LEARNING IN REPEATED INTERACTIONS ON NETWORKS

WANYING HUANG[†], PHILIPP STRACK[‡], AND OMER TAMUZ[§]

Abstract. We study how long-lived, rational, exponentially discounting agents learn in a social network. In every period, each agent observes the past actions of his neighbors, receives a private signal, and chooses an action with the objective of matching the state. Since agents behave strategically, and since their actions depend on higher order beliefs, it is difficult to characterize equilibrium behavior. Nevertheless, we show that regardless of the size and shape of the network, and the patience of the agents, the speed of learning in any equilibrium is bounded from above by a constant that only depends on the private signal distribution.

1. Introduction

We study the aggregation of information through interactions between agents who observe each others' actions on a social network. We show that information aggregation fails: the speed of learning stays bounded even in large networks, where efficient aggregation of all the private information would lead to arbitrarily fast learning. Methodologically, we introduce new techniques that allow us to relax commonly made assumptions and study a rational setting, on general networks, with forward-looking agents who interact repeatedly. Repeated interactions can, for example, describe the exchange of opinions and information among friends on social media or firms that learn from each others actions. Arguably, social learning driven by such interactions can be an important determinant in many choice domains, such as investments, health insurance, schools, technology adoption, or where to live and work.

We consider a general social network setting with repeated interactions between a group of rational agents. There is a fixed but unknown binary state of the world. In every period, each agent receives a private signal about the state and observes all past actions from a subset of society, to whom we refer as his neighbors. Based on this information, each agent updates his beliefs, chooses one of two possible actions, and receives a flow payoff if his action matches the state. The agents do not observe flow payoffs. We consider both myopic agents who maximize their instantaneous payoff, as well as strategic agents

Date: July 19, 2022.
[†]California Institute of Technology. Email: whhuang@caltech.edu.
[‡]Yale University. Email: philipp.strack@gmail.com.
[§]California Institute of Technology. Email: omertamuz@gmail.com. Omer Tamuz was supported by a Sloan fellowship, a BSF award (#2018397) and a National Science Foundation CAREER award (DMS-1944153).
We would like to thank Krishna Dasaratha, Ben Golub, Hamid Sabourian and Leeat Yariv for insightful comments.

who are forward-looking, and exponentially discount the future. Since strategic agents care about their future utilities, they may sacrifice their present flow utilities by changing their actions to affect the future actions of others, in order to induce them to reveal more information.

The constant influx of private information in our model allows every agent to eventually learn the truth and choose the correct action. Thus, we focus on how fast agents learn. If the number of agents doubles, the number of private signals available to society also doubles. Hence, if information were aggregated efficiently, the time that it would take for agents to choose correctly (say, with a given high probability) would decrease by a factor of two. In other words, when information aggregation is efficient, the speed of learning increases linearly with the number of agents. The question that we ask is: what is the speed of learning in equilibrium, and how does it depend on the number of agents and structure of the network?

We focus on *strongly connected* networks, where there is an observational path between every pair of agents, since otherwise efficient aggregation of information is excluded by the lack of informational channels. This is a mild assumption, and follows, for example, from the "six degrees of separation rule", that stipulates that there is a path of length at most six between every two members of a social network.[1] Our main result (Theorem 1) shows that the speed of learning does not increase linearly with the number of agents, and is, in fact, bounded from above by a constant that only depends on the private signal distribution, and is independent of the structure of the network.

For example, consider agents who, in each period, observe an independent binary signal that is equal to the state with a probability of 0.9. Then, regardless of the number of agents and network structure, the speed of learning never exceeds ten times the speed at which an agent learns on his own. This is despite the fact that if n agents shared their signals publicly, they would learn n times as fast. Thus, a society of 1,000 agents who observe their neighbors' past actions never learns faster than a society of ten agents in which information is efficiently aggregated. This means that in a society of 1,000 agents, at least 99% of information from the private signals is lost for *any* structure of the observational network (and *any* equilibrium played by the agents).

The mechanism behind this bounded rate of learning is as follows: First, in a strongly connected network all agents learn at the same rate, as each agent could guarantee himself the same learning rate as any of his neighbors. We establish the bound on the learning rate by contradiction. Suppose that agents learn at a rate that is higher than the rate that their individual private signals can support. This implies that social information, which consists of agents' actions, will become much more precise than each agent's private information. As a result, agents will ignore their private signals and only rely on the

[1] The science fiction writer Frigyes Karinthy proposed this rule in his 1929 short story "Láncszemek" (in English, "Chains"). The rule was confirmed empirically on a number of online social network data sets (Watts and Strogatz, 1998; Leskovec and Horvitz, 2008).

social information they observe from their neighbors. This implies that agents' actions no longer reveal any information about the state. Thus social information cannot grow to be too precise over time. This contradicts our previous hypothesis, and we conclude that agents cannot learn too fast. This mechanism has the same flavor of the classical information cascades that drive the failure of information aggregation in the herding literature (Bikhchandani et al., 1992).

We contribute to the social learning literature in three aspects. First, instead of focusing on short-lived agents who act only once, we consider a more realistic model in which agents are long-lived and repeatedly interact with each other. Second, we extend existing models of social learning by considering strategic agents who discount their future utilities at a common rate. The analysis of strategic agents is complicated since these agents may have an incentive to choose a sub-optimal action today to learn more information from the actions of others in the future. To avoid this difficulty, most of the social learning literature has focused on myopic agents who fully discount their future utilities, thus completely shutting down their incentives to experiment strategically. Third, we generalize previous work on learning speed to a general social network where agents only observe their neighbors' past actions. We will provide a detailed literature review next.

1.1. **Related Literature.** There are very few papers addressing repeated interaction between rational agents and its role in information aggregation.[2] This is because it is challenging to analyze the evolution of beliefs of long-lived agents, particularly when these beliefs are influenced by the interactions between them. Thus, most of the literature has focused on either short-lived agents who only act once (Banerjee, 1992; Bikhchandani et al., 1992; Acemoglu et al., 2011) or non-fully-rational agents such as quasi-Bayesian agents (Bala and Goyal, 1998; Molavi et al., 2018) and heuristic agents (DeGroot, 1974; Golub and Jackson, 2010; Dasaratha et al., 2020). Nevertheless, as many real-life situations involve repeated interactions, it is natural to study models that allow these interactions. It is likewise interesting to understand rational forward-looking agents as this is an important benchmark case to assess whether or to what extent failures of information aggregation are driven by a lack of patience or rationality of the agents.

Among models which consider repeated interactions, the network structure in which learning occurs has been a primary focus of study. For example, Bala and Goyal (1998) consider quasi-Bayesian agents who only use part of the available information.[3] They examine how the geometry of the social network affects learning outcomes. Instead of quasi-Bayesian agents, Mossel et al. (2014) consider rational but myopic agents and show that learning occurs for non-atomic private beliefs and undirected graphs. Later, Mossel

[2]See Vasal and Anastasopoulos (2016) for an exception where they study strategic players who act repeatedly. However, as agents' payoffs may depend on others' actions, the strategic interaction among their agents is not purely informational, which differs from the model we study here.

