just some of the latest. Moreover, an individual, when stating her own personal preference in a review site does not only care that her opinion coincides with the true quality, but also cares for the future agents to learn the truth. Therefore, when writing a review, individuals are *forward looking* in the sense that they are taking into account how their opinion will affect future readers. For example a new technology adopter may be tempted to be contrarian so as to enhance informational efficiency, for instance if she believes that choosing what currently appears the better technology is more likely to trigger a herd.

Another setting that infuses altruistic behaviour is that of a sequential electoral mechanism. Each voter will choose between two parties, based on her private information and after observing what others' have voted. A strategic voter recognizes that her decision affects the outcome of the election directly through the vote itself and indirectly by affecting subsequent voters. A voter whose interest is the best party to win the election must account for both the quality of her decision and the effect to future decisions.

Finally, consider an agent that at each point in time receives a signal that is relevant to the underlying state of the world and her goal is to eventually learn its' true value. It is reasonable to assume that the agent has finite memory. The number of bits of information that she can store is bounded and it is reasonable to assume that they represent a finite valued statistic of her information at the last K time instances.

This thesis studies an observational learning problem of this kind, with the constraint of finite memory; each agent can observe (or remember) only the last K actions. Our work contributes to both trends of the literature. First, we provide necessary and sufficient conditions for the possibility of information aggregation under this specific observation structure. Second, motivated by the discussion of this section, we analyze, under the same observation structure, the equilibrium learning behavior of *forward looking agents.*

1.2 Contributions

Consider a large number of agents that sequentially choose between two actions. There is a binary underlying state of the world. Each agent receives a noisy signal regarding the true state and observes the actions of the last K individuals. Their signals are assumed independent, identically distributed, conditional on the true state. The purpose of this thesis is to study asymptotic information aggregation-*learning*-in this environment, i.e., investigate whether there exist decision profiles, such that agents' actions converge almost surely or in probability, to the correct state of the world as the number of agents grows large. Two features turn out to be crucial in the study of learning. The first is whether the *Likelihood Ratio* implied by individual signals is always bounded away from 0 and infinity. We refer to this property as Bounded Likelihood Ratios. The second is the number K of observed actions.

This thesis makes two sets of fundamental contributions to this problem:

(i) As a benchmark, we start by analyzing the design limitations of this observation model. Specifically, we focus on conditions under which there exist decision profiles that guarantee convergence to the correct state. Our main results are summarized in two main theorems.

- We first prove that, under the Bounded Likelihood Ratio assumption, there does not exist *any* decision profile, for any value of K, that achieves almost sure convergence of agents' actions to the correct state of the world — *almost sure learning.* Our result, combined with a result from [7] establishes that Unbounded Likelihood Ratios are *necessary and sufficient* for the existence of a decision profile that achieves almost sure learning.
- Next, we investigate the existence of decision profiles under the Bounded Likelihood Ratio assumption, that achieve convergence in probability of agents' actions to the correct state of the world— *learning in probability.* There, a surprising dichotomy emerges. We prove that if $K = 1$ such a decision profile does not exist. On the other hand, if $K > 1$, we pro-

vide a decision profile that achieves learning in probability, underlining the delicate sensitivity of learning results to the observation structure.

(ii) The second part of the thesis focuses on forward looking agents. The observation structure remains unchanged but, in this framework, each agent makes a decision that maximizes the discounted sum of the probabilities of a correct decision from the future individuals, including herself. We study this model as an insightful approximation to the forward looking behaviors discussed in the previous section.

- We prove that, under the Bounded Likelihood Ratios assumption there exists no Perfect Bayesian Equilibrium of the corresponding game that achieves learning in probability, a result that contrasts to the existence of a decision profile that would achieve it.
- In contrast we construct, for the case of Unbounded Likelihood Ratios, a Perfect Bayesian Equilibrium that achieves learning in probability and one that does not.

1.3 Related literature

The literature on decentralized information aggregation is vast. Roughly, it can be separated in two main branches; one is the statistics/engineering literature and the other the economics literature.

1.3.1 Statistics/engineering literature

Research on decentralized information aggregation was initiated by [7] and [11]. An infinite population of sensors arranged in a tandem is considered, each of which receives a signal relevant to the underlying state of the world. Each sensor summarizes its information in a message that can take finitely many values, and prior to sending its own message, it receives the message of it' immediate predecessor. The issue of resolving the hypothesis testing problem asymptotically is studied, in the sense of convergence of a component of agents' messages to the correct state of the world. There, the sharp dichotomy between Bounded and Unbounded Likelihood Ratios is underlined, pointing out the importance of the signal structure to learning results. This dichotomy is evident throughout the literature in the field. [10] also studied this problem but focusing on the case where every agent uses the same decision rule.

Another branch of decentralized information aggregation problems was initiated in [19] with a problem to be discussed below. The main difference and novelty of the new setting is in the communication structure as well as in the evaluation criterion. There are two hypotheses on the state of the world and each one of a set of sensors receives a noisy signal regarding the true state. Each sensor summarizes its information in a message that can take finitely many values, and sends it to a fusion center. The fusion center solves a classical hypothesis testing problem and decides on one of the two hypotheses. The problem posed is the design of optimal decision profiles in terms of error probability of the fusion center's decision. A network structure, in which some sensors observe other sensors' messages before making a decision was introduced in [8] and [9]. Sensors, prior to sending their message on top of receiving a signal relevant to the underlying state observe the messages of a subset of sensors that have already communicated their information. Finding the optimal decision profiles is in general a non trivial task — it requires an optimization over a set of thresholds that sensors will use for their Likelihood Ratio test (see [20] for a survey on the topic).

Traditionally, the probability of error of the fusion center's decision converges to zero exponentially in the number of peripheral sensors. However, as the number of sensors increases, the communication burden for the fusion center grows unboundedly. For this reason, [14] returns to the tandem configuration and studies the learning behavior of the decision rule that minimizes the last agent's error probability. Moreover they introduce the myopic decision profile; sensors are choosing the action that maximizes the probability of a correct decision given their information, and study its' learning behavior. In both cases the result depends on the signal structure and the dichotomy between Bounded and Unbounded Likelihood Ratios reappears. With the same flavor, [12] studies sequential myopic decisions based on private signals and observation of ternary messages transmitted from a predecessor, under the tandem configuration.

Research in the field is vast and by no means the list of papers mentioned here is exhaustive; we list those that most closely relate to our analysis.

1.3.2 Economics literature

The references [4] and [3] started the literature on learning in situations in which individuals are Bayesian maximizers of the probability of a correct decision given the available information, assuming that agents, prior to making a decision observe the whole history of actions. They illustrate a "herding externality" through an example with an ex-post incorrect herd, after the two first individuals' signals by chance are misleading. The most complete analysis of this framework is [16], illustrating the dichotomy between Bounded and Unbounded Likelihood Ratios and establishing qualitative results to those in [14].

Generalizing to more general observation structures, [17] and [2] study the case where each individual observes an unordered sample of actions drawn from the pool of history and the observed statistic gives only the numbers of sampled predecessors who took the two possible actions. The most comprehensive and complete analysis of this environment, where agents are Bayesian but do not observe the full history, is provided in [1]. They study the learning problem for a set of agents connected via a general acyclic network. In their framework, agents observe the past actions of a stochastically-generated neighbourhood of individuals and provide conditions on the private signal structure and the social network's structure under which *asymptotic learning*, in the sense of convergence in probability to the correct state of the world, is achieved.

To the best of our knowledge, the first paper that studies forward looking agents is [18], where individuals minimize the discounted sum of error probabilities of all the subsequent agents including their own. They consider the case of observing the full history and show that when private signal distributions are characterized by unbounded informativeness, learning can be achieved while, for the case of bounded informativeness, the event of an incorrect herd has positive probability. Finally, [13] for a similar model and Bounded Likelihood Ratios, explicitly characterizes a simple and tractable equilibrium that generates a herd, showing that even with payoff interdependence and forward-looking incentives, observational learning can be susceptible to limited information aggregation.

1.4 Organization

The thesis is organized as follows. Chapter 2 introduces the model, distinguishes between Bounded and Unbounded Likelihood Ratios, and compares the engineered with the strategic setting. Chapter 3 studies the existence of decision profiles for almost sure convergence of actions to the correct state when Likelihood Ratios are Bounded. Chapter 4 explores the existence of decision profiles that achieve convergence in probability to the correct state. Chapter 5 introduces forward looking agents and focuses on their learning properties under Bounded Likelihood Ratios. Chapter 6 studies forward looking agents under Unbounded Likelihood Ratios, and Chapter

7 concludes the work.

Chapter 2

The model

In this chapter we introduce our sequential learning model, which specifies the agents' available information prior to making a decision. We define two important properties of signal structures, i.e., Bounded Likelihood Ratios and Unbounded Likelihood Ratios that will be key in establishing our learning results. We present the notions of learning under consideration and finally contrast the engineered to the strategic version of our problem.

2.1 Formulation

2.1.1 Observation model

Consider a countably infinite population of agents, indexed by $n \in \mathbb{N}$. Each agent in sequence is to make a binary decision.

There exists an underlying **state of the world** $\theta \in \{0, 1\}$ which is modeled as a random variable whose value is unknown by the agents. To simplify notation, we assume that both of the underlying states are a priori equally likely, that is, $\mathbb{P}(\theta = 0) = \mathbb{P}(\theta = 1) = \frac{1}{2}$. Each agent n forms beliefs about this state based on a **private signal** s_n that takes values in a set S and also by observing the actions of his K immediate predecessors. Note that we will denote by s_n the random variable of agent n's private signal while s will denote a specific value s in S. The **action** or **decision** of individual n is denoted by $x_n \in \{0, 1\}$ and will be a function of her available information.

Conditional on the state of the world θ, the private signals are independent random variables distributed according to a probability measure $\mathbb{F}_\theta$ on the set S. The pair of measures $(\mathbb{F}_0, \mathbb{F}_1)$ will be referred to as the **signal structure** of the model.

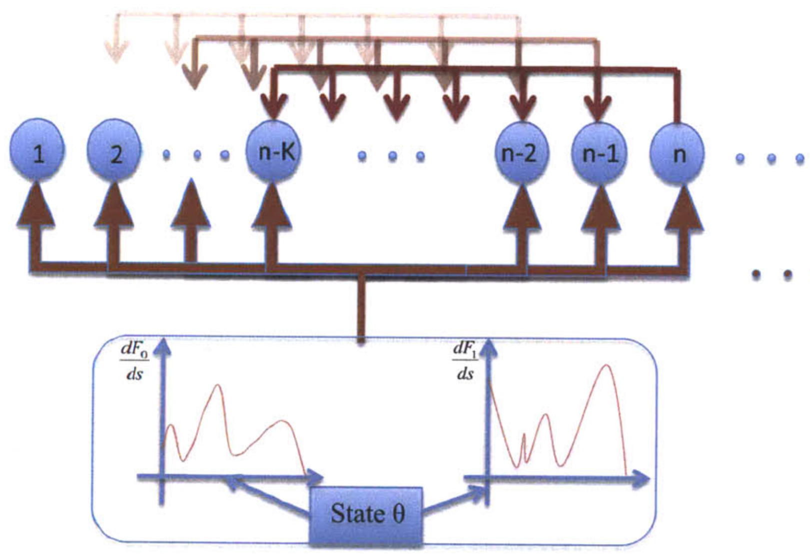


Figure 2-1: The observation model. Agents receive an independent private signal drawn from the distribution $\mathbb{F}_\theta$, and observe the last K immediate predecessors' actions. If agent n observes the decision of agent k, we draw an arrow pointing from n to k.

Throughout this thesis, the following two assumptions on the signal structure always remain in effect. First, $\mathbb{F}_0$ and $\mathbb{F}_1$ are absolutely continuous with respect to each other, implying that no signal is fully revealing about the correct state. Second, $\mathbb{F}_0$ and $\mathbb{F}_1$ are not identical, so that some signals are informative.

The **information set** I_n of agent n is defined as her private signal and the decisions of her last K predecessors. Let $D_n \triangleq \{n-K, \ldots, n-1\}$ denote the **neighbourhood** of agent n, that is, the agents whose actions are included in her information set. In other words,

$$I_n \triangleq \{s_n, x_k \text{ for all } k \in D_n\}. \tag{2.1}$$

A **decision rule** for agent n is a mapping $d_n : S \times \{0,1\}^K \to \{0,1\}$ that selects an action given the available information to agent n. A **decision profile** is a sequence of decision rules $d = \{d_n\}_{n \in \mathbb{N}}$. Given a decision profile, the sequence of decisions for all agents $\{x_n\}_{n \in \mathbb{N}}$ is a well defined stochastic process and induces a probability measure which will be denoted by $\mathbb{P}_d$.

2.1.2 Bounded and unbounded likelihood ratios

In this subsection, we discuss alternative assumptions on the signal structure, discriminating between two cases, of bounded and unbounded likelihood ratios. We first state the definition and discuss its implications.

Definition 1. *The signal structure has* ***Bounded Likelihood Ratios*** *if there exists some $m > 0$, $M < \infty$, such that the Radon-Nikodym derivative $\frac{d\mathbb{F}_0}{d\mathbb{F}_1}$ satisfies*

$$m < \frac{d\mathbb{F}_0}{d\mathbb{F}_1} < M,$$

for almost all $s \in S$ under the measure $(\mathbb{F}_0 + \mathbb{F}_1)/2$.

In order to understand the implications of this property observe that the Bayes rule yields

$$\mathbb{P}(\theta = 1 \mid s_n) = \frac{1}{1 + \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s_n)}.$$

Therefore, under the Bounded Likelihood Ratio assumption, we have :

$$\frac{1}{1+M} < \mathbb{P}(\theta = 1 \mid s_n = s) < \frac{1}{1+m} \text{ for all } s \in S.$$

Intuitively, under the Bounded Likelihood Ratio assumption, there is a maximum amount of information that an agent can extract from her private signal.

Definition 2. *The signal structure has* ***Unbounded Likelihood Ratios*** *if the essential infimum of* $d\mathbb{F}_0/d\mathbb{F}_1(s)$ *is* 0*, while the essential supremum of* $d\mathbb{F}_0/d\mathbb{F}_1(s)$ *is infinity, under the measure* $(\mathbb{F}_0+\mathbb{F}_1)/2$.

The Unbounded Likelihood Ratio assumption implies that agents may receive arbitrarily strong signals about the underlying state.

Differentiating between the cases of Bounded and Unbounded Likelihood Ratios is critical in the analysis that follows. We illustrate their difference by means of examples. The first example involves a biased coin.

Example 1 (Biased coin). *Consider a biased coin whose bias is unknown. Let* $p = \mathbb{P}(Heads)$ *take one of the two values, either* p_0 *or* p_1. *Assume that* $p_0 < \frac{1}{2} < p_1$*; the bias* p_0 *corresponds to* $\theta = 0$*, while* p_1 *corresponds to* $\theta = 1$. *A priori the two biases are equally likely.*

Each agent is allowed to privately flip the coin once and observe the outcome. Therefore our signal space is $S = \{H, T\}$. *Assume that agent* n *observes heads.*

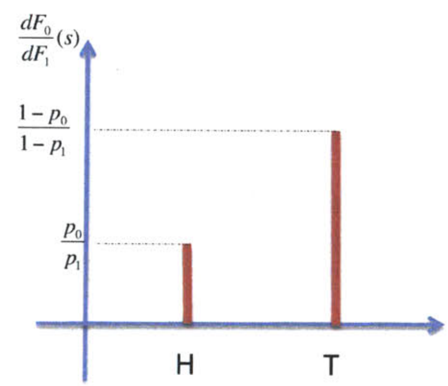


Figure 2-2: Likelihood ratios for the coin tossing example

Then,

$$\frac{d\mathbb{F}_0}{d\mathbb{F}_1}(H) = \frac{p_0}{p_1}.$$

Similarly,

$$\frac{d\mathbb{F}_0}{d\mathbb{F}_1}(T) = \frac{1-p_0}{1-p_1}.$$

Obviously, the likelihood ratios are bounded in this example and the conditions of Definition 1 are met with $M = (1-p_0)/(1-p_1)$ *and* $m = p_0/p_1$.

Example 2. *Consider a scenario where agents have an risky investment option that returns a reward from a normal distribution with unit variance, but unknown mean*

μ. Let μ take one of two possible values -1 or 1. Let the case $\mu = -1$ correspond to $\theta = 0$ and $\mu = 1$ to $\theta = 1$. The agents want to determine whether the investment is good or bad; namely whether the mean is $+1$ or -1 by just observing the payoffs from trying it.

In this case,

$$\frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s) = \frac{exp(-\frac{(s+1)^2}{2})}{exp(-\frac{(s-1)^2}{2})}.$$

One can observe that $\lim_{s\to\infty} d\mathbb{F}_0/d\mathbb{F}_1(s) = 0$ while $\lim_{s\to-\infty} d\mathbb{F}_0/d\mathbb{F}_1(s) = +\infty$ establishing that in this case the private signal distribution is characterized by unbounded likelihood ratios.

The fact that for signal structures characterized by Unbounded Likelihood Ratios an agent can receive arbitrarily strong signals about the underlying state, in an intuitive level, suggests that beliefs about the underlying state evolve faster and it is expected that agents who exploit their information reasonably can eventually learn the correct state of the world. At some points in time, arbitrarily strong signals arrive that indicate the underlying state with high certainty, leading to more informed decisions and eventually correct actions.

On the other hand, for signal structures characterized by Bounded Likelihood Ratios, it appears to be less intuitive that, especially with finite memory, there exists a way of exploiting available information that can lead to certainty about the underlying state. As we shall see later in this thesis, this intuition may, surprisingly, not be valid.

However, in order to deal with these issues, we need to provide a precise definition of "learning", which is done in the next section.

2.2 Almost sure learning versus learning in probability

In this section we shortly present the two modes of learning studied in this thesis. As we shall see, we obtain significantly different results based on the mode of learning under consideration.

Definition 3. *We say that the decision profile $d = \{d_n\}_{n=1}^{\infty}$ achieves **almost sure learning** if*

$$\lim_{n\to\infty} x_n = \theta, \;\; w.p.1.$$

We also investigate a looser mode of convergence, i.e., *learning in probability.*

Definition 4. *We say that the decision rule* $d = \{d_n\}_{n=1}^{\infty}$ *achieves* ***learning in probability*** *if*

$$\lim_{n\to\infty} \mathbb{P}_d(x_n = \theta) = 1.$$

2.3 Designed decision rules versus strategic agents

Chapters 3 and 4 will be devoted to providing a benchmark for the two modes of learning defined in the previous subsection. We study whether there are decision rules that achieve almost sure learning and learning in probability, respectively. For the case where it is possible to do so, learning relies on the presence of an engineer or a social planner who can design the agents' decision rules. This approach is applicable to a sensor network scenario, where sensors could just be programmed to act according to the programmer's will.

In a social network scenario, on the other hand, it is not reasonable to make such an assumption. There, agents are modeled as being strategic; each individual is assumed to take the action that maximizes her payoff. In this case, behaviors and strategies rise as equilibria of the corresponding games. Traditionally, in this trend of the literature, agents have been using the probability of making a correct decision as a metric for the quality of their decisions. In particular, the payoff of agent n typically is

$$u_n(x_n, \theta) = \begin{cases} 1, & \text{if } x_n = \theta \\ 0, & \text{if } x_n \neq \theta. \end{cases}$$

For the case $K = 1$, Cover [7] asserted that under the Bounded Likelihood Ratio assumption and myopic decision rules, there is no learning in probability. [1] A (rather complicated) proof was provided in [14]. It has been shown by Acemoglu *et al.* [1] that the strategy profile which emerges as a Perfect Bayesian Equilibrium of this game achieves learning in probability for the case of Unbounded Likelihood Ratios while learning in probability is not achieved if the private signal structure is characterized by Bounded Likelihood Ratios for any value of K.

This payoff structure fails to achieve learning in probability under Bounded Likelihood Ratios because of the creation of *herds*, a term introduced and studied in [3] and [4]. When the probability of error becomes large enough, agents start copying

[1]The exact statement reads: "In the Bayesian formulation, where prior probabilities are associated with the two hypotheses, the rule which stores the Bayes decision at each stage will not learn. Eventually the posterior probabilities will be such that no single observation will yield a change in the decision."

the actions of their predecessors just because they could not do better by trusting their private signal (which has bounded informativeness) if it indicated the opposite decision. This phenomenon prevents learning in probability, in contrast to the Unbounded Likelihood Ratio case.

We discuss another metric for the quality of the decisions, which assumes forward looking agents. This model is introduced by Smith and Sorensen [18] but under a different observation model; there agents observe the whole history of decisions. The payoff of agent n depends on the underlying state of the world θ, her decision, as well as the decision of all subsequent agents, and is given by

$$u_n(\mathbf{x}_n^+, \theta) = (1-\delta) \sum_{k=n}^{\infty} \delta^{k-n} \mathbf{1}_{x_k = \theta}, \tag{2.2}$$

where $\delta \in (0,1)$ is a discount factor, $\mathbf{x}_n^+ = \{x_{n+k}\}_{k=0}^{\infty}$, and $\mathbf{1}_A$ denotes the indicator random variable for the event A. In other words, each agent not only cares for herself to make a correct decision, but also takes into account the influence of this decision on the future actions.

One can directly observe that if $\delta < 1/2$, then the immediate payoff, namely the payoff for making a correct decision herself, overcomes the continuation payoff, the payoff that she gains if the future agents decide correctly. On the other hand if $\delta > 1/2$, then the continuation payoff dominates and thus we could hope for equilibria that would favour escaping from herds leading to learning in probability. Chapter 5 will prove this intuition wrong for any value of δ.

Chapter 3

Designing decision rules for almost sure learning

In this chapter we study almost sure learning and establish necessary and sufficient conditions for this to occur. Traditionally, in the existing literature, the dichotomy that arises, as far as learning is concerned is that between Bounded and Unbounded Likelihood Ratios ([1], [14]). In agreement with previous results, in this chapter we prove that almost sure learning is possible if and only if the signal structure is characterized by Unbounded Likelihood Ratios.

3.1 Unbounded likelihood ratios: Cover's construction

We first consider the case of Unbounded Likelihood Ratios, a problem that has been studied by Cover in [7].

Cover considers the case where $K = 1$ and provides a decision rule that achieves almost sure learning. That decision rule is as follows:

$$d_n(s_n, x_{n-1}) = \begin{cases} 1, & \text{if } \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s_n) > \bar{l}_n, \\ 0, & \text{if } \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s) < \underline{l}_n, \\ x_{n-1}, & \text{otherwise.} \end{cases}$$

where $\bar{l}_n$ and $\underline{l}_n$ are suitably chosen thresholds. These thresholds $\bar{l}_n$ and $\underline{l}_n$ are non-increasing and non-decreasing sequences, converging at a suitable rate to infinity and zero, respectively.

The fact that Cover's decision rule achieves almost sure learning is in agreement with the intuitive discussion from the previous section. Specifically, an agent makes a decision different from that of the previous agent only when her signal is strong enough in either direction. In the beginning, agents need not have a strong signal to change the decision and report what they observed. But as n increases, the thresholds approach the endpoints of the support of the likelihood ratios, namely zero and infinity, and one needs an extremely strong signal in order to switch. In other words, for large n, agents copy their predecessor except when they have an extremely strong indication for the opposite. Since arbitrarily strong signals in favour of the underlying state arrive almost surely, learning can be established.

Such a decision rule would not achieve learning in the case of Bounded Likelihood Ratios for the following simple reason. As $\bar{l}_n$ and $\underline{l}_n$ approach infinity and zero, respectively, at some point they go outside the interval $[m, M]$ (cf. Definition 1). Thus eventually, for some n^*, $\bar{l}_{n^*} > M$ and $\underline{l}_{n^*} < m$. For every $n > n^*$ agent n will be copying the decision of n^* and hence will be taking the wrong action with positive probability.

We provide a stronger result in the next section; no decision profile can achieve almost sure learning under Bounded Likelihood Ratios.

3.2 No almost sure learning for the case of Bounded Likelihood Ratios

If the signal structure is characterized by Bounded Likelihood Ratios, it is known from [11] that for the case of tossing a coin with unknown bias and if $K = 1$, then there does not exist a decision rule that achieves almost sure learning. Later, [12] proves that for agents who maximize a local cost function (essentially maximizing the probability of a correct decision) after observing their immediate predecessor, but are allowed to make three-valued decision, almost sure learning does not occur.

We provide a more general result arguing that almost sure learning cannot be achieved by *any* decision rule under *any* signal structure that is characterized by Bounded Likelihood Ratios, for *any* number of observed predecessors. Actually, our proof generalizes for any observation structure, allowing access to any subset of past agents' decisions.

This subsection will be devoted to proving the following Theorem.

Theorem 1. *Under the Bounded Likelihood Ratio assumption and for any number*

K of neighbours being observed, there **does not** *exist a decision rule* $d = \{d_n\}_{n=1}^{\infty}$ *that achieves almost sure learning.*[1]

Throughout this section, let D_n denote the set of actions observed by agent n. For the case of observing the last K predecessors, we have $D_n = \{n-K, \ldots, n-1\}$. For brevity, let $\mathbf{x}_{D_n}$ be the vector of observed actions, that is,

$$\mathbf{x}_{D_n} \triangleq \{x_k : k \in D_n\}.$$

A decision rule $d = \{d_n\}_{n=1}^{\infty}$ specifies for any observed vector of actions $\mathbf{x}_{D_n}$ a partition of the signal space S in two subsets $S_n^d(\mathbf{x}_{D_n})$ and $\bar{S}_n^d(\mathbf{x}_{D_n}) = S \setminus S_n^d(\mathbf{x}_{D_n})$ such that

$$d_n(s_n, \mathbf{x}_{D_n}) = \begin{cases} 1, & \text{if } s_n \in S_n^d(\mathbf{x}_{D_n}) \\ 0, & \text{if } s_n \in \bar{S}_n^d(\mathbf{x}_{D_n}) \end{cases}$$

For example, in Cover's decision rule (denoted by d^* here),

$$S_n^{d^*}(1) = \{s \in S : \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s) \geq \underline{l}_n\}$$

and

$$S_n^{d^*}(0) = \{s \in S : \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s) > \bar{l}_n\}.$$

From now on, let us fix a specific decision profile d so that we can drop the subscripts and superscripts related to a specific rule, for illustration purposes For notational convenience, let $\mathbb{P}^j(\cdot)$ denote conditioning on state of the world j, that is $\mathbb{P}^j(\cdot) = \mathbb{P}(\cdot|\theta = j)$.

