Learning Under Social Influence

Alireza Tahbaz-Salehi, Alvaro Sandroni, and Ali Jadbabaie

***Abstract*— In this paper, we study a model of social learning where individuals are under influence of others in their social clique. In our model, each agent receives private noisy signals about an unobservable, underlying state of the world. At the end of each time period, the belief of an individual is equal to the convex combination of her posterior beliefs derived from the signal observed, and the priors of her neighbors. Our model reduces to the well-known consensus model when private signals are non-informative. We show that if the network of social influences is strongly connected, then all agents will have asymptotically correct forecasts. In other words, all individuals will be able to asymptotically learn the true state of the world, as far as their observations are concerned. Finally, we show that all agents assign asymptotically equal beliefs to the true state of the world.**

I. Introduction

Individuals, for the most part, form and update their opinions based on their observations and signals. For example, a physician's personal experience with a new medical treatment might serve as a basis for her beliefs regarding its effectiveness. Such private signals, however, are not the only factors that determine the opinions, and hence, the behavior of an individual. In fact, a variety of evidence suggests that in many cases, an individual is highly influenced by the beliefs and actions of other agents in her social clique; her friends, colleagues, or family members. For instance, Hagerstrand [1] and Rogers [2] document such a phenomenon in the choice of new agricultural techniques by various farmers, while Kotler [3] shows the importance of social influences in the purchase of consumer products. Other settings where peers influence an agent's decisions include smoking and engagement in criminal behavior [4].

In this paper, we focus on understanding the evolution of beliefs in a society where agents are under the influence of others in their social clique. We base our study on a simple model according to which, at every time period, each agent receives a private noisy signal about the true state of the world and updates her beliefs as a convex combination of the Bayesian posterior belief generated by the observed signal and the prior beliefs of her neighbors. This second term is meant to capture the influence of other members of the society on the agent's beliefs. Agents in this world are boundedly rational in the sense that they fail to incorporate the information provided to them by their neighbors in a Bayesian manner. Instead, their opinion is simply swayed in the direction of the average belief in their neighborhood. Our model reduces to the model of discrete-time consensus algorithms over directed networks [5], when signals observed by agents are non-informative about the state of world.

First, we show that all agents will eventually hold correct forecasts about their signals, provided that the social network is *strongly connected*; that is, there exists either a direct or an indirect information path between any two agents. By the means of an example we show that the assumption of strong connectivity cannot be disposed of. We further show that these social interactions, not only do not prevent agents from accurately predicting their future observations, but also lead to information aggregation about the true underlying state of the world. In particular, we show that as long as the social network is strongly connected, all agents assign equal beliefs to the true state of the world. The difference between our model and consensus algorithms is that agents in our model not only reach asymptotic agreement, but also agree on the right parameter; in the sense that they will have correct predictions. The simplicity of our *local* update rule guarantees that agents avoid highly complex computations that are essential for full Bayesian learning over the network.

There exists a vast literature on learning over social networks, both boundedly and fully rational. The Bayesian learning literature mainly focuses on formulating the problem as a dynamic game over the network and characterizing its equilibria. However, since characterizing such equilibria in complex networks is generally an intractable problem, the existing literature focuses on relatively simple and stylized environments [6]. For example, Bikchandani, Hirshleifer, and Welch [7] and Banerjee [8] consider models where each individual takes a *single* action and observes all past actions. Similarly, Banerjee and Fudenberg [9] and Smith and Sørensen [10] focus on models where agents make decisions sequentially, but instead, only observe a representative sample of the past actions. In a more recent paper, Acemoglu *et al.* [11] generalize these results to an arbitrary network structure. Nevertheless, the assumption of a single decision for each agent remains in place. Another example of Bayesian learning over networks is [12], which shows how the complexity of a rational agent's decision-problem increases over time, as she has to hold beliefs about her neighbors' knowledge of their neighbors' actions and the private information they reveal.

This research is supported in parts by the following grants: DARPA/DSO SToMP, NSF ECS-0347285, ONR MURI N000140810747, and AFOSR: Complex Networks Program.

Alireza Tahbaz-Salehi is with Department of Economics and Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 (e-mail: atahbaz@seas.upenn.edu).

Alvaro Sandroni is with Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston IL 60208 (e-mail: sandroni@kellogg.northwestern.edu).

Ali Jadbabaie is with the General Robotics, Automation, Sensing and Perception (GRASP) Laboratory, Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 (e-mail: jadbabai@seas.upenn.edu).

Our work is closely related to another collection of works in the social learning literature that focus on non-Bayesian learning models; models such as [13]–[16] in which agents are not fully rational, and use simple rule-of-thumb methods to incorporate the information of their neighbors. In the same spirit are De Marzo *et al.* [17], Golub and Jackson [18] and Acemoglu, Ozdaglar, and ParandehGheibi [19], which are all based on the well-known consensus model of DeGroot [20]. These papers, unlike most of Bayesian learning literature, consider repeated communication and interactions among agents. More specifically, in DeGroot-style models, each individual initially receives *one* signal about the state of the world, and at the consequent time steps, updates her beliefs as a weighted average of the beliefs of her neighbors. Their main focus is on conditions under which individuals in the connected components of the social network converge to similar beliefs. Golub and Jackson further show that if the size of the network grows unboundedly, this asymptotic consensus belief converges to the true state of the world, provided that there are not overly influential agents in the society. One main feature that distinguishes our results from the works that are based on DeGroot's consensus model, such as [18], is the existence of time dynamics. Whereas in DeGroot's model each agent has only a single observation, the individuals in our model receive information in small bits over time, and therefore, need to incorporate new privately observed signals into their beliefs. This continuous flow of information, as well as the repeated social interaction and information exchange are also aspects that make our model different from the Bayesian learning literature.

