DIFFUSION SOCIAL LEARNING OVER WEAKLY-CONNECTED GRAPHS

Hawraa Salami, Bicheng Ying, and Ali H. Sayed

Department of Electrical Engineering
University of California, Los Angeles

ABSTRACT

In this paper, we study diffusion social learning over weakly-connected graphs. We show that the asymmetric flow of information hinders the learning abilities of certain agents regardless of their local observations. Under some circumstances that we clarify in this work, a scenario of *total influence* (or "mind-control") arises where a set of influential agents ends up shaping the beliefs of non-influential agents. We derive useful closed-form expressions that characterize this influence, and which can be used to motivate design problems to control it. We provide simulation examples to illustrate the results.

Index Terms— Weakly-connected networks, social learning, Bayesian update, diffusion strategy, leader-follower relationship.

1. INTRODUCTION AND RELATED WORKS

Social interactions among networked agents influence the beliefs of the agents about the state of nature in at least two ways: first, through the diffusion of information from neighboring agents and, second, through the sensing of local information by the agents. The nature of the graph topology over which these interactions occur plays an important role in determining the ultimate opinion formation of the agents. A critical difference in behavior arises between strongly-connected and weakly-connected graphs. In the latter case, a leader-follower relationship develops with some influential agents dictating the beliefs of other agents regardless of the observations that are sensed by these other agents. This situation leads to a *total influence* (or "mind-control") scenario. We examine the conditions under which this situation arises and derive closed-form expressions that characterize the influence behavior. The expressions can be used to motivate design problems to control the behavior.

Thus, consider a social network consisting of N agents connected by some graph. Let $\mathcal{N} = \{1, 2, \dots, N\}$ denote the indexes of the agents. We assign a pair of nonnegative weights, $\{a_{k\ell}, a_{\ell k}\}$, to the edge connecting any two agents k and ℓ. The scalar $a_{\ell k}$ represents the weight with which agent k scales the data arriving from agent ℓ and, similarly, for $a_{k\ell}$. Let $\mathcal{N}_k$ denote the neighborhood of agent k, which consists of all agents connected to k by edges. Each agent k scales data arriving from its neighbors in a convex manner, i.e.,

$$a_{\ell k} \geq 0, \quad \sum_{\ell \in \mathcal{N}_k} a_{\ell k} = 1, \quad a_{\ell k} = 0 \text{ if } \ell \notin \mathcal{N}_k \tag{1}$$

We denote by Θ the finite set of all possible states, and by $\theta^\circ \in \Theta$ the *unknown* true event that has happened. The objective of the network is to learn this true state. For this purpose, agents will be continually updating their beliefs about the true state through a localized cooperative process. Initially, at time $i = 0$, each agent k starts from some prior belief, denoted by the function $\mu_{k,0}(\theta) \in [0, 1]$. This function represents the probability distribution over the events $\theta \in \Theta$. For subsequent time instants $i \geq 1$, the private belief of agent k is denoted by $\mu_{k,i}(\theta) \in [0, 1]$. We assume that, at each time $i \geq 1$, every agent k observes a realization of some signal $\boldsymbol{\xi}_{k,i}$ that is generated according to a likelihood function $L_k(\cdot \,|\, \theta^\circ)$. We also assume that for each agent k, the signals $\{\boldsymbol{\xi}_{k,i}\}$ belong to a finite signal space denoted by Z_k and that these signals are temporally independent.