[3]Specifically, Bala and Goyal (1998) assume that agents do not make any inferences on how their neighbors' actions affect their neighbors' neighbors' actions and etc.

et al. (2015) further generalize their setting to allow for forward-looking agents. Unlike our model, agents in these models only receive one signal at the beginning of time; this precludes applications in which agents not only learn from others, but also have access to increasingly precise information that stems from their own experiences. They do not study the speed of learning, and instead focus on identifying the types of social networks in which learning always occurs.[4]

Complementing the previous literature, we take the next natural step: we ask how fast learning occurs and study its relationship with the size of the network. Furthermore, we consider agents with any discount factor $\delta \in [0, 1)$ so that myopic agents ($\delta = 0$) is a special case of the possible types of agents. To our best knowledge, this is the first paper to consider social learning in a network setting with fully-rational and forward-looking agents who interact repeatedly.

A recent paper that does consider rational agents in a repeated setting is by Harel, Mossel, Strack, and Tamuz (2021), who study the speed of learning on a network in which all agents are myopic and observe each other. They show that for any number of agents, the speed of learning from actions is bounded above by a constant. Their proofs rely crucially on the symmetry inherent in the complete network in which all agents observe each other, and the relative simplicity of myopic behavior. We generalize their result in two aspects. First, we consider a general social network rather than a complete observation structure. Second, we relax their assumption of myopic agents and consider strategic agents. We explore new techniques that enable us to make these generalizations, and provide new insight into the forces at work.

2. Model

Let $N = \{1, 2, \ldots, n\}$ be a finite set of agents. Time is discrete and the horizon is infinite, i.e. $t \in \{1, 2, \ldots\}$. In every time period t, each agent i has to choose an action $a_t^i \in \mathcal{A} = \{\mathfrak{b}, \mathfrak{g}\}$. There is an unknown state of the world $\Theta \in \{\mathfrak{b}, \mathfrak{g}\}$, with each state a priori equally likely.[5] The state does not change over time and agents are not informed about whether their previous actions match the state.[6] We assume a binary state and binary actions purely for expositional simplicity. In the conclusion we describe how the same results can be derived more generally, for finite state spaces, finite action spaces, and general utilities.

2.1. **Agents' Information.**

In each period t, agent i receives a private signal s_t^i drawn from a finite set Ω_t^i. Conditional on the state Θ, signals s_t^i are independent across agents

[4]They show that on connected infinite *egalitarian* graphs, in any equilibrium, all agents converge to the correct action with probability one.

[5]We assume a uniform prior purely to simplify notation. Relaxing it does not change the results of the paper.

[6]Suppose that agents receive noisy feedback about whether their previous actions match the state. Then, equivalently, we can think of this feedback as the agents' signals.

and time, with distribution $\mu^i_{t,\Theta} \in \Delta(\Omega^i_t)$. The distributions $\mu^i_{t,\mathfrak{g}}$ and $\mu^i_{t,\mathfrak{b}}$ are mutually absolutely continuous, so that no signal perfectly reveals the state. Furthermore, there exists a constant $M > 0$ that bounds the absolute value of the likelihood induced by any signal:

$$M = 2 \sup_{i,t,\omega\in\Omega^i_t} \left| \log \frac{\mu^i_{t,\mathfrak{g}}(\omega)}{\mu^i_{t,\mathfrak{b}}(\omega)} \right|. \tag{1}$$

We allow signals to depend on calendar time and the agents' identities to highlight the robustness of our results. However, to understand our main economic insight, it suffices to think of the setting where all signals are i.i.d. across agents and time, as is typically assumed in the literature.

For each agent i there is a subset of agents $N_i \subseteq N$ who are his *social network neighbors*, and whose actions he observes. We include $i \in N_i$ since agent i observes his own actions. The information available to agent i at time t, before taking his action a^i_t, thus consists of a sequence of private signals $(s^i_1, \cdots, s^i_t)$ and the history of actions observed by i, $H^i_t = \{a^j_s : s < t, j \in N_i\}$. Let $\mathcal{I}^i_t = \Omega^i_1 \times \cdots \times \Omega^i_t \times \mathcal{A}^{|N_i|\times(t-1)}$ be the space of information available to agent i at time t and thus $I^i_t = (s^i_1, \cdots, s^i_t, H^i_t)$ is an element of $\mathcal{I}^i_t$. We assume for most of the analysis that the social network is *strongly connected*: there is an observational path between every pair of agents (we relax this assumption in §6).[7] This assumption avoids situations in which efficient aggregation of information is precluded because there is no channel for information to travel from i to j.

2.2. **Speed of Learning.** We measure the speed of learning of agent i by the asymptotic rate at which he converges to the correct action – the action that matches the state (see, e.g., Vives, 1993; Molavi et al., 2018; Hann-Caruthers et al., 2018; Rosenberg and Vieille, 2019; Harel et al., 2021). Formally, the *speed of learning* of agent i is

$$\liminf_{t\to\infty} -\frac{1}{t} \log \mathbb{P}[a^i_t \neq \Theta]. \tag{2}$$

If this limit exists and is equal to r, then the probability of mistake at large times t is approximately e^{-rt}. As we explain below in §3, this is the case for the benchmark case of a single agent who receives conditionally i.i.d. private signals at each period.

2.3. **Strategies and Payoffs.** A pure strategy of agent i at time t is a function $\sigma^i_t : \mathcal{I}^i_t \to \mathcal{A}$. A pure strategy of agent i is a sequence of functions $\sigma^i = (\sigma^i_1, \sigma^i_2, \cdots)$ and a pure strategy profile is a collection of pure strategies of all agents, $\sigma = (\sigma^i)_{i\in N}$. We write $\sigma = (\sigma^i, \sigma^{-i})$ for any agent $i \in N$, where σ^{-i} denotes the pure strategies of all agents other than i. Given a pure strategy profile σ, the action of agent i at time t is $a^i_t = a^i_t(\sigma) = \sigma^i_t(I^i_t)$. The flow utility of agent i at time t is

$$U^i_t(\sigma) = \mathbb{1}(a^i_t(\sigma) = \Theta).$$

[7]Formally, for each $i, j \in N$ there is a sequence $i = i_1, i_2, \ldots, i_k = j$ such that $i_2 \in N_{i_1}, i_3 \in N_{i_2}, \ldots i_k \in N_{i_{k-1}}$.

The expected flow utility of agent i is

$$u_t^i(\sigma) = \mathbb{E}[U_t^i(\sigma)] = \mathbb{P}[a_t^i(\sigma) = \Theta].$$

We assume that agents discount their future utilities at a common rate $\delta \in [0,1)$. The expected utility of agent i under strategy profile σ is thus

$$u^i(\sigma) = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} u_t^i(\sigma) = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1}\mathbb{P}[a_t^i(\sigma) = \Theta],$$

where $(1-\delta)$ is the normalization factor so that $u^i(\sigma) \in [0,1]$ for any i and strategy profile σ. We call agents with $\delta = 0$ *myopic* and agents with $\delta > 0$ *strategic*. Myopic agents fully discount future payoffs and choose their actions in each period to maximize the expected flow utilities.