Another distinctive feature of our results is the absence of absolute continuity of the true measure with respect to *all* prior beliefs, as a requirement for social learning. In standard Bayesian learning literature, an agent will have accurate predictions, as long as she assigns a positive prior belief on the true parameter. On the other hand, in the absence of absolute continuity, no information can persuade a Bayesian agent to update her belief on a parameter from zero to some positive number, and hence, such an agent fails to predict future correctly [21], [22]. In contrast, in our model, we show that in order to achieve accurate predictions across the network, it is sufficient to have one agent with positive initial belief assigned to the true parameter.

The rest of the paper is organized as follows. The next section describes our model. In section III we formally define what we mean by learning. Section IV contains our results regarding learning in strongly connected social networks. In Section V we show how, in strongly connected networks, social influence leads to social agreement. Section VI concludes.

II. The Model

A. Agents and Observations

Let Θ denote a finite set of possible states of the world and let $\theta^* \in \Theta$ denote the true underlying state of the world. We consider a set $\mathcal{N} = \{1, 2, \ldots, n\}$ of agents interacting over a social network who do not know the true state of the world. Each agent i starts with a prior belief about the true state, represented by $\mu_{i,0} \in \Delta\Theta$ which is a probability distribution over the set Θ. $\mu_{i,0}(\theta)$ denotes the initial probability that agent i assigns to the state $\theta \in \Theta$ to be the true underlying state of the world.

Conditional on the state of the world θ, at each time period $t \geq 1$, an observation profile $s_t = (s_t^1, \ldots, s_t^n) \in S_1 \times \cdots \times S_n \equiv S$ is generated according to the likelihood function $\ell(s_t|\theta)$. $s_t^i \in S_i$ denotes a signal that is privately observed by agent i. These privately observed signals are independent over time, but might be correlated among agents at the same time period. Without much loss of generality we assume that $\ell(s|\theta) > 0$ for all $(s, \theta) \in S \times \Theta$. The signal space S_i is assumed to be finite for all i. We use $\ell_i(\cdot|\theta)$ to denote the ith marginal of $\ell(\cdot|\theta)$. We further assume that every agent i knows the conditional likelihood function $\ell_i(\cdot|\theta)$, known as her *signal structure*.

For a fixed $\theta \in \Theta$, we define a probability triple $(\Omega, \mathcal{F}, \mathbb{P}^\theta)$, where Ω is the space containing sequences of realizations of the signals $s_t \in S$ over time, and $\mathbb{P}^\theta$ is the probability measure induced over sample paths in Ω by the state θ. In other words, $\mathbb{P}^\theta = \otimes_{t=1}^\infty \ell(\cdot|\theta)$. We use $\mathbb{E}^\theta[\cdot]$ to denote the expectations with respect to the measure $\mathbb{P}^\theta$. Define $\mathcal{F}_{i,t}$ as the σ-field generated by the past history of agent i's observations up to time period t, and let $\mathcal{F}_t$ be the smallest σ-field containing all $\mathcal{F}_{i,t}$ for $1 \leq i \leq n$.

B. Social Structure

When updating their beliefs about the true state of the world, agents might be under the influence of other individuals in their social clique. We capture the social interaction structure between agents by a directed graph $G = (V, E)$, where each vertex in V corresponds to an agent, and an edge connecting vertex i to vertex j, denoted by the ordered pair $(i, j) \in E$, captures the fact that agent j is under the influence of agent i. Note that because of the way we have defined the social network, agent i might influence agent j, without being influenced by her.

For each agent i, define $\mathcal{N}_i = \{j \in V : (j, i) \in E\}$, called the set of *neighbors* of agent i. The elements of this set are agents influence the beliefs of agent i at each time period. A *directed path* in $G = (V, E)$ from vertex i to vertex j, is a sequence of vertices starting with i and ending with j such that each vertex is a neighbor of the next vertex in the sequence. The social network is *strongly connected*, if there exists a directed path from each vertex to any other vertex. We say the social network is *connected*, whenever such a path exists ignoring the direction of the edges. We refer to any network that is connected but not strongly connected as a *weakly connected* network.

C. Belief Updates Under Social Influence

Given the social structure described above, agents update their beliefs based on the signals they privately observe and under the influence of their neighbors. In particular, each agent's belief over Θ is a linear combination of her Bayesian posterior belief and her neighbors' priors. That is, if we let

$\mu_{i,t}(\theta)$ denote the belief that agent i assigns to parameter $\theta \in \Theta$ at time period t after observing $(s_1^i, s_2^i, \ldots, s_t^i)$, then

$$\mu_{i,t+1} = a_{ii} BU(\mu_{i,t}; s_{t+1}^i) + \sum_{j \in \mathcal{N}_i} a_{ij} \mu_{j,t}, \qquad (1)$$

where $a_{ij} \in \mathbb{R}_+$ captures the amount of influence that agent j has on agent i at each time period, $BU(\mu_{i,t}; s_{t+1}^i)(\cdot)$ is the Bayesian update of $\mu_{i,t}$ when signal s_{t+1}^i is observed, and a_{ii} is the weight assigned to the Bayesian update by agent i, which we call her *self-confidence*.[1] Note that the weights a_{ij} must satisfy $\sum_{j \in \mathcal{N}_i} a_{ij} = 1$, in order for the period $t+1$ beliefs to form a well-defined probability distribution. Therefore, at each time period, the posterior belief of agent i is a convex combination of her Bayesian update and the priors of her neighbors.

