2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Lower Bounds on the Rate of Learning in Social Networks

Ilan Lobel, Daron Acemoglu, Munther Dahleh and Asuman Ozdaglar
Massachusetts Institute of Technology
Cambridge, MA

Abstract

We study the rate of convergence of Bayesian learning in social networks. Each individual receives a signal about the underlying state of the world, observes a subset of past actions and chooses one of two possible actions. Our previous work [1] established that when signals generate unbounded likelihood ratios, there will be asymptotic learning under mild conditions on the social network topology—in the sense that beliefs and decisions converge (in probability) to the correct beliefs and action. The question of the speed of learning has not been investigated, however. In this paper, we provide estimates of the speed of learning (the rate at which the probability of the incorrect action converges to zero). We focus on a special class of topologies in which individuals observe either a random action from the past or the most recent action. We show that convergence to the correct action is faster than a polynomial rate when individuals observe the most recent action and is at a logarithmic rate when they sample a random action from the past. This suggests that communication in social networks that lead to repeated sampling of the same individuals lead to slower aggregation of information.

I. Introduction

In this paper, we study the speed of Bayesian learning in social networks. We start with the canonical sequential learning problem, where each member of a social group (network) takes one of two actions. Which action yields a higher payoff depends on an underlying unknown state of nature. Each individual receives a private signal correlated with the underlying state and observes some of the actions taken in the past. On the basis of this information, he makes a decision. The previous literature has investigated conditions under which there will be *asymptotic learning* in this environment. We say that there is asymptotic learning if individual beliefs and actions converge (in probability) to the correct beliefs and actions. Our previous work, Acemoglu, Dahleh, Lobel and Ozdaglar [1], characterized the general conditions under which such asymptotic learning takes place.[1] In particular, we showed that when signals generate unbounded likelihood ratios (meaning that there is no absolute bound on the informativeness of possible signals) and the (possibly stochastic) network topology determining which subset of past actions each individual observes satisfies some relatively weak regularity conditions, all perfect Bayesian equilibria of this sequential game generate asymptotic learning.

The rate at which learning happens is of considerable interest, since convergence to approximately correct beliefs and actions might take an arbitrarily long time (or require an arbitrarily large network). In particular, the results in [1] guarantee that a very minimal amount of communication is necessary between agents to obtain asymptotic learning (when signals have unbounded likelihood). Therefore, analyzing the speed of learning is important for understanding how network topology affects learning in social networks. Nevertheless, the question of the speed of learning has not be investigated in the social learning literature. In this paper, we investigate the rate of convergence of individual beliefs and actions to the correct beliefs and actions in a simplified version of our general setup from [1]. In particular, instead of general network topologies, we focus on situations in which individuals observe only one past action. We distinguish two cases. In the first, which we refer to as *random sampling*, each individual observes any one of the past actions with equal probability. In the second, which referred to as *immediate neighbor sampling*, each individual observes the most recent action.

We develop a new method of estimating a lower bound on the rate of convergence in both cases, based on approximating this lower bound with an ordinary differential equation. Our main results are as follows:

- With random sampling, the probability of the incorrect action converges to zero faster than a logarithmic rate [in the sense that this probability is no greater than

Ilan Lobel (lobel@mit.edu) is with the Operations Research Center at the Massachusetts Institute of Technology.

Daron Acemoglu (daron@mit.edu) is with the Department of Economics at the Massachusetts Institute of Technology.

Munther Dahleh (dahleh@mit.edu) and Asuman Ozdaglar (asuman@mit.edu) are with Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology.

[1] See Bikchandani, Hirshleifer and Welch [5], Banerjee [3], and Smith and Sorensen [12] for analyses of learning behavior of this model when individuals observe all past actions.

$(\log(n))^{-1/(K+1)}$ where n is the number of individuals in the network and K is a constant]. By means of an example, we show that this probability does *not* go to zero at a polynomial rate.