The following Lemma, is central for the analysis that ensues.

Lemma 1. *For any* $\mathbf{x}_{D_n}$*, we have*

$$m < \frac{\mathbb{P}^0(x_n = j \mid \mathbf{x}_{D_n})}{\mathbb{P}^1(x_n = j \mid \mathbf{x}_{D_n})} < M, \quad j \in \{0, 1\}. \tag{3.1}$$

where m *and* M *are scalars defined in Definition 1.*

Proof. We give the proof for the case where $j = 1$. The other case follows from a symmetrical argument.

[1] The proof of this theorem does not use anywhere the fact that agents only observe the last K immediate predecessors. The result can be directly generalized to agents who observe the actions of any subset of the first $n-1$ agents.

The probability of agent n choosing 1 when she observes $\mathbf{x}_{D_n}$ under state of the world 0 is the probability of her private signal lying in $S_n(\mathbf{x}_{D_n})$ under state 0, i.e.,

$$\mathbb{P}^0(x_n = 1 \mid \mathbf{x}_{D_n}) = \int_{s_n \in S_n(\mathbf{x}_{D_n})} d\mathbb{F}_0(s_n) = \int_{s_n \in S_n(\mathbf{x}_{D_n})} \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s_n) d\mathbb{F}_1(s_n).$$

Using Definition 1 we obtain

$$\begin{aligned} \mathbb{P}^0(x_n = 1 \mid \mathbf{x}_{D_n}) &= \int_{s_n \in S_n(\mathbf{x}_{D_n})} \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s_n)\, d\mathbb{F}_1(s_n) < M \int_{s_n \in S_n(\mathbf{x}_{D_n})} d\mathbb{F}_1(s_n), \\ &= M \cdot \mathbb{P}^1(x_n = 1 \mid \mathbf{x}_{D_n}), \end{aligned}$$

and similarly,

$$\begin{aligned} \mathbb{P}^0(x_n = 1 \mid \mathbf{x}_{D_n}) &= \int_{s_n \in S_n(\mathbf{x}_{D_n})} \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(s_n)\, d\mathbb{F}_1(s_n) > m \int_{s_n \in S_n(\mathbf{x}_{D_n})} d\mathbb{F}_1(s_n), \\ &= m \cdot \mathbb{P}^1(x_n = 1 \mid \mathbf{x}_{D_n}), \end{aligned}$$

from which the result follows. □

This Lemma says that the probabilities of making a decision given the observed actions are coupled between the two states of the world, under the BLR assumption. Therefore, if under one state of the world some agent n after observing $\mathbf{x}_{D_n}$ decides 0 with positive probability, then the same has to occur with proportional probability under the other state of the world. This proportional dependence of decision probabilities for the two possible underlying states is central to the proof of Theorem 1.

Before going to the main proof, we need two more lemmas. Consider a probability space $(\Omega, C, \mathbb{P})$ and a sequence of events $\{E_k\}$, $k = 1, 2, \ldots$. The upper limiting set of the sequence $\limsup_{k\to\infty} E_k$ is defined by

$$\limsup_{k\to\infty} E_k = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k.$$

The next, stronger version of Borel-Cantelli lemma that does not require independence of events will be used.

Lemma 2. *Consider a probability space $(\Omega, C, \mathbb{P})$ and a sequence of events $\{E_k\}$, $k = 1, 2, \ldots$. If*

$$\sum_{k=1}^{\infty} \mathbb{P}(E_k \mid E_1' \ldots E_{k-1}') = \infty,$$

then

$$\mathbb{P}(\limsup_{k\to\infty} E_k) = 1,$$

where E_k' denotes the complement of E_k.

Proof. See p. 48 of [5] or [6]. □

Finally, we prove an algebraic fact with a simple probabilistic interpretation.

Lemma 3. *Consider a sequence $\{a_n\}_{n\in\mathbb{N}}$ of real numbers $a_n \in [0,1]$, for all $n \in \mathbb{N}$. Then*

$$1 - \sum_{n\in Q} a_n \leq \prod_{n\in Q}(1-a_n) \leq e^{-\sum_{n\in Q} a_n},$$

for any $Q \subseteq \mathbb{N}$.

Proof. The second inequality is standard. For the first one, interpret the numbers $\{a_n\}_{n\in\mathbb{N}}$ as probabilities of independent events $\{E_n\}_{n\in\mathbb{N}}$. Then, clearly,

$$\mathbb{P}(\bigcup_{n\in Q} E_n) + \mathbb{P}(\bigcap_{n\in Q} E_n^c) = 1.$$

Observe that

$$\mathbb{P}(\bigcup_{n\in Q} E_n) = \prod_{n\in Q}(1-a_n),$$

and by the union bound,

$$\mathbb{P}(\bigcap_{n\in Q} E_n^c) \leq \sum_{n\in Q} a_n.$$

Combining the above yields the desired result. □

Now, we are ready to prove the main result of this chapter.

Proof of Theorem 1. Let V denote the set of all sequences that end up in ones, namely

$$V \triangleq \{v \in \{0,1\}^{\mathbb{N}} : \text{ there exists some } N \in \mathbb{N} \text{ such that } v_n = 1 \text{ for all } n > N\}.$$

Observe that V can be equivalently written as

$$V = \bigcup_{N\in\mathbb{N}} V_N, \text{ where } V_N \triangleq \{v \in \{0,1\}^{\mathbb{N}} : v_n = 1 \text{ for all } n \geq N\}.$$

Each of the sets V_N is finite, since it contains 2^N elements, and hence V is countable, as a countable union of countable sets. Therefore, we can enumerate the elements of V by the positive integers and write $V = \{v^i\}_{i \in \mathbb{N}}$.

We argue by contradiction. Suppose that d achieves learning with probability one. Then,

$$\mathbb{P}^1(\{x_k\}_{k=1}^\infty \in V) = 1,$$

or equivalently,

$$\mathbb{P}^1(\{x_k\}_{k=1}^\infty = v^i \text{ for some } i\ ) = 1. \tag{3.2}$$

That is, almost surely, there exists some (random) N after which all agents $n > N$ decide $x_n = 1$ under the state of the world one.

Let now $\hat{V}$ be defined as follows :

$$\hat{V} = \{v \in V : \mathbb{P}^1(\{x_k\}_{k=1}^\infty = v) > 0\}.$$

It follows from Equation (3.2) that $\hat{V} \neq \emptyset$. We will prove that $\hat{V} = \emptyset$, thus obtaining a contradiction. Since $\hat{V} \subset V$ and $\hat{V} \neq \emptyset$, we will look for elements of V from within $\hat{V}$.

Fix some $i \in \mathbb{N}$. Let $v^i \in \hat{V}$. Then,

$$\mathbb{P}^1(x_k = v^i_k \text{ for all } k \leq n) > 0, \text{ for all } n \in \mathbb{N}. \tag{3.3}$$

The above implies that

$$\mathbb{P}^0(x_k = v^i_k \text{ for all } k \leq n) > 0, \text{ for all } n \in \mathbb{N}. \tag{3.4}$$

Indeed, assume the contrary and let

$$\hat{n} = \min\{n \in \mathbb{N} : \mathbb{P}^0(x_k = v^i_k \text{ for all } k \leq n) = 0\}.$$

Then, $\mathbb{P}^0(x_k = v^i_k \text{ for all } k \leq \hat{n} - 1) > 0$ and

$$\mathbb{P}^0(x_{\hat{n}} = v^i_{\hat{n}} \mid x_k = v^i_k \text{ for all } k \leq \hat{n} - 1) = 0,$$

which in turn, using Lemma 1, implies that

$$\begin{aligned} 0 &\leq \mathbb{P}^1(x_{\hat{n}} = v^i_{\hat{n}} \mid x_k = v^i_k \text{ for all } k \leq \hat{n} - 1) \\ &\leq \frac{1}{m}\mathbb{P}^0(x_{\hat{n}} = v^i_{\hat{n}} \mid x_k = v^i_k \text{ for all } k \leq \hat{n} - 1) = 0, \end{aligned}$$

which contradicts (3.3).

Define a_n^i, b_n^i as follows:

$$a_n^i = \mathbb{P}^1(x_n \neq v_n^i \mid x_k = v_k^i \text{ for all } k < n),$$

$$b_n^i = \mathbb{P}^0(x_n \neq v_n^i \mid x_k = v_k^i \text{ for all } k < n),$$

and observe that Lemma 1 implies that

$$m \leq \frac{b_n^i}{a_n^i} \leq M, \tag{3.5}$$

because $\mathbb{P}^j(x_n \neq v_n^i \mid x_k = v_k^i \text{ for all } k < n) = \mathbb{P}^j(x_n \neq v_n^i \mid \mathbf{x}_{D_n} = \mathbf{v}_{D_n})$ for $j \in \{0,1\}$.

We claim that

$$\sum_{n=1}^{\infty} a_n^i = \infty \tag{3.6}$$

Indeed, assume the contrary, namely that

$$\sum_{n=1}^{\infty} a_n^i < \infty.$$

Then,

$$\sum_{n=1}^{\infty} b_n^i < M \cdot \sum_{n=1}^{\infty} a_n^i < \infty,$$

and

$$\lim_{N\to\infty} \sum_{n=N}^{\infty} b_n^i = \lim_{N\to\infty} \sum_{n=N}^{\infty} \mathbb{P}^0(x_n \neq v_n^i \mid x_k = v_k^i \text{ for all } k < n) = 0. \tag{3.7}$$

Choose some $\hat{N}$ such that

$$\sum_{n=\hat{N}}^{\infty} \mathbb{P}^0(x_n \neq v_n^i \mid x_k = v_k^i \text{ for all } k < n) < \frac{1}{2}.$$

Such a $\hat{N}$ exists because of (3.7). Then,

$$\begin{aligned}
&\mathbb{P}^0(\{x_k\}_{k\in\mathbb{N}} = v^i) \\
&\quad = \mathbb{P}^0(x_n = v_n^i \text{ for all } n \leq \hat{N}) \cdot \prod_{n=\hat{N}}^{\infty} (1 - \mathbb{P}^0(x_n \neq v_n^i \mid x_k = v_k^i \text{ for all } k < n)).
\end{aligned}$$

The first term of the right hand side is bounded away from zero by (3.4) while

$$\prod_{n=\hat{N}}^{\infty}(1-\mathbb{P}^0(x_n \neq v_n^i \mid x_k = v_k^i \text{ for all } k < n)) \geq \left(1 - \sum_{n=\hat{N}}^{\infty}\mathbb{P}^0(x_n \neq v_n^i \mid x_k = v_k^i \text{ for all } k < n)\right).$$

Combining the above, we obtain

$$\mathbb{P}^0(\{x_k\}_{k\in\mathbb{N}} = v^i) > 0,$$

which is a contradiction to the almost sure learning assumption, establishing (3.6).

We now show that if $\sum_{n=1}^{\infty} a_n^i = \infty$ then v^i cannot belong to $\hat{V}$. Indeed, using Lemma 2 we get that $\mathbb{P}^1(\limsup\{x_k \neq v_k^i\}) = 1$ showing that with probability one a deviation from the sequence v^i will happen under the state of the world one and therefore that v^i cannot belong to $\hat{V}$. This is a contradiction and concludes the proof. □

3.3 Discussion and Conclusions

In this chapter we completed the results of the existing literature as far as almost sure learning is concerned. It was known from [7] that there exists a decision profile which achieves almost sure learning for any $K \geq 1$ if the signal structure is characterized by Unbounded Likelihood Ratios. Our results strengthens this theorem making it an if and only if statement; there exists a decision profile that achieves almost sure learning for any $K \geq 1$ **if and only if** the signal structure is characterized by Unbounded Likelihood Ratios.

Chapter 4

Designing decision rules for learning in probability

In this chapter we discuss a looser learning mode, *learning in probability*, as defined in Section 2.2. We prove that learning in probability cannot be achieved when agents observe their immediate predecessor, i.e., $K = 1$. In contrast, we design a decision profile that achieves learning in probability when $K \geq 2$.

Koplowitz [11] showed that if $K = 1$ then, for the problem of tossing a coin with unknown bias, learning in probability cannot be achieved. Moreover [12] proves that if agents take a three valued action that maximizes their local payoff function, then learning in probability does not occur. In contrast to [11] we prove the result for any signal distribution that is characterized by Bounded Likelihood Ratios. On the other hand, our result does not imply that of [12] since we only allow binary actions but it applies in a broader setting since we consider all possible decision profiles.

4.1 No learning in probability when K=1

We start with the case where $K = 1$. Agents observe their immediate predecessor and the decision process can be described in terms of two Markov chains, one for each possible state of the world, as depicted in Figure 4-1. Indeed, consider a two-state Markov chain where the state corresponds to the observed action $x_{n-1} \in \{0, 1\}$. A transition from state i to state j for the Markov chain associated with $\theta = l$, where $i, j, l \in \{0, 1\}$, corresponds to agent n taking the action j given that her immediate predecessor $n-1$ decided i under the state $\theta = l$. Indeed, the Markov property is satisfied since agents' decisions, conditioned on the immediate predecessor's action, are independent from the history of actions. In other words, for any strategy profile

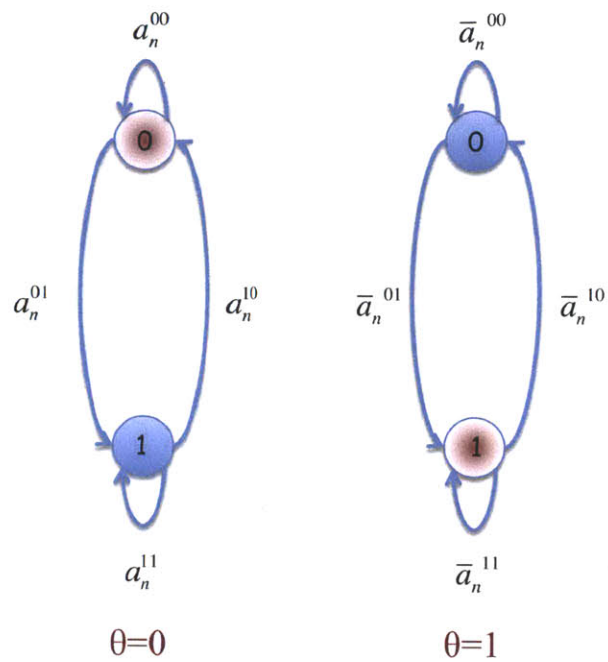


Figure 4-1: The Markov chains that model the decision process for $K = 1$. States correspond to observed actions while transitions correspond to agents' decisions.

d,

$$\mathbb{P}_d(x_n = j \mid x_{n-1}, \ldots, x_1, \theta = i) = \mathbb{P}_d(x_n = j \mid x_{n-1}, \theta = i), \quad i, j \in \{0, 1\}.$$

We denote the transition probabilities by

$$a_n^{ij}(d) \triangleq \mathbb{P}_d(x_n = j \mid x_{n-1} = i, \theta = 0), \quad i, j \in \{0, 1\}, \tag{4.1}$$

$$\bar{a}_n^{ij}(d) \triangleq \mathbb{P}_d(x_n = j \mid x_{n-1} = i, \theta = 1), \quad i, j \in \{0, 1\}. \tag{4.2}$$

Observe that using the notation from the previous chapter we have

$$a_n^{i1}(d) = \mathbb{P}^0(s_n \in S_n^d(i))$$

and a corresponding expression for all other transitions. Thus, we can deduce from Lemma 1 the following corollary.

Corollary 1. *Consider a strategy profile d and let $m > 0$ and $M < \infty$ be as defined* *in Definition 1. Then,*

$$m \leq \frac{a_n^{ij}(d)}{\bar{a}_n^{ij}(d)} \leq M, \quad \textit{for all } i, j \in \{0, 1\}.$$

In this subsection we establish the following impossibility result.

Theorem 2. *Assume that the signal structure is characterized by BLR and let agents observe their immediate predecessors ($K = 1$). Then, there does not exist a decision profile d that achieves learning in probability. Equivalently, for any decision profile $d = \{d_n\}_{n=1}^{\infty}$ we have*

$$\lim_{n \to \infty} \mathbb{P}_d(x_n = \theta) < 1.$$

Once more the coupling between the probabilities of taking an action given the observed decision under the two states of the world implied by Corollary 1 is key for this result. To simplify exposition our proof will be presented through a series of lemmas. For the rest of this section, we fix some decision profile d and for notational convenience we suppress the dependence on d.

The first lemma, which is directly obtained from the Bounded Likelihood Ratios property, couples the states of the Markov chains associated with the two states of the world, after a finite number of transitions.

Lemma 4. *Assume that the signal structure is characterized by Bounded Likelihood Ratios. Then, for any $j \in \{0, 1\}$ and $n \in \mathbb{N}$,*

$$m^n \leq \frac{\mathbb{P}^0(x_n = j)}{\mathbb{P}^1(x_n = j)} \leq M^n. \tag{4.3}$$

Proof. We will use induction on n. For $n = 1$, (4.3) holds from Lemma 1. Assume that (4.3) holds for some $n \in \mathbb{N}$ and for all $j \in \{0, 1\}$. Then,

$$\frac{\mathbb{P}^0(x_n = j)}{\mathbb{P}^1(x_n = j)} = \frac{\mathbb{P}^0(x_n = j \mid x_{n-1} = 0)\mathbb{P}^0(x_{n-1} = 0) + \mathbb{P}^0(x_n = j \mid x_{n-1} = 1)\mathbb{P}^0(x_{n-1} = 1)}{\mathbb{P}^1(x_n = j \mid x_{n-1} = 0)\mathbb{P}^1(x_{n-1} = 0) + \mathbb{P}^1(x_n = j \mid x_{n-1} = 1)\mathbb{P}^1(x_{n-1} = 1)}.$$

Using the induction hypothesis and Lemma 1, we obtain

$$\begin{aligned}
&\frac{\mathbb{P}^0(x_n = j)}{\mathbb{P}^1(x_n = j)} \\
&\quad \leq \frac{M\mathbb{P}^1(x_n = j \mid x_{n-1} = 0)M^n\mathbb{P}^1(x_{n-1} = 0) + M\mathbb{P}^1(x_n = j \mid x_{n-1} = 1)M^n\mathbb{P}^1(x_{n-1} = 1)}{\mathbb{P}^1(x_n = j \mid x_{n-1} = 0)\mathbb{P}^1(x_{n-1} = 0) + \mathbb{P}^1(x_n = j \mid x_{n-1} = 1)\mathbb{P}^1(x_{n-1} = 1)} \\
&\quad = M^{n+1}.
\end{aligned}$$

The lower bound follows using a similar argument concluding the induction and the proof of the lemma. □

We next establish that for learning in probability to occur, transitions between different states should not stop at some finite n. The intuitive argument for this result is expressed as follows. Assume that transitions stopped in favor of one of the states in some finite time under one state of the world. Then, the same would happen under the other state of the world, by Corollary 1, contradicting learning in probability. The next lemma formalizes that intuition.

Lemma 5. *Assume that learning in probability occurs. Let* $A_{ij} = \{n : a_n^{ij} > 0\}$ *and* $\bar{A}_{ij} = \{n : \bar{a}_n^{ij} > 0\}$. *Then,* $|A_{ij}| = |\bar{A}_{ij}| = \infty$, *for all* $i, j \in \{0, 1\}$, *with* $i \neq j$, *where* $|\cdot|$ *denotes the cardinality of a set.*

Proof. The first implication is straightforward. If for some $n \in \mathbb{N}$ we have $a_n^{ij} = 0$, then $0 \leq \bar{a}_n^{ij} \leq M \cdot 0$, establishing that $\bar{a}_n^{ij} = 0$ which yields $|A_{ij}| = |\bar{A}_{ij}|$.

Consider, first, the case where $|\bar{A}_{01}| < \infty$. Then, we can distinguish between the following cases.

(i) $|\bar{A}_{10}| = \infty$.

Then, there exists some $\hat{n} > |\bar{A}_{01}|$ for which

$$\mathbb{P}^1(x_{\hat{n}} = 0) > 0,$$

and for all $n > \hat{n}$,

$$\mathbb{P}^1(x_n = 1 \mid x_{n-1} = 0) = 0.$$

Therefore, for all $n > \hat{n}$,

$$\mathbb{P}^1(x_n = 0) \geq \mathbb{P}^1(x_n = 0 \mid x_{\hat{n}} = 0)\mathbb{P}^1(x_{\hat{n}} = 0) = \mathbb{P}^1(x_{\hat{n}} = 0),$$

which in turn yields

$$\liminf_{n\to\infty} \mathbb{P}(x_n \neq \theta) \geq \frac{1}{2}\mathbb{P}^1(x_{\hat{n}} = 0) > 0,$$

contradicting learning in probability.

(ii) $|A_{10}| = |\bar{A}_{10}| < \infty$

Without loss of generality, assume that $|A_{01} > A_{10}|$. In that case, transitions between states stop at some finite time $\hat{n} > \max\{|A_{10}|, |A_{01}|\}$ and therefore, by the learning in probability assumption,

$$\mathbb{P}^1(x_n = 1) = \mathbb{P}^1(x_{\hat{n}} = 1) = 1, \text{ for all } n > \hat{n}.$$

On the other hand, only finitely many transitions have occurred and thus, from Lemma 4,

$$m^{\hat{n}} \leq \frac{\mathbb{P}^0(x_{\hat{n}} = 1)}{\mathbb{P}^1(x_{\hat{n}} = 1)} \leq M^{\hat{n}},$$

from which we have

$$\mathbb{P}^0(x_n = 1) = \mathbb{P}^0(x_{\hat{n}} = 1) \geq m^{\hat{n}}\mathbb{P}^0(x_{\hat{n}} = 1) > 0, \text{ for all } n > \hat{n}.$$

contradicting the learning in probability assumption.

Symmetric arguments can be used for the case $|\bar{A}_{10}| < \infty$, concluding the proof. □

A consequence of the previous Lemma is the following corollary, which states that there cannot be learning in finite time, i.e., that the state of the Markov chain can take any of the two values infinitely many times with positive probability.

Corollary 2. *Assume that the decision d achieves learning in probability. For all $n_o \in N$, there exists some $n > n_o$ such that*

$$\mathbb{P}^\theta(x_n = j) > 0, \text{ for all } j, \theta \in \{0, 1\}$$

.

The next step is to derive some properties of the transition probabilities under the d that achieves learning. The first result extends the previous corollary and proves that transitions between states occur infinitely many times.

Lemma 6. *Assume that the decision d achieves learning in probability. Then,*

$$\sum_{n=1}^{\infty} a_n^{01} = \infty, \tag{4.4}$$

and

$$\sum_{n=1}^{\infty} a_n^{10} = \infty \tag{4.5}$$

Proof. For the sake of contradiction, assume that $\sum_{n=1}^{\infty} a_n^{01} < \infty$. Then,

$$\lim_{N\to\infty} \sum_{n=N}^{\infty} a_n^{01} = 0.$$

Therefore there exists some $\hat{N} \in \mathbb{N}$ such that

$$\sum_{n=\hat{N}}^{\infty} a_n^{01} < \frac{1}{2M}.$$

Corollary 2 guarantees that there exists some $\tilde{N} > \hat{N}$ such that $\mathbb{P}^1(x_{\tilde{N}} = 0) > 0$. Since $\tilde{N} > \hat{N}$, we have

$$\sum_{n=\hat{N}}^{\infty} a_n^{01} < \frac{1}{2M}.$$

Using Corollary 1 we get

$$\sum_{n=\hat{N}}^{\infty} \bar{a}_n^{01} \leq M \cdot \sum_{n=\hat{N}}^{\infty} a_n^{01} \leq \frac{1}{2}.$$

Moreover, for all $n > \tilde{N}$ we have

$$\mathbb{P}(x_n = 0 \mid x_{\tilde{N}} = 0) = \prod_{k=\tilde{N}+1}^{n} (1 - \bar{a}_k^{01}),$$

and thus,

$$\mathbb{P}^1(x_n = 0) \geq \mathbb{P}^1(x_{\tilde{N}+1} = 0) \cdot \prod_{k=\tilde{N}}^{n} (1 - \bar{a}_k^{01}).$$

On the other hand,

$$\prod_{k=\tilde{N}+1}^{n} (1 - \bar{a}_k^{01}) \geq 1 - \sum_{n=\tilde{N}+1}^{+\infty} \bar{a}_n^{01} \geq 1 - \frac{1}{2} = \frac{1}{2},$$

which leaves us with,

$$\mathbb{P}^1(x_n = 0) \geq \frac{1}{2} \cdot \mathbb{P}^1(x_{\tilde{N}} = 0),$$

for all $n > \tilde{N}$.

Therefore,

$$\liminf_{n\to\infty} \mathbb{P}^1(x_n = 0) \geq \frac{1}{2} \cdot \mathbb{P}^1(x_{\tilde{N}} = 0) > 0$$

contradicting the asymptotic learning assumption and thus concluding the proof. ☐

The last lemma states that transition probabilities between different states should asymptotically converge to zero, for a decision that achieves learning in probability. Intuitively, if this were not the case, even if agents' decisions converged to the correct state, infinitely often a transition to the other state would occur with positive probability, preventing learning in probability.

Lemma 7. *Assume that the decision d achieves learning in probability. Then*

$$\lim_{n\to\infty} a_n^{01} = 0. \tag{4.6}$$

Proof. Assume, to arrive to a contradiction that there exists some $\epsilon \in (0,1)$ and a subsequence $\{r_k\}_{k=1}^{\infty}$ such that

$$a_{r_k}^{01} = \mathbb{P}^0(x_{r_k} = 1 \mid x_{r_k-1} = 0) > \epsilon, \tag{4.7}$$

for all $k \in \mathbb{N}$.

The learning assumption implies that there exists $N \in \mathbb{N}$ such that for all $n > N$,

$$\mathbb{P}^0(x_n = 0) > 1 - \frac{\epsilon}{4}. \tag{4.8}$$

Let $\hat{k}$ be such that $r_{\hat{k}} > N$. Then,

$$\mathbb{P}^0(x_{r_{\hat{k}}} = 0) = \mathbb{P}^0(x_{r_{\hat{k}}} = 0 \mid x_{r_{\hat{k}}-1} = 0)\mathbb{P}^0(x_{r_{\hat{k}}-1} = 0) + \mathbb{P}^0(x_{r_{\hat{k}}} = 0 \mid x_{r_{\hat{k}}-1} = 1)\mathbb{P}^0(x_{r_{\hat{k}}-1} = 1).$$

Using (4.7) we get

$$\mathbb{P}^0(x_{r_{\hat{k}}} = 0 \mid x_{r_{\hat{k}}-1} = 0) \leq 1 - \epsilon,$$

while using (4.8) we get

$$\mathbb{P}^0(x_{r_{\hat{k}}-1} = 1) \leq \frac{\epsilon}{4}.$$

Combining the above we get that

$$\mathbb{P}^0(x_{r_{\hat{k}}} = 0) \leq 1 - \epsilon \cdot 1 + 1 \cdot \frac{\epsilon}{4} = 1 - \frac{3\epsilon}{4},$$

which contradicts (4.8), completing the proof.