Given agent i's beliefs at time period t, her time $t+1$ forecast of the next observation is given by

$$m_{i,t}(s_{t+1}^i) = \int_\Theta \ell_i(s_{t+1}^i | \theta) d\mu_{i,t}(\theta), \qquad (2)$$

where the forecasts $m_{i,t}(\cdot)$ form a probability measure on S_i. In other words, $m_{i,t}(s_{t+1}^i)$ is the subjective probability that agent i assigns to observing signal $s_{t+1}^i \in S_i$ at time step $t+1$ given her information up to time t. Therefore, the *law of motion* for the beliefs about the parameters can be written as

$$\mu_{i,t+1}(\theta) = a_{ii} \mu_{i,t}(\theta) \frac{\ell_i(s_{t+1}^i | \theta)}{m_{i,t}(s_{t+1}^i)} + \sum_{j \in \mathcal{N}_i} a_{ij} \mu_{j,t}(\theta), \qquad (3)$$

for all $\theta \in \Theta$. Note that in the special case that the signals observed by an agent are non-informative, equation (3) reduces to the update in benchmark discrete-time consensus algorithms.[2]

When analyzing the asymptotic behavior of beliefs, sometimes it is more convenient to use a matrix notation. Define A to be a real $n \times n$ matrix which captures the social interaction among agents as well as the amount of influence that each agent has on others. More specifically, we let the ij element of matrix A be a_{ij} when agent j is a neighbor of agent i, and zero otherwise. Thus, equation (3) can be rewritten as

$$\mu_{t+1}(\theta) = A \mu_t(\theta) + \operatorname{diag}\left(a_{ii} [\frac{\ell_i(s_{t+1}^i | \theta)}{m_{i,t}(s_{t+1}^i)} - 1] \right)_{i=1,\ldots,n} \mu_t(\theta) \qquad (4)$$

where $\mu_t(\cdot) = [\mu_{1,t}, \ldots, \mu_{n,t}]'(\cdot)$, and diag of a vector is a diagonal matrix which has the entries of the vector as its diagonal. Note that A is an irreducible matrix if and only if graph G is strongly connected.[3] Moreover, since at each time period the beliefs of all agents are convex combinations of their Bayesian posteriors and the priors of their neighbors, A is a stochastic matrix.[4] In the special case that A is the identity matrix, our model reduces to the benchmark Bayesian case, in which the society consists of n Bayesian agents who do not have access to the beliefs of other members of the society, and only observe their own private signals.

III. Social Learning

Given the model described above, we are interested in the question of what is learned in the long run. Learning may either signify learning the true parameter or learning to forecast future outcomes. These two notions of learning are distinct and might not occur simultaneously. We start this section by specifying what we exactly mean by either type.

Suppose that $\theta^* \in \Theta$ is the true state of the world and thus, the measure $\mathbb{P}^* = \otimes_{t=1}^{\infty} \ell(\cdot | \theta^*)$ is the probability law describing the process (s_t).

Definition 1: The forecasts of agent i are *eventually correct* on a path $\{s_t\}_{t=1}^{\infty}$ if, along that path,

$$m_{i,t}(\cdot) \to \ell_i(\cdot | \theta^*) \quad \text{as} \quad t \to \infty.$$

This notion of learning, which Kalai and Lehrer [25] call *weak merging* of opinions, captures the ability of agents to correctly forecast events in near future. It is well-known, that repeated applications of Bayes' rule leads to eventually correct forecasts with probability 1 under the truth, given suitable conditions, the key condition being absolute continuity of the true measure with respect to initial beliefs.[5] The implication is that the mere repetition of Bayes' rule eventually transforms the historical record into a near perfect guide for the future. However, predicting future observations accurately is not the same as learning the underlying state of the world. In fact, depending on the signal structure of each agent, there might be an "identification problem" which can potentially prevent the agent from learning the true (payoff-relevant) parameter θ^*. The other type of learning that we are concerned with, precisely captures this notion:

Definition 2: Agent $i \in \mathcal{N}$ *asymptotically learns* the true parameter θ^* on a path $\{s_t\}_{t=1}^{\infty}$ if, along that path,

$$\mu_{i,t}(\theta^*) \to 1 \quad \text{as} \quad t \to \infty.$$

Asymptotic learning occurs when the agent assigns probability one to the true parameter. As mentioned earlier, having eventually correct forecasts does not guarantee asymptotic learning of the true parameter. In general, the converse is not true either. However, it is straightforward to show that in the absence of time correlations, as in our model, asymptotically learning θ^* implies eventually correct forecasts.[6]

[1] One can modify this belief update model and assume that agent i's belief update also depends on his own beliefs at the previous time step, $\mu_{i,t}$. Such an assumption is equivalent to adding a prior-bias to the model, as stated in [23]. Since this generality does not change the results or the economic intuitions, we assume that agents have no prior bias.

[2] For more on consensus algorithms see e.g., [5] and [20].

[3] An $n \times n$ matrix A is said to be *reducible*, if for some permutation matrix P, the matrix $P'AP$ is block upper triangular. If a square matrix is not reducible, it is said to be *irreducible*. For more on this, see e.g., [24].

[4] A matrix is said to be *stochastic* if it is entry-wise non-negative and all its row sums are equal to one. A stochastic matrix is called *doubly stochastic* if all its column sums are equal to one, as well.

[5] Lehrer and Smorodinsky [26] show that an assumption weaker than absolute continuity, known as *accomodation*, is sufficient for weak merging of the opinion.