We note that regardless of whether agents are strategic or myopic, the sole benefit of observing others' past actions is to learn about the state. That is, others' actions reflect the signals they receive, and thus observing others' actions can help agents make better inferences about the state. This pure informational motive is an important feature of the model: each agent's stage utility depends only on his own action and the state, and it is independent of the actions of the others.

2.4. **Equilibrium.** This is a game of incomplete information, in which agents may have different information regarding the underlying unknown state. We use Nash equilibrium as our equilibrium concept and refer to it as equilibrium thereafter.[8] The existence of a (mixed) equilibrium is guaranteed in this game by standard arguments, since, in the product topology on strategies, the space of strategies is compact and utilities are continuous. We note here that every mixed equilibrium can be mapped to a behaviorally equivalent pure equilibrium by adding to each agent's private signal an additional component that is independent of the state and all other signals, and assuming that the agent uses this signal to randomly choose between actions.[9] As our results will only depend on the information about the state contained in the signal, it thus suffices to establish them for pure strategy equilibria, to show that they hold for all (pure and mixed) equilibria.

As usual, a pure strategy profile σ is an equilibrium if no agent can obtain a strictly higher expected utility by unilaterally deviating from σ. That is, a pure strategy profile σ is an *equilibrium* if for all agents i, and all strategies τ^i

$$u^i(\sigma^i, \sigma^{-i}) \geq u^i(\tau^i, \sigma^{-i})\,.$$

[8]Our results, which apply to all equilibria, thus also apply to any refinement of Nash equilibrium.

[9]Formally, for any signal space Ω_t^i we can consider $\tilde{\Omega}_t^i = \Omega_t^i \times \mathcal{A}^{|\mathcal{I}_t^i|}$ with the signal distributions $\tilde{\mu}_{t,\theta}^i$ equal to the product measure of $\mu_{t,\theta}^i$ and $|\mathcal{I}_t^i|$ independent Bernoulli random variables each taking the value $\mathfrak{g}$ with probability $\mathbb{P}[a_t^i = \mathfrak{g} | I_t^i]$. As this transformation does not affect the informational content of the signals it leaves the constant M unchanged. Furthermore, we can replicate any behaviorally mixed strategy by the pure strategy that takes the action $\mathfrak{g}$ if and only if the entry of the second component corresponding to the private history I_t^i equals $\mathfrak{g}$.

3. The Public Signals Benchmark

As a benchmark, we briefly discuss the case of *public* signals for a single or multiple agents. We also assume that signals are i.i.d. across time and agents, with $\mu_\theta = \mu^i_{t,\theta}$ for $\theta \in \{\mathfrak{b}, \mathfrak{g}\}$. In the single-agent case, it is known that the speed of learning is equal to

$$r_a = \sup_{z \geq 0} \left[-\log \sum_\omega \left(\frac{\mu_\mathfrak{b}(\omega)}{\mu_\mathfrak{g}(\omega)} \right)^z \mu_\mathfrak{g}(\omega) \right] > 0 \,.$$

This result follows from a standard application of large-deviation theory and is classic in the statistics and probability literature (see Cover and Thomas, 2006, 314-316).[10]

Next, consider the case where each of n agents observes all n independent public signals in each period, as well as their neighbors' past actions. As actions contain no additional information about the state relative to the signals, this situation is identical to the single-agent case, except that now each agent receives n independent signals at each period. An agent in period t will thus have observed exactly as many signals as a single agent in autarky in period $n \cdot t$. It thus follows from the single-agent case that when signals are public, the speed of learning in a social network with size n is $n \cdot r_a$.

These results for the case of public signals immediately bound the learning rate in the private signals case: In any social network, observed actions contain weakly less information than the private signals. Thus, $n \cdot r_a$ is an upper bound to the speed of learning for any network with n agents and private signals.

4. Results

We now state our main result. It turns out that in a strongly connected network, all agents learn at the same rate (by Lemma 2 in §5), and we call this common learning rate *the equilibrium speed of learning.* In contrast to the public signal case, our main result shows that regardless of the size of the network, the equilibrium speed of learning is bounded above by a constant. Recall that in (1) we defined $M/2$ to be the maximal log-likelihood ratio induced by any signal.

Theorem 1. *The equilibrium speed of learning is at most M, in any equilibrium, on any social network of any size, and for any discount factor $\delta \in [0, 1)$.*

Perhaps surprisingly, Theorem 1 shows that adding more agents (thus more information) to the network and expanding the network cannot improve the speed of learning beyond some bound, which is twice the strength of the strongest possible signal, as measured in log-likelihood ratios. Indeed, this upper bound on the learning speed implies that more and more information is lost as the size of the network increases.

For example, for independent binary signals that are equal to the state with probability 0.9, the speed of learning in a social network of any size is bounded by that of ten agents

[10] See also Appendix B in Harel et al. (2021) for more details and a complete proof.

who observe each other's signals directly.[11] Thus, a social network of 1,000 agents who observe their neighbors' past actions never learns faster than ten agents who share their private signals. Equivalently, they do not learn more than ten times as fast as a single agent. Thus almost all of the private information in large networks is lost, resulting in inefficient information aggregation.

The idea behind our proof of Theorem 1 is as follows. Intuitively, one might think that larger networks would boost the speed of learning as agents acquire more and more information from their neighbors, as well as indirectly from their neighbors' neighbors etc. However, we argue that the social information gathered from observing neighbors' past actions cannot be too precise. Indeed, if that were the case, agents would base their decisions only on the social information. As a result, their actions would no longer reveal any information about their private signals so that information aggregation would cease. Thus, social information cannot grow to be much more precise than private information. But if agents learn quickly, then their actions provide very precise social information. Hence, we conclude that agents cannot learn too quickly.

In sum, regardless of the size of the network, private information must continue to influence agents' decision-making, which can only happen if the social information is not too precise, which in turn can only happen if agents do not learn quickly. Moreover, as we state in the next section, M is an upper bound to how fast the precision of private information increases with time (see Lemma 4 in §5), and this bound, too, is independent of the network size. Combining these insights, we conclude that the speed of learning in a social network of any size is bounded by M.

5. Analysis

This section will provide a detailed analysis of the agents' beliefs and behavior and lead to a proof sketch for Theorem 1.