Diffusion social learning provides a mechanism by which agents can process the information they receive from their private signals and from their neighbors [1]. At every time $i \geq 1$, each agent k first updates its private belief, $\mu_{k,i-1}(\theta)$, based on its observed private signal $\xi_{k,i}$ by means of the Bayesian rule. This step leads to an intermediate belief $\psi_{k,i}(\theta)$. After learning from its observed signals, agent k can then learn from its social neighbors through cooperation. The combination of these two steps corresponds to the diffusion social learning model, written as (note that, in our notation, we use boldface letters to refer to random variables):

$$\begin{cases} \boldsymbol{\psi}_{k,i}(\theta) = \dfrac{\boldsymbol{\mu}_{k,i-1}(\theta) L_k(\boldsymbol{\xi}_{k,i}|\theta)}{\sum_{\theta' \in \Theta} \boldsymbol{\mu}_{k,i-1}(\theta') L_k(\boldsymbol{\xi}_{k,i}|\theta')} \\ \boldsymbol{\mu}_{k,i}(\theta) = \displaystyle\sum_{\ell \in \mathcal{N}_k} a_{\ell k}\, \boldsymbol{\psi}_{\ell,i}(\theta) \end{cases} \tag{2}$$

A consensus-based strategy can also be used, as studied in [2]. We focus on the diffusion learning scheme (2) due to its enhanced performance, as observed in [1] and as further discussed in [3,4]. Diffusion and consensus strategies are examples of the broad class of non-Bayesian learning models where agents communicate locally and aggregate beliefs across neighborhood — see also [5–9]. There are other Bayesian-learning models [10–15], where agents rely solely on Bayes' rule to update their beliefs.

The formulations in [1, 2] consider the case when the private signals (observations) of the agents do not hold enough information about the true state, so that agents are motivated to cooperate to identify θ°. This situation is modeled by assuming that each agent k has a subset of indistinguishable states $\Theta_k \subseteq \Theta$, such that:

$$L_k(\zeta_k|\theta) = L_k(\zeta_k|\theta^\circ) \tag{3}$$

for any $\zeta_k \in Z_k$ and $\theta \in \Theta_k$. However, through cooperation with their neighbors, agents are able to identify the true state by assuming the identifiability condition:

$$\bigcap_{k \in \mathcal{N}} \Theta_k = \{\theta^\circ\} \tag{4}$$

This work was supported in part by NSF grants CIF-1524250 and ECCS-1407712, DARPA project N66001–14–2-4029, and by a Visiting Professorship from the Leverhulme Trust, United Kingdom. Emails: {hsalami, ybc, sayed}@ucla.edu

We further assume that, for each agent k, there exists at least one prevailing signal ζ_k°, such that

$$L_k(\zeta_k^\circ|\theta^\circ) - L_k(\zeta_k^\circ|\theta) \geq 0, \quad \forall \theta \in \Theta \setminus \Theta_k \tag{5}$$

and that there exists at least one agent with a positive prior belief about the true state θ°, i.e.,

$$\boldsymbol{\mu}_{k,0}(\theta^\circ) > 0 \tag{6}$$

for some $k \in \mathcal{N}$. Under these conditions, it was shown in [1] that the agents are able to learn the true state asymptotically, i.e.,

$$\lim_{i\to\infty} \boldsymbol{\mu}_{k,i}(\theta^\circ) \overset{a.s.}{=} 1 \tag{7}$$

for any $k \in \mathcal{N}$. This result was derived for strongly-connected graphs where a path connecting any two arbitrary agents is always possible, in either direction, including self-loops around some nodes. Over such topology, all agents are able to learn the truth even when the local observations at the agents may carry information levels of varying quality with some agents being more informed than others.

In this work, we will examine how this result is affected over *weakly-connected* graphs, as opposed to strongly-connected graphs. Over a weak topology, information may only flow in one direction over a select number of edges, with information never flowing back from the receiving agents to the originating agents (even indirectly through nontrivial paths). Such configuration, whose influence on distributed inference strategies is examined in [16,17], is common in practice. For example, in Twitter networks, it is not unusual for some influential agents (e.g., celebrities) to have a multitude of followers, while the influential agent itself may not be receiving information from these followers. A similar effect arises when social networks operate in the presence of stubborn agents [6, 18]; these are agents that insist on their opinions regardless of the evidence provided by local observations or by neighboring agents. It turns out that weak graphs influence the evolution of the agents' beliefs in a critical way.