[6] See Lehrer and Smorodinsky [26], for an example of the case that learning the true parameter does not guarantee merging.

IV. Convergence Analysis in Strongly Connected Social Networks

We now turn to the main question of this paper: under what circumstances does learning occur over the social network, when agents update their beliefs according to (3)? Note that without loss of generality, we can limit our focus to two cases of strongly connected and weakly connected networks. Clearly, in the case that the network is disconnected, our results can be applied to each connected component separately.

In this section, we show that as long as the social network is strongly connected, all agents will eventually forecast their signals correctly, in spite of the fact that they are under the influence of their neighbors and update their beliefs in a non-Bayesian manner.

Before presenting our main theorem, we state and prove two lemmas, both of which are consequences of the martingale convergence theorem.[7] Our first lemma indicates that a weighted sum of the beliefs of the individuals is a bounded submartingale, and therefore, converges.

Lemma 1: If A is stochastic, then the sequence $\sum_{i=1}^n v_i\mu_{i,t}(\theta^*)$ converges $\mathbb{P}^*$-almost surely as $t \to \infty$, where v is any non-negative left eigenvector of A corresponding to its unit eigenvalue.

Proof: First, note that since A is stochastic, it always has at least one eigenvalue equal to 1. Moreover, there exists a non-negative left eigenvector corresponding to this eigenvalue.[8] We denote such a vector by v and its transpose by v'.

Evaluate equation (4) at the true parameter θ^* and multiply both sides by v' from left

$$v'\mu_{t+1}(\theta^*) = v'A\mu_t(\theta^*)+\sum_{i=1}^n v_i\mu_{i,t}(\theta^*)a_{ii}[\frac{\ell_i(s_{t+1}^i|\theta^*)}{m_{i,t}(s_{t+1}^i)}-1].$$

Thus,

$$\mathbb{E}^*\left[\sum_{i=1}^n v_i\mu_{i,t+1}(\theta^*)|\mathcal{F}_t\right] = \sum_{i=1}^n v_i\mu_{i,t}(\theta^*)+ \\ +\sum_{i=1}^n v_i a_{ii}\mu_{i,t}(\theta^*)\mathbb{E}^*\left[\frac{\ell_i(s_{t+1}^i|\theta^*)}{m_{i,t}(s_{t+1}^i)}-1|\mathcal{F}_t\right], \quad (5)$$

where $\mathbb{E}^*$ denotes expectation with respect to the measure $\mathbb{P}^*$. Since the function $f(x) = 1/x$ is convex, Jensen's inequality implies

$$\mathbb{E}^*\left[\frac{\ell_i(s_{t+1}^i|\theta^*)}{m_{i,t}(s_{t+1}^i)}|\mathcal{F}_t\right] \geq \left(\mathbb{E}^*\left[\frac{m_{i,t}(s_{t+1}^i)}{\ell_i(s_{t+1}^i|\theta^*)}|\mathcal{F}_t\right]\right)^{-1} = 1,$$

and therefore,

$$\mathbb{E}^*\left[\sum_{i=1}^n v_i\mu_{i,t+1}(\theta^*)|\mathcal{F}_t\right] \geq \sum_{i=1}^n v_i\mu_{i,t}(\theta^*).$$

Note that this last inequality is due to the fact that v is element-wise non-negative. As a result, $\sum_{i=1}^n v_i\mu_{i,t}(\theta^*)$ is a submartingale with respect to the filtration $\mathcal{F}_t$, which is also bounded above by $\|v\|_1$. Hence, it converges $\mathbb{P}^*$-almost surely. ■

Our second lemma shows that whenever the social network is strongly connected, the weighted sum of the logarithms of the belief assigned to the true parameter converges.

Lemma 2: Suppose that there exists an agent i such that $\mu_{i,0}(\theta^*) > 0$. Then, whenever A is stochastic and irreducible, the sequence $\sum_{i=1}^n v_i \log\mu_{i,t}(\theta^*)$ converges $\mathbb{P}^*$-almost surely as $t \to \infty$, where v is any non-negative left eigenvector of A corresponding to its unit eigenvalue.

Proof: Similar to the proof of the previous lemma, we show that $\sum_{i=1}^n v_i \log\mu_{i,t}(\theta^*)$ is a bounded submartingale and invoke the martingale convergence theorem to obtain almost sure convergence.

First, note that since A is a stochastic matrix, the right hand side of equation (3) is a convex combination for all i. Therefore, by evaluating the law of motion at θ^* and taking log from both sides, we obtain

$$\log\mu_{i,t+1}(\theta^*) \geq a_{ii}\log\mu_{i,t}(\theta^*)+ \\ +a_{ii}\log\left(\frac{\ell_i(s_{t+1}^i|\theta^*)}{m_{i,t}(s_{t+1}^i)}\right)+\sum_{j\in\mathcal{N}_i} a_{ij}\log\mu_{j,t}(\theta^*),$$

where we have used the concavity of the logarithm function. Since A is irreducible, the social network is strongly connected. Thus, (3) implies that the existence of one agent with a positive prior on θ^* guarantees the fact that after at most n time periods, all agents assign a strictly positive probability to the true parameter, which means that $\log\mu_{i,t}(\theta^*)$ is well-defined for large enough t for all i.