- In contrast, with immediate neighbor sampling, the probability of the incorrect action converges to zero faster than a polynomial rate [in the sense that this probability is no greater than $n^{-1/(K+1)}$].

These results are intuitive. The immediate neighbor sampling enables faster aggregation of information, because each individual is sampled only once. In contrast, with random sampling, each individual is sampled *infinitely often*. Thus, with a random sampling, there is slower rate of arrival of new information into the system. While real-world social networks are much more complex than the two stylized topologies we study, our results already suggest some general insights. In particular, they suggest that social networks in which the same individuals are the (main or only) source of the information of many others will lead to slower learning than networks in which the opinions and information of new members are incorporated into the "social belief" more rapidly.

In addition to our previous work and to the papers on Bayesian social learning mentioned above, our paper is related to Tay, Tsitsiklis and Win [13], who study the problem of information aggregation over sensor networks. They focus on threshold rules for learning (different from the perfect Bayesian equilibrium characterized here) and provide a lower bound on the rate of convergence of posteriors generated by decentralized agents. The formal model is similar to our model of social learning with immediate neighbor sampling, though our bounds on the rate of convergence with immediate neighbors sampling are not present in this paper. In addition, to the best of our knowledge, our results on the speed of learning in the random sampling case are also not present in any previous paper.

The next section introduces our model, and Section III provides the main results. Section IV illustrates our main results using two examples and Section V concludes.

II. MODEL

In this section, we introduce our model of social learning. The environment is a special case of that presented in our previous work, Acemoglu, Dahleh, Lobel and Ozdaglar [1].

A countably infinite number of agents, indexed by $n \in \mathbb{N}$, make decisions sequentially. The payoff of agent n depends on an underlying state of the world θ and his decision. To simplify the notation and the exposition, we assume that both the underlying state and decisions are binary. In particular, the decision of agent n is denoted by $x_n \in \{0,1\}$ and the underlying state is $\theta \in \{0,1\}$. The payoff of agent n is

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

Again to simplify notation, we assume that both values of the underlying state are equally likely, so that $P(\theta=0) = P(\theta=1) = 1/2$.

Each agent $n \in \mathbb{N}$ forms beliefs about the state θ using a *private signal*, $s_n \in S$ (where S is a metric space or simply a Euclidean space) and his observation of the actions of other agents. Conditional on the state of the world θ, the signals are independently generated according to a probability measure F_θ. We refer to the pair of measures (F_0,F_1) as the *signal structure* of the model. We assume that F_0 and F_1 are *absolutely continuous* with respect to each other, which immediately implies that no signal is fully revealing about the underlying state. We also assume that F_0 and F_1 are not identical, so that some signals are *informative*.

Each agent n observes the actions of a subset of agents, denoted by $B(n) \subseteq \{1,\dots,n-1\}$, according to the social network. In this paper, we study two particular network structures. First, we analyze the immediate neighbor sampling network, where $B(n) = \{n-1\}$. That is, each agent observes the action of the agent that acted most recently. The second topology we study is the random sampling network, where for each agent n and each $b \in \{1,...,n-1\}$, $B(n) = \{b\}$ with probability $1/(n-1)$. In this topology, each agent samples uniformly the action of an agent from the past. We assume that agent n knows the identity of the agent he is observing, that is, he knows $B(n)$, but no other agent does. For a detailed discussion of these informational assumptions, please refer to [1].

The set of all possible information sets of agent n is denoted by I_n. A strategy for individual n is a mapping $\sigma_n : I_n \to \{0,1\}$ that selects a decision for each possible information set. A strategy profile is a sequence of strategies $\sigma = \{\sigma_n\}_{n\in\mathbb{N}}$. We use the standard notation $\sigma_{-n} = \{\sigma_1,\dots,\sigma_{n-1},\sigma_{n+1},\dots\}$ to denote the strategies of all agents other than n and also (σ_n,σ_{-n}) for any n to denote the strategy profile σ.