2. WEAKLY-CONNECTED GRAPHS

We first review the model for weakly-connected graphs following [16, 17]. In simple terms, weakly-connected networks consist of multiple sub-networks where at least one sub-network feeds information forward to some other sub-network but never receives information back from this sub-network — see Fig.1 for an example involving four sub-networks.

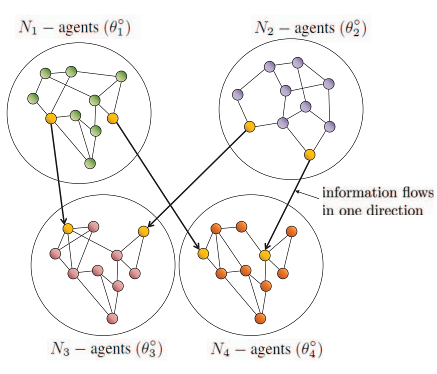

Fig. 1: An example of a weakly-connected network.

The agents in each sub-network observe signals related to their own true states denoted by $\{\theta_1^\circ, \theta_2^\circ, \theta_3^\circ, \theta_4^\circ\}$, which may not be necessarily the same (a special case of a weak topology was considered in [9], where it was assumed that all sub-networks have the *same* true state). Each of the two sub-networks on top is strongly-connected. Therefore, if their agents follow the model of diffusion social learning (2), they can asymptotically learn their true states. The third and fourth sub-networks in the bottom are seen to receive information from the other sub-networks without feeding information back to them. We would like to examine whether this external influence limits the truth learning of non-influential agents, and under what circumstances.

For more general networks, we adopt the same notation and terminology from [16, 17], which we repeat here briefly for convenience. We consider networks that consist of two types of sub-networks: S strongly-connected sub-networks and R connected sub-networks. The interaction between S and R sub-networks is not symmetric: information can flow from S ("sending") sub-networks to R ("receiving") sub-networks but not the other way around. We index each strongly-connected sub-network by s where $s = \{1, 2 \cdots, S\}$. Similarly, we index each receiving sub-network by r where $r = \{S+1, \ldots, S+R\}$. Each sub-network s has N_s agents. Similarly, each sub-network r has N_r agents. We denote by N_{gS} and N_{gR} the total number of agents in the S sub-networks and R sub-networks, respectively. We continue to denote by $\mathcal{N} = \{1, 2, \cdots, N\}$ the indexes of the agents. We assume that the agents are numbered such that the indexes of $\mathcal{N}$ represent first the agents from the S sub-networks, followed by those from the R sub-networks. In this way, the structure of the network can be represented by a large $N \times N$ combination matrix A of the following upper-block triangular form [16, 17, 19]:

$$\begin{array}{c}
\overbrace{\hspace{8em}}^{\text{Subnetworks:}1,2,\ldots,S} \quad \overbrace{\hspace{14em}}^{\text{Subnetworks:}S+1,S+2,\ldots,S+R} \\
\left[\begin{array}{cccc|cccc}
A_1 & 0 & \ldots & 0 & A_{1,S+1} & A_{1,S+2} & \cdots & A_{1,S+R} \\
0 & A_2 & \ldots & 0 & A_{2,S+1} & A_{2,S+2} & \cdots & A_{2,S+R} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & A_S & A_{S,S+1} & A_{S,S+2} & \cdots & A_{S,S+R} \\
\hline
0 & 0 & \ldots & 0 & A_{S+1} & A_{S+1,S+2} & \cdots & A_{S+1,S+R} \\
0 & 0 & \ldots & 0 & 0 & A_{S+2} & \cdots & A_{S+2,S+R} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0 & 0 & 0 & \cdots & A_{S+R}
\end{array}\right]
\end{array} \tag{8}$$