☐

At this point we are ready to prove the main theorem of this section.

Since the sum of transition probabilities from state 0 to state 1 is infinite, we can divide the agents into blocks so that the corresponding sums over each block are approximately constant. If during a block the sum of transition probabilities from state 1 to state 0 is small, then under state of the world 0, there is high probability of starting the block at state 0 and ending at state 1. If on the other hand the sum of the transition probabilities from state 1 to state 0 is large, then under the state of the world one, there is high probability of starting the corresponding block at state 1 and ending at state 0. Both cases prevent actions' convergence in probability to the correct state. The main idea is illustrated in Figure 4-2.

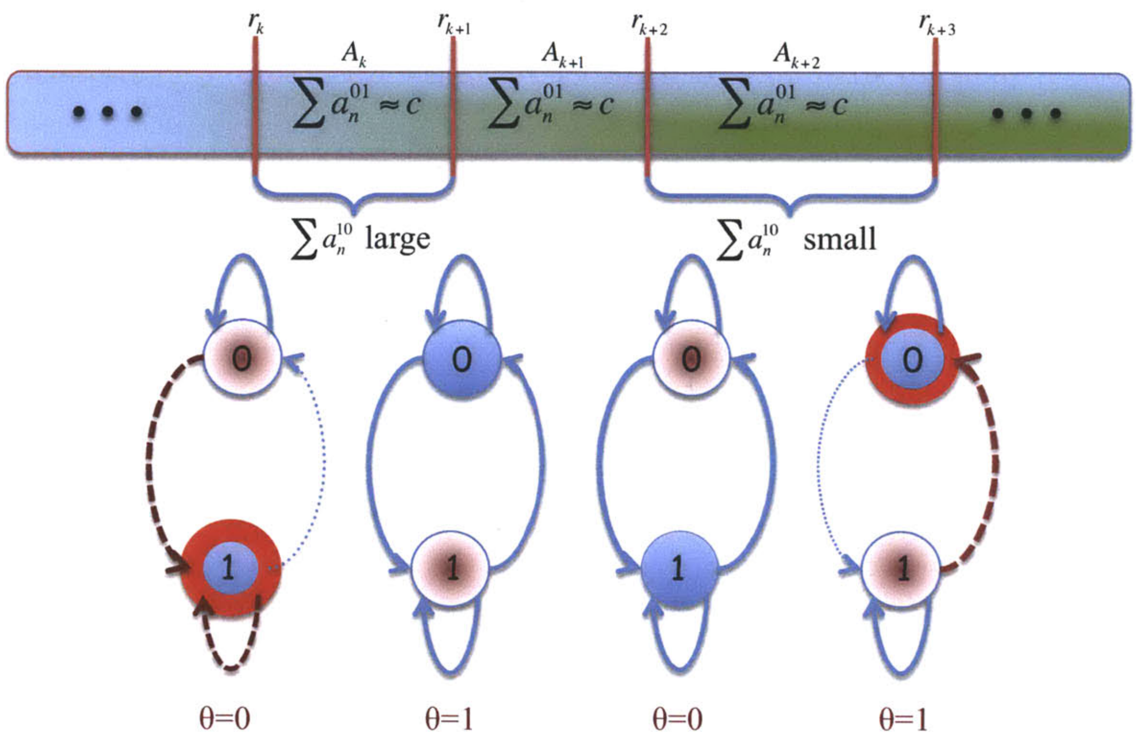


Figure 4-2: Proof sketch for theorem 2. Divide agents into blocks so that the sum, in each block, of transition probabilities from 0 to 1 are constant. If during such a block the sum of transition probabilities from 1 to 0 is small there is positive probability of "getting stuck" at state 1 under $\theta = 0$. Similarly, if it is large there is positive probability of getting stuck at state 0 under $\theta = 1$.

Proof of Theorem 2. We prove the result by contradiction. Assume that a decision profile d achieves learning in probability. From Lemma 7, $\lim_{n\to\infty} a_n^{01} = 0$ and therefore there exists a $\hat{N} \in \mathbb{N}$ such that for all $n > \hat{N}$,

$$a_n^{01} < \frac{1}{6M}. \tag{4.9}$$

Moreover, by the learning in probability assumption, there exists some $\tilde{N} \in \mathbb{N}$ such that for all $n > \tilde{N}$,

$$\mathbb{P}^0(x_n = 0) > \frac{1}{2}, \tag{4.10}$$

and

$$\mathbb{P}^1(x_n = 1) > \frac{1}{2}. \tag{4.11}$$

Let $N = \max\{\hat{N}, \tilde{N}\}$ so that for all $n > N$ Eqs. (4.9)-(4.11) all hold.

The next step is to divide the agents in blocks such that in each block the sum of the transition probabilities from state 0 to state 1 can be simultaneously bounded from above and below. Define the last agents of each block using the following recursive procedure.

$$\begin{aligned} r_1 &= N, \\ r_k &= \min\left\{ l : \sum_{n=r_{k-1}+1}^{l} a_n^{01} \geq \frac{1}{2M} \right\}. \end{aligned}$$

From Lemma 6 we have that $\sum_{n=N}^{\infty} a_n^{01} = \infty$. This fact together with Equation (4.9) guarantee that the sequence r_k is well defined and strictly increasing.

Let A_k denote the block that ends with agent r_{k+1}, i.e., $A_k \triangleq \{r_k + 1, \ldots, r_{k+1}\}$. The construction of the sequence $\{r_k\}_{k=1}^{\infty}$ yields

$$\sum_{n \in A_k} a_n^{01} \geq \frac{1}{2M}.$$

On the other hand, r_{k+1} is the first agent for which the sum is at least $1/2M$ and since by (4.9) $a_{r_{k+1}} < 1/6M$, we get that

$$\sum_{n \in A_k} a_n^{01} \leq \frac{1}{2M} + \frac{1}{6M} = \frac{2}{3M}.$$

Using Corollary 1, we conclude that blocks $\{A_k\}_{k=1}^{\infty}$ satisfy

$$\frac{1}{2M} \leq \sum_{n \in A_k} a_n^{01} \leq \frac{2}{3M}, \text{ and} \tag{4.12}$$

$$\frac{m}{2M} \leq \sum_{n \in A_k} \bar{a}_n^{01} \leq \frac{2}{3}. \tag{4.13}$$

Consider the following two cases for the summation of transition probabilities from state 1 to state 0 during block A_k:

(i) Assume that

$$\sum_{n \in A_k} a_n^{10} > \frac{1}{2}.$$

Using Corollary 1, we obtain

$$\sum_{n \in A_k} \bar{a}_n^{10} > \sum_{n \in A_k} m \cdot a_n^{10} > \frac{m}{2} \tag{4.14}$$

The assumption of learning in probability suggests that at the beginning of the block, under the state of the world $\theta = 1$, the chain is more likely to be at state one as (4.10) indicates, namely $\mathbb{P}^1(x_{r_k+1} = 1) > 1/2$. The probability of a transition to state zero during the block A_k can be computed as

$$\mathbb{P}^1(\cup_{n \in A_k}\{x_n = 0\} \mid x_{r_k+1} = 1) = \left[1 - \prod_{n \in A_k}(1 - \bar{a}_n^{10})\right].$$

By (4.14) the right-hand side can be bounded from below using the inequality

$$\prod_{n \in A_k}(1 - \bar{a}_n^{10}) \leq e^{-\sum_{n \in A_k} \bar{a}_n^{10}} \leq e^{-\frac{m}{2}},$$

which yields

$$\mathbb{P}^1(\cup_{n \in A_k}\{x_n = 0\} \mid x_{r_k+1} = 1) \geq 1 - e^{-\frac{m}{2}}.$$

After a transition to state 0 occurs the probability of staying at that state until the end of the block is bounded below as follows,

$$\mathbb{P}^1(x_{r_{k+1}} = 0 \mid \cup_{n \in A_k}\{x_n = 0\}) \geq \prod_{n \in A_k}(1 - \bar{a}_n^{01}).$$

The right-hand side can be bounded using Equation (4.13) as follows:

$$\prod_{n \in A_k} (1 - \bar{a}_n^{01}) \geq 1 - \sum_{n \in A_k} \bar{a}_n^{01} \geq \frac{1}{3}.$$

Combining the above, we conclude that

$$\begin{aligned}
&\mathbb{P}^1(x_{r_{k+1}} = 0) \\
&\quad \geq \mathbb{P}(x_{r_{k+1}} = 0 \mid \cup_{n \in A_k}\{x_n = 0\})\mathbb{P}(\cup_{n \in A_k}\{x_n = 0\} \mid x_{r_k+1} = 1)\mathbb{P}^1(x_{r_k+1} = 1) \\
&\quad \geq \frac{1}{6}(1 - e^{\frac{m}{2}}).
\end{aligned}$$

(ii) Assume

$$\sum_{n \in A_k} a_n^{10} \leq \frac{1}{2}$$

The assumption of learning in probability implies that at the beginning of the block, under the state of the world $\theta = 0$, the chain is highly likely to be at state zero as (4.11) indicates, namely $\mathbb{P}^0(x_{r_k+1} = 0) > 1/2$. The probability of a transition to state one from any agent during the block A_k is

$$\mathbb{P}^0(\cup_{n \in A_k}\{x_n = 1\} \mid x_{r_k+1} = 0) = \left[1 - \prod_{n \in A_k} (1 - a_n^{01})\right].$$

The right-hand side can be bounded from below using the inequality

$$\prod_{n \in A_k} (1 - a_n^{01}) \leq e^{-\sum_{n \in A_k} \bar{a}_n^{01}} \leq e^{-\frac{1}{2M}},$$

which yields

$$\mathbb{P}^0(\cup_{n \in A_k}\{x_n = 1\} \mid x_{r_k+1} = 0) \geq 1 - e^{-\frac{1}{2M}}.$$

After a transition to state one occurs, the probability of staying at that state until the end of the block is bounded from below as follows:

$$\mathbb{P}^0(x_{r_{k+1}} = 1 \mid \cup_{n \in A_k}\{x_n = 1\}) \geq \prod_{n \in A_k} (1 - a_n^{10}).$$

The right-hand side can be bounded using Equation (4.13) as follows:

$$\prod_{n\in A_k}(1-a_n^{10}) \geq 1-\sum_{n\in A_k} a_n^{10} \geq \frac{1}{2}.$$

Combining the above, we conclude that

$$\begin{aligned}
&\mathbb{P}^0(x_{r_{k+1}}=1)\\
&\geq \mathbb{P}^0(x_{r_{k+1}}=1 \mid \cup_{n\in A_k}\{x_n=0\})\mathbb{P}^0(\cup_{n\in A_k}\{x_n=1\} \mid x_{r_k+1}=0)\mathbb{P}^0(x_{r_k+1}=0)\\
&\geq \frac{1}{4}(1-e^{\frac{1}{2M}})
\end{aligned}$$

Combining those two cases we conclude that

$$\liminf_{n\to\infty} \mathbb{P}_d(x_n \neq \theta) \geq \frac{1}{2}\min\left\{\frac{1}{6}(1-e^{-\frac{m}{2}}), \frac{1}{4}(1-e^{\frac{1}{2M}})\right\} > 0 \tag{4.15}$$

which contradicts the learning in probability assumption and concludes the proof. □

The coupling between the two states of the world is central in the proof of Theorem 2. The importance of the Bounded Likelihood Ratio assumption is highlighted by the observation that if either $m=0$ or $M=\infty$ then the lower bound obtained in (4.15) is zero, making our proof to fail.

The main idea underlying this proof is the following. For learning in probability to occur the state of the Markov chain should converge to 0 or 1 depending on the state of the world. In order for this to happen, that is, for convergence to the correct state, the chain should be given enough time to explore between the possible states; transitions out of each of the desired states should happen infinitely often, otherwise the chain could get "stuck" to the wrong state with positive probability. For the case $K=1$, transitions out of one of the two desired states correspond to transitions towards the other. Therefore, the "experimentation" necessity leads to a positive probability of transitions between states 0 and 1 during certain finite intervals, under both states of the world because of the Bounded Likelihood Ratio assumption. Then, even if the state of the Markov chain is at the correct state, after some time it will make a transition to the wrong one with positive probability and thus will not learn.

An analogous proof technique was expected to hold for the case of $K \geq 2$. The next section explores this case and, surprisingly, constructs a decision that achieves learning in probability when agents observe the last 2 immediate predecessors.

4.2 Learning in probability when $K \geq 2$

In this section we discuss learning in probability when agents observe two or more of their immediate predecessors and the private signal structure is characterized by Bounded Likelihood Ratios.

4.2.1 A first approach

A first attempt to tackle the problem when $K = 2$ would be to mimic the proof techniques of the previous section. The Markov chains that correspond to the new problem are illustrated in Figure 4-3. The Bounded Likelihood ratio assumption would imply

$$m < \frac{a_n^i}{\bar{a}_n^i} < M,$$

for all $i \in \{1, \ldots, 6\}$.

The "experimentation" necessity would imply that for a decision profile which achieves learning in probability, we should have

$$\sum_{n \in \mathbb{N}} a_n^i = \infty,$$

and

$$\sum_{n \in \mathbb{N}} \bar{a}_n^i = \infty,$$

for $i = 1, 6$ while we **should not** have this property for the rest values of i. Finally, for learning in probability to occur, we should have $a_n^i \to 0$ as $n \to \infty$, for $i = 1, 6$.

In words, for learning in probability to occur, the state of the Markov chain should converge in probability to 00 and 11 when $\theta = 0$ or $\theta = 1$, respectively. Observe now that the experimentation necessity implies that transition out of those desired states should happen infinitely often. In contrast to the case $K = 1$ a transition out of state 00 **does not** imply a transition to state 11. Therefore for those intervals that either $\sum a_n^1$ is large or $\sum a_n^6$ is large, it is not necessarily implied that a visit to state 11 or 00 will occur with non-negligible probability. In order to make such a statement, one should establish properties for the intermediate transition probabilities a_n^i, for $i \in \{2, \ldots, 5\}$.

For example, if one could argue that transitions between the intermediate stages 01 and 10 happen at most L times then the transition probabilities between state 00 and 11 would be bounded from below and above by m^{L+2} and M^{L+2} respectively

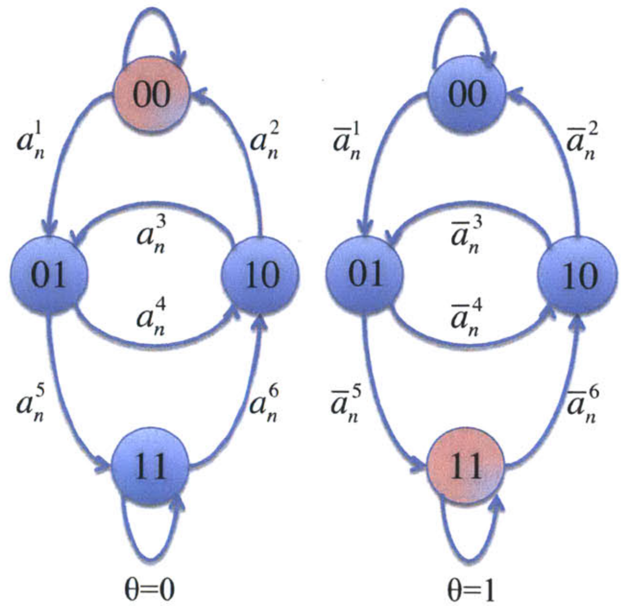


Figure 4-3: The Markov chains that correspond to the decision process when K=2.

and the arguments from the previous section would go through. Such guarantees, on the other hand, cannot be obtained. Specifically, consider a decision rule for which if the process visits state 10 at time n it oscillates between 10 and 01, q_n times where $\{q_n\}$ is a strictly increasing sequence. Then, transition probabilities between the desired states 00 and 11 cannot be bounded by numbers independent of n and the proof techniques used so far fail. This is exactly the property that we exploit in our construction that follows.

4.2.2 Biased coin observation model

Consider the biased coin example that we introduced in Section 2.1.2. Denote by p_0 and p_1 the two possible values for the bias of the coin with $p_0 < p_1$, corresponding to the cases of $\theta = 0$ and $\theta = 1$ respectively, and let p denote the true value. Define p_m to be the average bias, namely $p_m \triangleq (p_0 + p_1)/2$. Resolving the uncertainty between the two possible states of the world is equivalent to deciding whether $p > p_0$ or $p < p_1$. The agents' private information consists of the outcome of a coin toss; their private signal can take values $s_n \in \{H, T\}$ and the private signal distributions are Bernoulli with $\mathbb{P}(s_n = H) = p$ and $\mathbb{P}(s_n = T) = 1 - p$.

Observe that the general two-hypothesis testing problem, with $\tilde{s}_n$ drawn from any private signal structure $(\mathbb{F}_0, \mathbb{F}_1)$ characterized by Bounded Likelihood Ratios may be put in this framework under the correspondence

$$s_n = \begin{cases} H, \text{ if } \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(\tilde{s}_n) \geq 1 \\ T, \text{ if } \frac{d\mathbb{F}_0}{d\mathbb{F}_1}(\tilde{s}_n) < 1 \end{cases}$$

Therefore, constructing a decision that achieves learning in probability for the case of the biased coin can be extended to any private signal structure. For this reason, we focus on constructing a decision that achieves learning in probability for the biased coin signal structure.

4.2.3 Cover's decision and Koplowitz's modification

In this framework, two papers by Cover [7] and by Koplowitz [11] are closely related to our work. In their papers, agents are not constrained to make a binary decision. Cover exhibits a decision profile under which the hypothesis $p < p_m$ vs $p > p_m$ is resolved with limiting error probability of zero under either hypothesis.

We first present the decision profile proposed by Cover. Agents' decisions can take four values represented by the pair (T_n, Q_n) where $T_n, Q_n \in \{0, 1\}$. Consider two sequences $\{k_i\}_{i\in\mathbb{N}}$ and $\{r_i\}_{i\in\mathbb{N}}$ of appropriately chosen positive integers. Divide agents into blocks according to the following rule. The first k_1 agents will be block S_1, the next r_1 agents will be block R_1, the next k_2 agents will be block S_2, the next r_2 agents will be block R_2, etc.

The decision profile is defined as follows:

(i) During an S block:

(a) At the beginning of the block if the initial observation is 1 (Heads) set $Q_n = 1$.

(b) Subsequently, in that block let

$$Q_n = \begin{cases} 0, \text{ if } s_n = 0, \\ Q_{n-1}, \text{ otherwise,} \end{cases}$$

$$T_n = T_{n-1}.$$

(c) For the last agent of the block

$$T_n = \begin{cases} 1, & \text{if } Q_n = 1, \\ T_{n-1}, & \text{otherwise.} \end{cases}$$

(ii) During an R-block:

(a) At the beginning of the block, if the initial observation is 0 (Heads), set $Q_n = 1$.

(b) Subsequently, in that block let

$$Q_n = \begin{cases} 0, & \text{if } s_n = 1, \\ Q_{n-1}, & \text{otherwise,} \end{cases}$$

$$T_n = T_{n-1}.$$

(c) For the last agent of the block

$$T_n = \begin{cases} 0, & \text{if } Q_n = 1, \\ T_{n-1}, & \text{otherwise.} \end{cases}$$

Thus, R_i checks for r_i consecutive zeros while S_i checks for k_i consecutive ones, and $Q_n = 1$ at the end of the block if and only if the desired run has occurred. After the occurrence of such a run, the T component of the last agent's decision changes to 0 and 1, respectively, indicating the success.

By choosing appropriately the lengths $\{k_i\}_{i\in\mathbb{N}}$ and $\{r_i\}_{i\in\mathbb{N}}$ of the blocks, the component T_n of agents' actions converges in probability to the correct state of the world.

Later Koplowitz [11] observed that during an S-block, if the currently favored hypothesis is $T_n = 1$, then regardless of the observations throughout the block the favored hypothesis remains $T_n = 1$. Only if the favored hypothesis is $T_n = 0$ it is necessary to check for consecutive ones. Using this fact, he proposed a decision profile where agents are allowed to take a three-valued decision (in contrast to Cover's four-valued). It turns out that the sequence of $\{T_{q_k}\}_{k\in\mathbb{N}}$, where $\{q_k\}_{k\in\mathbb{N}}$ denotes the sequence of the first agents of each block converges to the correct state of the world, in probability.

Those decision profiles do not fit in our framework for the following two reasons:

(i) Agents' actions are allowed to take more than two values.

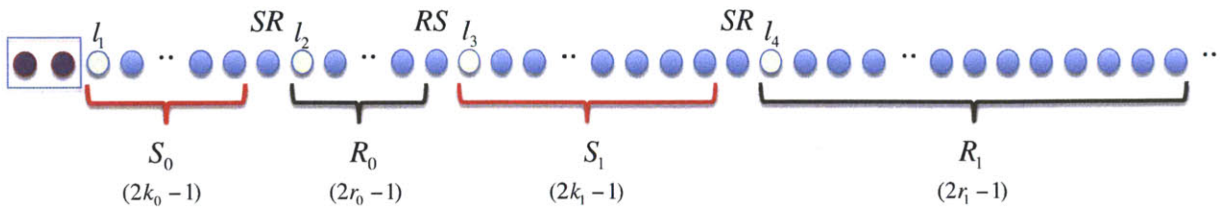


Figure 4-4: Dividing agents in R-blocks and S-blocks

(ii) More importantly for Cover's decision profile just one component of agents' actions converges in probability to the correct state of the world. Similarly, for Koplowitz's decision profile only one subsequence of agents, that of the blocks' leaders learns in probability.

In the next section we provide a decision for $K > 1$ requiring agents to take binary decisions, under which agents' actions converge to the correct state in probability (not just along a subsequence).

4.2.4 Learning in probability when observing two or more of the immediate predecessors

In this section we construct a decision profile that achieves learning in probability. For brevity we define $q \triangleq 1 - p$, $q_i \triangleq 1 - p_i$ for $i \in \{0, 1\}$ and $q_m = 1 - p_m$. Observe that $p > p_m$ if and only if $q < q_m$. Let $\{k_i\}_{i\in\mathbb{N}}$ and $\{r_i\}_{i\in\mathbb{N}}$ be sequences of integers that we will define later in this section.

Divide agents in S-blocks, R-blocks and transient agents as follows.

(i) The first two agents do not belong to any block.

(ii) The next $2k_0 - 1$ agents belong to the block S_0.

(iii) The next agent is an SR transient agent.

(iv) The next $2r_0 - 1$ agents belong to the block R_0.

(v) The next agent is an RS transient agent.

(vi) Division continues as illustrated in Figure 4-4.

Agents' information set consists of the outcome of their coin tossing as well as the last two decisions, denoted by $T_n = (x_{n-2}, x_{n-1})$.

Our decision profile is as follows:

(i) Agents 1 and 2 choose 1 irrespective of their private signal.

(ii) During block S_m:

(a) If the first agent observes 11 she chooses 1 irrespective of her private signal. If she observes 00 and her private signal is 1 then

$$x_n = \begin{cases} 1, & \text{with probability } \frac{1}{m+\bar{k}}, \\ 0, & \text{otherwise.} \end{cases}$$

If the first event is realized we say that the **searching phase is initiated.** In all other cases she decides 0.

(b) For the rest of the agents in the block:

- Agents who observe 01 decide 0 for all private signals.
- Agents who observe 10 decide 0 if and only if their private signal is 0.
- Agents who observe 00 decide 0 for all private signals.
- Agents who observe 00 decide 0 for all private signals.

(iii) During block R_m :

(a) If the first agent observes 00 she chooses 0 irrespective of her private signal. If she observes 11 and her private signal is 0, then

$$x_n = \begin{cases} 0, & \text{with probability } \frac{1}{m+\bar{r}}, \\ 0, & \text{otherwise.} \end{cases}$$

If the first event is realized, we say that the **searching phase is initiated.** In all other cases she decides 0.

(b) For the rest of the agents in the block:

- Agents who observe 10 decide 1 for all private signals.
- Agents who observe 01 decide 1 if and only if their private signal is 1.
- Agents who observe 00 decide 0 for all private signals.
- Agents who observe 11 decide 1 for all private signals.

(iv) SR transient agents decide 1 if and only if they observe 10 for any private signal.

(v) RS transient agents decide 0 if and only if they observe 01 for any private signal.

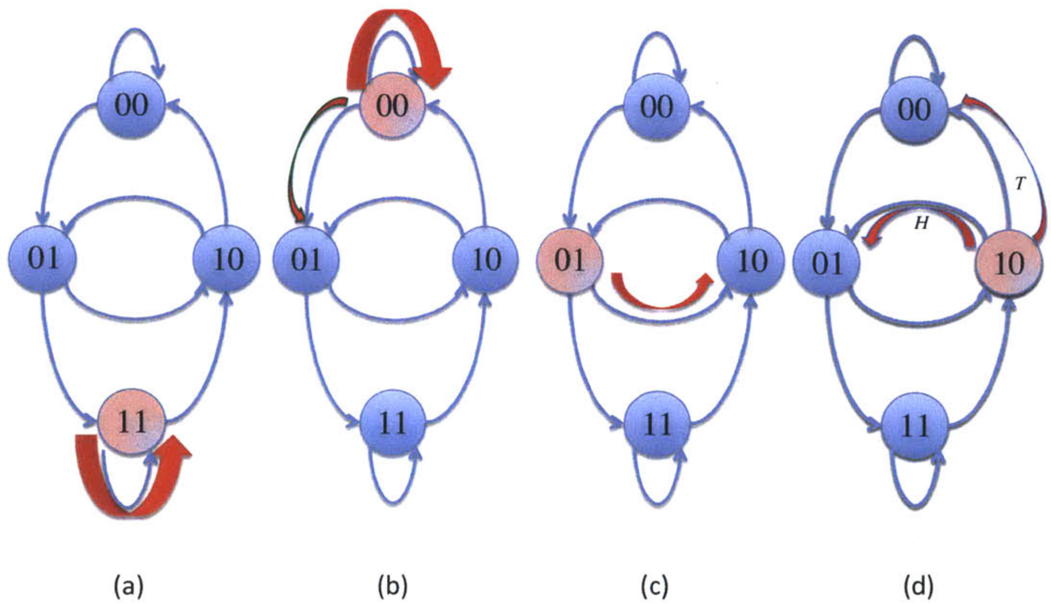


Figure 4-5: The decision profile during an S-block. (a) Current state is 11 anywhere in the block (b) The state at the beginning of the block is 00 (c) The current state is 01. (d) The current state is 10.