Definition 1: A strategy profile σ^* is a pure-strategy *Perfect Bayesian Equilibrium* of this game of social learning if for each $n \in \mathbb{N}$, σ_n^* maximizes the expected payoff of agent n given the strategies of other agents σ_{-n}^*.

A pure-strategy Perfect Bayesian Equilibrium always exists in this game. Given a strategy profile σ, the expected payoff of agent n from action $x_n = \sigma_n(I_n)$ is simply $P(x_n = \theta | I_n)$. Therefore, for any equilibrium σ^*, we have

$$\sigma_n^*(I_n) \in \arg\max_{y\in\{0,1\}} P(y = \theta \mid I_n).$$

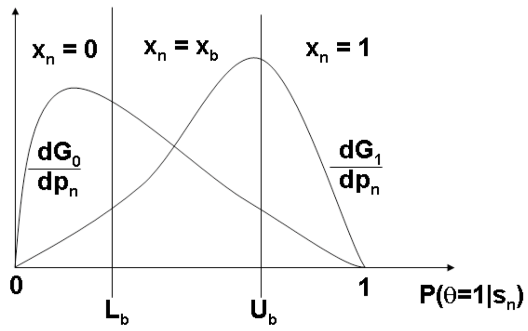

Fig. 1. Equilibrium decision rule.

We say that *asymptotic learning occurs in a Perfect Bayesian Equilibrium* if x_n converges to θ in probability. In this paper, our goal is to characterize the speed of learning for the special classes of network topologies considered here.

The key property of the signal structure that determines learning in social networks is the set of possible beliefs about the state that it induces. We call $p_n(s_n) = P(\theta = 1|s_n)$ the private belief of agent n and it is equal to

$$p_n(s_n) = P(\theta = 1|s_n) = \left(1 + \frac{dF_0}{dF_1}(s_n)\right)^{-1}, \quad (1)$$

where $\frac{dF_0}{dF_1}$ is the Radon-Nikodym derivative of F_0 by F_1, which is finite and positive everywhere since F_0 and F_1 are absolutely continuous with respect to each other.

The decision rule that agents use in equilibrium is given by the following proposition from [1].

Proposition 1: Let $B(n) = \{b\}$ for some agent n. In equilibrium, there exist deterministic L_b and U_b such that

$$x_n = \begin{cases} 0, & \text{if } p_n(s_n) < L_b \\ x_b, & \text{if } p_n(s_n) \in (L_b, U_b) \\ 1, & \text{if } p_n(s_n) > U_b. \end{cases}$$

The thresholds L_b and U_b are indexed by b because they are functions of the probability that agent b's action is correct in each possible state of the world, i.e., $P(x_b = \theta|\theta = 0)$ and $P(x_b = \theta|\theta = 1)$.

The state-conditional distribution of private belief is represented by G_j, for each $j \in \{0,1\}$, i.e.,

$$G_j(r) = P(p_1(s_1) \leq r|\theta = j). \quad (2)$$

In this paper, we assume G_0 and G_1 are continuous distributions in order to guarantee that there is a unique Perfect Bayesian equilibrium. This assumption could be relaxed at the expense of more notation.

The following lemma from [1] establishes a lower bound on the amount of improvement in the ex-ante probability of making the correct decision between an agent and his neighbor. This bound will be key in the subsequent convergence rate analysis.

Lemma 1: Let $B(n) = \{b\}$ for some n and denote $\alpha = P(x_b = \theta)$. In equilibrium,

$$P(x_n = \theta|B(n) = \{b\}) \geq \alpha + \frac{1}{8}(1-\alpha)^2 \min\left\{G_1\left(\frac{1-\alpha}{2}\right), 1 - G_0\left(\frac{1+\alpha}{2}\right)\right\}.$$

III. Rate of Convergence

In this section, we prove the results about the speed of learning in social networks. We first introduce an important property of the private belief distributions that will impact the rate of convergence results.