The matrices $\{A_1, \cdots, A_S\}$ on the upper left corner are left-stochastic primitive matrices corresponding to the S strongly-connected sub-networks. Likewise, the matrices $\{A_{S+1}, \cdots, A_{S+R}\}$ in the lower right-most block correspond to the internal weights of the R sub-networks. These matrices are not necessarily left-stochastic because they do not include the coefficients over the links that connect the R sub-networks to the S sub-networks. We denote the block structure of A in (8) by:

$$A \triangleq \left[\begin{array}{c|c} T_{SS} & T_{SR} \\ \hline 0 & T_{RR} \end{array}\right] \tag{9}$$

and introduce the $N_{gS} \times N_{gR}$ matrix

$$W \triangleq T_{SR}(I - T_{RR})^{-1} \tag{10}$$

which can be shown to be left-stochastic [16].

3. DIFFUSION LEARNING OVER WEAK GRAPHS

We now examine the belief evolution of agents in weakly-connected networks. We still denote by Θ the set of all possible states, and we assume that Θ is uniform across all sub-networks. However, we allow each sub-network to have its own true state, which may differ from one sub-network to another. We denote by θ_s° the true state of sending sub-network s and by θ_r° the true state of receiving sub-network r, where both θ_s° and θ_r° are in Θ. If agent k belongs to sub-network s, then according to (7) and the results in [1], it holds that:

$$\lim_{i\to\infty} \boldsymbol{\mu}_{k,i}(\theta_s^\circ) \overset{a.s.}{=} 1 \tag{11}$$

The question that we want to examine is how the beliefs of the agents in the receiving sub-networks are affected. These agents are now influenced by the beliefs of the $S-$type groups. Since this external influence carries information not related to the true state of each receiving sub-network, the receiving agents may not be able to learn their own true states.

Thus, in a manner similar to (3), for any agent k that belongs to sub-network r, we assume that there exists a subset of states $\Theta_k \subseteq \Theta$ such that:

$$L_k(\zeta_k|\theta) = L_k(\zeta_k|\theta_r^\circ), \quad \forall\, \zeta_k \in Z_k,\, \theta \in \Theta_k \tag{12}$$

Moreover, we assume a scenario in which the private signals of agents in the receiving sub-networks are not informative enough to let their agents discover that the true states of the sending sub-networks do not represent their own truth. That is, we are assuming for now that the true state θ_s°, of each sub-network s, belongs to the indistinguishable set Θ_k:

$$\theta_s^\circ \in \Theta_k \tag{13}$$

for any $s \in \{1, 2\ldots, S\}$ and agent k belongs to group R. Under (13), it will hold that the interaction with the S sub-networks ends up forcing the receiving agents to focus their beliefs on the true states of the $S-$type. The exact result can be stated as follows. We first let

$$\boldsymbol{\mu}_{\mathcal{S},i}(\theta) \;\triangleq\; \text{col}\left\{\boldsymbol{\mu}_{1,i}(\theta), \boldsymbol{\mu}_{2,i}(\theta), \ldots, \boldsymbol{\mu}_{N_{gS},i}(\theta)\right\} \tag{14}$$

$$\boldsymbol{\mu}_{\mathcal{R},i}(\theta) \;\triangleq\; \text{col}\left\{\boldsymbol{\mu}_{N_{gS}+1,i}(\theta), \boldsymbol{\mu}_{N_{gS}+2,i}(\theta), \ldots, \boldsymbol{\mu}_{N,i}(\theta)\right\} \tag{15}$$

collect all beliefs from all agents in the S and R sub-networks respectively.