The decision profile is summarized in Table 4.1.

Observe that by the construction of the decision , at the beginning of each block the state can be either 00 or 11. If at the beginning of an S-block the state is 11, it does not change. On the contrary, if the state is 00, then we consider two cases. If the searching phase is not initiated, the state remains 00 until the end of the block. If the searching phase is initiated, the state at the beginning of the next block becomes 11 if and only if k_m ones are observed. Otherwise it returns at state 00. The same holds for the R-blocks.

4.2.5 Proof

The following fact is used in the proof that follows.

Lemma 8. *If $\alpha > 1$, then the series*

$$\sum_{i=L}^{+\infty} \frac{1}{i(\log(i))^{\alpha}}, \qquad \text{with } L \geq 2$$

converges; if $\alpha \leq 1$, then the series diverges. Here $\log(\cdot)$ denotes the natural logarithm.

Decision that achieves learning under BLR

$x_1 = 1$ and $x_2 = 1$

During S_i block	**During R_i block**
If agent n is the **first**	**If** agent n is the **first**
$d_n(s_n, 0, 0) = \begin{cases} s_n, & \text{with probability } \frac{1}{k+i} \\ 0, & \text{otherwise} \end{cases}$	$d_n(s_n, 1, 1) = \begin{cases} s_n, & \text{with probability } \frac{1}{\bar{r}+i} \\ 1, & \text{otherwise} \end{cases}$
$d_n(s_n, 1, 1) = 1$	$d_n(s_n, 0, 0) = 0$
If agent n is **not** the **first**	**If** agent n is **not** the **first**
$d_n(s_n, 1, 1) = 1$	$d_n(s_n, 0, 0) = 0$
$d_n(s_n, 0, 0) = 0$	$d_n(s_n, 1, 1) = 0$
$d_n(s_n, 0, 1) = 0$	$d_n(s_n, 1, 0) = 1$
$d_n(s_n, 1, 0) = s_n$	$d_n(s_n, 0, 1) = s_n$
If agent n is **SR** transient	**If** agent n is **RS** transient
$d_n(s_n, 0, 0) = 0$	$d_n(s_n, 1, 1) = 0$
$d_n(s_n, 0, 1) = 1$	$d_n(s_n, 1, 0) = 0$

Table 4.1: A decision profile that achieves learning in probability. Note that this is a randomized decision . Our analysis in the previous chapters assumes deterministic decision s but can be extended to randomized by standard arguments that would just make notation harder.

Proof. See Theorem 3.29 of [15]. □

Define $\bar{k} = \max\left\{2, \lceil e^{\frac{1}{p_m}} \rceil\right\}$ and $\bar{r} = \max\left\{2, \lceil e^{\frac{1}{q_m}} \rceil\right\}$.
Moreover let $\{k_i\}_0^\infty$, $\{r_i\}_0^\infty$ be sequences of integers defined as

$$k_i = \left\lceil \log_{p_m}\left(\frac{1}{\log(i+\bar{k})}\right)\right\rceil, \tag{4.16}$$

and

$$r_i = \left\lceil \log_{q_m}\left(\frac{1}{\log(i+\bar{r})}\right)\right\rceil. \tag{4.17}$$

Theorem 3. *Let the agents decide according to the decision described in Section 4.2.4. Then,*

$$\lim_{n\to\infty} \mathbb{P}(x_n = \theta) = 1.$$

Proof. Let $\{l_i\}_{i=0}^{+\infty}$ denote the sequence of the first agents of consecutive R-blocks or S-blocks. In other words l_1 is the first agent in the first S-block , l_2 is the first agent in the first R-block etc. Moreover, remember that $T_n = \{x_{n-2}, x_{n-1}\}$ denotes the string of observed actions for agent n. From the definition of the decision profile, we have $T_{l_i} \in \{00, 11\}$ with probability one for all $i \in \mathbb{N}$.

Observe that if for some i_0 , $T_{l_{i_0}} = 00$, then $T_{l_{i_0+1}} = 11$ if and only if the following three conditions hold. First, the block starting from l_{i_0} is an S-block; denote it by S_{m_o}, $m_o = \frac{i_0+1}{2}$. Second, the searching phase did start; which happens with probability $\frac{1}{m_o+\bar{k}}$. Finally, $s_n = 1$ for all agents in S_{m_o} who are an odd number away from the first; which happens with probability $p^{s_{m_o}}$. Similarly, if for some i_0, $T_{l_{i_0}} = 11$, then $T_{l_{i_0+1}} = 00$ if and only if the following three conditions hold. First, the block starting from l_{i_0} is an R-block; denote it by R_{m_o}, $m_o = \frac{n_o}{2}$. Second, the searching phase did start; which happens with probability $\frac{1}{m_o+\bar{r}}$. Finally, $s_n = 0$ for all n in S_{m_o} that are an even number away from the first; which happens with probability $q^{r_{k_o}}$. Therefore,

$$\mathbb{P}(T_{l_{i+1}} = 11 \mid T_{l_i} = 00) = p^{s_{m_o}} \frac{1}{\bar{k}+m_o}, \text{ where } m_o = \frac{i+1}{2} \text{ and } i \text{ is odd}, \tag{4.18}$$

$$\mathbb{P}(T_{l_{i+1}} = 00 \mid T_{l_i} = 11) = q^{r_{m_o}} \frac{1}{\bar{r}+m_o}, \text{ where } m_o = \frac{i}{2} \text{ and } i \text{ is even}. \tag{4.19}$$

It follows, by the independence of the blocks and the Borel-Cantelli Lemma, that

$T_{l_i} \to 11$, with probability one if

$$\begin{aligned} \sum_{i=0}^{+\infty} p^{k_i} \frac{1}{i+\bar{k}} &= \infty, \text{ and} \\ \sum_{i=0}^{+\infty} q^{r_i} \frac{1}{i+\bar{r}} &< \infty, \end{aligned} \tag{4.20}$$

since if both conditions hold, transitions from the state 00 to the state 11 happen infinitely many times, while transitions from state 11 to state 00 happen only finitely many times.

Similarly, $T_{l_i} \to 00$ with probability one if

$$\begin{aligned} \sum_{i=0}^{+\infty} p^{k_i} \frac{1}{i+\bar{k}} &< \infty, \text{ and} \\ \sum_{i=0}^{+\infty} q^{r_i} \frac{1}{i+\bar{r}} &= \infty, \end{aligned} \tag{4.21}$$

The definition of the sequence $\{k_i\}$ implies that

$$\log_{p_m}(\frac{1}{i+\bar{k}}) \leq k_i < \log_{p_m}(\frac{1}{i+\bar{k}}) + 1$$

or, equivalently,

$$p \cdot p^{\log_{p_m}(\frac{1}{i+\bar{k}})} < p^{k_i} \leq p^{\log_{p_m}(\frac{1}{i+\bar{k}})}$$

Consequently,

$$p\frac{1}{(i+\bar{k})\log_{p_m}(i+\bar{k})^{\alpha}} < p^{k_i}\frac{1}{i+\bar{k}} \leq \frac{1}{(i+\bar{k})\log_{p_m}(i+\bar{k})^{\alpha}}, \quad \alpha = \log_{p_m}(p) \tag{4.22}$$

Similarly,

$$q\frac{1}{(i+\bar{r})\log_{q_m}(i+\bar{k})^{\beta}} < q^{k_i}\frac{1}{i+\bar{r}} \leq \frac{1}{(i+\bar{r})\log_{q_m}(i+\bar{k})^{\beta}}, \quad \beta = \log_{q_m}(q) \tag{4.23}$$

Now, observe that if $\theta = 1$ or equivalently $p > p_m$ then $\alpha > 1$ and $\beta < 1$. In that case, Lemma 8 implies that the conditions (4.20) hold and thus $T_{l_n} \to 11$ with probability one. Similarly, if $\theta = 0$ or equivalently $p < p_m$ then $\alpha < 1$ and $\beta > 1$. In that case, Lemma 8 implies that the conditions (4.21) hold and thus $T_{l_n} \to 00$ with probability one. In other words, under the proposed algorithm, $T_{l_n} \to \{\theta, \theta\}$ with probability one.

Fix some $\epsilon > 0$. Without loss of generality assume that $\theta = 0$. Note that almost sure convergence implies convergence in probability and therefore there exists $\tilde{N} \in \mathbb{N}$ such that

$$\mathbb{P}(T_{l_n} = 00 \mid \theta = 0) > 1 - \frac{\epsilon}{2} \text{ for all } n > \tilde{N}$$

Let n_0 be the first agent of an R block, and assume $T_{n_o} = 00$. Then with probability one the state will not change throughout the block. On the other hand if at the beginning (n_o) of an S block $T_{n_o} = 00$ then the searching mode will initiate with probability $\frac{1}{\bar{k}+no}$. Let $c(n) = \text{argmax}\{k : l_{2k} < n, \text{k even}\}$ denote the closest to agent n S-block, then for all $n > \tilde{N}$,

$$\mathbb{P}(T_n \neq 00 \mid \theta = 0) < \frac{\epsilon}{2} + \frac{1}{\bar{k} + c(n)},$$

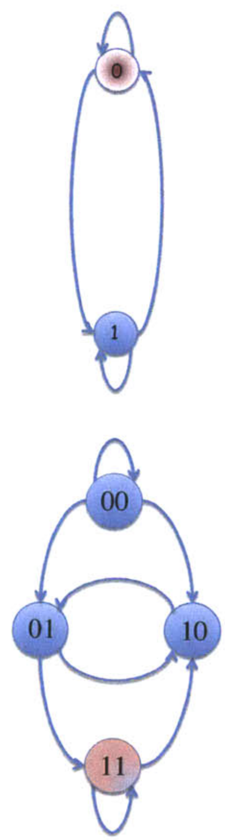
where the first term on the right-hand side corresponds to initiating a block from a state other than 00, while the second term corresponds to the case of initiating the search phase in the closest S block.

Let $K = \max\{0, \frac{2}{\epsilon} - \bar{k}\}$ and observe that for all $k > K$, $\frac{1}{\bar{k}+k} < \frac{\epsilon}{2}$. Let $\hat{N} = \max\{\tilde{N}, l_{2K}\}$. Then, for all $n > \hat{N}$,

$$\mathbb{P}(T_n \neq 00 \mid \theta = 0) < \frac{\epsilon}{2} + \frac{1}{\bar{k} + c(n)} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

which concludes the proof. □

4.3 Discussion and Conclusions



The dichotomy that this chapter presents is unexpected, especially combined with the results of Chapter 3. First, we should explain the difference between the cases $K = 1$ and $K > 1$. In order to do that we should go back to the Markov chain representation where the core difference can be illustrated. Observe that the property of Bounded Likelihood Ratios couples the transition probabilities between the states 0 and 1,at which, if learning was achieved, the state would converge under either state of the world. This fact, as illustrated, is central in the impossibility proof presented in Section 4.1. In contrast, observe the Markov chain for the case $K = 2$. The Bounded Likelihood Ratio property does couple the individual transition probabilities, but the same does not hold for transitions between states 00 and 11. Indeed, a transition from state 00 to state 11 would go through state 01. Then, the chain could cycle arbitrarily many times among the intermediate states before reaching state 11. In that case, the multi-step transition probabilities consist of a product of a growing number of terms. Individually, those terms are coupled between the two states of the world but the bounds for the ratio of their products asymptotically reach zero and infinity respectively. Allowing a growing number of cycles between intermediate states essentially decouples the transition probabilities between the two desired states 00 and 11 making the problem equivalent to a private signal structure with Unbounded Likelihood Ratios.

Finally, Theorem 3 can be contrasted with Theorem 1. In the latter we prove that there does not exist a decision that achieves almost sure learning. On the other hand, in the former, we provide a decision that achieves learning in probability. This contrast highlights the delicate difference between those two modes of convergence. For almost sure convergence we require the existence of a finite time/agent after which all actions are equal to the true state of the world. This restricts our attention to a countable subset of possible realizations of the decision process. Our proof exhausted all those possible realizations. On the other hand, for learning in probability to occur, all we require is the probability of a wrong decision to asymptotically go to zero. This allows for a larger set of realizations of the decision process. We made use of this relaxation by instructing the first agent of each block to initiate a searching phase with a probability $\Theta(1/n)$, thus constraining agents' actions to deviate from the final state with limiting probability of zero.

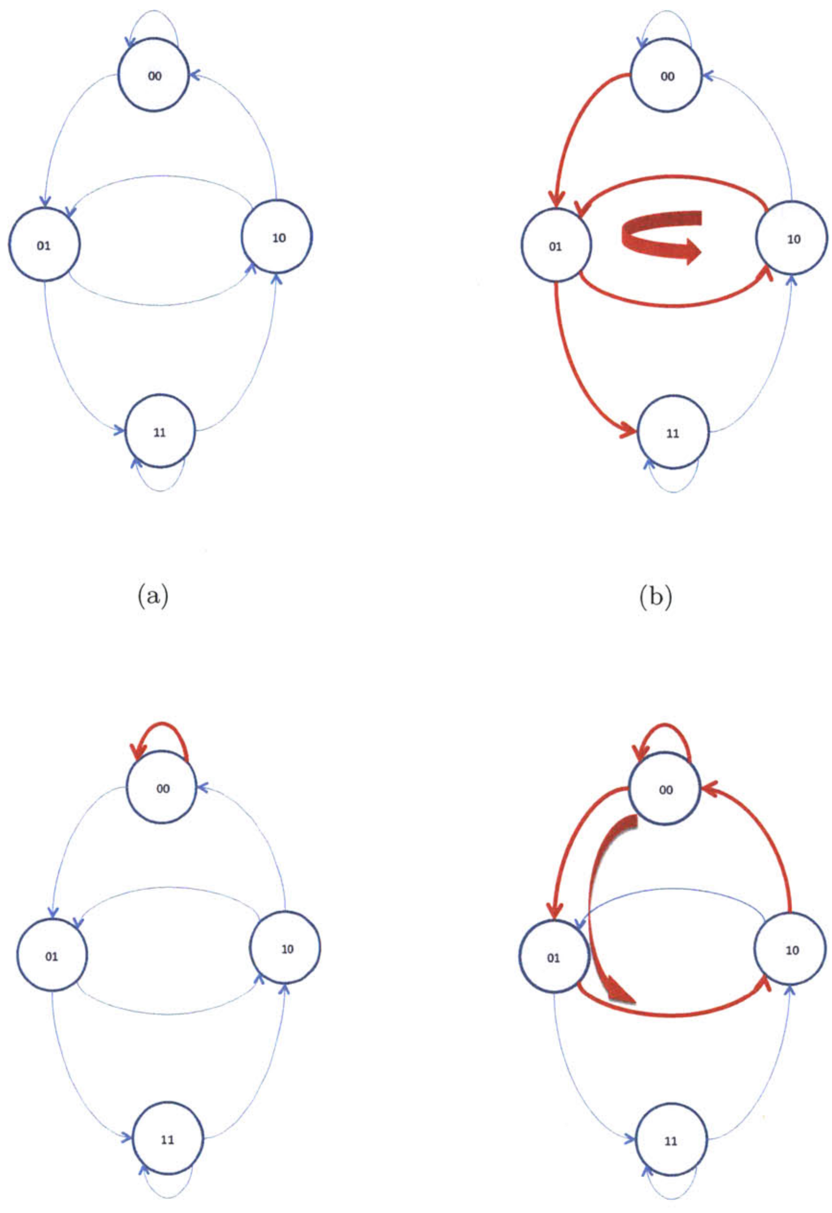


Figure 4-6: An illustration of the algorithm. (a) A state automaton that represents the possible histories observed. (b) The decision process for of a signal realization of all $H * H * H * H * H * ..$, where $*$ represents any possible private signal, during an S block that begins from 00 and for which the *search phase is initiated.* (c) An R block that begins with 00. (d) The decision process during an S-block where the search phase is initiated in case of signal realization of $H * T * * * * \ldots$

Chapter 5

Forward looking agents: Bounded Likelihood Ratios

In this chapter we consider forward looking agents, i.e., agents with a utility function given by (5.2). We study the resulting dynamic game and in particular we characterize its Perfect Bayesian Equilibria. Eventually we prove that there does not exist a Perfect Bayesian Equilibrium that achieves learning in probability. Throughout this chapter we focus on the case of private signal structures that are characterized by Bounded Likelihood Ratios.

5.1 Basic definitions

We adopt the same observation model as in Section 2.1.1. The information available to agents prior to making their decisions consists of their private signal and the actions of their K immediate predecessors,

$$I_n \triangleq \{s_n, x_k \text{ for all } k \in D_n\}, \tag{5.1}$$

where $D_n = n-K, \ldots, n-1$.

A *strategy* for individual n is a mapping $\sigma_n : S \times \{0,1\}^K \to \{0,1\}$ that selects a decision for each possible information set. Let Σ denote the pure strategy space for each player n , i.e., the set of all mappings $\sigma_n : S \times \{0,1\}^K \to \{0,1\}$.

A *strategy profile* is a sequence of strategies $\sigma = \{\sigma_n\}_{n\in\mathbb{N}}$. We will use the standard notation $\sigma_{-n} = \{\sigma_1, \ldots, \sigma_{n-1}, \sigma_{n+1}, \ldots\}$ to denote the strategies of all agents other than n, and $\sigma = \{\sigma_n, \sigma_{-n}\}$.

Given a strategy profile σ, the sequence of decisions $\{x_n\}_{n\in\mathbb{N}}$ is a well defined stochastic process and we denote the measure generated by this stochastic process by $\mathbb{P}_\sigma$.

We call the resulting game the *altruistic learning game* and it consists of the following components.

(i) The set of players is $\mathbb{N}$.

(ii) The set of states of the world is $\{0, 1\}$.

(iii) The set of actions of each player is $\{0, 1\}$.

(iv) The set of types for each player is $S \times \{0, 1\}^K$.

(v) The payoff function $u_i : \{0, 1\} \times \{0, 1\}^{\mathbb{N}} \to \infty$ for each player i given by

$$u_n(\theta, \mathbf{x}) = (1 - \delta) \sum_{k=n}^{+\infty} \delta^{k-n} \mathbf{1}_{x_k=\theta}, \tag{5.2}$$

where δ is a discount factor with $\delta \in (0, 1)$, $\mathbf{x} = \{x_n\}_{n=0}^{\infty}$ and $\mathbf{1}_A$ denotes the indicator random variable for the event A.

(vi) Each player has a uniform prior distribution over the state of the world.

Definition 5. *A strategy profile σ^* is a pure-strategy Perfect Bayesian Equilibrium of the altruistic learning game if for each $n \in \mathbb{N}$, σ_n^* maximizes the expected payoff of agent n given the strategies of other agents σ_{-n}^*.*

In the rest of this thesis we focus on pure-strategy Perfect Bayesian Equilibria, and simply refer to them as *equilibria* (without the other qualifiers).

Given a strategy profile σ, the expected payoff of agent n from action $x_n = y \in \{0, 1\}$ is given by

$$\bar{U}_n(y; I_n) = (1 - \delta) \sum_{k=n}^{\infty} \delta^{n-k} \mathbb{P}_\sigma(x_k = \theta \mid x_n = y, I_n)$$

and therefore, for an equilibrium σ^* we have,

$$\sigma_n^*(I_n) \in \operatorname*{argmax}_{y \in \{0,1\}} \sum_{k=n}^{\infty} \delta^{k-n} \mathbb{P}_{\sigma_{-n}^*}(x_k = \theta \mid x_n = y, I_n).$$

Typically, in the Bayesian learning literature, learning fails when private signals are characterized by Bounded Likelihood Ratios because of the creation of herds;

when the probability of a correct decision by copying the observed action becomes large, then it is suboptimal for individuals to trust their private signal and decide accordingly. This behaviour allows the creation, with positive probability, of an incorrect herd. It should be evident, though, that the creation of herds is strongly dependent on the selfish decisions of agents, who copy the myopically optimal action disregarding their private information.

In our model though, the payoff structure does not necessarily encourage myopic decisions. Consider the case of large discount factors; making a suboptimal decision yields an immediate payoff of zero in the worst case but, if this suboptimal action instructs better decisions for future agents it could yield a continuation payoff of $1/(1-\delta)$ and therefore our payoff structure potentially encourages *altruistic* behaviour. Unfortunately, it turns out that myopic decisions may arise under the forward looking model as well, as the following equilibrium illustrates for the special case of coin tossing.

Example 3 (A naive herding equilibrium for symmetric coin tossing). *Consider the standard example of coin tossing and let the bias take the following possible values. If $\theta = 0$, then $\mathbb{P}(Heads) = 0.75$, otherwise $\mathbb{P}(Heads) = 0.25$. Fix some $K > 0$ and let $\delta \in (0,1)$. Let agents decide according to the following strategy profile σ:*

$$\sigma_n(s_n, x_{D_n}) = \begin{cases} 1, & \text{if } n = 1 \text{ and } s_n = \text{Heads}, \\ 0, & \text{if } n = 1 \text{ and } s_n = \text{Tails}, \\ x_{n-1}, & \text{otherwise.} \end{cases}$$

The strategy profile σ is a Perfect Bayesian Equilibrium for the corresponding game. Indeed, consider the decision problem for the first agent and assume that the result of his coin tossing is $s_1 = Heads$. The Bayes rule yields,

$$\mathbb{P}(\theta = 1 \mid s_1 = Heads) = \frac{\mathbb{P}(Heads \mid \theta = 1)\frac{1}{2}}{\mathbb{P}(Heads \mid \theta = 1)\frac{1}{2} + \mathbb{P}(Heads \mid \theta = 0)\frac{1}{2}} = \frac{0.75}{0.75 + 0.25} = 0.75. \tag{5.3}$$

Clearly, $\mathbb{P}(\theta = 0 \mid s_1 = Heads) = 0.25$. By the definition of the strategy profile σ it follows that

$$\mathbb{P}_\sigma(x_k = y \mid x_1 = y) = 1 \text{ for all } k \geq 1 \text{ and } y \in \{0, 1\}.$$

Therefore, the expected payoff for the first agent from action 1*is*

$$\bar{U}_1(1; \textit{Heads}) = (1-\delta)\left[0.75(1+\delta+\delta^2+\ldots) + 0.25(0+0+\ldots)\right] = 0.75,$$

where the first term corresponds to the expected payoff conditioned on $\theta = 1$ *and the second term to* $\theta = 0$.

Similarly the expected payoff from action 0 *is*

$$\bar{U}_1(0; \textit{Heads}) = (1-\delta)\left[0.75(0+0+\ldots) + 0.25(1+\delta+\delta^2+\ldots)\right] = 0.25.$$

Therefore, the optimal action is to choose 1 *if the private signal is Heads. A symmetric argument holds for the case of Tails.*

The next step is to consider a typical agent $n > 1$. *We prove that conditioned on the fact that the rest of the agents decide according to* σ_{-n} *it is optimal for agent* n *to follow the strategy* σ_n.

Indeed, assume that $x_{n-1} = 1$. *First, consider the case* $s_n =$ *Heads . Observe that by the definition of the strategy profile,*

$$\mathbb{P}_\sigma(x_{n=1} = 1 \mid \theta = j) = \mathbb{P}(s_1 = \textit{ Heads } \mid \theta = j), \;\; j \in \{0,1\},$$

and therefore,

$$\begin{aligned}\mathbb{P}_\sigma(\theta = 1 \mid I_n) &= \frac{\mathbb{P}_\sigma(s_n = \textit{Heads}, x_{n-1} = 1 \mid \theta = 1)\frac{1}{2}}{\mathbb{P}_\sigma(s_n = \textit{Heads}, x_{n-1} = 1 \mid \theta = 1)\frac{1}{2} + \mathbb{P}_\sigma(s_n = \textit{Heads}, x_{n-1} = 1 \mid \theta = 1)\frac{1}{2}} \\ &= \frac{0.75 \cdot 0.75}{0.75 \cdot 0.75 + 0.25 \cdot 0.25} = 0.9. \qquad (5.4)\end{aligned}$$

Arguing as before, the expected payoffs from the two possible actions are equal to $\bar{U}(1; \textit{Heads}, 1) = 0.9$ *and* $\bar{U}(0; \textit{Heads}, 1) = 0.1$ *respectively. Clearly, it is optimal for agent* n *to choose* 1.

The last case to consider is $x_{n-1} = 1$ *but* $s_n =$ *Tails . In that case,*

$$\begin{aligned}\mathbb{P}_\sigma(\theta = 1 \mid I_n) &= \frac{\mathbb{P}_\sigma(s_n = \textit{Tails}, x_{n-1} = 1 \mid \theta = 1)\frac{1}{2}}{\mathbb{P}_\sigma(s_n = \textit{Tails}, x_{n-1} = 1 \mid \theta = 1)\frac{1}{2} + \mathbb{P}_\sigma(s_n = \textit{Tails}, x_{n-1} = 1 \mid \theta = 1)\frac{1}{2}} \\ &= \frac{0.75 \cdot 0.25}{0.25 \cdot 0.75 + 0.75 \cdot 0.25} = 0.5. \qquad (5.5)\end{aligned}$$

It follows that $\bar{U}(1; \textit{Tails}, 1) = 0.5$ *and* $\bar{U}(0; \textit{Tails}, 1) = 0.5$ *and therefore in this case agent* n *is indifferent between the two actions. Therefore, deciding according to* σ *is indeed weakly optimal, proving that* σ *is a Perfect Bayesian Equilibrium.*

The equilibrium strategy profile that we analyze in this example does not achieve learning in probability because for all $n > 1$,

$$\begin{aligned}\mathbb{P}_\sigma(x_n = \theta) &= \mathbb{P}_\sigma(x_n = 1 \mid \theta = 1)\frac{1}{2} + \mathbb{P}_\sigma(x_n = 0 \mid \theta = 0)\frac{1}{2} \\ &= \mathbb{P}(s_1 = \textit{Heads} \mid \theta = 1)\frac{1}{2} + \mathbb{P}(s_1 = \textit{Tails} \mid \theta = 0)\frac{1}{2} = 0.75 < 1\end{aligned}$$

The above equilibrium illustrates how forward looking agents can fail to achieve learning in probability, by following myopic strategies. On the other hand, equilibria need not be unique and therefore there may exist other, more sophisticated equilibria with strategy profiles that achieve learning, such as the decision rule of Section 4.2.