5.1. **Agents' Beliefs.** Let p_t^i denote the posterior belief of agent i assigned to state $\Theta = \mathfrak{g}$ after observing I_t^i, i.e. $p_t^i = \mathbb{P}[\Theta = \mathfrak{g} | I_t^i]$. The log-likelihood ratio of agent i's posterior beliefs at time t is

$$L_t^i = \log \frac{p_t^i}{1 - p_t^i} = \log \frac{\mathbb{P}[\Theta = \mathfrak{g} | s_1^i, \ldots, s_t^i, H_t^i]}{\mathbb{P}[\Theta = \mathfrak{b} | s_1^i, \ldots, s_t^i, H_t^i]}. \tag{3}$$

Then, it follows from Bayes' rule that the log-likelihood ratio of agent i's posterior belief p_t^i at time t is:

$$L_t^i = \log \frac{\mathbb{P}[\Theta = \mathfrak{g}]}{\mathbb{P}[\Theta = \mathfrak{b}]} + \log \frac{\mathbb{P}[H_t^i | \Theta = \mathfrak{g}]}{\mathbb{P}[H_t^i | \Theta = \mathfrak{b}]} + \log \frac{\mathbb{P}[s_1^i, \ldots, s_t^i | H_t^i, \Theta = \mathfrak{g}]}{\mathbb{P}[s_1^i, \cdots, s_t^i | H_t^i, \Theta = \mathfrak{b}]}.$$

[11] To see this, given the signal distribution, we calculate the learning rate in the single-agent case, which is approximately equal to 0.51. As discussed in §3, the learning rate in a network of ten agents with public signals is ten times that of the single-agent case. From Theorem 1, the upper bound M to the equilibrium speed of learning on any social network of any size is approximately 4.4, which is less than ten times 0.51.

We call

$$S_t^i = \log \frac{\mathbb{P}[H_t^i|\Theta = \mathfrak{g}]}{\mathbb{P}[H_t^i|\Theta = \mathfrak{b}]},$$

the *social likelihood* of agent i at time t. This is the log-likelihood ratio of the social information observed by agent i. Intuitively, S_t^i measures the inference the agent would draw from the observations of his neighbors actions if he had not observed any private signals himself. Similarly, we call

$$P_t^i = \log \frac{\mathbb{P}[s_1^i, \ldots, s_t^i|H_t^i, \Theta = \mathfrak{g}]}{\mathbb{P}[s_1^i, \cdots, s_t^i|H_t^i, \Theta = \mathfrak{b}]},$$

the *private likelihood* agent i at time t.

Notice that given our uniform prior assumption, the log-likelihood of agent i's prior belief $L_0^i = \log \frac{\mathbb{P}[\Theta=\mathfrak{g}]}{\mathbb{P}[\Theta=\mathfrak{b}]}$ is equal to zero. Thus, we can write the log-likelihood ratio of agent i's posterior belief at time t as

$$L_t^i = S_t^i + P_t^i, \tag{4}$$

which is the sum of his social likelihood S_t^i and his private likelihood P_t^i. We call L_t^i the *posterior likelihood* of agent i at time t.

5.2. **Agents' Behavior.** In the context of a strategy profile σ, the *myopic action* of agent i at time t is

$$m_t^i = \begin{cases} \mathfrak{g} & \text{if } p_t^i \geq \frac{1}{2} \\ \mathfrak{b} & \text{otherwise.} \end{cases}$$

This is the action that maximizes the expected flow utility, and hence it is the action that a myopic agent ($\delta = 0$) would take.[12]

In contrast to a myopic agent, a strategic agent may not always choose the myopic action in equilibrium. Indeed, since a strategic agent is forward-looking, in each period he faces a trade-off between choosing the myopic action and strategically experimenting by choosing the non-myopic action. On the one hand, he needs to bear the immediate cost associated with the non-myopic action. On the other hand, choosing the non-myopic action may allow him to elicit more information from his neighbors' future actions, which he could then use to make better choices in the future. Hence, when the informational gain from experimenting exceeds the current loss caused by the non-myopic action, a strategic agent has an incentive to experiment.

Nevertheless, we show that when an agent's belief is close enough to zero or one, he chooses the myopic action in equilibrium. This holds despite the fact that the agent is forward-looking. Intuitively, if a strategic agent is very confident about the state, he is expected to pay a high cost if he chooses the non-myopic action and experiments.

[12]When $p_t^i = 1/2$ the agent is indifferent, and we assume that he chooses $m_i^t = \mathfrak{g}$. Our results do not depend on this choice and would follow for any tie breaking rule that is common knowledge.

Consequently, as his expected future gain will not exceed the expected current loss from experimenting, he has no incentive to experiment.

Lemma 1 (Myopic and Strategic Behavior)**.**

(i) In equilibrium, if for some time t it holds that $|L_t^i| \geq -\log(1-\delta)$ then $a_t^i = m_t^i$.

(ii) There exists a random time $T < \infty$ such that in equilibrium, all agents behave myopically after T, i.e. $t \geq T \Rightarrow a_t^i = m_t^i$ for all i almost surely.

The first part of Lemma 1 implies that as agents become more patient, i.e., δ increases, the threshold of the posterior likelihood $-\log(1-\delta)$ for choosing the myopic action becomes closer to certainty. Indeed, as δ approaches 1, agents value their future utilities more and the incentive of experimenting becomes stronger. For these agents to forgo their potential expected future informational gains and choose the myopic action, they must be fairly confident about the state. The second part of Lemma 1 states that in finite time the belief of all agents will be sufficiently precise such that they all behave myopically in all future periods.

The next lemma shows that in equilibrium, each agent learns weakly faster than any of his neighbors. Denote the equilibrium speed of learning of agent i by r_i.

Lemma 2 (All agents learn at the same rate)**.**

(i) If agent i can observe agent j, i.e. $j \in N_i$, then in equilibrium i learns at a (weakly) higher rate, i.e. $r_i \geq r_j$.

(ii) All agents learn at the same rate, i.e. $r_i = r_j$ for all i, j, in any strongly connected network.

We will henceforth call the common rate of learning in a strongly connected network *the equilibrium speed of learning*. The proof of this lemma relies on an extension of the *imitation principle* from myopic agents to strategic agents. For myopic agents, the imitation principle states that if i observes j, then i's actions are not worse than j's:

$$\mathbb{P}[a_t^i \neq \Theta] \leq \mathbb{P}[a_{t-1}^j \neq \Theta],$$

since i can always imitate j (Gale and Kariv, 2003; Golub and Sadler, 2017). We show that for strategic agents, in equilibrium, i's actions are never much worse than j's, even though i may choose myopically sub-optimal actions:

$$\mathbb{P}[a_t^i \neq \Theta] \leq \frac{1}{1-\delta}\mathbb{P}[a_{t-1}^j \neq \Theta].$$

5.3. **Social and Private Beliefs.** We analyze the agents' beliefs by decomposing their likelihoods into the private and social parts. Recall that by (4), the posterior likelihood of agent i's at time t, L_t^i, is equal to $S_t^i + P_t^i$, the sum of the social and the private likelihoods. We are interested in the sign of L_t^i as it determines the corresponding myopic action: m_t^i equals $\mathfrak{g}$ if $L_t^i \geq 0$, and $\mathfrak{b}$ otherwise. Let us first focus on the first component of L_t^i: agent i's social likelihood S_t^i. In the following lemma, we establish a relationship between the equilibrium speed of learning and the precision of the social likelihood, which is crucial in proving our main theorem.