Our next step to show that $\mathbb{E}^*\left[\log\frac{\ell_i(s_{t+1}^i|\theta^*)}{m_{i,t}(s_{t+1}^i)}|\mathcal{F}_t\right] \geq 0$. To obtain this,

$$\begin{aligned}\mathbb{E}^*\left[\log\frac{\ell_i(s_{t+1}^i|\theta^*)}{m_{i,t}(s_{t+1}^i)}|\mathcal{F}_t\right] &= -\mathbb{E}^*\left[\log\frac{m_{i,t}(s_{t+1}^i)}{\ell_i(s_{t+1}^i|\theta^*)}|\mathcal{F}_t\right] \\ &\geq -\log\mathbb{E}^*\left[\frac{m_{i,t}(s_{t+1}^i)}{\ell_i(s_{t+1}^i|\theta^*)}|\mathcal{F}_t\right] \\ &= 0.\end{aligned}$$

Thus,

$$\mathbb{E}^*\left[\log\mu_{i,t+1}(\theta^*)|\mathcal{F}_t\right] \geq a_{ii}\log\mu_{i,t}(\theta^*)+\sum_{j\in\mathcal{N}_i} a_{ij}\log\mu_{j,t}(\theta^*)$$

which can be rewritten in matrix form as $\mathbb{E}^*\left[\log\mu_{t+1}(\theta^*)|\mathcal{F}_t\right] \geq A\log\mu_t(\theta^*)$, where by the logarithm of a vector, we mean its entry-wise logarithm. Multiplying both sides by the A's non-negative left eigenvector v' leads to

$$\mathbb{E}^*\left[\sum_{i=1}^n v_i\log\mu_{i,t+1}(\theta^*)|\mathcal{F}_t\right] \geq \sum_{i=1}^n v_i\log\mu_{i,t}(\theta^*)$$

which means that $\sum_{i=1}^n v_i\log\mu_{i,t}(\theta^*)$ is a submartingale with respect to the filtration $\mathcal{F}_t$. Moreover, the sum is non-positive, and therefore, it converges with $\mathbb{P}^*$-probability one. ■

[7] For the statement of martingale convergence theorem and related definitions, see [27].

[8] This is a consequence of the Perron-Frobenius theorem. For more on the properties of non-negative and stochastic matrices, see [24].

We now present our main result regarding learning in strongly connected social networks.

Theorem 1: Suppose that the network of social interactions is strongly connected, and all agents have strictly positive self-confidence. Then, all agents eventually forecast their private observations accurately with $\mathbb{P}^*$-probability one, provided that there exists an agent in the social network with strictly positive prior on the true parameter θ^*.

The above theorem relies on three main assumptions: strictly positive weights on new observations by all agents, a strongly connected social network, and finally, a strictly positive prior belief on the true parameter by at least one agent in the network. If either of these assumptions are dropped, the forecasts might either diverge or converge to wrong values. Before presenting the proof of Theorem 1, we briefly discuss each assumption.

The first assumption on strictly positive self-confidences is quite intuitive: it prohibits agents from completely discarding information provided to them through their observations. Clearly, if all agents discard their private signals, no new information is incorporated to their beliefs, and (3) simply turns into a diffusion of prior beliefs. An interesting case is when only some of them discard their private signals. In such a case, only agents with no self-confidence fail to predict their signals accurately. This failure, however, does not prevent the ones with positive self-confidence from forecasting their own observations correctly.

The second requirement for accurate predictions is strong connectivity of the social network. To understand why strong connectivity is the key, we present a simple example of a weakly connected social network, in which an agent fails in forecasting future correctly.

Example 1: Consider a society consisting of two agents, $\mathcal{N} = \{1, 2\}$, and assume that $\Theta = \{\theta_1, \theta_2\}$ with the true state being $\theta^* = \theta_1$. Both agents have non-degenerate prior beliefs over Θ. The signals observed by agents are independent conditional on the state of the world, and belong to the set $S_1 = S_2 = \{H, T\}$. We further assume that the signals observed by agent 2 are non-informative, while agent 1's observations are perfectly informative about the state; that is, $\ell_1(H|\theta_1) = \ell_1(T|\theta_2) = 1$, and $\ell_2(s|\theta_1) = \ell_2(s|\theta_2)$ for $s \in \{H, T\}$. As for the social structure, we assume that agent 1's beliefs are influenced by agent 2, while agent 2 is not under the influence of agent 1. Therefore, the social interaction matrix is given by

$$A = \begin{bmatrix} 1-\alpha & \alpha \\ 0 & 1 \end{bmatrix},$$

where $\alpha \in (0, 1)$ captures the level of influence of agent 2 on the beliefs of agents 1, when she updates her beliefs using equation (3).

Since the private signals observed by agent 2 are non-informative, her beliefs, at all times, remain equal to her prior. Clearly, she has correct forecasts at all times. Agent 1, on the other hand, will not have eventually correct forecasts. To see this, notice that agent 1 has eventually correct forecasts, only if she eventually assigns probability 1 to the true state, θ_1. This is due to the fact that her observations are perfectly informative. However, the belief she assigns to θ_2 follows the law of motion

$$\mu_{1,t+1}(\theta_2) = (1-\alpha)\mu_{1,t}(\theta_2)\frac{\ell_1(s^1_{t+1}|\theta_2)}{m_{1,t}(s^1_{t+1})} + \alpha\mu_{2,t}(\theta_2)$$

which cannot converge to zero, due to the fact that $\mu_{2,t}(\theta_2) = \mu_{2,0}(\theta_2)$ is strictly positive.