Definition 2: The private belief distributions have *polynomial shape* if there exist some constant $C' > 0$ and $K > 0$ such that

$$G_1(\alpha) \geq C'\alpha^K \text{ and } 1 - G_0(1-\alpha) \leq 1 - C'\alpha^K \quad (3)$$

for all $\alpha \in [1/2, 1]$. If Eq. (3) holds only for $\alpha \in [1-\varepsilon, 1]$ for some $\varepsilon > 0$, then the private belief distributions have *polynomial tails*.

For simplicity, the results in this paper assume polynomial shape, but they extend immediately to private belief distributions with polynomial tails. Note that polynomial shape implies there exist some constants $C > 0$ and $K > 0$ such that

$$\min\left\{G_1\left(\frac{1-\alpha}{2}\right), 1 - G_0\left(\frac{1+\alpha}{2}\right)\right\} \geq C(1-\alpha)^K$$

for all $\alpha \in [1/2, 1]$.

When the private beliefs have polynomial shape, we obtain from Lemma 1 that for some $C > 0$ and $K > 0$,

$$P(x_n = \theta|B(n) = \{b\}) \geq \alpha + \frac{C}{8}(1-\alpha)^{K+2}, \quad (4)$$

where $\alpha = P(x_b = \theta)$. Note that the right-hand side of Eq. (4) is increasing in α over $[1/2, 1]$ if

$$C < \frac{2^{K+1}}{K+2}. \quad (5)$$

Note that if Eq. (4) holds for some $C > 0$, then it also holds for any $C' \in (0, C)$. So, we can assume without loss of generality that Eq. (5) holds and, therefore, the right-hand side of Eq. (4) is increasing in α.

Proposition 2: Suppose agents sample immediate neighbors and the private beliefs have polynomial shape, then

$$P(x_n \neq \theta) = O\left(n^{\frac{-1}{K+1}}\right).$$

Proof: To prove this bound, we construct a pair of functions ϕ and $\tilde{\phi}$, where ϕ is a difference equation and $\tilde{\phi}$ is a differential equation and show that for all $t \in \mathbb{N}$, $P(x_n = \theta) \geq \phi(t) \geq \tilde{\phi}(t)$.

When the conditions of the proposition hold, we have from Eq. (4) that for all $n \in \mathbb{N}$,

$$P(x_{n+1} = \theta) \geq P(x_n = \theta) + \frac{C}{8}\left(1 - P(x_n = \theta)\right)^{K+2}. \quad (6)$$

Construct a sequence $\{\phi(n)\}_{n\in\mathbb{N}}$ where $\phi(1)=P(x_1=\theta)$ and, recursively, for all $n\in\mathbb{N}$,

$$\phi(n+1)=\phi(n)+\frac{C}{8}(1-\phi(n))^{K+2}.$$

Now, we show by induction that for all $n\in N$,

$$P(x_n=\theta)\geq\phi(n). \tag{7}$$

The relation holds for $n=1$ by construction. Suppose the relation holds for some n. Then,

$$\begin{aligned} P(x_{n+1}=\theta) &\geq P(x_n=\theta)+\frac{C}{8}(1-P(x_n=\theta))^{K+2} \\ &\geq \phi(n)+\frac{C}{8}(1-\phi(n))^{K+2} \\ &= \phi(n+1), \end{aligned}$$

where the second inequality holds because we assumed that Eq. (4) is increasing in α. Therefore, Eq. (7) holds for all n.