Theorem 1 (Limiting Beliefs for Receiving Agents) *Under (13), it holds that*

$$\lim_{i\to\infty} \boldsymbol{\mu}_{\mathcal{R},i}(\theta) = W^{\mathsf{T}} \left(\lim_{i\to\infty} \boldsymbol{\mu}_{\mathcal{S},i}(\theta)\right) \tag{16}$$

Proof: Omitted for brevity. ■

We expand (16) to clarify its meaning. Let

$$e_{\theta,\theta_s^\circ} \;\triangleq\; \begin{cases} \mathbb{1}_{N_s}, & \text{if} \quad \theta = \theta_s^\circ \\ \mathbf{0}_{N_s}, & \text{otherwise} \end{cases} \tag{17}$$

where $\mathbb{1}_{N_s}$ denotes a column vector of ones of length N_s. Similarly, $\mathbf{0}_{N_s}$ denotes a column vector of zeros of length N_s. Now, let w_k^{T} denote the row of W^{T} that corresponds to agent k in sub-network[1] r. We partition it into

$$w_k^{\mathsf{T}} = \left[\; w_{k,N_1}^{\mathsf{T}} \quad w_{k,N_2}^{\mathsf{T}} \quad \ldots, w_{k,N_S}^{\mathsf{T}} \;\right] \tag{18}$$

[1] The real index of the row of W^{T} that corresponds to agent k is $k-N_{gS}$.

By examining (16), we conclude that the distribution for each agent k in an $R-$type sub-network converges to a combination of the various vectors $\{e_{\theta,\theta_s^\circ}\}$, namely,

$$\lim_{i\to\infty} \boldsymbol{\mu}_{k,i}(\theta) \;=\; q_k(\theta) \;\triangleq\; \sum_{s=1}^{S} w_{k,N_s}^{\mathsf{T}} e_{\theta,\theta_s^\circ} \tag{19}$$

This result shows that the beliefs of receiving agents converge to a distribution whose support is limited to the true states of the sending sub-networks, i.e., $q_k(\theta) = 0$ for any $\theta \neq \theta_s^\circ$ and any $s \in \{1, 2, \cdots, S\}$. Moreover, to obtain the value of $q_k(.)$ at any true state θ_s°, the elements of the corresponding blocks in w_k will need to be summed. Note that this is a valid probability measure since W is left-stochastic. This "total influence" or "mind-control" scenario arises in the presence of assumption (13), under which the private signals of receiving agents are not informative enough so that agents are naturally driven to be under the influence of the sending sub-networks. This phenomenon is ubiquitous in the current information era; huge amounts of information are easily available to individuals whose limited experiences may not enable them to identify truthful information with confidence.

4. DIFFUSION LEARNING WITH SELF-AWARENESS

Now, assumption (13) is not always guaranteed to occur. It is easy to construct a counter-example that shows that when (13) is not met, the sending agents cannot fully control the receiving agents, and, moreover, the beliefs of the receiving agents will not be able to approach a *fixed* asymptotic distribution. We are interested in knowing whether the total influence situation can be restored even when assumption (13) is not satisfied anymore. We show next that this is possible by incorporating an element of self-awareness into the learning process.

We modify the diffusion strategy (2) by incorporating a non-negative scalar $\gamma_{k,i}$ into the first step. This factor enables agents to assign more or less weight to their local information in comparison to the information received from their neighbors. Specifically, we modify (2) as follows:

$$\begin{cases} \boldsymbol{\psi}_{k,i}(\theta) = (1-\gamma_{k,i})\, \boldsymbol{\mu}_{k,i-1}(\theta) \\ \qquad\qquad + \gamma_{k,i} \dfrac{\boldsymbol{\mu}_{k,i-1}(\theta) L_k(\boldsymbol{\xi}_{k,i}|\theta)}{\sum_{\theta'\in\Theta} \boldsymbol{\mu}_{k,i-1}(\theta') L_k(\boldsymbol{\xi}_{k,i}|\theta')} \\ \boldsymbol{\mu}_{k,i}(\theta) = \displaystyle\sum_{\ell\in\mathcal{N}_k} a_{\ell k} \boldsymbol{\psi}_{\ell,i}(\theta) \end{cases} \tag{20}$$