The intuition for the failure of learning in the above example is simple. First of all, notice that the two agents have different signal structures, which means that they interpret the states differently. Moreover, agent 1, in essence, is following the beliefs of agent 2 without considering the fact that agent 2 is less informed than herself, while at the same time being unable of influencing her back. This one-way influence and non-identical interpretations of signals (due to non-identical signal structures) result in confusion on the part of agent 1, and hence incorrect forecasts. Clearly, if agent 1 were Bayesian and capable of incorporating the information provided to her by agent 2 rationally, she would have learned the true parameter. As a final remark on the second assumption, note that in the special case that all social interactions are bidirectional, the social network is trivially strongly connected.

Finally, in order to have accurate predictions, Theorem 1 requires the existence of at least one agent with a positive prior belief on the true state θ^*. Note that correct forecasts are achieved, even if the true measure, $\mathbb{P}^*$, is not absolutely continuous with respect to the prior beliefs of many agents in the network. This is in contrast to the standard Bayesian learning literature, which requires absolute continuity in order to guarantee accurate predictions. This feature of our model is significant: as long as some agent i assigns a positive prior belief to the true state, all agents in the network are able to correctly forecast their observations, even if agent i is located on the fringe of the society and has very small influence on her neighbors. Clearly, if $\mu_{i,0}(\theta^*) = 0$ for all $i \in \mathcal{N}$, then the belief assigned to the true parameter by all agents will remain equal to zero over time, and no learning can happen.

We now present the proof of Theorem 1.

Proof of Theorem 1: According to Lemma 1, the term $\sum_{i=1}^n v_i\mu_{i,t}(\theta^*)$ converges with $\mathbb{P}^*$-probability one, where v is the non-negative left eigenvector of A corresponding to its unit eigenvalue. Therefore, equation (5) implies that

$$\sum_{i=1}^n v_i a_{ii}\mu_{i,t}(\theta^*)\left(\mathbb{E}^*\left[\frac{\ell_i(s^i_{t+1}|\theta^*)}{m_{i,t}(s^i_{t+1})}|\mathcal{F}_t\right] - 1\right) \to 0 \quad \mathbb{P}^*\text{-a.s.}$$

Since for every individual i in the social network, the term $v_i a_{ii}\mu_{i,t}(\theta^*)\mathbb{E}^*\left[\ell_i(s^i_{t+1}|\theta^*)/m_{i,t}(s^i_{t+1}) - 1|\mathcal{F}_t\right]$ is non-negative, each such term converges to zero with $\mathbb{P}^*$-probability one. Moreover, the assumption that all the diagonal entries of A are strictly positive and the fact that A is irreducible (and hence, v is entry-wise positive) lead to

$$\mu_{i,t}(\theta^*)\left(\mathbb{E}^*\left[\frac{\ell_i(s^i_{t+1}|\theta^*)}{m_{i,t}(s^i_{t+1})}|\mathcal{F}_t\right] - 1\right) \to 0 \quad \forall i \quad \mathbb{P}^* - \text{a.s.} \tag{6}$$

Furthermore, Lemma 2 guarantees that $\sum_{i=1}^{n} v_i \log \mu_{i,t}(\theta^*)$ converges almost surely, meaning that $\mu_{i,t}(\theta^*)$ is uniformly bounded away from zero for all i with probability one. Note that, again we are using the fact that v is a positive vector. Hence, $\mathbb{E}^*\left[\frac{\ell_i(s_{t+1}^i|\theta^*)}{m_{i,t}(s_{t+1}^i)}|\mathcal{F}_t\right] \to 1$ almost surely. Thus,

$$\mathbb{E}^*\left[\frac{\ell_i(s_{t+1}^i|\theta^*)}{m_{i,t}(s_{t+1}^i)}|\mathcal{F}_t\right] - 1 = \sum_{s\in S_i} \ell_i(s|\theta^*)\left(\frac{\ell_i(s|\theta^*)}{m_{i,t}(s)} - 1\right)$$
$$= \sum_{s\in S_i}\left(\ell_i(s|\theta^*)\frac{\ell_i(s|\theta^*) - m_{i,t}(s)}{m_{i,t}(s)} + m_{i,t}(s) - \ell_i(s|\theta^*)\right)$$
$$= \sum_{s\in S_i}\frac{\left[\ell_i(s|\theta^*) - m_{i,t}(s)\right]^2}{m_{i,t}(s)} \to 0 \qquad \mathbb{P}^* - \text{a.s.},$$

where the second equality is due to the fact that both $\ell_i(\cdot|\theta^*)$ and $m_{i,t}(\cdot)$ are measures on S_i, and therefore, $\sum_{s\in S_i}\ell_i(s|\theta^*) = \sum_{s\in S_i} m_{i,t}(s) = 1$.

In the last expression, the term in the braces and the denominator are always non-negative and therefore,

$$m_{i,t}(s) \to \ell_i(s|\theta^*) \qquad \mathbb{P}^* - \text{a.s.}$$

for all $s \in S_i$ and all $i \in \mathcal{N}$. This proves the theorem. ■

V. Asymptotic Agreement: Information Aggregation

In this section, we show that the social influence among individuals, not only does not prevent them from having correct forecasts, but also can potentially provide them with a better degree of learning. In particular, in the next theorem we show that if the social network is strongly connected, then all agents assign equal beliefs to the true state of the world. This implies that if different agents have different signal structures, it is possible for them to achieve a better degree of learning when they are under social influence, relative to the case that they ignore the information provided by their neighbors. Characterizing the extent of this information aggregation occurring through the network is the focus of our future research.

Theorem 2: Suppose that the social network is strongly connected, $a_{ii} > 0$ for all i, and there exists an agent with a positive prior belief on the true state θ^*. Then for all i, with $\mathbb{P}^*$-probability one, $\mu_{i,\infty}^* := \lim_{t\to\infty}\mu_{i,t}(\theta^*)$ exists and does not depend on i.