Let us now define $\phi(t)$ on non-integer values of t by linear interpolation of the integer values. That is, for any $t\in[1,\infty)$,

$$\phi(t)=\phi(\lfloor t\rfloor)+\frac{C}{8}(t-\lfloor t\rfloor)(1-\phi(\lfloor t\rfloor))^{K+2}.$$

Finding such a ϕ is equivalent to solving the following differential equation (where derivatives are defined only at non-integer times):

$$\frac{d\phi(t)}{dt}=\frac{C}{8}(1-\phi(\lfloor t\rfloor))^{K+2}.$$

Let us now construct yet another continuous time function $\tilde{\phi}(t)$ to bound $\phi(t)$. Let $\tilde{\phi}(1)=\phi(1)$ and

$$\frac{d\tilde{\phi}(t)}{dt}=\frac{C}{8}\left(1-\tilde{\phi}(t)\right)^{K+2}. \tag{8}$$

Note that both ϕ and $\tilde{\phi}$ are increasing functions. We now show that

$$\tilde{\phi}(t)\leq\phi(t) \text{ for all } t\in[1,\infty). \tag{9}$$

Let t^* be some value such that $\tilde{\phi}(t^*)=\phi(t^*)$. Then, for any $\varepsilon\in[0,1-(t^*-\lfloor t^*\rfloor)]$,

$$\begin{aligned} \phi(t^*+\varepsilon) &= \phi(t^*)+\varepsilon\frac{C}{8}(1-\phi(\lfloor t^*\rfloor))^{K+2} \\ &\geq \phi(t^*)+\varepsilon\frac{C}{8}(1-\phi(t^*))^{K+2} \\ &= \tilde{\phi}(t^*)+\varepsilon\frac{C}{8}\left(1-\tilde{\phi}(t^*)\right)^{K+2} \\ &\geq \tilde{\phi}(t^*)+\int_{t^*}^{t^*+\varepsilon}\frac{C}{8}\left(1-\tilde{\phi}(t)\right)^{K+2}dt \\ &= \tilde{\phi}(t^*+\varepsilon), \end{aligned}$$

where the first equality comes from ϕ's piecewise linearity, the following inequality from the monotonicity of ϕ, the second equality from $\tilde{\phi}(t^*)=\phi(t^*)$ and the second inequality from the monotonicity of $\tilde{\phi}$. Therefore, for any t^* such that $\tilde{\phi}(t^*)=\phi(t^*)$, we have that for all $\varepsilon\in[0,1-(t^*-\lfloor t^*\rfloor)]$, $\tilde{\phi}(t^*+\varepsilon)\leq\phi(t^*+\varepsilon)$. Hence, $\tilde{\phi}$ does not cross ϕ at any point and, since both functions are continuous, it implies Eq. (9) holds.

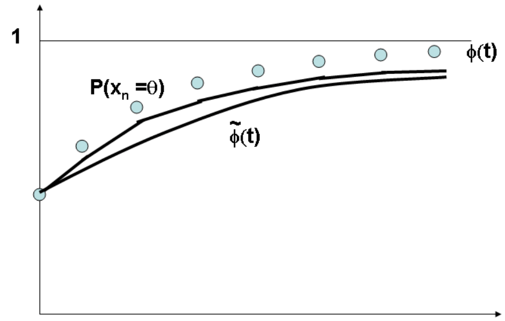

Fig. 2. Bound used to prove Propositions 1 and 2.

Combining Eqs. (7) and (9) we obtain that for any $n\in\mathbb{N}$,

$$P(x_n=\theta)\geq\phi(n)\geq\tilde{\phi}(n).$$

See Figure 2 for graphical description of this bound. We can solve the ODE of Eq. (8) exactly and calculate $\tilde{\phi}(n)$ for every n. The solution is that there exists some other constant $\overline{C}$ (which is determined by the boundary value $P(x_1=\theta)=\tilde{\phi}(1)$) such that for each n,

$$\tilde{\phi}(n)=1-\left(\frac{1}{8(K+1)C(n+\overline{C})}\right)^{\frac{1}{K+1}}.$$

Therefore,

$$P(x_n\neq\theta)\leq\left(\frac{1}{8(K+1)C(n+\overline{C})}\right)^{\frac{1}{K+1}}$$

and the desired result follows. ■

Proposition 3: Suppose that agents use random sampling and the private beliefs have polynomial shape, then

$$P(x_n\neq\theta)=O\left((\log n)^{\frac{-1}{K+1}}\right).$$

Proof: To prove this result, we also construct a pair of functions ϕ and $\tilde{\phi}$ and show that for all $t\in\mathbb{N}$, $P(x_n=\theta)\geq\phi(t)\geq\tilde{\phi}(t)$.