where $\gamma_{k,i} \in [0, 1]$ is a scalar variable. Observe that the intermediate belief $\boldsymbol{\psi}_{k,i}(\theta)$ of agent k is now a combination of its prior belief $\boldsymbol{\mu}_{k,i-1}(\theta)$ and the Bayesian update. The scalar $\gamma_{k,i}$ represents the amount of trust that agent k gives to its private signal and how it is balancing this trust between the new observation and its own past belief. This weight can also model the lack of an observational signal at time instant i. This model was studied for *single* stand-alone agents in [7, 20] and was motivated as a mechanism for self-control and temptation. We analyze this model over graphs now, where coupling exists among agents. Specifically, we consider weakly-connected graphs and establish two results (their proofs are omitted for brevity). The first result is related to the sending agents and the second result is related to the receiving agents.

Theorem 2 (Truth Learning by Self-Aware Sending Agents) *Assume that* $\lim_{i\to\infty} \gamma_{k,i} \neq 0$ *for any sending agent k. Then, under assumptions (4)–(6), self-aware sending agents learn the truth* *asymptotically and condition (7) continues to hold for any sending agent k.* ■

Sending agents in a strongly-connected sub-network share together information related to one common parameter and cooperate together to find their true state. Therefore, whether sending agents are self-aware or not, they can always learn the truth.

We write for each agent k in a receiving sub-network r, $\gamma_{k,i} = \tau_{k,i}\gamma_{\max}$, where $\tau_{k,i}$ and $\gamma_{\max}$ are both positive scalars less than one.

Theorem 3 (Learning by Self-Aware Receiving Agents) *The beliefs of self-aware receiving agents are confined as follows:*

$$\limsup_{i\to\infty} \boldsymbol{\mu}_{\mathcal{R},i}(\theta) \preceq W^{\mathsf{T}} \left(\lim_{i\to\infty} \boldsymbol{\mu}_{\mathcal{S},i}(\theta)\right) + \gamma_{\max} C \mathbb{1}_{N_{gR}} \quad (21a)$$

$$\liminf_{i\to\infty} \boldsymbol{\mu}_{\mathcal{R},i}(\theta) \succeq W^{\mathsf{T}} \left(\lim_{i\to\infty} \boldsymbol{\mu}_{\mathcal{S},i}(\theta)\right) - \gamma_{\max} C \mathbb{1}_{N_{gR}} \quad (21b)$$

where $C = (I - T_{RR}^{\mathsf{T}})^{-1}$ is an $N_{gR} \times N_{gR}$ matrix. ■

Results (21a)–(21b) coincides with that of Theorem 1, but with an additional $O(\gamma_{\max})$ term. This means that if each receiving agent chooses the $\gamma-$coefficient to be small enough, then its belief converges to the same distribution (16). When agent k gives a small weight to its Bayesian update, it means that it is giving its current signal $\boldsymbol{\xi}_{k,i}$ a small role to play in affecting its belief formation at time i. In this way, its belief update is mainly affected by its interaction with influential agents and with neighbors that are also under the influence of sending agents. Therefore, with time, these circumstances will promote a leader-follower relationship in the network. In other words, when nodes follow the influential agents in a blind manner, without employing self-awareness or critical analysis, the receiving sub-networks will be driven away from the truth and come under total indoctrination by the influential agents.

5. SIMULATION RESULTS

We illustrate the previous results for weakly-connected networks by using the same numerical network structure from [16, 17]. We assume that the social network has $N = 8$ agents interconnected as shown in Fig. 2, with the following combination matrix:

$$A = \left[\begin{array}{ccccc|ccc} 0.2 & 0.2 & 0.8 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.4 & 0.1 & 0 & 0 & 0.2 & 0 & 0.4 \\ 0.3 & 0.4 & 0.1 & 0 & 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 0.3 & 0.3 & 0 & 0 \\ 0 & 0 & 0 & 0.6 & 0.7 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0.2 & 0.3 & 0.2 \\ 0 & 0 & 0 & 0 & 0 & 0.1 & 0.5 & 0.3 \\ 0 & 0 & 0 & 0 & 0 & 0.1 & 0.2 & 0.1 \end{array}\right] \quad (22)$$

We assume that there are 3 possible states $\Theta = \{\theta_1^\circ, \theta_2^\circ, \theta_3^\circ\}$ that represent respectively the true events of the first, second and third sub-networks. We further assume that the observational signals of each agent k are binary and belong to $Z_k = \{H, T\}$ where H denotes "head" and T denotes "tail".