The calculations (5.3)-(5.5) suggest the following definition.

Definition 6. *We refer to the probability*

$$p_n \triangleq \mathbb{P}(\theta = 1 \mid s_n)$$

as the ***private belief*** *of agent n.*

The following Proposition is a direct application of the Bayes' rule.

Proposition 1. *For any $n \in \mathbb{N}$ and any $s_n \in S$ the private belief p_n of agent n is given by*

$$p_n(s_n) = \left[1 + \frac{dF_0}{dF_1}(s_n)\right]^{-1}.$$

The assumption of Bounded Likelihood Ratios for the private signal structure implies bounded support for the private beliefs as the Proposition below implies.

Proposition 2. *Assume that the private signal structure is characterized by Bounded Likelihood Ratios. Then, there exist $\underline{\beta} > 0$ and $\bar{\beta} < 1$, where*

$$\underline{\beta} = \inf\{r \in [0,1] \mid \mathbb{P}(p_n \le r) > 0\} \text{ and } \bar{\beta} = \sup\{r \in [0,1] \mid \mathbb{P}(p_n \le r) < 1\}$$

Proof. Clearly $\underline{\beta} = \frac{1}{1+M}$ and $\bar{\beta} = \frac{1}{1+m}$ satisfy the conditions in the proposition. □

Similar to the private belief we define the *social belief* and the *posterior belief.*

Definition 7. *Given a strategy profile σ, the* ***social belief*** *$\pi_n(x_k, k \in D_n)$ of agent n given a set of observations $\{x_k, k \in D_n\}$ is defined as*

$$\pi_n(x_k, k \in D_n) = \mathbb{P}_\sigma(\theta = 1 \mid x_k, k \in D_n). \tag{5.6}$$

Definition 8. *Given a strategy profile σ, the* **posterior belief** f_n *of agent n given a set of observations $\{x_k, k \in D_n\}$ and a private signal s_n, is defined as*

$$f_n(s_n, x_k, k \in D_n) = \mathbb{P}_\sigma(\theta = 1 \mid s_n, x_k, k \in D_n). \tag{5.7}$$

The Bayes' rule connects the different types of beliefs with the following relation:

$$f_n(s_n, x_k, k \in D_n) = \frac{\pi_n(x_k, k \in D_n)p_n(s_n)}{\pi_n(x_k, k \in D_n)p_n(s_n) + (1 - \pi_n(x_k, k \in D_n))(1 - p_n(s_n))}. \tag{5.8}$$

For illustration purposes, whenever the arguments follow directly from the context, we will denote the social beliefs by just writing π_n and the posterior beliefs by f_n.

5.2 Characterization of Best Responses

In this section we characterize agents' best responses assuming fixed strategies for the rest and express them in terms of thresholds in the private or posterior belief space.

Specifically, assume that all agents other than n follow the strategy profile σ_{-n}. By the definition of agents' utilities, the expected payoff of agent n from taking action $y \in \{0, 1\}$, given that her private signal is $\tilde{s}_n$ and that she has observed decisions $\tilde{x}_k$ for $k \in D_n$ is equal to

$$\bar{U}_n(y; \tilde{s}_n, \tilde{x}_k, k \in D_n) = \mathbb{E}_{\sigma_{-n}}\left[(1-\delta)\left(\sum_{k=n}^{\infty} \delta^{k-n}\mathbf{1}_{\tilde{x}_k = \theta}\right) \,\middle|\, \tilde{I}_k, x_n = y\right],$$

where the expectation is taken over the state of the world θ and the process of all subsequent decisions. For compactness, we wrote $\tilde{I}_k = \{\tilde{s}_n, \tilde{x}_k, k \in D_n\}$.

The law of total expectation, conditioning on the value of θ, yields

$$\begin{aligned}
&\bar{U}_n(y; \tilde{s}_n, \tilde{x}_k, k \in D_n) \\
&= \mathbb{P}(\theta = 1 \mid \tilde{s}_n, \tilde{x}_k, k \in D_n)\mathbb{E}_{\sigma_{-n}}\left[(1-\delta)\left(\sum_{l=n}^{\infty} \delta^{l-n}\mathbf{1}_{x_l = 1}\right) \,\middle|\, \tilde{I}_k, x_n = y, \theta = 1\right] \\
&\quad + \mathbb{P}(\theta = 0 \mid \tilde{s}_n, \tilde{x}_k, k \in D_n)\mathbb{E}_{\sigma_{-n}}\left[(1-\delta)\left(\sum_{l=n}^{\infty} \delta^{l-n}\mathbf{1}_{x_l = 0}\right) \,\middle|\, \tilde{I}_k, x_n = y, \theta = 0\right],
\end{aligned}$$

where expectations are now taken only with respect to the decision process $\{x_l\}_{l \geq n}$.

Note that the private signal of agent n and the action of individual $k > n$ are independent conditioned on x_n, since private signals are not observable by subsequent agents. Using the latter property and the linearity of expectations we obtain,

$$
\begin{aligned}
&\bar{U}_n(y; \tilde{s}_n, \tilde{x}_k, k \in D_n) \\
&\quad = f_n(\tilde{s}_n, \tilde{x}_k, k \in D_n)(1-\delta)\left(\sum_{l=n}^{\infty} \delta^{l-n}\mathbb{P}_{\sigma_{-n}}(x_l = 1 \mid x_n = y, \tilde{x}_k, k \in D_n, \theta = 1)\right) \\
&\quad + (1 - f_n(\tilde{s}_n, \tilde{x}_k, k \in D_n))(1-\delta)\left(\sum_{l=n}^{\infty} \delta^{l-n}\mathbb{P}_{\sigma_{-n}}(x_l = 0 \mid x_n = y, \tilde{x}_k, k \in D_n, \theta = 0)\right).
\end{aligned}
$$

The probabilities that appear on the right hand side of each of the terms in the above expression can be computed given the private signal distribution and the strategies σ_{-n}. For ease of exposition we write,

$$A_n^{\sigma}(x_k, k \in D_n) \triangleq \sum_{j=n+1}^{+\infty} \delta^{j-(n+1)}\mathbb{P}_{\sigma}(x_j = 1 \mid x_n = 1, x_k, k \in D_n, \theta = 1), \tag{5.9}$$

$$B_n^{\sigma}(x_k, k \in D_n) \triangleq \sum_{j=n+1}^{+\infty} \delta^{j-(n+1)}\mathbb{P}_{\sigma}(x_j = 0 \mid x_n = 1, x_k, k \in D_n, \theta = 0), \tag{5.10}$$

$$C_n^{\sigma}(x_k, k \in D_n) \triangleq \sum_{j=n+1}^{+\infty} \delta^{j-(n+1)}\mathbb{P}_{\sigma}(x_j = 1 \mid x_n = 0, x_k, k \in D_n, \theta = 1), \tag{5.11}$$

$$F_n^{\sigma}(x_k, k \in D_n) \triangleq \sum_{j=n+1}^{+\infty} \delta^{j-(n+1)}\mathbb{P}_{\sigma}(x_j = 0 \mid x_n = 0, x_k, k \in D_n, \theta = 0). \tag{5.12}$$

Using Eqs. (5.9)-(5.12) we can calculate the expected payoffs from each action, for a specific private signal $\tilde{s}_n$ and observations $\tilde{x}_k$, $k \in D_n$, as follows:

$$
\begin{aligned}
\bar{U}(1; \tilde{s}_n, \tilde{x}_k \text{ for all } k \in D_n) &= (1-\delta) f_n(\tilde{s}_n, \tilde{x}_k, k \in D_n) A_n^{\sigma}(\tilde{x}_k, k \in D_n) \\
&+ (1-\delta)(1 - f_n(\tilde{s}_n, \tilde{x}_k, k \in D_n)) B_n^{\sigma}(\tilde{x}_k, k \in D_n),
\end{aligned}
$$

and

$$
\begin{aligned}
\bar{U}(0; \tilde{s}_n, \tilde{x}_k \text{ for all } k \in D_n) &= (1-\delta) f_n(\tilde{s}_n, \tilde{x}_k, k \in D_n) C_n^{\sigma}(\tilde{x}_k, k \in D_n) \\
&+ (1-\delta)(1 - f_n(\tilde{s}_n, \tilde{x}_k, k \in D_n)) F_n^{\sigma}(\tilde{x}_k, k \in D_n).
\end{aligned}
$$

Clearly, expected payoffs from each action are affine with respect to the posterior belief. The best response $\sigma^*(\tilde{I}_n)$ of agent n given her information set $\tilde{I}_n = \{\tilde{s}_n, \tilde{x}_k, k \in$

$D_n\}$, i.e. the action that maximizes her expected payoff, is the outcome of the comparison of the corresponding expected payoffs. Specifically,

$$\sigma^*(\tilde{I}_n) = \underset{y\in\{0,1\}}{\operatorname{argmax}} \{U_n(y; \tilde{x}_k, k \in D_n, \tilde{s}_n)\}.$$

Let $L_n^\sigma(\tilde{x}_k, k \in D_n)$ the indifference point, i.e., the value of $f_n(\tilde{s}_n, \tilde{x}_k, k \in D_n)$ for which the expected payoff from each action is the same. A simple calculation yields,

$$\begin{aligned} &L_n^\sigma(\tilde{x}_k, k \in D_n) \\ &\triangleq \frac{1 + \delta(F_n^\sigma(\tilde{x}_k, k \in D_n) - B_n^\sigma(\tilde{x}_k, k \in D_n))}{2 + \delta(A_n^\sigma(\tilde{x}_k, k \in D_n) + F_n^\sigma(\delta, \tilde{x}_k, k \in D_n) - B_n^\sigma(\tilde{x}_k, k \in D_n) - C_n^\sigma(\tilde{x}_k, k \in D_n))} \end{aligned} \quad (5.13)$$

An agent's best response is the outcome of the comparison of two affine functions that meet at $L_n^\sigma(\tilde{x}_k, k \in D_n)$. Therefore her best response can be described as in the following lemma.

Lemma 9. *Given the strategies σ_{-n} of all agents other than n a best response $\sigma^*(\tilde{I}_n)$ of individual n, given $\tilde{I}_n$ is, modulo an arbitrary tie braking rule, of the form*

$$\sigma^*(\tilde{I}_n) = \mathbf{1}_{f_n(\tilde{s}_n, \tilde{x}_k, k\in D_n) \geq L_n^\sigma(x_k, k\in D_n)} \quad \textit{or} \quad \sigma^*(\tilde{I}_n) = \mathbf{1}_{f_n(\tilde{s}_n, \tilde{x}_k, k\in D_n) < L_n^\sigma(x_k, k\in D_n)}$$

The above characterizes the best response of agent n in terms of a threshold in the space of posterior beliefs. We translate a best response in the space of private beliefs by defining for each possible set of observations $\{x_k\}_{k\in D_n}$,

$$\begin{aligned} &\mu_n(\tilde{x}_k, k \in D_n) \\ &\triangleq \frac{L_n(\tilde{x}_k, k \in D_n)(1 - \pi_n(\tilde{x}_k, k \in D_n))}{L_n(\tilde{x}_k, k \in D_n) - 2L_n(\tilde{x}_k, k \in D_n)\pi_n(\tilde{x}_k, k \in D_n) + \pi_n(\tilde{x}_k, k \in D_n)}. \end{aligned} \quad (5.14)$$

Using (5.8) we immediately get the equivalent characterization of the best response of agent n given her observations set $\tilde{I}_n$:

$$\sigma^*(\tilde{I}_n) = \mathbf{1}_{p_n(\tilde{s}_n) \geq \mu_n^\sigma(\tilde{x}_k, k\in D_n)} \quad \text{or} \quad \sigma^*(I_n) = \mathbf{1}_{p_n(\tilde{s}_n) < \mu_n^\sigma(\tilde{x}_k, k\in D_n)}.$$

The discussion of this section can be summarized in the following lemma.

Lemma 10. *Let $\sigma = \{\sigma_n\}_{n\in\mathbb{N}}$ be a Perfect Bayesian Equilibrium of the altruistic learning game. Then σ, modulo an arbitrary tie breaking rule, is of the form*

$$\sigma(I_n) = \mathbf{1}_{p_n(s_n) \geq \mu_n^\sigma(x_k, k\in D_n)} \quad \textit{or} \quad \sigma(I_n) = \mathbf{1}_{p_n(s_n) < \mu_n^\sigma(x_k, k\in D_n)}$$

or equivalently,

$$\sigma(I_n) = \mathbf{1}_{f_n(s_n, x_k, k \in D_n) \geq L_n^\sigma(x_k, k \in D_n)} \quad or \quad \sigma(I_n) = \mathbf{1}_{f_n(s_n, x_k, k \in D_n) < L_n^\sigma(x_k, k \in D_n)}$$

where $\mathbf{1}$ *denotes the indicator function and* $L_n^\sigma(x_k, k \in D_n)$, $\mu_n^\sigma(x_k, k \in D_n)$ *are as defined in (5.13) and (5.14).*

5.3 No learning in probability for forward looking agents

In this section we show the main result of this chapter; there is no Perfect Bayesian Equilibrium of the altruistic learning game that achieves learning in probability.

Throughout the rest of the chapter we assume that the private signal structure is characterized by Bounded Likelihood Ratios. Moreover we consider some fixed Perfect Bayesian Equilibrium σ and we drop the corresponding indices associated with σ.

Assume that the Perfect Bayesian Equilibrium σ achieves learning in probability. Then, under state of the world $\theta = 1$, since all agents will eventually be choosing 1 with high probability, blocks of size K with all agents choosing 1 will also occur with high probability. Simultaneously, this implies that the posterior probability of the state of the world being 0 after observing $x_k = 1$ for all $k \in D_n$ is negligible. Therefore, the social belief inferred from observing $x_k = 1$ for all $k \in D_n$ will eventually be arbitrarily close to one. Our first lemma formalizes this argument. For brevity, we define

$$\tilde{\pi}_n \triangleq \pi(x_k = 1, k \in D_n).$$

Lemma 11. *Assume that learning in probability occurs for equilibrium* σ*. Then :*

(i) $\lim_{n\to\infty} \mathbb{P}(x_k = 1, k \in D_n \mid \theta = 1) = 1$*, and*

(ii) $\lim_{n\to\infty} \tilde{\pi}_n = 1$.

Proof. (i) Fix some $\epsilon > 0$. We have assumed that learning in probability occurs. Since

$$\mathbb{P}(x_n = \theta) = \frac{1}{2}\mathbb{P}(x_n = 1 \mid \theta = 1) + \frac{1}{2}\mathbb{P}(x_n = 0 \mid \theta = 0),$$

we obtain

$$\lim_{n\to\infty} \mathbb{P}(x_n = 1 \mid \theta = 1) = 1, \tag{5.15}$$

$$\lim_{n\to\infty} \mathbb{P}(x_n = 0 \mid \theta = 0) = 1. \tag{5.16}$$

Hence, there exists some $\tilde{K} \in \mathbb{N}$ such that for all $n \geq \tilde{K}$, the following hold simultaneously:

$$\mathbb{P}(x_k = 1 \mid \theta = 1) > 1 - \frac{\epsilon}{K}, \tag{5.17}$$

$$\mathbb{P}(x_k = 1 \mid \theta = 0) < \epsilon. \tag{5.18}$$

Using the union bound and Equation (5.17) we obtain, for all $n > \tilde{K} + K$,

$$\begin{aligned}
\mathbb{P}(x_k = 1, k \in D_n \mid \theta = 1) &\geq 1 - \sum_{k=n-K}^{n-1} \mathbb{P}(x_k = 0 \mid \theta = 1) \qquad (5.19)\\
&= 1 - \sum_{k=n-K}^{n-1} (1 - \mathbb{P}(x_k = 1 \mid \theta = 1)) > 1 - \epsilon.
\end{aligned}$$

(ii) Continuing from above, using Equation (5.18) we have

$$\mathbb{P}(x_k = 1, k \in D_n \mid \theta = 0) \leq P(x_{n-1} = 1 \mid \theta = 0) < \epsilon. \tag{5.20}$$

Direct application of Bayes rule yields:

$$\tilde{\pi}_n = \left[1 + \frac{\mathbb{P}(x_k = 1, k \in D_n \mid \theta = 0)}{\mathbb{P}(x_k = 1, k \in D_n \mid \theta = 1)}\right]^{-1}. \tag{5.21}$$

It is immediate from the above equation that $\tilde{\pi}_n$ is increasing with respect to $\mathbb{P}(x_k = 1, k \in D_n \mid \theta = 1)$ and decreasing with respect to $\mathbb{P}(x_k = 1, k \in D_n \mid \theta = 0)$. Therefore, for all $n > \tilde{K} + K$

$$\tilde{\pi}_n \geq \left(1 + \frac{\epsilon}{1-\epsilon}\right)^{-1} = 1 - \epsilon,$$

which completes the proof.

□

The next lemma follows directly from the Bounded Likelihood Ratio assumption. Since agents' decisions are coupled between the states of the world, if a specific outcome of decisions for a finite subset of agents has positive probability under one state of the world, the same should hold under the other.

Lemma 12. *Assume that learning in probability occurs for equilibrium σ. Then there exists some $N \in \mathbb{N}$ such that for all $n > N$*

$$\mathbb{P}(x_k = 1, k \in D_n \mid \theta = 0) > 0 \tag{5.22}$$

Proof. From part (i) of Lemma 11 we know that there exists some $N \in \mathbb{N}$ such that for all $n > N$,

$$\mathbb{P}(x_k = 1 \text{ for all } k \in D_n \mid \theta = 1) > \frac{1}{2}. \tag{5.23}$$

We claim that for all $n > N$, we have $\mathbb{P}(x_k = 1 \text{ for all } k \in D_n \mid \theta = 0) > 0$. For the purpose of contradiction assume the contrary and let

$$\hat{n} = \min\{n > N : \mathbb{P}(x_k = 1, k \in D_n \mid \theta = 0) = 0\}.$$

Then, $\mathbb{P}(x_k = 1 \text{ for all } k \in D_{\hat{n}-1} \mid \theta = 0) > 0$ and

$$\mathbb{P}_0(x_{\hat{n}} = 1 \mid x_k = 1, k \in D_{\hat{n}} \mid \theta = 0) = 0,$$

which, using Lemma 1, implies that

$$0 \leq \mathbb{P}(x_{\hat{n}} = 1 \mid x_k = 1, k \in D_{\hat{n}}, \theta = 1) \leq \frac{1}{m}\mathbb{P}(x_{\hat{n}} = 1 \mid x_k = 1, k \in D_{\hat{n}}, \theta = 0) = 0,$$

and thus

$$\mathbb{P}(x_k = 1, k \in D_{\hat{n}+1} \mid \theta = 1) = 0,$$

which is a contradiction by (5.23).

□

Since Lemma 11 shows that the social belief converges to one for agents who observe ones from their K predecessors, and since the private signal of an agent cannot be strong enough to overcome the change of bias caused by strong social beliefs, the posterior belief must converge to one as well, as stated in the next Lemma. For brevity, define

$$\tilde{f}_n \triangleq f_n(s_n, x_k = 1, \text{ for all } k \in D_n).$$

Lemma 13. *Assume that learning in probability occurs for the equilibrium σ. Then,*

$$\lim_{n\to\infty} \tilde{f}_n = 1$$

Proof. Observe from Equation (5.8) that $\tilde{f}_n$ is increasing with respect to $\tilde{\pi}_n$, as $\tilde{\pi}_n$

ranges in $[0, 1]$. From Proposition 2, $s_n \geq \underline{\beta}$, with probability one. Therefore,

$$\tilde{f}_n \triangleq f_n(\delta, x_k = 1, k \in D_n) = \frac{\tilde{\pi}_n \sigma_n}{\tilde{\pi}_n \sigma_n + (1 - \tilde{\pi}_n)(1 - \sigma_n)} \geq \frac{\tilde{\pi}_n \underline{\beta}}{\tilde{\pi}_n \underline{\beta} + (1 - \tilde{\pi}_n)(1 - \underline{\beta})}.$$

Lemma 11 guarantees that under the assumption of learning in probability,

$$\lim_{n \to \infty} \tilde{\pi}_n = 1.$$

Therefore, taking limits on both sides of the above, we get

$$1 \geq \lim_{n \to \infty} \tilde{f}_n \geq \liminf_{n \to \infty} \tilde{f}_n \geq 1,$$

which completes the proof. □

If learning in probability occurs for the equilibrium σ, under the state of the world $\theta = 1$ agents' decisions are 1 with high probability. If there were agents who deviate with sufficiently positive probability from choosing 1 when observing all their predecessors' decisions to be 1, this would immediate contradict learning. The next lemma formalizes this argument.

Lemma 14. *Assume that learning in probability occurs for the equilibrium σ. Then,*

$$\lim_{n \to \infty} \mathbb{P}(x_n = 1 \mid x_k = 1, k \in D_n, \theta = 1) = 1. \tag{5.24}$$

Proof. Assume for the purpose of deriving a contradiction that there exists an $\epsilon > 0$ and a subsequence $\{r_i\}_{i=1}^{\infty}$ of agents for which for all $i \in \mathbb{N}$,

$$\mathbb{P}(x_{r_i} = 1 \mid x_k = 1, k \in D_{r_i}, \theta = 1) < 1 - \epsilon. \tag{5.25}$$

Part (i) of Lemma 11 suggests that

$$\lim_{n \to \infty} \mathbb{P}(x_k = 1, k \in D_n \mid \theta = 1) = 1.$$

Therefore, there exists some $\hat{N} \in \mathbb{N}$ such that, for all $n > \hat{N}$

$$\mathbb{P}(x_k = 1, k \in D_n \mid \theta = 1) > 1 - \frac{\epsilon}{4}, \tag{5.26}$$

Let $H \triangleq \{0, 1\}^K \setminus \{\mathbf{e}\}$, where $\mathbf{e}$ denotes the K-dimensional vector of ones. In words, H consists of all K-dimensional vectors for which at least one element is not one.

Then,

$$\sum_{h \in H} \mathbb{P}(x_{D_n} = h \mid \theta = 1) = \mathbb{P}(x_{D_n} \in H \mid \theta = 1) \tag{5.27}$$
$$= 1 - \mathbb{P}(x_k = 1 \text{ for all } k \in D_n \mid \theta = 1) < \frac{\epsilon}{4}$$

for all $n > \hat{N}$.

Define the subsequence $\{w_i\}_{i=1}^{+\infty} = \{r_i : r_i > \hat{N}, i \in \mathbb{N}\}$. The probability of agent w_i choosing one is equal to

$$\begin{aligned} \mathbb{P}(x_{w_i} = 1 \mid \theta = 1) &= \mathbb{P}(x_{D_{w_i}} = \mathbf{e} \mid \theta = 1)\mathbb{P}(x_{w_i} = 1 \mid x_{D_{w_i}} = \mathbf{e}, \theta = 1) \\ &+ \sum_{h \in H} \mathbb{P}(x_{D_{w_i}} = h \mid \theta = 1)\mathbb{P}(x_{w_i} = 1 \mid x_{D_{w_i}} = h, \theta = 1) \end{aligned}$$

We can bound the first component of the first term by one and the second by $1-\epsilon$ using Equation (5.25). Similarly, for the second term,

$$\sum_{h \in H} \mathbb{P}(x_{D_{w_i}} = h \mid \theta = 1)\mathbb{P}(x_{w_i} = 1 \mid x_{D_{w_i}} = h, \theta = 1) \leq \sum_{h \in H} \mathbb{P}(x_{D_{w_i}} = h \mid \theta = 1) \leq \frac{\epsilon}{4},$$

where we bound $\mathbb{P}(x_{w_i} = 1 \mid x_{D_{w_i}} = h, \theta = 1)$ by one for the first inequality and we use (5.27) for the second. Therefore, we obtain

$$\mathbb{P}(x_{w_i} = 1 \mid \theta = 1) \leq 1 - \epsilon + \frac{\epsilon}{4} = 1 - \frac{3}{4}\epsilon,$$

which contradicts (5.26), completing the proof. □

The sequence of lemmas that follow, are the core of our proof and essentially prove a stronger version of Lemma 14, establishing that the thresholds with which agents compare their private beliefs converge to zero. This fact, combined with the boundedness of the support of the private beliefs established in Proposition 2, proves that there exists some finite time after which agents start choosing one irrespective of their private signal, when observing ones from their predecessors, for any state of the world. This is acceptable and expected when $\theta = 1$ but not, as we will shortly prove, if $\theta = 0$.

For brevity, let

$$\tilde{A}_n = A_n(x_k = 1, k \in D_n) \text{ and } \tilde{B}_n = B_n(x_k = 1, k \in D_n),$$

where $A_n(x_k = 1, k \in D_n)$ is defined in (5.9).

Lemma 15. *Assume that learning in probability occurs for equilibrium σ. Then,*

(i) $\lim_{n\to\infty} \tilde{A}_n = 1/(1-\delta)$ *and*

(ii) $\lim_{n\to\infty} \tilde{B}_n = 0$.

Proof. (i) From the definition of $\tilde{A}_n$ we have

$$\tilde{A}_n = \sum_{k=n+1}^{\infty} \delta^{k-(n+1)} \mathbb{P}(x_k = 1 \mid x_j = 1, j \in D_n, \theta = 1) \leq \sum_{k=n+1}^{\infty} \delta^{k-(n+1)} = \frac{1}{1-\delta}.$$

Therefore

$$\limsup_{n\to\infty} \tilde{A}_n \leq \frac{1}{1-\delta}. \tag{5.28}$$

Getting a lower bound for $\liminf_{n\to\infty} \tilde{A}_n$ is more involved. From the definition of $\tilde{A}_n$ obtain

$$\begin{aligned} \tilde{A}_n &= \sum_{k=1}^{\infty} \delta^{k-(n+1)} \mathbb{P}(x_k = 1 \mid x_j = 1, j \in D_n, x_n = 1, \theta = 1) \qquad (5.29) \\ &\geq \mathbb{P}(x_{n+1} = 1 \mid x_j = 1, j \in D_n, x_n = 1, \theta = 1) \\ &\quad \cdot \left(1 + \delta \sum_{k=n+2}^{\infty} \delta^{k-(n+2)} \mathbb{P}(x_k = 1 \mid x_j = 1, j \in D_n, x_n = 1, x_{n+1} = 1, \theta = 1)\right) \end{aligned}$$

where the inequality follows from the fact that when expanding $\tilde{A}_n$ we only consider the case where $x_{n+1} = 1$.