When the conditions of the proposition hold, we have for all $n\in\mathbb{N}$,

$$\begin{aligned} P(x_{n+1}=\theta)&=\frac{1}{n}\sum_{b=1}^{n}P(x_{n+1}=\theta|B(n+1)=\{b\}) \\ &=\frac{1}{n}\left[P(x_{n+1}=\theta|B(n+1)=\{n\})+(n-1)P(x_n=\theta)\right], \end{aligned}$$

because conditional on observing the same $b<n$, agents n and $n+1$ have identical probabilities of making an optimal decision. From Eq. (4), we obtain that

$$P(x_{n+1}=\theta)\geq P(x_n=\theta)+\frac{C}{8n}(1-P(x_n=\theta))^{K+2}. \tag{10}$$

Note that the right-hand side of Eq. (10) is increasing in $P(x_n = \theta)$ for any $n \geq 1$ because we assumed Eq. (4) is increasing in α.

Let's recursively construct a sequence $\phi(n)$ with $\phi(1) = P(x_1 = \theta)$ and, for all $n \in \mathbb{N}$,

$$\phi(n+1) = \phi(n) + \frac{C}{8n}(1-\phi(n))^{K+2}.$$

Now, we show by induction that

$$P(x_n = \theta) \geq \phi(n). \tag{11}$$

The relation holds for $n = 1$ by construction. Suppose the relation holds for some n. Then,

$$\begin{aligned} P(x_{n+1} = \theta) &\geq P(x_n = \theta) + \frac{C}{8n}(1 - P(x_n = \theta))^{K+2} \\ &\geq \phi(n) + \frac{C}{8n}(1-\phi(n))^{K+2} \\ &= \phi(n+1). \end{aligned}$$

where the second inequality holds because the right-hand side of Eq. (10) is increasing in $P(x_n = \theta)$. Therefore, Eq. (11) holds for all n.

Let us now define $\phi(t)$ on non-integer values of t by linear interpolation of the integer values. That is, for any $t \in [1,\infty)$,

$$\phi(t) = \phi(\lfloor t \rfloor) + \frac{C}{8\lfloor t \rfloor}(t - \lfloor t \rfloor)(1-\phi(\lfloor t \rfloor))^{K+2}.$$

Finding such a ϕ is equivalent to solving the following differential equation (where derivatives are defined only at non-integer times):

$$\frac{d\phi(t)}{dt} = \frac{C}{8\lfloor t \rfloor}(1-\phi(\lfloor t \rfloor))^{K+2}.$$

Let us now construct yet another continuous time function $\tilde{\phi}(t)$ to bound $\phi(t)$. Let $\tilde{\phi}(1) = \phi(1)$ and

$$\frac{d\tilde{\phi}(t)}{dt} = \frac{C}{8t}\left(1-\tilde{\phi}(t)\right)^{K+2}. \tag{12}$$

Note that both ϕ and $\tilde{\phi}$ are increasing functions. We now show that

$$\tilde{\phi}(t) \leq \phi(t) \text{ for all } t \in [1,\infty). \tag{13}$$

Let t^* be some value such that $\tilde{\phi}(t^*) = \phi(t^*)$. Then, for any $\varepsilon \in [0, 1-(t^* - \lfloor t^* \rfloor)]$,

$$\begin{aligned} \phi(t^*+\varepsilon) &= \phi(t^*) + \varepsilon\frac{C}{8\lfloor t^* \rfloor}(1-\phi(\lfloor t^* \rfloor))^{K+2} \\ &\geq \phi(t^*) + \varepsilon\frac{C}{8t^*}(1-\phi(t^*))^{K+2} \\ &= \tilde{\phi}(t^*) + \varepsilon\frac{C}{8t^*}\left(1-\tilde{\phi}(t^*)\right)^{K+2} \\ &\geq \tilde{\phi}(t^*) + \int_{t^*}^{t^*+\varepsilon}\frac{C}{8t}\left(1-\tilde{\phi}(t)\right)^{K+2}dt \\ &= \tilde{\phi}(t^*+\varepsilon), \end{aligned}$$