Proof: Theorem 1 from the previous section shows that under the given assumptions, all agents will have asymptotically correct forecasts almost surely. Therefore, we first show that eventually correct forecasts for all agents implies that $\lim_{t\to\infty}|\mu_{i,t}(\theta^*) - \mu_{j,t}(\theta^*)| = 0$ for any $i,j \in \mathcal{N}$. By equation (4), on any path that all agents have eventually correct forecasts, $\mu_{t+1}(\theta^*) - A\mu_t(\theta^*) \to 0$. That is, for any $\epsilon > 0$ there exists a large enough time T such that for all $t \geq T$,

$$|\mu_{i,t+1}(\theta^*) - \sum_{k=1}^{n} a_{ik}\mu_{k,t}(\theta^*)| < \frac{\epsilon}{2} \qquad \forall i \in \mathcal{N}$$

Therefore, given any two agents i and j,

$$|\left(\mu_{i,t+1}(\theta^*) - \mu_{j,t+1}(\theta^*)\right) - \sum_{k=1}^{n}\mu_{k,t}(\theta^*)(a_{ik} - a_{jk})| < \epsilon,$$

and hence,

$$|\mu_{i,t+1}(\theta^*) - \mu_{j,t+1}(\theta^*)| < \epsilon + |\sum_{k=1}^{n}\mu_{k,t}(\theta^*)(a_{ik} - a_{jk})|.$$

Since A is a stochastic matrix, $\sum_{k=1}^{n}(a_{ik} - a_{jk}) = 0$. Therefore, we can use Paz's inequality to find an upper bound for the right hand side of the above inequality:[9]

$$|\mu_{i,t+1}(\theta^*) - \mu_{j,t+1}(\theta^*)| < \epsilon + \\ +\frac{1}{2}\max_{p,q}|\mu_{p,t}(\theta^*) - \mu_{q,t}(\theta^*)|\sum_{k=1}^{n}|a_{ik} - a_{jk}|.$$

Thus,

$$\max_{i,j}|\mu_{i,t+1}(\theta^*) - \mu_{j,t+1}(\theta^*)| < \epsilon + \\ +\tau(A)\max_{i,j}|\mu_{i,t}(\theta^*) - \mu_{j,t}(\theta^*)|,$$

where $\tau(A) = \frac{1}{2}\max_{i,j}\sum_{k=1}^{n}|a_{ik} - a_{jk}|$, is the *coefficient of ergodicity* of matrix A. It is an easy exercise to show that the coefficient of ergodicity of any stochastic matrix lies in the interval $[0,1]$. This along with the fact that $\epsilon > 0$ is arbitrary imply that $\max_{i,j}|\mu_{i,t}(\theta^*) - \mu_{j,t}(\theta^*)|$ is a non-increasing sequence and hence, converges. We claim that the limit is in fact zero. To show this, we consider two cases.

First suppose that $\tau(A) < 1$. In that case, for any positive integer p, we have

$$\begin{aligned}\max_{i,j}|\mu_{i,t+p}(\theta^*) - \mu_{j,t+p}(\theta^*)| &< \sum_{k=0}^{p-1}[\tau(A)]^k\epsilon + \\ &+[\tau(A)]^p\max_{i,j}|\mu_{i,t}(\theta^*) - \mu_{j,t}(\theta^*)| \\ &= \frac{1-[\tau(A)]^p}{1-\tau(A)}\epsilon + [\tau(A)]^p\max_{i,j}|\mu_{i,t}(\theta^*) - \mu_{j,t}(\theta^*)| \\ &\to \frac{\epsilon}{1-\tau(A)} \quad \text{as } p\to\infty.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, the right hand side can be made arbitrarily small, and as a result, $\max_{i,j}|\mu_{i,t}(\theta^*) - \mu_{j,t}(\theta^*)|$ must converge to zero.

We now consider the case that $\tau(A) = 1$. In this case, as shown by [29], since A is irreducible and corresponds to a strongly connected graph, there exists a positive integer r such that $\tau(A^r) < 1$. Therefore, using a similar argument as above for the matrix A^r, one can show that the convergent sequence $\{\max_{i,j}|\mu_{i,t}(\theta^*) - \mu_{j,t}(\theta^*)|\}_{t=1}^{\infty}$ has a subsequence that converges to zero. Therefore, on any connected network, if all agents have eventually correct forecasts, then $\mu_{i,t}(\theta^*) - \mu_{j,t}(\theta^*) \to 0$ for all $i,j \in \mathcal{N}$.

In order to complete the proof, we need to show that $\lim_{t\to\infty}\mu_{i,t}(\theta^*)$ exists. Since $|\mu_{i,t}(\theta^*) - \mu_{j,t}(\theta^*)| \to 0$ as

[9]Paz's inequality states that if d is a vector with an entry-wise sum of zero, then for any arbitrary vector z of the same size, $|d'z| \leq \frac{1}{2}\|d\|_1\max_{i,j}|z_i - z_j|$. This inequality can be found in [28].