Lemma 3. *Suppose that the equilibrium speed of learning is strictly greater than r. Then, conditioned on $\Theta = \mathfrak{b}$,*

$$\limsup_{t\to\infty} \frac{1}{t} S_t^i \leq -r \quad \textit{almost surely,}$$

and conditioned on $\Theta = \mathfrak{g}$,

$$\liminf_{t\to\infty} \frac{1}{t} S_t^i \geq r \quad \textit{almost surely.}$$

This lemma states that a high learning rate implies that the social information inferred from a given agent i and his neighbors' actions must become precise at a high rate. Intuitively, if agents learn quickly, then their actions provide very precise social information.

The proof of Lemma 3 uses the idea of a fictitious outside observer who observes the same social information as agent i and nothing else. Since he observes i's actions, he can achieve the same learning rate as i. This implies that his posterior likelihood increases fast. Hence, as this outside observer's posterior likelihood coincides with agent i's social likelihood, the precision of i's social information increases at a rate of at least r.

Next, we focus on the second component of L_t^i: agent i's private likelihood, P_t^i. As agents receive more independent private signals over time, their private information about the state becomes more precise. However, the precision of their private information cannot increase without bound, as shown in the following lemma.

Lemma 4. *For any agent i at time t, the absolute value of the private likelihood is at most $t \cdot M$, i.e.*

$$\frac{1}{t}|P_t^i| \leq M \quad \textit{almost surely.}$$

This lemma states that that at any given time t, there is an upper bound to the precision of agents' private information, which only depends on the private signals distribution and is independent of the structure of the network and the history of observed actions. Notice that since $P_t^i = L_t^i - S_t^i$ by (4), it captures the difference between what i knows about the state and what an outside observer who observes i's actions and his neighbors' actions would know about the state. Thus, the bound Mt assigned to i's private signals also applies to the difference between the posterior likelihood L_t^i and the social likelihood S_t^i.

5.4. **Proof sketch for Theorem 1.** We end this section by providing a sketch of the proof of Theorem 1 using our earlier results. Suppose to the contrary that in equilibrium, agents learn at a rate that is strictly higher than M, where M is twice the log-likelihood ratio of the strongest signal. Then, by Lemma 3, the social information would become precise at a rate that is also strictly higher than M. Meanwhile, at any given time t, the precision of the private information is at most Mt, as shown in Lemma 4. Hence, by (4) the sign of L_t^i would be determined purely by the social information after some (random) time. By Lemma 1, there exists a time T such that from T onward, in equilibrium, all agents act only based on the social information, and furthermore they would choose the myopic action even though they are forward-looking. Consequently, their actions would no longer reveal any information about their private signals and information would cease to be aggregated. This contradicts our hypothesis that the precision of the social information grows at such a high rate. Therefore, we conclude that the equilibrium speed of learning in networks does not increase beyond M.

6. Networks which are not Strongly Connected

So far, we have focused on strongly connected networks where there is an observational path between every pair of agents. While on a strongly connected network all agents learn at the same rate, this is not true for general networks. For example, consider a simple star network where there is a single agent at the center who observes everyone, and where the remaining peripheral agents observe no one. Here, the peripheral agents' actions are independent conditional on the state. These actions supply the central agent with $n-1$ additional independent signals, and he thus learns at a rate that increases linearly with the number of agents.[13] In contrast, all peripheral agents learn at a constant rate r_a, as in the single-agent case. Hence, for general networks, depending on the structure of the network, some agents can learn faster than others.

More importantly, this star network example implies that the bound obtained in Theorem 1 does not hold for all agents in a non-strongly connected network. The intuitive reason is that when the network is not strongly connected, some agents might remain unobservable to others, e.g. the central agent in the star network. These agents can thus learn very fast from observing others since their own past actions do not affect the actions of others, rendering others' past actions more informative. This cannot happen in a strongly connected network where every agent (potentially indirectly) learns about the actions of every other agent.

Nevertheless, even though it is not necessarily true that all agents learn slowly, our next result establishes that it is still true that some agents will learn slowly in any network.

Proposition 1. *Consider an arbitrary network and let r_i be the speed of learning of agent i. We have that $\min_i r_i \leq M$.*

The proof of Proposition 1 relies on the idea that within any general network, there is always a strongly connected sub-network, say E, in which no agent observes any agent outside of this sub-network. Thus, the learning process at E is independent of the agents outside of E. Since E is strongly connected, Theorem 1 implies that the speed of learning on E is bounded by a constant.

[13] See Theorem 5 in Harel et al. (2021).

7. Conclusion

In this paper, we show that information aggregation is highly inefficient for large groups of agents who learn from private signals and by observing their social network neighbors. We measure the efficiency of information aggregation by *the speed of learning* and show that regardless of the size of the network, the speed of learning is bounded above by a constant, which only depends on the private signal distribution. To overcome the difficulty of constructing equilibria explicitly, we focus on the asymptotic rate of learning, allowing us to prove results that apply to all equilibria.

To simplify notation we made several assumptions which could be relaxed without major changes to the proofs or results. Allowing for a non-symmetric prior (or for a more general utility function over actions and state) simply changes the cut-off that separates the regions where the agent takes either action. This corresponds to a shift of the log-likelihood by a constant which does not influence the long-run rate of learning and thus does not affect our results. Similarly, allowing for more than two actions does not change our results if one assumes that there is a unique optimal action in each state. In this case one can again define the speed of learning as the rate at which agents choose the action that is optimal given the state. Our results can also be extended to a random network setting in which agents do not know the exact structure of the realized social network, but only observe their local neighborhood.

Finally, one can easily extend our analysis to the case where agents observe each others' actions not in every period but only at random times, as long as the number of consecutive unobserved actions is uniformly bounded. For example, we could allow for the time between periods in which an agent observes another along an observational path to be uniformly drawn between 0 and 10 periods. As this finite delay in observation does not change the rates of learning it does not affect our proofs or results.

One natural substantive extension is to allow the underlying state to change over time (see Moscarini et al., 1998; Frongillo et al., 2011; Dasaratha et al., 2020). For example, the underlying unknown state could capture the quality of a local restaurant or school, which might fluctuate gradually. In such a setting, one could replace the speed of learning metric with the long-run probability of making the correct choice and study whether information gets aggregated in this case and, if so, whether the information aggregation process is efficient.