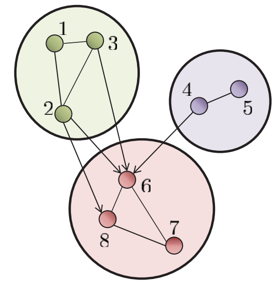

Fig. 2: A weakly-connected network [16, 17].

The likelihood of the head signals for each agent k is given by the following 3×8 matrix:

$$L(H) = \begin{bmatrix} 5/8 & 3/4 & 1/3 & 7/8 & 5/8 & 1/3 & 1/4 & 5/8 \\ 5/8 & 1/4 & 1/6 & 7/8 & 2/3 & 1/3 & 1/4 & 5/8 \\ 1/4 & 3/4 & 1/6 & 1/3 & 2/3 & 1/3 & 1/4 & 5/8 \end{bmatrix}$$

where each (j, k)-th element of this matrix corresponds to $L_k(H/\theta_j)$, i.e., each column corresponds to one agent and each row to one network state. The likelihood of the tail signal is $L(T) = \mathbb{1}_{3\times8} - L(H)$. We observe from $L(H)$ that assumption (13) is met here where for agent k in the receiving sub-network ($k > 5$) we have $L_k(\zeta_k|\theta_1^\circ) = L_k(\zeta_k|\theta_2^\circ) = L_k(\zeta_k|\theta_3^\circ)$ for both cases in which ζ_k is either head or tail. We further assume that each agent starts at time $i = 0$ with an initial belief that is uniform over Θ and then updates it over time according to the model described in (2). Then, we know from [1] that $\lim_{i\to\infty} \boldsymbol{\mu}_{k,i}(\theta_1^\circ) = 1$ for $k = 1, 2, 3$ and $\lim_{i\to\infty} \boldsymbol{\mu}_{k,i}(\theta_2^\circ) = 1$ for $k = 4, 5$. Now for the agents of the receiving sub-network, we compute W using (10):

$$W^{\mathsf{T}} = \left[\begin{array}{ccc|cc} 0 & 0.4045 & 0.1489 & 0.4466 & 0 \\ 0 & 0.5267 & 0.1183 & 0.3550 & 0 \\ 0 & 0.7099 & 0.0725 & 0.2176 & 0 \end{array}\right] \quad (23)$$

The rows of W^{T} correspond respectively to agents 6, 7 and 8. Each row is partitioned into two blocks: the first block is of length $N_1 = 3$ that corresponds to θ_1° and the second block is of length $N_2 = 2$ that corresponds to θ_2°. Then, by (19), we compute the limiting beliefs at θ_1° and θ_2°, by summing the elements of the corresponding blocks:

$$\begin{bmatrix} q_6(\theta_1^\circ) \\ q_7(\theta_1^\circ) \\ q_8(\theta_1^\circ) \end{bmatrix} = \begin{bmatrix} 0.5534 \\ 0.6450 \\ 0.7824 \end{bmatrix}, \quad \begin{bmatrix} q_6(\theta_2^\circ) \\ q_7(\theta_2^\circ) \\ q_8(\theta_2^\circ) \end{bmatrix} = \begin{bmatrix} 0.4466 \\ 0.3550 \\ 0.2176 \end{bmatrix} \quad (24)$$

We run this example for 5000 time iterations. Figure 3 shows the evolution of $\boldsymbol{\mu}_{k,i}(\theta_1^\circ)$ and $\boldsymbol{\mu}_{k,i}(\theta_2^\circ)$ of agents in the receiving sub-network ($k = 6, 7, 8$), and their convergence to the same probability distribution (24) computed according to (19).