$t \to \infty$, for any $\delta > 0$, there exists a large enough t such that $-\delta < \mu_{i,t}(\theta^*) - \mu_{j,t}(\theta^*) < \delta$ uniformly for all i and j. Thus,

$$\mu_{j,t}(\theta^*) - \delta < \sum_{i=1}^{n} a_{ji}\mu_{i,t}(\theta^*) < \mu_{j,t}(\theta^*) + \delta.$$

Note that the term $\sum_{i=1}^{n} a_{ji}\mu_{i,t}(\theta^*)$ can be made arbitrarily close to $\mu_{j,t+1}(\theta^*)$ for large enough t, implying that $-\delta < \mu_{j,t+1}(\theta^*) - \mu_{j,t}(\theta^*) < \delta$. Therefore, $\{\mu_{j,t}(\theta^*)\}_{t=1}^{\infty}$ is a Cauchy sequence for all j and hence, converges. ∎

VI. Conclusions

In this paper, we studied the problem of learning over social networks when individuals observe informative signals about the state of the world, but at the same time, are influenced by other members of the society. More specifically, we assumed that the belief of each individual is a convex combination of her posterior beliefs and the priors of her neighbors. Our model is a generalization of consensus algorithms to the case that every agent receives an informative signal about the true state of the world at ever time period. It reduces to the benchmark consensus model if the observations of all agents are non-informative. In fact, our model of semi-rational agents can be considered as a combination of Bayesian learning models (which consist of fully rational agents) and consensus algorithms (with naïve agents). The simplicity of our *local* update rule guarantees that agents avoid highly complex computations that are essential for full Bayesian learning over the network.

We explored the effect of the social structure on the asymptotic beliefs of the individuals. We showed that, in general, agents might be misled due to the bias they exhibit towards the opinion held by their neighbors. Such a phenomenon arises in the presence of "one-way" influences, where an individual is influenced by a group of other individuals without being able to influence them back. On the other hand, we also showed that as long as the social network is strongly connected, all agents will eventually predict their signals accurately. Furthermore, we showed that in our model, all agents will assign asymptotically equal beliefs to the true parameter. We believe that this result suggests that social influences lead to some form of information aggregation in the network; that is, agents achieve a higher degree of learning through sharing their information. Exact characterization of this phenomenon is the focus of our ongoing research.

References

[1] T. Hagerstrand, *Innovation Diffusion as a Spatial Process*. Chicago, IL: University of Chicago Press, 1969.

[2] E. M. Rogers, *Diffusion of Innovations*, 3rd ed. New York: Free Press, 1983.

[3] P. Kotler, *The Principles of Marketing*, 3rd ed., 1986.

[4] A. Calvo-Armengol and M. O. Jackson, "Peer pressure," *Journal of the European Economic Association*, 2008, (Forthcoming).

[5] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988–1001, 2003.

[6] D. Acemoglu, A. Nedić, and A. Ozdaglar, "Convergence of Rule-of-Thumb Learning Rules in Social Networks," in *Proceedings of Conference in Decision and Control*, Dec. 2008.

[7] S. Bikchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as information cascades," *Journal of Political Economy*, vol. 100, pp. 992–1026, 1992.

[8] A. Banerjee, "A simple model of herd behavior," *Quarterly Journal of Economics*, vol. 107, pp. 797–817, 1992.

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

[10] L. Smith and P. Sørensen, "Rational social learning with random sampling," 1998, unpublished Manuscript.

[11] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian Learning in Social Networks," 2008, lIDS Working Paper ♯2780, Massachusetts Institute of Technology.

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

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

[14] ——, "Word-of-mouth communication and social learning," *The Quarterly Journal of Economics*, vol. 110, pp. 93–126, 1995.

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

[16] ——, "Conformism and diversity under social learning," *Economic Theory*, vol. 17, no. 1, pp. 101–120, 2001.

[17] P. M. DeMarzo, D. Vayanos, and J. Zwiebel, "Persuasion Bias, Social Influence, and Unidimensional Opinions," *The Quarterly Journal of Economics*, vol. 118, no. 3, pp. 909–968, 2003.

[18] B. Golub and M. O. Jackson, "Naïve learning in social networks: Convergence, influence, and the wisdom of crowds," 2007, unpublished Manuscript.

[19] D. Acemoglu, A. Ozdaglar, and A. ParandehGheibi, "Spread of (mis)information in social networks," Massachusetts Institute of Technology, LIDS Report 2812, May 2009, submitted for Publication.

[20] M. H. DeGroot, "Reaching a Consensus," *Journal of American Statistical Association*, vol. 69, no. 345, pp. 118–121, Mar. 1974.

[21] R. I. Miller and C. W. Sanchirico, "The role of absolute continuity in "merging of opinions" and "rational learning"," *Games and Economic Behavior*, vol. 29, pp. 170–190, 1999.

[22] A. Sandroni, "Necessary and sufficient conditions for convergence to Nash equilibrium: The almost absolute continuity hypothesis," *Games and Economic Behavior*, vol. 22, pp. 121–147, 1998.

[23] L. G. Epstein, J. Noor, and A. Sandroni, "Supplementary Appendix for 'Non-Bayesian Updating: A Theoretical Framework'," 2008, working Paper 08-017, Penn Institute for Economic Research, University of Pennsylvania.

[24] A. Berman and R. J. Plemmons, *Nonnegative matrices in the mathematical sciences*. New York: Academic Press, 1979.

[25] E. Kalai and E. Lehrer, "Weak and strong merging of opinions," *Journal of Mathematical Economics*, vol. 23, pp. 73–86, 1994.

[26] E. Lehrer and R. Smorodinsky, "Merging and learning," *Statistics, Probability and Game Theory*, vol. 30, pp. 147–168, 1996.

[27] R. Durrett, *Probability: Theory and Examples*, 3rd ed. Belmont, CA: Duxbury Press, 2005.

[28] A. Paz, *Introduction to Probabilistic Automata*. New York: Academic Press, 1971.

[29] E. Seneta, *Non-negative Matrices and Markov Chains*, 2nd ed. New York: Springer, 1981.