Now, we can use the following Markovian property of the decision process: an agent's decision is independent of the past decisions, conditioned on the observed actions of agents in D_n. More precisely,

$$\mathbb{P}(x_{n+1} = 1 \mid x_j = 1, j \in D_n, x_n = 1, \theta = 1) = \mathbb{P}(x_{n+1} = 1 \mid x_j = 1, j \in D_{n+1}, \theta = 1),$$

and

$$\mathbb{P}(x_k = 1 \mid x_j = 1, j \in D_n, x_n = 1, x_{n+1} = 1, \theta = 1) = \mathbb{P}(x_k = 1 \mid x_j = 1, j \in D_k, \theta = 1),$$

for all $k \geq n+2$.

Rewriting (5.29) by using the equalities above, we obtain

$$\begin{aligned}\tilde{A}_n &\geq \mathbb{P}(x_{n+1} = 1 \mid x_j = 1, j \in D_{n+1}, \theta = 1) \\ &\cdot \left(1 + \delta \sum_{k=n+2}^{\infty} \delta^{k-(n+2)} \mathbb{P}(x_k = 1 \mid x_j = 1, j \in D_{n+2}, \theta = 1)\right).\end{aligned}$$

We recognize the second term in the parenthesis as $\tilde{A}_{n+1}$, so that

$$\tilde{A}_n \geq \mathbb{P}(x_{n+1} = 1 \mid x_j = 1, j \in D_{n+1}, \theta = 1)(1 + \delta \tilde{A}_{n+1}). \tag{5.30}$$

Lemma 14 implies that for any $\epsilon > 0$ there exists some N such that for all $n > N$

$$\mathbb{P}(x_{n+1} = 1 \mid x_j = 1, j \in D_n, x_n = 1, \theta = 1) > 1 - \epsilon,$$

and thus, for those agents

$$\tilde{A}_n \geq (1 - \epsilon)(1 + \delta \bar{A}_{n+1}).$$

Therefore,

$$\liminf_{n \to \infty} \tilde{A}_n \geq (1 - \epsilon)(1 + \delta \liminf_{n \to \infty} \tilde{A}_n)$$

and consequently,

$$\liminf_{n \to \infty} \tilde{A}_n \geq \frac{1 - \epsilon}{1 - \delta(1 - \epsilon)}, \quad \text{for all } \epsilon > 0,$$

which in turn implies that

$$\liminf_{n \to \infty} \tilde{A}_n \geq \frac{1}{1 - \delta}.$$

The above, together with (5.28) proves the desired result.

(ii) Arguing similar to part (i), we get

$$\begin{aligned}\tilde{B}_n &\leq \mathbb{P}(x_{n+1} = 1 \mid x_k = 1, k \in D_n, x_n = 1, \theta = 0)(0 + \delta \tilde{B}_{n+1}) \\ &+ \mathbb{P}(x_{n+1} = 0 \mid x_k = 1, k \in D_n, x_n = 1, \theta = 0)\frac{1}{1 - \delta}\end{aligned}$$

where the inequality follows from the obvious bound

$$1+\delta\sum_{j=n+1}^{\infty}\delta^{j-(n+1)}\mathbb{P}(x_d=1\mid x_k=1,k\in D_m,x_n=1,x_{n+1}=0)\leq\frac{1}{1-\delta}.$$

Lemma 14 implies that for any $\epsilon>0$ there exists some N such that for all $n>N$,

$$\mathbb{P}(x_{n+1}=0\mid x_k=1,k\in D_{n+1},\theta=1)<\frac{1}{M}\epsilon.$$

By the Bounded Likelihood Ratio assumption we get that for all $n>N$,

$$\mathbb{P}(x_{n+1}=0\mid x_k=1,k\in D_{n+1},\theta=0)<\epsilon$$

and thus,

$$\tilde{B}_n\leq\delta\tilde{B}_{n+1}+\epsilon\frac{1}{1-\delta},$$

from which we conclude that

$$\limsup_{n\to\infty}\tilde{B}_n\leq\frac{\epsilon}{(1-\delta)^2}.$$

Letting $\epsilon\to 0$ yields the desired result.

□

The above is an important step for the proof of our main theorem. It states that the thresholds $L_n(x_k=1,k\in D_n)$, to which agents compare their posterior belief (cf. Equation 5.13), do not converge to one as shown in the following lemma. For brevity we use the following notation :

$$\tilde{L}_n\triangleq L_n(\delta,x_k=1,k\in D_n).$$

Lemma 16. *Assume that learning in probability occurs for the equilibrium σ. Then,*

$$\limsup_{n\to\infty}\tilde{L}_n<1.$$

Proof. For convenience let $\tilde{A}_n=A_n(x_k=1,k\in D_n)$, $\tilde{B}_n=B_n(x_k=1,k\in D_n)$, $\tilde{C}_n=C_n(x_k=1,k\in D_n)$, $\tilde{F}_n=F_n(x_k=1,k\in D_n)$. Moreover, for compactness define,

$$a_n=1+\delta(\tilde{F}_n-\tilde{B}_n)$$

and

$$b_n = 1 + \delta(\tilde{A}_n - \tilde{C}_n).$$

From part (i) of Lemma 15 we get that

$$\liminf_{n \to \infty} b_n \geq 1, \tag{5.31}$$

while from part (ii) of the same Lemma we conclude that

$$\liminf_{n \to \infty} a_n \geq 1. \tag{5.32}$$

By the definition (5.13) of the thresholds in the posterior belief space,

$$\tilde{L}_n = \frac{a_n}{a_n + b_n}.$$

From (5.31) and (5.32) we conclude that there exists a $N \in \mathbb{N}$ such that for all $n > N$ $a_n > 1/2$ and $b_n > 1/2$. Then, for all $n > N$, $\tilde{L}_n$ is increasing with respect to b_n and therefore,

$$\tilde{L}_n \leq \frac{a_n}{a_n + \frac{1}{2}}.$$

A trivial upper bound for a_n can be obtained by observing that

$$a_n = 1 + \delta(\tilde{F}_n - \tilde{B}_n) \leq 1 + \delta\frac{1}{1-\delta} + 0 = \frac{1}{1-\delta},$$

from which we conclude that for all $n > N$,

$$\tilde{L}_n \leq \frac{\frac{1}{1-\delta}}{\frac{1}{1-\delta} + \frac{1}{2}},$$

implying,

$$\limsup_{n \to \infty} \tilde{L}_n \leq \frac{\frac{1}{1-\delta}}{\frac{1}{1-\delta} + \frac{1}{2}} < 1,$$

which concludes the proof.

□

The next Lemma establishes that for learning in probability to occur only one of the two possible strategies of Lemma 9 should eventually "survive", because the contrary would contradict learning.

Lemma 17. *Assume that learning in probability occurs for the equilibrium σ. Then* *there exists some N such that for all $n > N$, agents, when they observe $x_k = 1$ for all $k \in D_n$, decide according to*

$$f_n \underset{x_n=1}{\overset{x_n=0}{\lesseqqgtr}} L_n(x_k = 1, k \in D_n) \text{ or equivalently } p_n \underset{x_n=1}{\overset{x_n=0}{\lesseqqgtr}} \mu_n(x_k = 1, k \in D_n).$$

Proof. Assume to arrive at a contradiction that there exists a sequence $\{r_i\}_{i\in\mathbb{N}}$ that decides according to

$$f_{r_i} \underset{x_{r_i}=0}{\overset{x_{r_i}=1}{\lesseqqgtr}} \tilde{L}_{r_i},$$

that is, they choose zero for large posterior beliefs.

From Lemma 16, there exists some $\hat{L} < 1$ and some $\hat{N} \in \mathbb{N}$ such that for all $n > N$, $\tilde{L}_{r_i} < \hat{L}$.

From Corollary 13 $\lim_{i\to\infty} \tilde{f}_{r_i} = 1$ and therefore there exists a $\tilde{N} \in \mathbb{N}$ such that for all $i > \tilde{N}$, $\tilde{f}_{r_i} > (1 + \hat{L})/2$. In that case, for all $i > \max\{\hat{N}, \tilde{N}\}$,

$$\mathbb{P}(x_{r_i} = 0 \mid x_k = 1, k \in D_{r_i}, \theta = 1) = 1,$$

contradicting Lemma 14 and concluding the proof. □

The last step is to prove that the threshold with which agents compare their private belief converges to zero. What we need, in order to prove that for an equilibrium which achieves learning in probability, agents eventually copy blocks of ones (under either state of the world), is to bound $\limsup_{n\to\infty} \mu(x_k = 1, k \in D_n)$ by a number below $\underline{\beta}$, the lower end of the support of the private beliefs, but this lemma proves something stronger.

Lemma 18. *Assume that, for the equilibrium σ, learning in probability occurs. Then,*

$$\lim_{n\to\infty} \mu(\delta, x_k = 1 \text{ for all } k \in D_n) = 0.$$

Proof. Assume, to arrive at a contradiction, that there exists a sequence $\{r_i\}_{i\in\mathbb{N}}$ such that

$$\lim_{i\to\infty} \mu(x_k = 1, k \in D_{r_i}) = \zeta,\ \zeta \neq 0.$$

Rearranging (5.14), yields

$$\tilde{L}_{r_i} = \frac{\mu(x_k = 1, k \in D_{r_i})}{1 - \tilde{\pi}_{r_i} - \mu(x_k = 1, k \in D_{r_i}) + 2\mu(x_k = 1, k \in D_{r_i})\tilde{\pi}_{r_i}}.$$

Using Lemma 11 and taking the limit as $i \to \infty$, we get

$$\limsup_{i\to\infty} \bar{L}_{r_i} = \frac{\zeta}{\zeta} = 1 \tag{5.33}$$

which contradicts Lemma 16, completing the proof. □

At this point we have established that under either state of the world, at some finite time, the threshold with which agents compare their private belief given that they observe 1 from their predecessors, falls below the lower end of the support of their private belief. Therefore, irrespective of their private signal, after that point in time agents will choose one with high probability. On the other hand, with positive probability, under state $\theta = 0$, K agents may choose 1. In that case, all subsequent agents will be choosing 1, contradicting the learning in probability assumption. The next theorem formalizes this argument.

Theorem 4. *Assume that the signal structure is characterized by Bounded Likelihood Ratios and that agents are forward looking. Then, for any $\delta \in [0, 1)$ there does not exist a Perfect Bayesian Equilibrium of the altruistic learning game under which learning in probability is achieved. In other words, if σ is a Perfect Bayesian Equilibrium of the altruistic learning game, then*

$$\lim_{n\to\infty} \mathbb{P}_\sigma(x_n = \theta) < 1.$$

Proof. Assume that learning in probability did occur under the perfect Bayesian Equilibrium σ. We prove the result by contradiction.

By Lemma 18, there exists a $N_1 \in \mathbb{N}$ such that for all $n > N_1$,

$$\mu(x_k = 1, k \in D_n) < \frac{1}{2}\underline{\beta}.$$

where $\underline{\beta} > 0$ is the left end of the support of the private beliefs as defined in Proposition 2. Therefore, using Lemma 17, we conclude that for all $n > N_1$

$$\mathbb{P}(x_n = 0 \mid x_k = 1, k \in D_n, s_n, \theta = j) = 0 \text{ for all } j \in \{0, 1\}. \tag{5.34}$$

On the other hand, Lemma 12 guarantees the existence of a $N_2 \in \mathbb{N}$ such that for all $n > N_2$,

$$\mathbb{P}(x_k = 1, k \in D_n, \mid \theta = 0) > 0.$$

Choose some $\hat{n} > \max\{N_1, N_2\}$ such that

$$\mathbb{P}(x_k = 1, k \in D_{\hat{n}} \mid \theta = 0) > 0.$$

Then for all $n > \hat{n}$, using (5.34),

$$\begin{aligned}\mathbb{P}(x_n = 1 \mid \theta = 0) &\geq \mathbb{P}(x_k = 1, k \in D_{\hat{n}} \mid \theta = 0)\mathbb{P}(x_n = 1 \mid x_k = 1, k \in D_{\hat{n}}, \theta = 0) \\ &= \mathbb{P}(x_k = 1, k \in D_{\hat{n}} \mid \theta = 0) > 0,\end{aligned}$$

and thus,

$$\liminf_{n\to\infty} \mathbb{P}(x_n = 1 \mid \theta = 0) \geq \mathbb{P}(x_k = 1, k \in D_{\hat{n}} \mid \theta = 0) > 0,$$

contradicting the learning in probability assumption and concluding the proof of the Theorem. □

5.4 Discussion and Conclusions

The main result of this chapter is the impossibility of learning in probability by agents who are forward looking, maximizing the discounted sum (over themselves and subsequent agents) of the probabilities of a correct decision. This is in contrast to the results of Chapter 4 where we showed that there exists a decision rule that guarantees learning in probability. One would have hoped that, since agents are altruistic and as we argued earlier in this chapter could sacrifice myopically optimal decisions for the welfare of future agents, learning in probability might be possible.

Surprisingly, this is not the case. One could argue that this happens because of the fact that in individuals' payoffs the payoff from a correct decision is multiplied by 1 while correct decisions for subsequent agents return payoffs $\delta, \delta^2, \ldots$ and therefore myopic behaviour dominates. On the other hand, if we change our model slightly starting the summation in agents' payoffs from agent $n + l$ for some $l > 0$ the same proof and therefore the same results would go through. (But, if individuals' payoffs are given by the limit itself of the probability of correct decisions, then the decision rule designed in Chapter 4 is an equilibrium.)

Note that in he course of our proof we did not at any point characterize equilibrium strategy profiles for the altruistic learning game. We just stated necessary conditions for a strategy profile to be a Perfect Bayesian Equilibrium and built on them to prove our result. As we discuss in the next chapter, an explicit characterization of equilibria is very involved and for this reason it is harder to establish positive results.

Concluding, it is an open and especially interesting problem to provide a natural payoff structure that would allow equilibria that achieve learning in probability. On the other hand, decision rules such as the one in Chapter 4 involve peculiar actions (for example, always choosing one when the observed actions are 01 for any private signal) that are unlikely to be supported as best responses of strategic agents.

Chapter 6

Forward looking agents: Unbounded Likelihood Ratios

In this chapter we focus on forward looking agents and study the learning properties of the equilibria of the altruistic game under the Unbounded Likelihood Ratios assumption. The first section explores some useful properties of *monotone equilibria*, that will be used in the analysis that follows. The second section presents an equilibrium for the case of symmetric private signal distributions that achieves learning in probability. Next, we construct from this equilibrium an equilibrium that does not achieve learning in probability. A key property of the non-learning equilibrium is its *non monotonicity*; i.e., there exist agents who decide zero when their posterior belief is large, which draws our attention to monotone equilibria. Finally, we prove the existence of monotone equilibria for the altruistic learning game and we state a conjecture about necessary and sufficient conditions under which equilibria achieve learning in probability.

6.1 Characterization and properties of monotone equilibria.

Throughout this chapter we consider forward looking agents, introduced in Section 2.3. We concentrate on the case $K = 1$, i.e. on agents who just observe the decision of their immediate predecessor. We start this section by defining monotone equilibria.

Definition 9. *We say that a Perfect Bayesian Equilibrium σ is* **monotone** *if there exist sequences $\{\underline{\mu}_n\}_{n\in\mathbb{N}}$, $\{\bar{\mu}_n\}_{n\in\mathbb{N}}$ of real numbers such that the strategy of each player* *can be described as*

$$\sigma_n(s_n, x_{n-1}) = \begin{cases} 1, & \text{if } p_n(s_n) \geq \bar{\mu}_n \\ x_{n-1}, & \text{if } \bar{\mu}_n > p_n(s_n) \geq \underline{\mu}_n \\ 0, & \text{if } p_n(s_n) < \underline{\mu}_n \end{cases}$$

Equivalently, for a specific monotone Perfect Bayesian Equilibrium σ one could specify a sequence of thresholds $\{l_n\}_{n\in\mathbb{N}}$ such that agents' strategies are given by

$$\sigma_n(s_n, x_{n-1}) = \begin{cases} 1, & \text{if } f_n(s_n, x_{n-1}) \geq l_n, \\ 0, & \text{otherwise.} \end{cases}$$

Once a sequence of thresholds $\{l_n\}_{n\in\mathbb{N}}$ is specified, a sequence of $\{\underline{\mu}_n\}_{n\in\mathbb{N}}, \{\bar{\mu}_n\}_{n\in\mathbb{N}}$ thresholds is also defined through,

$$\underline{\mu}_n = \min_{j\in\{0,1\}} \left\{ \frac{l_n(1 - \pi_n(x_{n-1} = j))}{\pi_n(x_{n-1} = j) - 2l_n\pi_n(x_{n-1} = j) + l_n} \right\}, \tag{6.1}$$

and

$$\bar{\mu}_n = \max_{j\in\{0,1\}} \left\{ \frac{l_n(1 - \pi_n(x_{n-1} = j))}{\pi_n(x_{n-1} = j) - 2l_n\pi_n(x_{n-1} = j) + l_n} \right\}. \tag{6.2}$$

Therefore, whenever we refer to agents' strategies through the thresholds $\{l_n\}_{n\in\mathbb{N}}$ we essentially refer to the corresponding sequences of thresholds $\{\underline{\mu}_n\}_{n\in\mathbb{N}}, \{\bar{\mu}_n\}_{n\in\mathbb{N}}$.

In order to study the properties of monotone equilibria we make use of a fact about the private belief distribution that we present in the following Lemma from [1]. We denote the conditional distribution of private beliefs, given the underlying state by $\mathbb{G}_j$ for each $j \in \{0, 1\}$, i.e.,

$$\mathbb{G}_j(r) = \mathbb{P}(p_1 \leq r \mid \theta = j).$$

Lemma 19. *Assume that the private signal structure is characterized by Unbounded Likelihood Ratios. Then, the following relation holds:*

$$\mathbb{G}_0(r) > \mathbb{G}_1(r), \quad \text{for all } r \in \mathbb{R}.$$

This lemma establishes an intuitively obvious fact, that it is more likely to get smaller private beliefs under state of the world $\theta = 0$ than under state of the world $\theta = 1$. Using the above, we establish a series of properties for monotone equilibria.

Lemma 20. *Consider a monotone equilibrium σ, described by the sequence of thresholds $\{l_n\}_{n\in\mathbb{N}}$. Then, for all $n \in \mathbb{N}$, $i, j \in \{0,1\}$, and $k > n$, we have:*

(i) $\mathbb{P}_\sigma(x_n = j \mid \theta = j) > \mathbb{P}_\sigma(x_n = j \mid \theta = 1-j)$,

(ii) $\mathbb{P}_\sigma(x_n = i \mid x_{n-1} = i, \theta = j) > \mathbb{P}_\sigma(x_n = i \mid x_{n-1} = 1-i, \theta = j)$,

(iii) $\mathbb{P}_\sigma(x_k = i \mid x_n = i, \theta = j) > \mathbb{P}_\sigma(x_k = i \mid x_n = 1-i, \theta = j)$.

Proof.

(i) We prove the result by induction. Indeed, for $n = 1$

$$\mathbb{P}_\sigma(x_1 = 1 \mid \theta = 1) = \mathbb{G}_1(p_n > l_1) > \mathbb{G}_0(p_n > l_1) = \mathbb{P}_\sigma(x_1 = 1 \mid \theta = 0),\ j \in \{0,1\},$$

and

$$\mathbb{P}_\sigma(x_1 = 0 \mid \theta = 0) = \mathbb{G}_0(p_n < l_1) < \mathbb{G}_1(p_n > l1) = \mathbb{P}_\sigma(x_1 = 0 \mid \theta = 1),\ j \in \{0,1\}.$$

Assume that

$$\mathbb{P}_\sigma(x_{n-1} = j \mid \theta = j) > \mathbb{P}_\sigma(x_{n-1} = j \mid \theta = 1-j),$$

for all $j \in \{0,1\}$. We prove the result for agent n.

The social belief induced from the observed decision $x_{n-1} = \hat{x}_{n-1}$ is

$$\mathbb{P}_\sigma(\theta = 1 \mid x_{n-1} = \hat{x}_{n-1}) = \frac{1}{1 + \frac{\mathbb{P}(x_{n-1} = \hat{x}_{n-1} \mid \theta = 0)}{\mathbb{P}(x_{n-1} = \hat{x}_{n-1} \mid \theta = 1)}}.$$

Using the induction hypothesis we can conclude that

$$\frac{\mathbb{P}(x_{n-1} = 1 \mid \theta = 0)}{\mathbb{P}(x_{n-1} = 1 \mid \theta = 1)} < 1 < \frac{\mathbb{P}(x_{n-1} = 0 \mid \theta = 0)}{\mathbb{P}(x_{n-1} = 0 \mid \theta = 1)},$$

and therefore

$$\mathbb{P}_\sigma(\theta = 1 \mid x_{n-1} = 1) > \mathbb{P}_\sigma(\theta = 1 \mid x_{n-1} = 0),$$

establishing that the social belief induced by decisions of 1 is higher, namely $\pi(x_{n-1} = 1) > \pi(x_{n-1} = 0)$. Therefore,

$$\mu_n(x_{n-1} = 1) = \frac{l_n(1 - \pi_n(x_{n-1} = 1))}{\pi_n(x_{n-1} = 1) - 2l_n\pi_n(x_{n-1} = 1) + l_n} < \mu_n(x_{n-1} = 0)$$

and thus,

$$1-\mathbb{P}\left(p_n<\mu_n(x_{n-1}=1)\mid\theta=j\right)>1-\mathbb{P}\left(p_n<\mu_n(x_{n-1}=0)\mid\theta=j\right),$$

which in turn implies

$$\mathbb{P}_\sigma(x_n=1\mid x_{n-1}=1,\theta=j)>\mathbb{P}_\sigma(x_n=1\mid x_{n-1}=0,\theta=j). \tag{6.3}$$

A symmetric argument yields

$$\mathbb{P}_\sigma(x_n=0\mid x_{n-1}=0,\theta=j)>\mathbb{P}_\sigma(x_n=0\mid x_{n-1}=1,\theta=j). \tag{6.4}$$

Assume that agent n observes the action $\hat{x}_{n-1}$ from her immediate predecessor. Then, she computes her social belief $\hat{\pi}_n=\pi_n(\hat{x}_{n-1})$ and decides according to

$$p_n \underset{x_n=1}{\overset{x_n=0}{\lesseqgtr}} \frac{l_n(1-\hat{\pi}_n)}{\hat{\pi}_n-2l_n\hat{\pi}_n+l_n}.$$

Therefore,

$$\mathbb{P}_\sigma(x_n=1\mid x_{n-1}=\hat{x}_{n-1},\theta=1)=1-\mathbb{P}_\sigma\left(p_n<\frac{l_n(1-\hat{\pi}_n)}{\hat{\pi}_n-2l_n\hat{\pi}_n+l_n}\;\middle|\;\theta=1\right)$$

On the other hand,

$$1-\mathbb{P}_\sigma\left(p_n<\frac{l_n(1-\hat{\pi}_n)}{\hat{\pi}_n-2l_n\hat{\pi}_n+l_n}\;\middle|\;\theta=1\right)>1-\mathbb{P}_\sigma\left(p_n<\frac{l_n(1-\hat{\pi}_n)}{\hat{\pi}_n-2l_n\hat{\pi}_n+l_n}\;\middle|\;\theta=0\right),$$

from Lemma 19 and thus,

$$\mathbb{P}_\sigma(x_n=1\mid x_{n-1}=\hat{x}_{n-1},\theta=1)>\mathbb{P}_\sigma(x_n=1\mid x_{n-1}=\hat{x}_{n-1},\theta=0). \tag{6.5}$$

Hence,

$$\begin{aligned}
&\mathbb{P}_\sigma(x_n=1\mid\theta=1)\\
&\quad=\mathbb{P}_\sigma(x_n=1\mid x_{n-1}=1,\theta=1)\mathbb{P}_\sigma(x_{n-1}=1\mid\theta=1)\\
&\quad+\mathbb{P}_\sigma(x_n=1\mid x_{n-1}=0,\theta=1)\mathbb{P}_\sigma(x_{n-1}=0\mid\theta=1)\\
&\quad=(\mathbb{P}_\sigma(x_n=1\mid x_{n-1}=1,\theta=1)-\mathbb{P}_\sigma(x_n=1\mid x_{n-1}=0,\theta=1))\mathbb{P}_\sigma(x_{n-1}=1\mid\theta=1)\\
&\quad+\mathbb{P}_\sigma(x_n=1\mid x_{n-1}=0,\theta=1).
\end{aligned}$$

Equation (6.3) guarantees that $\mathbb{P}_\sigma(x_n = 1 \mid x_{n-1} = 1, \theta = 1) - \mathbb{P}_\sigma(x_n = 1 \mid x_{n-1} = 0, \theta = 1) \geq 0$ from which, using the induction hypothesis we obtain

$$\begin{aligned}
&\mathbb{P}_\sigma(x_n = 1 \mid \theta = 1) \\
&= (\mathbb{P}_\sigma(x_n = 1 \mid x_{n-1} = 1, \theta = 1) - \mathbb{P}_\sigma(x_n = 1 \mid x_{n-1} = 0, \theta = 1))\mathbb{P}_\sigma(x_{n-1} = 1 \mid \theta = 0) \\
&+ \mathbb{P}_\sigma(x_n = 1 \mid x_{n-1} = 0, \theta = 1) \\
&= \mathbb{P}_\sigma(x_n = 1 \mid x_{n-1} = 1, \theta = 1)\mathbb{P}_\sigma(x_{n-1} = 1 \mid \theta = 0) \\
&+ \mathbb{P}_\sigma(x_n = 1 \mid x_{n-1} = 0, \theta = 1)\mathbb{P}_\sigma(x_{n-1} = 0 \mid \theta = 0).
\end{aligned}$$

The last step is to use (6.5) and conclude that

$$\begin{aligned}
&\mathbb{P}_\sigma(x_n = 1 \mid \theta = 1) \\
&= \mathbb{P}_\sigma(x_n = 1 \mid x_{n-1} = 1, \theta = 1)\mathbb{P}_\sigma(x_{n-1} = 1 \mid \theta = 0) \\
&+ \mathbb{P}_\sigma(x_n = 1 \mid x_{n-1} = 0, \theta = 1)\mathbb{P}_\sigma(x_{n-1} = 0 \mid \theta = 0) \\
&> \mathbb{P}_\sigma(x_n = 1 \mid \theta = 0).
\end{aligned}$$

The other case ($j = 0$) follows similarly.