Appendix A. Proofs

Proof of Lemma 1. (i) First notice that we can write the agent's expected utility at some fixed time t as the sum $U_{<t} + \delta^t U_{\geq t}$, where

$$U_{<t} = (1-\delta) \sum_{k=1}^{t} \delta^{k-1} \mathbb{P}[a_k^i = \Theta | I_t^i]$$

is the sum of the expected flow utilities until time t, and $U_{\geq t}$ is the expected continuation utility at time t given by

$$U_{\geq t} = (1-\delta)\sum_{k=0}^{\infty} \delta^k \mathbb{P}[a^i_{t+k} = \Theta | I^i_t].$$

Now fix t and without loss of generality suppose $p^i_t \geq 1/2$. As $\mathbb{P}[\mathfrak{b} = \Theta | I^i_t] = 1 - p^i_t$ and $\mathbb{P}[a^i_{t+k+1} = \Theta | I^i_t] \leq 1$, the expected continuation utility when taking the suboptimal action $a^i_t = \mathfrak{b}$ is at most

$$(1-\delta)\mathbb{P}[\mathfrak{b} = \Theta | I^i_t] + \delta(1-\delta)\sum_{k=0}^{\infty} \delta^k \mathbb{P}[a^i_{t+k+1} = \Theta | I^i_t] \leq (1-\delta)(1-p^i_t) + \delta\,.$$

As $\mathbb{P}[a^i_{t+k+1} = \Theta | I^i_t] \geq \mathbb{P}[\mathfrak{g} = \Theta | I^i_t] = p^i_t$ we get that the payoff when taking the optimal action $a^i_t = \mathfrak{g}$ is at least

$$(1-\delta)\mathbb{P}[\mathfrak{g} = \Theta | I^i_t] + \delta(1-\delta)\sum_{k=0}^{\infty} \delta^k \mathbb{P}[a^i_{t+k+1} = \Theta | I^i_t] \geq (1-\delta)p^i_t + \delta p^i_t = p^i_t\,.$$

Thus, taking the myopically optimal action $\mathfrak{g}$ is optimal if

$$p^i_t \geq (1-\delta)(1-p^i_t) + \delta \Leftrightarrow \log \frac{p^i_t}{1-p^i_t} \geq -\log(1-\delta)\,.$$

By a symmetric argument for the case $p^i_t \leq 1/2$ it follows that $\log \frac{p^i_t}{1-p^i_t} \leq \log(1-\delta)$ is necessary.

(ii) Fix a discount factor $\delta \in [0,1)$ and let σ be an equilibrium. Let a^i_t be the action taken by i at time t under σ. Since the entire sequence of private signals reveals the state, the sequence of posterior beliefs $(p^i_1, p^i_2, \cdots)$ converges to either zero or one depending on the underlying state. Hence for all t large enough,

$$|L^i_t| = \left| \log \frac{p^i_t}{1-p^i_t} \right| \geq -\log(1-\delta).$$

By part (i) the agent will choose the myopic action in equilibrium. Since there are finitely many agents, this will hold for all i for all t larger than some (random) T. This means that from T onward, in equilibrium all agents will choose the myopic action, which is the action that is most likely to match the state. □

Proof of Lemma 2. Suppose i observes j. Let σ be an equilibrium and let a^i_t be the action taken by i at time t under σ. We claim that for $t > 1$,

$$\mathbb{P}[a^i_t \neq \Theta] \leq \frac{1}{1-\delta}\mathbb{P}[a^j_{t-1} \neq \Theta]. \tag{5}$$

To see this, consider the following deviation $(\bar{a}^i_\tau)_{\tau=1}^{\infty}$ for agent i: at all times $\tau < t$ there is no deviation, so that $\bar{a}^i_\tau = a^i_\tau$, and in all times $\tau \geq t$, i imitates j's action at time $t-1$ and chooses $\bar{a}^i_\tau = a^j_{t-1}$ at every history. Under this deviation, the continuation utility from time t on is simply

$$(1-\delta)\sum_{\tau=t}^{\infty}\delta^{\tau-t}\mathbb{P}[\bar{a}^i_\tau=\Theta]=(1-\delta)\sum_{\tau=t}^{\infty}\delta^{\tau-t}\mathbb{P}[\bar{a}^j_{t-1}=\Theta]=\mathbb{P}[a^j_{t-1}=\Theta].$$

Thus this deviation yields a gain in utility from time t on which is equal to

$$\Delta:=\mathbb{P}[a^j_{t-1}=\Theta]-(1-\delta)\sum_{\tau=t}^{\infty}\delta^{t-\tau}\mathbb{P}[a^i_\tau=\Theta].$$

Writing this in terms of the mistake probabilities, we get

$$\begin{aligned}\Delta&=-\mathbb{P}[a^j_{t-1}\neq\Theta]+(1-\delta)\sum_{\tau=t}^{\infty}\delta^{t-\tau}\mathbb{P}[a^i_\tau\neq\Theta]\\&=(1-\delta)\mathbb{P}[a^i_t\neq\Theta]-\mathbb{P}[a^j_{t-1}\neq\Theta]+(1-\delta)\sum_{\tau=t+1}^{\infty}\delta^{t-\tau}\mathbb{P}[a^i_\tau\neq\Theta]\\&\geq(1-\delta)\mathbb{P}[a^i_t\neq\Theta]-\mathbb{P}[a^j_{t-1}\neq\Theta].\end{aligned}$$

In equilibrium, $\bar{a}^i_\tau$ cannot be a profitable deviation. Hence $\Delta\leq 0$, and so (5) holds.

As a consequence,

$$\begin{aligned}r_i=\liminf_{t\to\infty}-\frac{1}{t}\log\mathbb{P}[a^i_t\neq\Theta]&\geq\liminf_{t\to\infty}-\frac{1}{t}\log\frac{1}{1-\delta}\mathbb{P}[a^j_{t-1}\neq\Theta]\\&=\liminf_{t\to\infty}-\frac{1}{t}\log\frac{1}{1-\delta}+\liminf_{t\to\infty}-\frac{1}{t}\log\mathbb{P}[a^j_{t-1}\neq\Theta]\\&=\liminf_{t\to\infty}-\frac{1}{t}\log\mathbb{P}[a^j_{t-1}\neq\Theta]=r_j.\end{aligned}$$

This proves part (i) of the claim. Part (ii) follows as in any strongly connected network, there is an observational path from each agent i to each other agent j and the monotonicity of learning rates shown in (i) applied along this path, implies that $r_i\geq r_j$. The opposite inequality holds by the same argument. □

Proof of Lemma 3. Consider an observer x who observes H^i_t, and chooses an action $a^x_t\in\mathcal{A}$ to maximize the probability that $a^x_t=\Theta$. Denote by $p^x_t=\mathbb{P}[\Theta=\mathfrak{g}|H^i_t]$ the belief of x conditional on what he observes, so that his corresponding posterior likelihood is:

$$\log\frac{p^x_t}{1-p^x_t}=\log\frac{\mathbb{P}[H^i_t|\Theta=\mathfrak{g}]}{\mathbb{P}[H^i_t|\Theta=\mathfrak{b}]}+\log\frac{\mathbb{P}[\Theta=\mathfrak{g}]}{\mathbb{P}[\Theta=\mathfrak{b}]}=S^i_t.$$