where the first equality comes from ϕ's piecewise linearity, the following inequality from the monotonicity of ϕ, the second equality from $\tilde{\phi}(t^*) = \phi(t^*)$ and the second inequality from the monotonicity of $\tilde{\phi}$. Therefore, for any t^* such that $\tilde{\phi}(t^*) = \phi(t^*)$, we have that for all $\varepsilon \in [0, 1-(t^* - \lfloor t^* \rfloor)]$, $\tilde{\phi}(t^*+\varepsilon) \leq \phi(t^*+\varepsilon)$. Hence, $\tilde{\phi}$ does not cross ϕ at any point and, since both functions are continuous, it implies Eq. (13) holds.

Combining Eqs. (11) and (13) we obtain that for any $n \in \mathbb{N}$,

$$P(x_n = \theta) \geq \phi(n) \geq \tilde{\phi}(n).$$

We can solve the ODE of Eq. (12) exactly and calculate $\tilde{\phi}(n)$ for every n. The solution is that there exists some other constant $\overline{C}$ (which is determined by the boundary value $P(x_1 = \theta) = \tilde{\phi}(1)$) such that for each n,

$$\tilde{\phi}(n) = 1 - \left(\frac{1}{8(K+1)C(\log n + \overline{C})}\right)^{\frac{1}{K+1}}.$$

Therefore,

$$P(x_n \neq \theta) \leq \left(\frac{1}{8(K+1)C(\log n + \overline{C})}\right)^{\frac{1}{K+1}}$$

and the desired result follows. ■

IV. An Example

This section presents an example that demonstrates a signal structure for which learning with random sampling occurs at a logarithmic rate, hence establishing that the bound of Proposition 3 is tight.

Example 1. For each n, let the signal $s_n \in [0,1]$ be generated by the following signal structure:

$$F_0(s_n) = 2s_n - s_n^2 \text{ and } F_1(s_n) = s_n^2.$$

For this example, the signals are identical to the private beliefs, i.e., $p_n(s_n) = s_n$ [cf. Eq. (1)]. Therefore, the state-conditional distributions of private signals [cf. Eq. (2)] are

$$G_0(r) = 2r - r^2 \text{ and } G_1(r) = r^2.$$

Therefore, this signal structure has polynomial shape. We now show that under this signal structure, for any n,

$$P(x_{n+1} = \theta | B(n+1) = \{n\}) = P(x_n = \theta) + (1 - P(x_n = \theta))^2. \tag{14}$$

From Lemma 4 of [1], we obtain a recursion on $P(x_n = \theta)$. Since the signal structure is symmetric, we obtain that all the terms used in Lemma 4 of [1] are the same, i.e.,

$$P(x_n = \theta) = N_n = Y_n = U_n = 1 - L_n.$$

To simplify the notation, denote $P(x_n = \theta) = z$. The recursion can be written as

$$P(x_{n+1} = \theta | B(n+1) = \{n\}) = \frac{1}{2}[G_0(1-z) +$$

$$(G_1(z) - G_0(1-z))z + (1 - G_1(z)) + (G_1(z) - G_1(1-z))z],$$

which reduces to

$$P(x_{n+1}=\theta|B(n+1)=\{n\})=z+(1-z)^2,$$

and, thus proves Eq. (14). Using the same ODE bounding methods from the two propositions, the probability $P(x_n = \theta)$ can be approximated by $w(t)$ in the immediate neighbor sampling case $(B(n)=\{n-1\})$ and