(ii) This has already been proven in (i).

(iii) We use induction to prove the result. For brevity we write $\mathbb{P}_\sigma^j$ to denote conditioning on the event $\theta = j$. The first step of the induction, namely for $k = n+1$ is part (ii). Assume that the result holds for k, i.e.,

$$\mathbb{P}_\sigma^j(x_k = i \mid x_n = i) > \mathbb{P}_\sigma^j(x_k = i \mid x_n = 1 - i)$$

where $i, j \in \{0, 1\}$. We prove it for $k + 1$. Indeed,

$$\begin{aligned}
&\mathbb{P}_\sigma^j(x_{k+1} = i \mid x_n = i) \\
&= \mathbb{P}_\sigma^j(x_{k+1} = i \mid x_k = i)\mathbb{P}_\sigma^j(x_k = i \mid x_n = i) \\
&+ \mathbb{P}_\sigma^j(x_{k+1} = i \mid x_k = 1 - i)\mathbb{P}_\sigma^j(x_k = 1 - i \mid x_n = i)
\end{aligned}$$

from which, using $\mathbb{P}_\sigma^j(x_k = i \mid x_n = i) + \mathbb{P}_\sigma^j(x_k = 1 - i \mid x_n = i) = 1$, we conclude that

$$
\begin{aligned}
&\mathbb{P}^j_\sigma(x_{k+1} = i \mid x_n = i) \\
&\quad = (\mathbb{P}^j_\sigma(x_{k+1} = i \mid x_k = i) - \mathbb{P}^j_\sigma(x_{k+1} = i \mid x_k = 1-i))\mathbb{P}^j_\sigma(x_k = i \mid x_n = i) \\
&\quad + \mathbb{P}^j_\sigma(x_k = 1-i \mid x_k = i).
\end{aligned}
$$

From the previous part, we get that $\mathbb{P}^j_\sigma(x_{k+1} = i \mid x_k = i) - \mathbb{P}^j_\sigma(x_{k+1} = i \mid x_k = 1-i) > 0$ and thus, by the induction hypothesis,

$$
\begin{aligned}
&\mathbb{P}^j_\sigma(x_{k+1} = i \mid x_n = i) \\
&\quad \geq (\mathbb{P}^j_\sigma(x_{k+1} = i \mid x_k = i) - \mathbb{P}^j_\sigma(x_{k+1} = i \mid x_k = 1-i))\mathbb{P}^j_\sigma(x_k = i \mid x_n = 1-i) \\
&\quad + \mathbb{P}^j_\sigma(x_k = 1-i \mid x_k = i) \\
&\quad = \mathbb{P}^j_\sigma(x_{k+1} = i \mid x_n = 1-i).
\end{aligned}
$$

□

Part (i) of the above lemma establishes a monotonicity property, namely, that it is more likely to make the correct decision than the wrong one when all agents decide according to a monotone strategy profile. Part (ii) is a weak positive correlation property establishing that it is more likely for the subsequent agent choose the observed action than to switch it. Finally, part (iii) establishes a stronger positive correlation property showing that it is more likely for any agent in the future to chose the same decision as agent n than to switch it.

6.2 Myopic behaviour as equilibrium for symmetric distributions

In this section we prove that when private signal distributions are symmetric, the myopic strategy profile that was described in Section 2.3 is a Perfect Bayesian Equilibrium for the altruistic learning game. First, we define *symmetric distributions.* Intuitively, the private signal distributions are symmetric when the probability of inferring a private belief above a specific threshold under state of the world one is the same as inferring a private belief below the symmetric threshold under state of the world zero.

Definition 10. *We say that the private signal structure is symmetric if for all $r \in (0,1)$,*

$$\mathbb{G}_0(1-r) + \mathbb{G}_1(r) = 1.$$

Another important definition is that of the *myopic* strategy profile. It is essentially each agent's best response assuming that she is just maximizing her own probability of a correct decision.

Definition 11. *We say that agents are* **myopic** *when their strategies can be described by*

$$\sigma_n(s_n, x_{n-1}) = \begin{cases} 1, & \text{if } f_n(s_n, x_{n-1}) \geq \frac{1}{2}, \\ 0, & \text{otherwise.} \end{cases}$$

It is clear that the myopic strategy profile is monotone and therefore the properties in Lemma 20 are applicable. Simultaneously, under the assumption of symmetric private signal distributions, we can deduce some properties of the myopic strategy profile which would not otherwise hold. These are summarized in the following lemma.

Lemma 21. *Assume that private signal distributions are symmetric and that agents are myopic. Then for all $k \geq n+1$, $i, j \in \{0,1\}$, and $k > n$, we have:*

(i) $\mathbb{P}(x_n = j \mid \theta = 1) = \mathbb{P}(x_n = 1 - j \mid \theta = 0)$,

(ii) $\mathbb{P}^j(x_{n+1} = j \mid x_n = i) = \mathbb{P}^{1-j}(x_{n+1} = 1 - j \mid x_n = 1 - i)$,

(iii) $\mathbb{P}^j(x_k = j \mid x_n = i) = \mathbb{P}^{1-j}(x_k = 1 - j \mid x_n = 1 - i)$.

Proof. (i) We use induction to prove the first part. For the first agent,

$$\mathbb{P}(x_1 = 1 \mid \theta = 1) = \mathbb{P}(p_1 \geq \frac{1}{2} \mid \theta = 1) = 1 - \mathbb{G}_1(1/2) = \mathbb{G}_0(1/2) = \mathbb{P}(x_1 = 0 \mid \theta = 0),$$

where the third equality follows from the symmetry assumption and

$$\mathbb{P}(x_1 = 0 \mid \theta = 1) = \mathbb{P}(p_1 \leq \frac{1}{2} \mid \theta = 1) = \mathbb{G}_1(1/2) = 1 - \mathbb{G}_0(1/2) = \mathbb{P}(x_1 = 1 \mid \theta = 0).$$

Assume that the result holds for n; we prove that it also does for $n+1$. Indeed, The social belief induced by the action $x_n = 1$ is calculated as follows

$$\begin{aligned} \pi_{n+1}(x_n = 1) &= \frac{\mathbb{P}(x_n = 1 \mid \theta = 1)}{\mathbb{P}(x_n = 1 \mid \theta = 0) + \mathbb{P}(x_n = 1 \mid \theta = 1)} \\ &= \frac{\mathbb{P}(x_n = 0 \mid \theta = 0)}{\mathbb{P}(x_n = 0 \mid \theta = 1) + \mathbb{P}(x_n = 0 \mid \theta = 0)}, \end{aligned}$$

where we used the induction hypothesis to obtain the second equality.

Observe that the second term is essentially the posterior probability of the state of the world being zero given $x_n = 0$,

$$\frac{\mathbb{P}(x_n = 0 \mid \theta = 0)}{\mathbb{P}(x_n = 0 \mid \theta = 1) + \mathbb{P}(x_n = 0 \mid \theta = 0)} = \mathbb{P}(\theta = 0 \mid x_n = 0) = 1 - \pi_{n+1}(x_n = 0)$$

Concluding, $\pi_{n+1}(x_n = 1) = 1 - \pi_{n+1}(x_n = 0)$. Then, by (6.1)

$$\mu_{n+1}(x_n = i) = (1 - \pi(i)) = \pi(1 - i) = 1 - \mu_{n+1}(x_n = 1 - i)$$

from which we conclude that

$$\begin{aligned}
\mathbb{P}^1(x_{n+1} = 1 \mid x_n = i) &= 1 - \mathbb{G}_1(\mu_{n+1}(x_n = i)) = \mathbb{G}_0(1 - \mu_{n+1}(x_n = i)) \\
&= \mathbb{G}_0(\mu_{n+1}(x_n = 1 - i)) = \mathbb{P}^0(x_{n+1} = 0 \mid x_n = 1 - i),
\end{aligned}$$

and similarly,

$$\begin{aligned}
\mathbb{P}^1(x_{n+1} = 0 \mid x_n = i) &= \mathbb{G}_1(\mu_{n+1}(x_n = i)) = 1 - \mathbb{G}_0(1 - \mu_{n+1}(x_n = i)) \\
&= 1 - \mathbb{G}_0(\mu_{n+1}(x_n = 1 - i)) = \mathbb{P}^0(x_{n+1} = 1 \mid x_n = 1 - i).
\end{aligned}$$

Clearly,

$$\begin{aligned}
&\mathbb{P}(x_{n+1} = j \mid \theta = 1) \\
&\quad = \mathbb{P}^1(x_{n+1} = j \mid x_n = 1)\mathbb{P}^1(x_n = 1) + \mathbb{P}^1(x_{n+1} = j \mid x_n = 1)\mathbb{P}^1(x_n = 1)
\end{aligned}$$

and by using the above, we get

$$\begin{aligned}
&\mathbb{P}^1(x_{n+1} = j \mid x_n = 1)\mathbb{P}^1(x_n = 1) + \mathbb{P}^1(x_{n+1} = j \mid x_n = 1)\mathbb{P}^1(x_n = 1) \\
&\quad = \mathbb{P}^0(x_{n+1} = 1 - j \mid x_n = 0)\mathbb{P}^0(x_n = 0) + \mathbb{P}^0(x_{n+1} = 1 - j \mid x_n = 0)\mathbb{P}^0(x_n = 0) \\
&\quad = \mathbb{P}(x_{n+1} = 1 - j \mid \theta = 0)
\end{aligned}$$

completing the proof.

(ii) This has already been proven in (i).

(iii) We use induction to prove this part. For $k = n + 1$, the result holds from part (ii). Assume that it holds for some k. Then,

$$\begin{aligned}
\mathbb{P}^j(x_{k+1} = j \mid x_n = i) &= \mathbb{P}^j(x_{k+1} = j \mid x_k = j)\mathbb{P}^j(x_k = j \mid x_n = i) \\
&\quad + \mathbb{P}^j(x_{k+1} = j \mid x_k = 1-j)\mathbb{P}^j(x_k = 1-j \mid x_n = i) \\
&= \mathbb{P}^{1-j}(x_{k+1} = 1-j \mid x_k = 1-j)\mathbb{P}^{1-j}(x_k = 1-j \mid x_n = 1-i) \\
&\quad + \mathbb{P}^{1-j}(x_{k+1} = 1-j \mid x_k = j)\mathbb{P}^{1-j}(x_k = j \mid x_n = 1-i) \\
&= \mathbb{P}^{1-j}(x_{k+1} = 1-j \mid x_n = 1-i)
\end{aligned}$$

□

Combining the properties of monotone equilibria established in Section 6.1 and the properties that we derived for symmetric private signal distributions, we prove the main result of this subsection, i.e., that the myopic strategy profile is a Perfect Bayesian Equilibrium for the forward looking game.

Lemma 22. *Let σ^m be the myopic strategy profile. Then, for all $\delta \in (0,1)$, σ^m is a Perfect Bayesian equilibrium of the altruistic learning game.*

Proof. Assume that the updated posterior belief of agent n given her private information is $f_n(I_n)$. Note that because none of the subsequent agents observe the action of $n-1$, the private information of agent n will appear in her expected payoff only through the updated posterior belief. The expected payoff from choosing action 1 is equal to

$$\begin{aligned}
\overline{U}(1; I_n) =& f_n(I_n)\left(1 + \sum_{k=n+1}^{\infty} \delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k = 1 \mid x_n = 1)\right) \\
&+ (1 - f_n(I_n)) \sum_{k=n+1}^{\infty} \delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k = 0 \mid x_n = 1)
\end{aligned}$$

while the expected payoff from choosing action 0 is equal to

$$\begin{aligned}
\overline{U}(0; I_n) =& f_n(I_n) \sum_{k=n+1}^{\infty} \delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k = 1 \mid x_n = 0) \\
&+ (1 - f_n(I_n))\left(1 + \sum_{k=n+1}^{\infty} \delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k = 0 \mid x_n = 0)\right).
\end{aligned}$$

Using part (iii) of Lemma 21 we can rewrite the above as

$$\begin{aligned}\overline{U}(1; I_n) =& f_n(I_n)\left(1 + \sum_{k=n}^{\infty} \delta^{k-n} \mathbb{P}^1_{\sigma^m}(x_k = 1 \mid x_n = 1)\right) \\ &+ (1 - f_n(I_n)) \sum_{k=n}^{\infty} \delta^{k-n} \mathbb{P}^1_{\sigma^m}(x_k = 1 \mid x_n = 0),\end{aligned}$$

and

$$\begin{aligned}\overline{U}(0; I_n) =& f_n(I_n) \sum_{k=n}^{\infty} \delta^{k-n} \mathbb{P}^1_{\sigma^m}(x_k = 1 \mid x_n = 0) \\ &+ (1 - f_n(I_n)) \left(1 + \sum_{k=n}^{\infty} \delta^{k-n} \mathbb{P}^1_{\sigma^m}(x_k = 1 \mid x_n = 1)\right),\end{aligned}$$

and therefore

$$\begin{aligned}&\overline{U}(1; I_n) - \overline{U}(0; I_n) = (2f_n(I_n) - 1) \\ &\cdot \left(1 + \sum_{k=n}^{\infty} \delta^{k-n} \left(\mathbb{P}_{\sigma^m}(x_k = 1 \mid x_n = 1, \theta = 1) - \mathbb{P}_{\sigma^m}(x_k = 1 \mid x_n = 0, \theta = 1)\right)\right).\end{aligned}$$

Since the myopic strategy profile is monotone, we get from Lemma 20 that for all $k \geq n + 1$,

$$\mathbb{P}_{\sigma^m}(x_k = 1 \mid x_n = 1, \theta = 1) - \mathbb{P}_{\sigma^m}(x_k = 1 \mid x_n = 0, \theta = 1) \geq 0$$

and thus, $\overline{U}(1; I_n) \geq \overline{U}(0; I_n)$ if and only if $f_n(I_n) \geq 1/2$ concluding the proof.

□

It is well known (Athans and Papastavrou [14] and Acemoglu et al. [1]) that when agents are myopic and the private signal structure is characterized by Unbounded Likelihood Ratios, then learning in probability occurs. Therefore, we have constructed an equilibrium for the altruistic learning game that achieves learning in probability. If $\delta < 1/2$ and for any private signal distribution it is straightforward to show that the myopic strategy profile is the unique Perfect Bayesian Equilibrium of the forward looking game and therefore all equilibria achieve learning in probability when $\delta < 1/2$. One conjecture could be that this is the case for all equilibria of this game and for all values of the discount factors. We show in the next section that this is not correct, by constructing an equilibrium which does not achieve learning in probability.

6.3 Construction of a non-learning equilibrium

We established that for the case of symmetric private signal distributions, the myopic strategy profile, denoted by σ^m, is a Perfect Bayesian Equilibrium that achieves learning in probability. In this section we construct another strategy profile σ^c which we show is a Perfect Bayesian Equilibrium that does not achieve learning in probability.

Fix $M \in \mathbb{N}$ and define $\mathcal{M} = \{kM\}_{k\in\mathbb{N}}$. Consider the following decision rule:

(i) For $n \notin \mathcal{M}$

$$\sigma_n(s_n, x_{n-1}) = \begin{cases} 1, & \text{if } f_n(s_n, x_{n-1}) \geq \frac{1}{2} \\ 0, & \text{otherwise.} \end{cases}$$

(ii) For $n \in \mathcal{M}$

$$\sigma_n(s_n, x_{n-1}) = \begin{cases} 0, & \text{if } f_n(s_n, x_{n-1}) \geq \frac{1}{2} \\ 1, & \text{otherwise.} \end{cases}$$

We prove that for large values of M this strategy profile is an equilibrium for the altruistic learning game (Proposition 3). The main idea is captured in Figure 6-1.

Proposition 3. *Consider the strategy profile σ^c and let $\delta \in (1/2, 1)$. Then, there exist values of M for which σ^c is a Perfect Bayesian Equilibrium of the altruistic learning game.*

Proof. We prove the result by considering the optimization problem of the two types of agents: the myopic and the counter-myopic.

(i) Consider a myopic agent $n \notin \mathcal{M}$. Note that for all $k \notin \mathcal{M}$

$$\mathbb{P}_{\sigma^c}(x_k = j \mid x_n = i, \theta = l) = \mathbb{P}_{\sigma^m}(x_k = j \mid x_n = i, \theta = l) \tag{6.6}$$

while if $k \in \mathcal{M}$

$$\mathbb{P}_{\sigma^c}(x_k = j \mid x_n = i, \theta = l) = \mathbb{P}_{\sigma^m}(x_k = 1 - j \mid x_n = i, \theta = l) \tag{6.7}$$

for all $i, j, l \in \{0, 1\}$.

Using Eqs. (6.6) and (6.7), we can write the expected payoffs from each action,

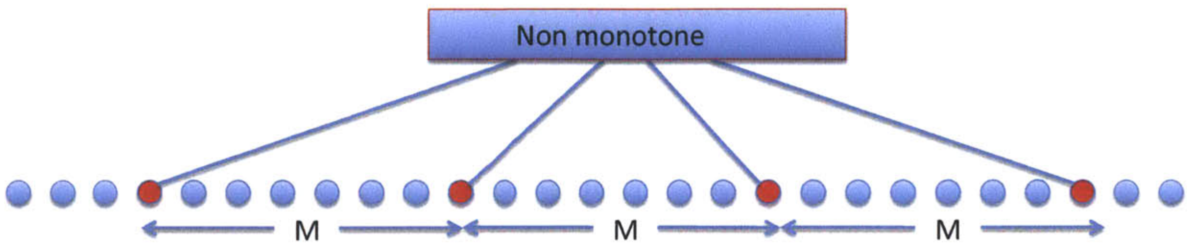


Figure 6-1: In our construction, most agents behave myopically, except for those that are multiples of M. Those behave counter-myopically; they are choosing zero when they believe that the state of the world is one. If M is large enough, we prove that this strategy profile is a Perfect Bayesian Equilibrium of the forward looking game for $\delta \in (1/2, 1)$. Indeed, each of the myopic agents, by acting myopically, drives most future agents (the myopic ones) towards her favored decision and some (the counter-myopic ones) towards the opposite decision, since according to the monotone strategy profiles properties agents are more likely to copy than to switch the observed actions. As long as counter-myopic agents are placed in the tandem sparsely enough, myopic decisions remain optimal. Similarly, for counter-myopic agents it is optimal to choose non-monotonically since the subsequent agents expect them to do so. Clearly, such a construction is possible only when the continuation payoff can be greater than the immediate payoff, which happens only if $\delta > 1/2$.

given the information set I_n as

$$\overline{U}_n(1; I_n)$$
$$= f_n(I_n)\left(1+\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=1\mid x_n=1)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=0\mid x_n=1)\right)$$
$$+(1-f_n(I_n))\left(\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=0\mid x_n=1)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=1\mid x_n=1)\right)$$

and

$$\overline{U}_n(0; I_n)$$
$$= f_n(I_n)\left(\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=1\mid x_n=0)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=0\mid x_n=0)\right)$$
$$+(1-f_n(I_n))\left(1+\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=0\mid x_n=0)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=1\mid x_n=0)\right)$$

Observe that, using part (iii) of Lemma 21 we obtain

$$1+\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=0\mid x_n=0)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=1\mid x_n=0)$$
$$=1+\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=1\mid x_n=1)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=0\mid x_n=1),$$

and

$$\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=0\mid x_n=1)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=1\mid x_n=1)$$
$$=\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=1\mid x_n=0)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=0\mid x_n=0).$$

It should now be obvious that if I_n is such that $f_n(I_n) = 1/2$, $\overline{U}_n(1; I_n) = \overline{U}_n(0; I_n)$.

From the discussion in Section 5.2 we know that expected payoffs from each action are affine with respect to the updated posterior belief. Therefore in order to prove optimality of σ_n^c we need to show that if I_n is such that $f_n(I_n) = 1$ it is optimal for agent n to choose 1. To this end we first show that I_n is such that $f_n(I_n) = 1$ then, for all $i, j \in \{0, 1\}$ and $k > n$,

$$\mathbb{P}(x_k = j \mid x_n = j, \theta = i) = 1.$$

Indeed, assume that $f_n(s_n, x_{n-1}) = 1$. Then, it has to be the case that $x_{n-1} = 1$. Assume the contrary. Since

$$\pi_n(x_{n-1} = y) = \frac{1}{1 + \frac{\mathbb{P}_{\sigma^m}(x_{n-1}=y|\theta=0)}{\mathbb{P}_{\sigma^m}(x_{n-1}=y|\theta=0)}},$$

it follows that $\mathbb{P}_{\sigma^m}(x_n = 0 \mid \theta = 0) = 0$. But, from the properties of monotone equilibria, $\mathbb{P}_{\sigma^m}(x_{n-1} = 0 \mid \theta = 1) \leq \mathbb{P}_{\sigma^m}(x_{n-1} = 0 \mid \theta = 1) = 0$ and therefore $\mathbb{P}_{\sigma^m}(x_{n-1} = 0) = 0$ which is a contradiction.

At this point we make use of the well known monotonicity property of the myopic decision rule (Acemoglu et al. [1]), namely,

$$\mathbb{P}_{\sigma^m}(x_{n+1} = j \mid \theta = j) \geq \mathbb{P}_{\sigma^m}(x_n = j \mid \theta = j). \tag{6.8}$$

Using the above we can show that if $f_n(s_n, x_{n-1} = 1) = 1$ then $f_{n+1}(s_{n+1}, x_n = 1) = 1$ while $f_{n+1}(s_{n+1}, x_n = 0) = 0$ for all $s_{n+1} \in S$. Indeed, if $f_n(s_n, x_{n-1} = 1) = 1$ then $\mathbb{P}_{\sigma^m}(x_{n-1} = 1 \mid \theta = 0) = 0$ and using (6.8) we obtain $\mathbb{P}_{\sigma^m}(x_n = 1 \mid \theta = 0) = 0$. But then $f_{n+1}(s_{n+1}, x_n = 1) = 1$. Similarly, $\mathbb{P}_{\sigma^m}(x_n = 0 \mid \theta = 1) \leq \mathbb{P}_{\sigma^m}(x_{n-1} = 0 \mid \theta = 1) = \mathbb{P}_{\sigma^m}(x_{n-1} = 1 \mid \theta = 0) = 0$, where we used part (i) of Lemma 21 to obtain the last equality. The latter yields $f_{n+1}(x_{n+1} = 0) = 0.$. Therefore, by the definition of the myopic strategy profile, for all $i, j \in \{0, 1\}$

$$\mathbb{P}(x_k = j \mid x_n = j, \theta = i) = 1.$$

Using a trivial induction argument our claim follows.

Given the above, if the information set I_n of agent n is such that $f_n(I_n) = 1$, we can compute the expected payoffs from each action, as follows

$$\overline{U}_n(1; I_n) = 1 + \sum_{\substack{k>n \\ k \notin \mathcal{M}}} \delta^{k-n} 1 + \sum_{\substack{k>n \\ k \in \mathcal{M}}} \delta^{k-n} 0,$$

and

$$\overline{U}_n(0; I_n) = \sum_{\substack{k>n \\ k \notin \mathcal{M}}} \delta^{k-n} 0 + \sum_{\substack{k>n \\ k \in \mathcal{M}}} \delta^{k-n} 1.$$

Therefore,

$$\overline{U}_n(1; I_n) - \overline{U}_n(0; I_n) = 1 + \sum_{\substack{k>n \\ k \notin \mathcal{M}}} \delta^{k-n} - \sum_{\substack{k>n \\ k \in \mathcal{M}}} \delta^{k-n} = \frac{1}{1-\delta} - 2 \sum_{\substack{k>n \\ k \in \mathcal{M}}} \delta^{k-n}.$$

The worst case is when $n = l \cdot M - 1$ for some $l \in \mathbb{N}$ and in that case.

$$\overline{U}_n(1; I_n) - \overline{U}_n(0; I_n) = \frac{1}{1-\delta} - 2\frac{\delta}{1-\delta^M},$$

which yields

$$\lim_{M \to \infty} (\overline{U}_n(1; I_n) - \overline{U}_n(0; I_n)) = \frac{1}{1-\delta} - 2\delta > 0$$

and therefore there exists some $\hat{M}$ such that for all $M > \hat{M}$, $\overline{U}_n(1; I_n) - \overline{U}_n(0; I_n) > 0$ concluding the proof for the myopic agents.

(ii) Consider a counter-myopic agent $n \in \mathcal{M}$. We follow similar reasoning as for the previous case. Note that for all $k \notin \mathcal{M}$

$$\mathbb{P}_{\sigma^c}(x_k = j \mid x_n = i, \theta = l) = \mathbb{P}_{\sigma^m}(x_k = j \mid x_n = 0, \theta = l)$$

while if $k \in \mathcal{M}$

$$\mathbb{P}_{\sigma^c}(x_k = j \mid x_n = i, \theta = l) = \mathbb{P}_{\sigma^m}(x_k = 1 - j \mid x_n = 0, \theta = l)$$

for all $i, j, l \in \{0, 1\}$.

Given the above we get

$$\begin{aligned} \overline{U}_n(1; I_n) &= f_n(I_n) \left(1 + \sum_{\substack{k>n \\ k \notin \mathcal{M}}} \delta^{k-n} \mathbb{P}^1_{\sigma^m}(x_k = 1 \mid x_n = 0) + \sum_{\substack{k>n \\ k \in \mathcal{M}}} \delta^{k-n} \mathbb{P}^1_{\sigma^m}(x_k = 0 \mid x_n = 0) \right) \\ &+ (1 - f_n(I_n)) \left(\sum_{\substack{k>n \\ k \notin \mathcal{M}}} \delta^{k-n} \mathbb{P}^0_{\sigma^m}(x_k = 0 \mid x_n = 0) + \sum_{\substack{k>n \\ k \in \mathcal{M}}} \delta^{k-n} \mathbb{P}^0_{\sigma^m}(x_k = 1 \mid x_n = 0) \right) \end{aligned}$$

and

$$\overline{U}_n(0;I_n) = f_n(I_n)\left(\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=1\mid x_n=1)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=0\mid x_n=1)\right)$$
$$+ (1-f_n(I_n))\left(1+\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=0\mid x_n=1)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=1\mid x_n=1\right.$$

Observe that using part (iii) of Lemma 21 we get

$$1+\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=0\mid x_n=1)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=1\mid x_n=1)$$
$$=1+\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=1\mid x_n=0)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=0\mid x_n=0)$$

and similarly,

$$\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=0\mid x_n=0)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^0_{\sigma^m}(x_k=1\mid x_n=0)$$
$$=\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=1\mid x_n=1)+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}\mathbb{P}^1_{\sigma^m}(x_k=0\mid x_n=1).$$

It should now be obvious that if I_n is such that $f_n(I_n) = 1/2$, $\overline{U}_n(1;I_n) = \overline{U}_n(0;I_n)$.