Then, since x observes i's actions up to time $t-1$, $\mathbb{P}[a^x_t\neq\Theta]\leq\mathbb{P}[a^i_{t-1}\neq\Theta]$ for all t. By assumption and the definition of speed of learning in (2), $\mathbb{P}[a^i_t\neq\Theta]\leq\mathrm{e}^{-rt}$ for all t large enough. Thus, it follows that $\mathbb{P}[a^x_t\neq\Theta]\leq\mathrm{e}^{-r(t-1)}$ for all t large enough. Moreover, since

$$\mathbb{P}[a^x_t\neq\Theta]=\frac{1}{2}\mathbb{P}[a^x_t=\mathfrak{g}|\Theta=\mathfrak{b}]+\frac{1}{2}\mathbb{P}[a^x_t=\mathfrak{b}|\Theta=\mathfrak{g}],$$

it holds for all t large enough that

$$\mathbb{P}[a^x_t=\mathfrak{g}|\Theta=\mathfrak{b}]\leq 2\mathrm{e}^{-r(t-1)}\quad\text{and}\quad\mathbb{P}[a^x_t=\mathfrak{b}|\Theta=\mathfrak{g}]\leq 2\mathrm{e}^{-r(t-1)}.$$

It follows the Borel-Cantelli lemma that conditioned on $\Theta = \mathfrak{b}$, the probability that the event $\{a_t^x = \mathfrak{g}\}$ occurs infinitely many times is zero. Since the event $\{a_t^x = \mathfrak{b}\}$ implies $p_t^x \leq 1/2$, it follows that conditioned on $\Theta = \mathfrak{b}$,

$$\limsup_{t \to \infty} p_t^x \leq 1/2 \quad \text{almost surely.} \tag{6}$$

Notice that $\min\{p_t^x, 1 - p_t^x\}$ equals the probability of choosing the wrong action from x's perspective, that is, $\min\{p_t^x, 1 - p_t^x\} = \mathbb{P}[a_t^x \neq \Theta | H_t^i]$. Thus by assumption, for all t large enough

$$\mathbb{E}[\min\{p_t^x, 1 - p_t^x\}] = \mathbb{P}[a_t^x \neq \Theta] \leq \mathrm{e}^{-r(t-1)}.$$

Thus, by Markov's inequality, for any $\varepsilon > 0$ and all t large enough

$$\mathbb{P}\left[\min\{p_t^x, 1 - p_t^x\} \geq \mathrm{e}^{-(r-\varepsilon)t}\right] \leq \mathrm{e}^{-\varepsilon t + r}.$$

It follows from the Borel-Cantelli lemma that the event

$$\left\{ \min\{p_t^x, 1 - p_t^x\} \leq \mathrm{e}^{-(r-\varepsilon)t} \text{ for all } t \text{ large enough} \right\}$$

occurs almost surely, and hence

$$\liminf_{t \to \infty} -\frac{1}{t} \log \frac{\min\{p_t^x, 1 - p_t^x\}}{1 - \min\{p_t^x, 1 - p_t^x\}} \geq r \quad \text{almost surely.} \tag{7}$$

Given that $S_t^i = \log \frac{p_t^x}{1 - p_t^x}$, it follows from (6) and (7) that conditioned on $\Theta = \mathfrak{b}$,

$$\limsup_{t \to \infty} \frac{1}{t} S_t^i \leq -r \quad \text{almost surely.}$$

An analogous argument applied to $\Theta = \mathfrak{g}$ shows the second claim. $\square$

Proof of Lemma 4. Recall that at time t, H_t^i is the history of actions observed by i and $s_1^i, \ldots, s_t^i$ is the sequence of private signals received by i. Given a *pure* strategy profile σ, agent i chooses a unique action $a_t^i \in \mathcal{A}$ at time t: $a_t^i = \sigma(s_1^i, \cdots, s_t^i, H_t^i)$. It follows that for each history H_t^i there is a set $\Omega_{H_t^i}^i \subseteq \Omega_1^i \times \cdots \times \Omega_{t-1}^i$ of possible private signal realizations $s_1^i, \ldots, s_{t-1}^i$ that are consistent with H_t^i. In other words, if we imagine an outside observer who sees only H_t^i—i.e. sees i's actions and his neighbors' actions—then $\Omega_{H_t^i}^i$ is the set of private signal realizations $(s_1^i, \ldots, s_{t-1}^i)$ to which this observer assigns positive probability.

Consider the numerator $\mathbb{P}[s_1^i, \ldots, s_t^i | H_t^i, \Theta = \mathfrak{g}]$ of P_t^i. Since for each history of actions H_t^i, there is a set of $\Omega_{H_t^i}^i$ of possible signals $s_1^i, \cdots, s_{t-1}^i$ that are consistent with H_t^i, it follows that

$$\mathbb{P}[s_1^i, \ldots, s_t^i | H_t^i, \Theta = \mathfrak{g}] = \mathbb{P}[s_1^i, \ldots, s_t^i | (s_1^i, \ldots, s_{t-1}^i) \in \Omega_{H_t^i}^i, \Theta = \mathfrak{g}] \quad \text{almost surely.}$$

Since the signals are independent conditioned on the state, we can write

$$\mathbb{P}[s_1^i, \dots, s_t^i | (s_1^i, \dots, s_{t-1}^i) \in \Omega_{H_t^i}^i, \Theta = \mathfrak{g}] = \frac{\mu_{\mathfrak{g}}^{i,t}(s_1^i, \dots, s_t^i)}{\mu_{\mathfrak{g}}^{i,t-1}(\Omega_{H_t^i}^i)},$$

where $\mu_{\mathfrak{g}}^{i,t} \in \prod_{s \le t} \Delta(\Omega_s^i)$ denotes the product of $\mu_{s,\mathfrak{g}}^i$ for $s \le t$. We thus have that

$$P_t^i = \log \frac{\mu_{\mathfrak{g}}^{i,t}(s_1^i, \dots, s_t^i)}{\mu_{\mathfrak{b}}^{i,t}(s_1^i, \dots, s_t^i)} + \log \frac{\mu_{\mathfrak{b}}^{i,t-1}(\Omega_{H_t^i}^i)}{\mu_{\mathfrak{g}}^{i,t-1}(\Omega_{H_t^i}^i)} \quad \text{almost surely.} \tag{8}$$

Then the first term of (8), which is equal to

$$\sum_{n=1}^{t} \log \frac{\mu_{n,\mathfrak{g}}^i(s_n^i)}{\mu_{n,\mathfrak{b}}^i(s_n^i)},$$

is at most $\frac{1}{2}Mt$. The second term of (8), which is equal to

$$\log \frac{\sum_{(s_1^i, \cdots, s_{t-1}^i) \in \Omega_{H_t^i}^i} \prod_{n=1}^{t-1} \mu_{n,\mathfrak{b}}^i(s_n^i)}{\sum_{(s_1^i, \cdots, s_{t-1}^i) \in \Omega_{H_t^i}^i} \prod_{n=1}^{t-1} \mu_{n,\mathfrak{g}}^i(s_n^i)}$$

is also at most $\frac{1}{2}M(t-1)$. Thus, it follows that P_t^i is at most Mt almost surely. By an analogous argument P_t^i is at least $-Mt$ almost surely, and so $|P_t^i|$ is at most Mt almost surely. $\square$

Proof of Theorem 1. Fix a discount factor $\delta \in [0, 1)$. Let σ be an equilibrium and a_t^i be the action taken by i at time t under σ. Now consider an outside observer x who observes everybody's actions so that the information available to him at time t is $H_t = \{a_s^i, i \in N, s \le t\}$ and at time infinity is $H_\infty = \cup_t H_t$. Thus at any time t, this observer can calculate the social likelihood S_t^i for all i.