$$\frac{dw(t)}{dt}=(1-w(t))^2.$$

Solving this ODE yields that for some constant $\overline{C}$,

$$w(t)=1-\frac{1}{t+\overline{C}}.$$

Thus, $1-w(t)=\Theta\left(t^{-1}\right)$, which implies that

$$P(x_n \neq \theta)=\Theta\left(n^{-1}\right).$$

For the random sampling case, the ODE that bounds the probability of optimal decision is

$$\frac{dw(t)}{dt}=\frac{(1-w(t))^2}{t}.$$

Solving this ODE we obtain that for some constant $\overline{C}$,

$$w(t)=1-\frac{1}{\log(t)+\overline{C}}$$

and thus obtain

$$P(x_n \neq \theta)=\Theta\left((\log n)^{-1}\right).$$

Therefore, we conclude that learning is significantly slower with random sampling than with immediate neighbor sampling, at least for a class of signal structures.

V. Conclusion

In this paper we studied the rate of convergence of Bayesian learning in a class of simple social networks. The environment is as follows: each individual receives a signal about the underlying state of the world, observes a subset of past actions and chooses one of two possible actions.

This is a special case of the general environment studied in previous work [1]. Although previous work has characterized the conditions on signals and network topologies that ensure asymptotic learning, the question of the speed of learning has not been investigated. In this paper, we focus on a special class of topologies in which individuals observe either a random action from the past or the most recent action and provide estimates of the speed of learning.

Our main results show that convergence to the correct action is faster than a polynomial rate when individuals observe the most recent action and is at a logarithmic rate when they sample a random action from the past. These results suggest that communication in social networks that lead to repeated sampling of the same individuals lead to slower aggregation of information. As a byproduct, we develop a new method of determining lower bounds on the speed of learning. The insights from the special networks studied in this paper can be generalized to more realistic network topologies. The analysis of speed of Bayesian learning in more general networks is part of our ongoing work.

References

[1] Acemoglu D., Dahleh M., Lobel I., Ozdaglar A., "Bayesian Learning in Social Networks,", MIT LIDS Working Paper 2780, 2008.

[2] Bala V. and Goyal S., "Learning from Neighbours," *The Review of Economic Studies*, vol. 65, no. 3, pp. 595-621, 1998.

[3] Banerjee A., "A Simple Model of Herd Behavior," *Quarterly Journal of Economics,* vol. 107, pp. 797-817, 1992.

[4] Banerjee A. and Fudenberg D., "Word-of-mouth Learning," *Games and Economic Behavior*, vol. 46, pp. 1-22, 2004.

[5] Bikchandani S., Hirshleifer D., and Welch I., "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades," *Journal of Political Economy*, vol. 100, pp. 992-1026, 1992.

[6] Cover T., "Hypothesis Testing with Finite Statistics," *The Annals of Mathematical Statistics*, vol. 40, no. 3, pp. 828-835, 1969.

[7] Ellison G., Fudenberg D., "Rules of Thumb for Social Learning," *Journal of Political Economy*, vol. 101, no. 4, pp. 612-643, 1993.

[8] Ellison G. and Fudenberg, D., "Word-of-mouth communication and social learning," *The Quarterly Journal of Economics*, vol. 110, pp. 93-126, 1995.

[9] Gale D. and Kariv S., "Bayesian Learning in Social Networks," *Games and Economic Behavior*, vol. 45, no. 2, pp. 329-346, 2003.

[10] Papastavrou J.D. and Athans M., "Distributed Detection by a Large Team of Sensors in Tandem," *Proceedings of IEEE CDC*, 1990.

[11] Smith L. and Sorensen P., "Rational Social Learning with Random Sampling," unpublished manuscript, 1998.

[12] Smith L. and Sorensen P., "Pathological Outcomes of Observational Learning," *Econometrica*, vol. 68, no. 2, pp. 371-398, 2000.

[13] Tay W.P., Tsitsiklis J.N. and Win M.Z., "On the Sub-Exponential Decay of Detection Error Probabilities in Long Tandems," *Proceedings of ICASSP*, 2007.