From the discussion in Section 5.2 we know that expected payoffs from each action are affine with respect to the updated posterior belief. Therefore in order to prove optimality of σ^c_n we need to show that if I_n is such that $f_n(I_n) = 1$ it is optimal for agent n to choose 0.

Arguing similarly to the previous case, the expected payoffs from each action for the case of an information set I_n for which $f_n(I_n) = 1$ is given by

$$\overline{U}_n(1;I_n) = 1+\sum_{\substack{k>n\\k\notin\mathcal{M}}}\delta^{k-n}0+\sum_{\substack{k>n\\k\in\mathcal{M}}}\delta^{k-n}1,$$

and

$$\overline{U}_n(0; I_n) = \sum_{\substack{k>n \\ k\notin \mathcal{M}}} \delta^{k-n}1 + \sum_{\substack{k>n \\ k\in \mathcal{M}}} \delta^{k-n}0.$$

Therefore,

$$\overline{U}_n(1; I_n) - \overline{U}_n(0; I_n) = 1 + \sum_{\substack{k>n \\ k\in \mathcal{M}}} \delta^{k-n} - \sum_{\substack{k>n \\ k\notin \mathcal{M}}} \delta^{k-n} = 1 - \frac{\delta}{1-\delta} + 2\delta^M \frac{1}{1-\delta^M}.$$

Thus,

$$\lim_{M\to\infty} (\overline{U}_n(1; I_n) - \overline{U}_n(0; I_n)) = 1 - \frac{\delta}{1-\delta} < 0$$

for $\delta > 1/2$ and therefore there exists some $\tilde{M}$ such that for all $M > \tilde{M}$, $\overline{U}_n(1; I_n) - \overline{U}_n(0; I_n) > 0$ concluding the proof for the counter-myopic agents.

Therefore if $M > \max\{\hat{M}, \tilde{M}\}$, the strategy profile σ^c is a Perfect Bayesian Equilibrium.

□

After establishing that σ^c is a Perfect Bayesian Equilibrium, we now prove that it does not achieve learning in probability. This should be intuitively clear by the fact that infinitely often there are those counter-myopic agents who are flipping the correct decision, preventing learning in probability.

Proposition 4. *Assume that each agent's strategy is given by σ^c. Then,*

$$\lim_{n\to\infty} \mathbb{P}(x_n = \theta) < 1$$

Proof. Assume that σ^c achieves learning in probability. Then, by Lemma 11 and the learning in probability assumption, there exists an $\hat{N} \in \mathbb{N}$ such that for all $n > \hat{N}$ the social belief $\pi(x_{n-1} = 1) > 1 - \gamma$ and $\mathbb{P}(x_{n-1} = 1 \mid \theta = 1) > 0$. Moreover, from our assumptions on the private signal distributions, there exists some γ for which $\mathbb{P}^j(p_n > \gamma) > 0$ for all $j \in \{0, 1\}$. In that case, consider a counter-myopic agent $n > \hat{N}$. If she observes $x_{n-1} = 1$ and receives a private signal that induces private belief greater than γ, then her posterior belief will be

$$f_n(s_n, x_{n-1} = 1) = \frac{\pi(x_{n-1} = 1)\gamma}{\pi(x_{n-1} = 1)\gamma + (1 - \pi(x_{n-1} = 1))(1-\gamma)} > \frac{(1-\gamma)\gamma}{(1-\gamma)\gamma + (1-\gamma)\gamma} = \frac{1}{2}$$

and therefore for all counter-myopic agents after $\hat{N}$,

$$\mathbb{P}_{\sigma^c}(x_n = 0 \mid \theta = 1) \geq \mathbb{P}_1(s_n > \gamma)\mathbb{P}^1_{\sigma^c}(x_{n-1} = 1) > 0,$$

contradicting the learning assumption and concluding the proof. $\square$

This construction shows that not all equilibria of the forward looking game achieve learning in probability. A characteristic property of the non-learning equilibrium is that there are infinitely many agents who choose non-monotonically, deciding zero for large updated posterior beliefs. Such a strategy can be a best response as we saw, exactly because of the incorporation of future agents' probability of a correct decision in each individual's payoff. In other words, individuals may act non-monotonically because future agents expect them to. Our conjecture is that monotone Perfect Bayesian Equilibria achieve learning in probability but we were not able to prove it.

Conjecture 1. *Let σ be a monotone Perfect Bayesian equilibrium of the forward looking game. Then σ achieves learning in probability or equivalently*

$$\lim_{n\to\infty} \mathbb{P}_\sigma(x_n = \theta) = 1.$$

In the next section we prove the existence of such equilibria.

6.4 Existence of monotone Perfect Bayesian Equilibria

We start by presenting a continuity result, namely that if for two different strategy profiles the thresholds according to which agents decide are arbitrarily close, then so are the probabilities of making the corresponding decisions. Throughout this section we are assuming atomless private signal distributions and therefore continuity of $\mathbb{G}^j(r)$ for all $j \in \{0, 1\}$.

Proposition 5. *Assume that $\mathbb{G}_1(\cdot)$ and $\mathbb{G}_0(\cdot)$ are continuous. Let σ and σ' be two different strategy profiles and let $\{\underline{\mu}_n\}_{n\in\mathbb{N}}$, $\{\bar{\mu}_n\}_{n\in\mathbb{N}}$, $\{\underline{\mu}'_n\}_{n\in\mathbb{N}}$, $\{\bar{\mu}'_n\}_{n\in\mathbb{N}}$ be the corresponding thresholds. If $\sup_{n\in\mathbb{N}} \max\{|\bar{\mu}_n - \bar{\mu}'_n|, |\underline{\mu}_n - \underline{\mu}'_n|\} < \epsilon$, then there exist $\delta_{nn}(\epsilon)$, $\delta_n(\epsilon)$, and $\delta_{nl}(\epsilon)$, with $\delta_{nn}(\epsilon) \to 0$, $\delta_n(\epsilon) \to 0$, and $\delta_{nl}(\epsilon) \to 0$ as $\epsilon \to 0$, such that for all $i, j, k \in \{0, 1\}$, $n, l \in \mathbb{N}$:*

(i) $|\mathbb{P}_\sigma(x_n = k \mid x_{n-1} = i, \theta = j) - \mathbb{P}_{\sigma'}(x_n = k \mid x_{n-1} = i, \theta = j)| < \delta_{nn}(\epsilon)$,

(ii) $|\mathbb{P}_\sigma(x_n = k \mid \theta = j) - \mathbb{P}_{\sigma'}(x_n = k \mid \theta = j)| < \delta_n(\epsilon)$,

(iii) $|\mathbb{P}_\sigma(x_l = k \mid x_{n-1} = i, \theta = j) - \mathbb{P}_{\sigma'}(x_n = k \mid x_{n-1} = i, \theta = j)| < \delta_{nl}(\epsilon)$.

Proof.

All results are direct consequences of the continuity of $\mathbb{G}_j(r)$ with respect to r for all $j \in \{0, 1\}$.

□

6.4.1 Existence of a monotone Perfect Bayesian Equilibrium for the finite horizon altruistic learning game

Let M denote the number of agents in a finite horizon forward looking game. The utility of each gent is now be given by

$$u_n(\mathbf{x}_n, \theta) = (1 - \delta) \sum_{k=n}^{M} \delta^k \mathbf{1}_{x_n = \theta}, \quad \text{for } n \leq M.$$

Observe that the properties of monotone equilibria established in Lemma 20 still hold, despite the modification of the game. We will exploit those properties to establish the following. If all other agents use monotone strategies then it is optimal for any other agent to also use monotone strategies, as the following lemma shows.

Lemma 23. *Assume that agents* $\{1, \ldots, n-1, n+1, \ldots\}$ *use monotone strategies described by* $\{\underline{\mu}_k\}_{k \in \mathbb{N} \setminus \{n\}}, \{\bar{\mu}_k\}_{k \in \mathbb{N} \setminus \{n\}}$. *Let* $\alpha(f_n \mid \underline{\mu}, \bar{\mu}) : [0, 1] \to \{0, 1\}$ *be a best response of agent* n *given her posterior belief* f_n. *Then* $\alpha(f_n \mid \underline{\mu}, \bar{\mu}) : [0, 1] \to \{0, 1\}$ *is non decreasing in* f_n.

Proof. Let I_n be the information set of agent n and let $f_n \triangleq f_n(I_n)$ denote the posterior belief inferred from I_n. The information set appears in the calculation of the expected payoffs of agent n only through the posterior belief $f(I_n)$. Therefore, without loss of generality we write $\bar{U}_n(j; f_n)$ to denote $\bar{U}(j; I_n)$ for all $j \in \{0, 1\}$. The expected payoff for agent n by choosing action one can be evaluated as follows:

$$\begin{aligned}\overline{U}_n(1;f_n) &= f_n\left(1+\sum_{k=n}^{M}\delta^{k-n}\mathbb{P}_{\mathbf{l}}^1(x_k=1\mid x_n=1)\right)+(1-f_n)\sum_{k=n}^{M}\delta^{(k-n)}\mathbb{P}_{\mathbf{l}}^1(x_k=1\mid x_n=0)\\ &= f_n\left(1+\sum_{k=n}^{M}\delta^{k-n}(\mathbb{P}_{\mathbf{l}}^1(x_k=1\mid x_n=1)-\mathbb{P}_{\mathbf{l}}^1(x_k=1\mid x_n=0))\right)\\ &\quad+\sum_{k=n}^{M}\delta^{(k-n)}\mathbb{P}_{\mathbf{l}}^1(x_k=1\mid x_n=0).\end{aligned}$$

From Lemma 20 we get that for all $k>n$,

$$\mathbb{P}_{\mathbf{l}}^1(x_k=1\mid x_n=1)-\mathbb{P}_{\mathbf{l}}^1(x_k=1\mid x_n=0)\geq 0,$$

establishing that the expected payoff from choosing action 1 is non-decreasing in the updated posterior belief. Similarly, we get that the expected payoff from choosing 0 is non-increasing in the updated posterior belief, establishing that the best response is non-decreasing in f_n. □

The next Corollary formalizes our intuition that the best response to monotone strategies is a monotone strategy.

Corollary 3. *Assume that agents $\{1,\ldots,n-1,n+1,\ldots\}$ use monotone strategies. Then there exist two thresholds $\underline{\mu}_n, \bar{\mu}_n$ such that a best response of agent n is given by*

$$\sigma_n(s_n,x_{n-1})=\begin{cases}1, & \text{if } p_n(s_n)\geq\bar{\mu}_n,\\ x_{n-1}, & \text{if } \bar{\mu}_n>p_n(s_n)\geq\underline{\mu}_n,\\ 0, & \text{if } p_n(s_n)<\underline{\mu}_n.\end{cases}\tag{6.9}$$

Proof. Let l_n be the point of indifference between the expected payoffs from choosing one and zero respectively. Then from the previous lemma, the best response of agent n is $\mathbf{1}_{f_n>l_n}$ which translates to (6.9), using (6.1) and (6.2).

□

The last step before we proceed to the main result is another continuity result, which follows directly from Proposition 5. Since transition probabilities are continuous with respect to the thresholds so are agents' expected payoffs from each action.

Corollary 4. *Let σ and σ' be two different monotone strategy profiles and let $\{\underline{\mu}_n\}_{n\in\{1,\ldots,M\}\setminus\{n\}}$, $\{\bar{\mu}_n\}_{n\in\{1,\ldots,M\}\setminus\{n\}}$, $\{\underline{\mu'}_n\}_{n\in\{1,\ldots,M\}\setminus\{n\}}$, $\{\bar{\mu'}_n\}_{n\in\{1,\ldots,M\}\setminus\{n\}}$ be the corresponding thresholds. Then if $\sup_{n\in\{1,\ldots,M\}\setminus\{n\}}\max\{|\bar{\mu}_n-\bar{\mu'}_n|,|\underline{\mu}_n-\underline{\mu'}_n|\}<\epsilon$, and $\underline{\mu}_n, \bar{\mu}_n, \underline{\mu'}_n, \bar{\mu'}_n$ are the thresholds associated with the best responses of agent n we have*

$$\lim_{\epsilon\to 0}|\underline{\mu}_n-\underline{\mu'}_n|=0$$

and

$$\lim_{\epsilon\to 0}|\bar{\mu}_n-\bar{\mu'}_n|=0.$$

Proof. From Proposition 5 it follows that, for a given updated posterior belief the expected payoffs from choosing one and zero respectively satisfy ,

$$\lim_{\epsilon\to 0}|\bar{U}_n^{\sigma}(1;f_n)-\bar{U}_n^{\sigma'}(1;f_n)|=0$$

and

$$\lim_{\epsilon\to 0}|\bar{U}_n^{\sigma}(0;f_n)-\bar{U}_n^{\sigma'}(0;f_n)|=0.$$

Therefore, if l_n, l'_n denote the indifference points, we get that

$$\lim_{\epsilon\to 0}|l_n-l'_n|=0$$

and thus since the corresponding private belief space thresholds $\underline{\mu}_n, \bar{\mu}_n$ are continuous functions of l_n and the posterior beliefs we get the desired result. □

Using this property we are able to prove the existence of a Perfect Bayesian Equilibrium for the finite horizon forward looking game. In order to do that we define the correspondence

$$R_n(\underline{\mu},\bar{\mu})=\{(r,l)\in[0,1]:\alpha(p_n,x_n-1\mid\underline{\mu},\bar{\mu})=\mathbf{1}_{p_n>r}+(1-\mathbf{1}_{p_n<l})+x_{n-1}\mathbf{1}_{l\le p_n\le r}\}. \tag{6.10}$$

of the best responses of agent n to monotone strategies of the rest of the agents characterized by thresholds $\{\underline{\mu}_n\}_{n\in\{1,\ldots,M\}\setminus\{n\}}$, $\{\bar{\mu}_n\}_{n\in\{1,\ldots,M\}\setminus\{n\}}$. We prove that the correspondence $R=(R_1,\ldots,R_M)$ has a fixed point using Kakutani's fixed point theorem.

Theorem 5. *Let $M>0$. There exists a monotone Perfect Bayesian equilibrium for the M-horizon forward looking game.*

Proof. We prove that the correspondence defined in (6.10) has a fixed point using Kakutani's fixed point theorem. We need to check the following:

(i) $[0,1]^M$ is a compact, convex subset of $\mathbb{R}^M$.

(ii) Since each agent has two possible actions her optimization problem has a solution. This, together with Corollary 3 establish the non emptiness of $R(\underline{\mu}, \bar{\mu})$.

(iii) $R(\underline{\mu}, \bar{\mu})$ is convex, since there is a unique best response to other agents' strategies (comparison of two linear functions).

(iv) $R(\underline{\mu}, \bar{\mu})$ has a closed graph. This follows from Corollary 4.

□

6.4.2 Existence of a monotone Perfect Bayesian Equilibrium for the forward looking game

In this last subsection we prove the existence of a monotone Perfect Bayesian Equilibrium.

Lemma 24. *There exists a monotone Perfect Bayesian Equilibrium for the infinite horizon forward looking game.*

Proof. Consider the sequence of M-horizon games for $M \in \mathbb{N}$. From the previous section, for each M there exists a sequence of thresholds $\{\underline{\mu}_n^M\}_{n=1}^M$, $\{\bar{\mu}_n^M\}_{n=1}^M$ such that the strategy profile

$$\sigma^M = \begin{cases} 1, & \text{if } p_n > \bar{\mu}_n^M, \\ x_{n-1}, & \text{if } \underline{\mu}_n^M < p_n \leq \bar{\mu}_n^M, \\ 0, & \text{otherwise,} \end{cases}$$

is a Perfect Bayesian equilibrium. Since $[0,1]$ is a compact subset of $\mathbb{R}$ there exists a subsequence $\{M_k\}_{k\in\mathbb{N}}$ such that the corresponding sequences $\{\underline{\mu}_n^{M_k}\}$, $\{\bar{\mu}_n^{M_k}\}$ converge, and let $\{\underline{\mu}_n\}_{n=1}^\infty$, $\{\bar{\mu}_n\}_{n=1}^\infty$ be the corresponding limits.

Without loss of generality we prove that if $p_n > \bar{\mu}_n$ and $x_{n-1} = 0$, deciding one is optimal. The other cases follow similarly. Indeed, let $\epsilon > 0$. Consider an agent $n \in \mathbb{N}$. Choose a $K \in \mathbb{N}$ such that for all $k > K$, $\delta^{M_k} < \epsilon$. From Proposition 5 there exists an $\epsilon' > 0$ such that

$$\max_{k=n,\ldots,M_K} \delta_{nk}(\epsilon') < \frac{\epsilon}{1-\delta}.$$

Let K' be such that for all $k > K'$, $sup_{m\in\mathbb{N}} \max\{|\underline{\mu}_m^{M_k} - \underline{\mu}_m|, |\bar{\mu}_m^{M_k} - \bar{\mu}_m|\} < \epsilon'$. There exists some $K'' \in \mathbb{N}$ such that for all $k > K''$, $p_n > \bar{\mu}_n^{M_k}$.

In that case for all $k > \hat{K} = \max\{K, K', K''\}$, using Proposition 5, we get that

$$\overline{U}_\sigma(1; f_n(p_n, x_{n-1} = 0)) \geq \overline{U}_\sigma^{M_k}(1; f_n(p_n, x_{n-1} = 0)) - \frac{\epsilon}{1-\delta}(1-\delta) + \delta^{M_{\hat{K}}}$$

and

$$\overline{U}_\sigma(1; f_n(p_n, x_{n-1} = 0)) \leq \overline{U}_\sigma^{M_k}(0; f_n(p_n, x_{n-1} = 0)) + \frac{\epsilon}{1-\delta}(1-\delta) + \delta^{M_{\hat{K}}}.$$

Therefore,

$$\overline{U}_\sigma(1; f_n(p_n, x_{n-1} = 0)) \geq \overline{U}_\sigma(1; f_n(p_n, x_{n-1} = 0)) - 4\epsilon$$

for all $\epsilon > 0$ and, letting $\epsilon \to 0$ we get the desired result concluding the proof.

□

6.5 Discussion and Conclusions

This chapter studies some equilibria that arise from the altruistic learning game. Specifically, for symmetric distributions, we construct an equilibrium that achieves learning in probability, when the private signal structure is characterized by Unbounded Likelihood Ratios. On the other hand, we can also construct a Perfect Bayesian Equilibrium that does not and therefore we cannot expect all equilibria of the game to achieve learning in probability. Therefore, different equilibria of the game may have different learning properties. We conjecture that all monotone equilibria of the forward looking game achieve learning in probability. Finally, the existence of such monotone equilibria is established.

Chapter 7

Conclusions

7.1 Summary

Table 7.1 summarizes the results that are currently known for the observational learning problem under consideration, including the contributions of this thesis.

In this thesis, we studied different aspects of the problem of sequential decision making under uncertainty. A large literature, pioneered by Cover [7] and Koplowitz [11] from the statistics/engineering literature, and Bikhchandani, Hirshleifer and Welch [4], Banerjee [3] and Smith and Sorensen [16] from the economics literature, has studied such problems of decentralized information aggregation. Our focus has been dual. First, to find necessary and sufficient conditions under which information aggregation is possible and second to study whether equilibria may lead to such information aggregation and thus to learning.

In many situations in engineering systems there are memory or communication constraints. In social settings individuals obtain their information not by observing all past actions but the agents who made a decision most recently. These motivate our observation model according to which each agent receives a signal about the underlying state of the world and observes the past actions of her K immediate predecessors. The signal structure determines the conditional distributions of the signals received by each individual as a function of the underlying state. Each individual then chooses one of two possible actions. Learning corresponds to agents' decisions converging (almost surely or in probability) to the right action as the number of agents becomes large.

We first studied the possibility of learning under arbitrary decision rules. Two concepts turn out to be crucial in determining whether there will be learning. The first whether the informativity of private signals is bounded or unbounded, distinguishing between private signal structures with Bounded and Unbounded Likelihood Ratios. The second is the size K of the observed neighbourhood. We prove that almost sure learning is possible if and only if the private signal structure is characterized by Unbounded Likelihood Ratios, irrespective of K. Our second main result is sensitive to the size of the neighborhood. Specifically, under the Bounded Likelihood Ratios assumption, if $K = 1$, learning in probability is impossible under any decision rule. On the other hand, if $K > 2$, we construct a decision rule that achieves learning in probability.

We then focus on equilibrium learning behavior. We assume forward looking agents, who make a decision that maximizes the discounted sum of probabilities of correct decisions over all subsequent agents, including theirs. We prove that, in contrast to the existence of a decision rule that achieves learning in probability, there is no equilibrium of the corresponding game that achieves information aggregation. The latter is surprising since it suggests that even with payoff interdependence and forward-looking incentives, observational learning can be susceptible to limited information aggregation.

Concluding, we believe that contrasting between the engineering/designed learning and equilibrium learning illustrates the importance of strategic interactions. In simple words, if agents followed decision rules that were predefined by a social planner they could eventually learn the truth, even with boundedly informative signals. On the other hand, strategic interaction, even with altruistic incentives and collective preferences, leads to limited information aggregation.

7.2 Research directions—Open problems

This thesis has explored many aspects of the learning problem under a specific information structure. Nevertheless, there are questions that remain unanswered. A quick look at Table 7.1 reveals an obvious one; whether almost sure learning is possible under Unbounded Likelihood Ratios for the myopic strategy profile. Similarly, an interesting problem concerns almost sure learning along the equilibria of the altruistic learning game that do achieve learning in probability.

For the benchmark case of engineered systems, we proved that there exist decision profiles that achieve learning in probability when $K > 2$. In order to achieve this, for some observed actions agents are requires to take non-monotone actions. For example, during S-blocks agents decide 0 when observing 01 irrespective of their private signal and the order of the block. We believe that if we focus on monotone strategies, in the sense that each agent's decision rule is monotone in the number of observed ones, then learning in probability does not occur.

Along the same lines, we believe that for payoff structures that satisfy assumptions (A1) and (A2) of [13], i.e., that preferences are collective and agents prefer higher actions when the state is 1, there exists no Perfect Bayesian Equilibrium that achieves learning in probability ([13] constructs one such equilibrium but of course it is not unique).

	B.L.R.			UB.L.R.		
	engineered	myopic	forward looking	engineered	myopic	forward looking
›st sure learning	✗ [DOT]	✗[7][1]	✗ [DOT]	✓[7]	?	?/✗ [DOT]
ng in probability	$K = 1$	K=1	✗ [DOT]	✓[7]	$K = 1$	✓/ ✗ [DOT]
	✗[DOT]	✗ [7]			✓[14]	
	$K > 1$	$K > 1$			$K > 1$	
	✓ [DOT]	✗ [1]			✓[1]	

7.1: Summary of results. ✓ denotes positive results while ✗ denotes negative results. The references next to ‹
:ite the paper who obtained it. [DOT] stands for this thesis. Theorem 1 of this thesis establishes that if the private si
ıre is characterized by Bounded Likelihood Ratios then there does not exist a decision rule that achieves almost
.g. Therefore there cannot exist an equilibrium strategy profile that achieves almost sure learning for the forward loo
Theorems 2 and 3 of this thesis establish that if the private signal structure is characterized by Bounded Likelił
then there does not exist a decision rule that achieves learning in probability if $K = 1$ but there exists one that achi
.g in probability when $K > 1$. (The first result for the special case of coin tossing has been established by Koplo
). If the private signal structure is characterized by Unbounded Likelihood Ratios then Cover in [7] has construct
n rule that achieves almost sure learning and consequently learning in probability. For the case of Unbounded Likelił
and forward looking agents we were able to show that there exist equilibria that learn and others that do not. There
nnot give specific answers for this case.

Bibliography

[1] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar. Bayesian Learning in Social Networks. *National Bureau of Economic Research Working Paper Series*, pages 14040+, May 2008.

[2] A. Banerjee and D. Fudenberg. Word-of-mouth learning. *Games and Economic Behavior*, 46(1):1 - 22, 2004.

[3] A. V. Banerjee. A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, 107:797–817, 1992.

[4] S. Bikhchandani, D. Hirshleifer, and I. Welch. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *The Journal of Political Economy*, 100(5):992–1026, October 1992.

[5] E. Borel. *Traité du calcul des probabilités et de ses applications*, volume 2 of *Applications a l' arithmétique et a la théorie des fonctions*. 1926.

[6] K.L. Chung and P. Erdos. On the Application of the Borel-Cantelli Lemma. 1952.

[7] T. M. Cover. Hypothesis Testing with Finite Statistics. *Ann. Math. Statist.*, 40, 1969.

[8] L.K. Ekchian. *Optimal Design of Distributed Detection Networks*. PhD thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1982.

[9] L.K. Ekchian and R.R. Tenney. Detection Networks. *Proc. 21st IEEE Conf. Decision Control*, 1982.

[10] M.E. Hellman and T.M. Cover. Learning with Finite Memory. *The Annals of Mathematical Statistics*, 41, 1970.

[11] J. Koplowitz. Necessary and Sufficient Memory size for m-Hypothesis Testing. *IEEE Transactions on Infromation Theory*, 21, 1975.

[12] D. Manal. *On decision making in tandem networks*. Master's thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 2009.

[13] S. Nageeb and A. Kartik. Observational Learning with Collective Preferences , 2010.

[14] J.D. Papastavrou and M. Athans. Distributed Detection by a Large Team of Sensors in Tandem. *IEEE Transactions on Aerospace Electronic Systems*, 1992.

[15] W. Rudin. *Principles of Mathematical Analysis*, volume 1. McGraw-Hill, Reading, Massachusetts, third edition.

[16] L. Smith and P. Sorensen. Pathological Outcomes of Observational Learning. *Econometrica*, 68(2):371–398, 2000.

[17] L. Smith and P. N. Sorensen. Rational Social Learning with Random Sampling. *Unpublished manuscript*, 1998.

[18] L. Smith and P. N. Sorensen. Informational Herding and Optimal Experimentation. Cowles foundation discussion papers, Cowles Foundation for Research in Economics, Yale University, January 2006.

[19] R. Tenney and N.R.Jr. Sandell. Detection with Distributed Sensors. *IEEE Trans. Aero. Electron. Syst.*, 17, 1981.

[20] J. N. Tsitsiklis. Decentralized detection. In *Advances in Statistical Signal Processing*, pages 297–344. JAI Press, 1993.