Suppose that the social likelihood is high, and in particular $S_t^i > Mt - \log(1-\delta)$ at some t. Since the private likelihood P_t^i cannot be less than $-Mt$ (Lemma 4), the posterior likelihood $L_t^i = S_t^i + P_t^i > -\log(1-\delta)$. In this case, by part (i) of Lemma 1, in equilibrium, the agent will choose the myopic action $a_t^i = m_t^i = \mathfrak{g}$. Likewise, if $S_t^i < -(Mt - \log(1-\delta))$, in equilibrium, the agent will choose $a_t^i = m_t^i = \mathfrak{b}$. Thus, when $|S_t^i| > Mt - \log(1-\delta)$ the outside observer will know which action the agent will choose in equilibrium, and will not learn anything (in particular, about the agent's signals or the state) from observing this action.

Suppose towards a contradiction that the equilibrium speed of learning r is strictly higher than M, i.e. $r = M + 2\varepsilon$ for some $\varepsilon > 0$. Then it follows from Lemma 3 that $|S_t^i| \ge (M+\varepsilon)t > Mt - \log(1-\delta)$ for all t large enough almost surely. Since there are finitely many agents, this will hold for all i, for all t larger than some (random) T. Since the outside observer x observes H_∞ at time infinity, he knows when T has occurred and that from T onward, what all agents will do based on what he has observed. Thus x learns nothing more from the agents' actions.

Let a_t^x be the action that x would choose to maximize the probability of matching the state at time t. Since no new information is gained after time T, $a_T^x = a_\infty^x$. Hence $p_\infty := \mathbb{P}[a_\infty^x \neq \Theta] > 0$.

Since x observes everyone's actions, by the imitation principle

$$\mathbb{P}[a_t^i \neq \Theta] \geq \mathbb{P}[a_\infty^x \neq \Theta] = p_\infty \tag{9}$$

for all agents i and all times t. But since the equilibrium speed of learning is $M + 2\varepsilon$, by definition, $\mathbb{P}[a_t^i \neq \Theta] \leq \mathrm{e}^{-(M+\varepsilon)t} < p_\infty$ for all t large enough, in contradiction with (9).

□

Proof of Proposition 1. Let $L = \{i \colon r_i = \min_j r_j\}$ be the set of agents who learn with the slowest rate. These agents can not observe any agent outside of L, i.e. $\cup_{i\in L} N_i \subseteq L$ as otherwise, by Lemma 2, they could achieve a higher rate by imitating that agent. As a consequence the rate at which agents in the set L learn is independent of agents outside of L and we henceforth ignore them. Now take a minimal subset E of L such that no agent in E observes any agent in $L \setminus E$, i.e. $\cup_{i\in E} N_i \subseteq E$. Such as set exists as we have established that $\cup_{i\in L} N_i \subseteq L$ and L is finite. Note, that every strict subset of E' of E (apart from the empty set) observes at least one agent not in E', i.e. $\cup_{i\in E'} N_i \not\subseteq E'$ as otherwise E is not minimal. Consequently, E is strongly connected.[14] Again the learning rate of agents in E remains unchanged when ignoring the agents in $L \setminus E$ and hence by Theorem 1 the rate of learning for agents in E is less than M. □

References

Acemoglu, D., M. A. Dahleh, I. Lobel, and A. Ozdaglar (2011): "Bayesian learning in social networks," *The Review of Economic Studies*, 78, 1201–1236.

Bala, V. and S. Goyal (1998): "Learning from neighbours," *The Review of Economic Studies*, 65, 595–621.

Banerjee, A. V. (1992): "A simple model of herd behavior," *The Quarterly Journal of Economics*, 107, 797–817.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of Political Economy*, 992–1026.

Cover, T. M. and J. A. Thomas (2006): *Elements of information theory*, John Wiley & Sons.

Dasaratha, K., B. Golub, and N. Hak (2020): "Learning from neighbors about a changing state," *Available at SSRN 3097505.*

DeGroot, M. H. (1974): "Reaching a consensus," *Journal of the American Statistical Association*, 69, 118–121.

[14]Indeed, E can be equivalently defined to be any strongly connected component in which no agent observes agents not in the component.

Frongillo, R. M., G. Schoenebeck, and O. Tamuz (2011): "Social learning in a changing world," in *International Workshop on Internet and Network Economics*, Springer, 146–157.

Gale, D. and S. Kariv (2003): "Bayesian learning in social networks," *Games and economic behavior*, 45, 329–346.

Golub, B. and M. O. Jackson (2010): "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, 2, 112–49.

Golub, B. and E. Sadler (2017): "Learning in social networks," *Available at SSRN 2919146*.

Hann-Caruthers, W., V. V. Martynov, and O. Tamuz (2018): "The speed of sequential asymptotic learning," *Journal of Economic Theory*, 173, 383–409.

Harel, M., E. Mossel, P. Strack, and O. Tamuz (2021): "Rational groupthink," *The Quarterly Journal of Economics*, 136, 621–668.

Leskovec, J. and E. Horvitz (2008): "Planetary-scale views on a large instant-messaging network," in *Proceedings of the 17th international conference on World Wide Web*, 915–924.

Molavi, P., A. Tahbaz-Salehi, and A. Jadbabaie (2018): "A theory of non-Bayesian social learning," *Econometrica*, 86, 445–490.

Moscarini, G., M. Ottaviani, and L. Smith (1998): "Social learning in a changing world," *Economic Theory*, 11, 657–665.

Mossel, E., A. Sly, and O. Tamuz (2014): "Asymptotic learning on bayesian social networks," *Probability Theory and Related Fields*, 158, 127–157.

——— (2015): "Strategic learning and the topology of social networks," *Econometrica*, 83, 1755–1794.

Rosenberg, D. and N. Vieille (2019): "On the efficiency of social learning," *Econometrica*, 87, 2141–2168.

Vasal, D. and A. Anastasopoulos (2016): "Decentralized Bayesian learning in dynamic games," in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 264–273.

Vives, X. (1993): "How fast do rational agents learn?" *The Review of Economic Studies*, 60, 329–347.

Watts, D. J. and S. H. Strogatz (1998): "Collective dynamics of 'small-world'networks," *nature*, 393, 440–442.