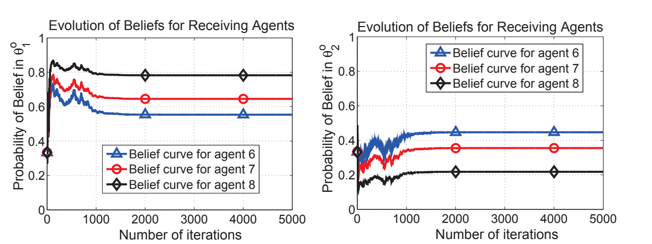

Fig. 3: Evolution of agent k beliefs $\boldsymbol{\mu}_{k,i}(\theta_1^\circ)$ and $\boldsymbol{\mu}_{k,i}(\theta_2^\circ)$ over time for $k = 6, 7, 8$.

6. REFERENCES

[1] X. Zhao and A. H. Sayed, "Learning over social networks via diffusion adaptation," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, 2012, pp. 709–713.

[2] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi, "Non-Bayesian social learning," *Games and Economic Behavior*, vol. 76, no. 1, pp. 210–225, 2012.

[3] A. H. Sayed, "Adaptive networks," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 460–497, 2014.

[4] A. H. Sayed, "Adaptation, learning, and optimization over networks," *Foundations and Trends in Machine Learning*, vol. 7, no. 4-5, pp. 311–801, 2014.

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

[6] D. Acemoglu, A. Ozdaglar, and A. ParandehGheibi, "Spread of (mis)information in social networks," *Games and Economic Behavior*, vol. 70, no. 2, pp. 194–227, 2010.

[7] L. Epstein, J. Noor, and A. Sandroni, "Non-Bayesian learning," *The BE Journal of Theoretical Economics*, vol. 10, no. 1, 2010.

[8] B. Golub and M. Jackson, "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, vol. 2, no. 1, pp. 112–49, 2010.

[9] P. Molavi, A. Jadbabaie, K. R. Rad, and A. Tahbaz-Salehi, "Reaching consensus with increasing information," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 358–369, 2013.

[10] C. Chamley, *Rational Herds: Economic Models of Social Learning*, Cambridge University Press, 2004.

[11] L. Smith and P. Sorensen, "Pathological outcomes of observational learning," *Econometrica*, vol. 68, no. 2, pp. 371–398, 2000.

[12] D. Acemoglu, M. Dahleh, A. Ozdaglar, and A. Tahbaz-Salehi, "Observational learning in an uncertain world," in *Proc. IEEE Conf. Decision and Control*, Atlanta, Georgia, Dec 2010, pp. 6645–6650.

[13] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *The Review of Economic Studies*, vol. 78, no. 4, pp. 1201–1236, 2011.

[14] V. Krishnamurthy and H. V. Poor, "Social learning and Bayesian games in multiagent signal processing," *IEEE Signal Processing Magazine*, vol. 33, no. 3, pp. 43–757, 2013.

[15] V. Krishnamurthy, O. N. Gharehshiran, and M. Hamdi, "Interactive sensing and decision making in social networks," *Foundatons and Trends in Signal Processing*, vol. 7, no. 1-2, pp. 1–196, April 2014.

[16] B. Ying and A. H. Sayed, "Information exchange and learning dynamics over weakly-connected adaptive networks," *arXiv:1412.1523*, Dec 2014.

[17] B. Ying and A. H. Sayed, "Learning by weakly-connected adaptive agents," in *Proc. IEEE ICASSP*, Brisbane, Australia, April 2015, pp. 5788–5792.

[18] E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione, "Binary opinion dynamics with stubborn agents," *ACM Trans. Econ. Comput.*, vol. 1, no. 4, pp. 19:1–19:30, Dec. 2013.

[19] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, PA, 2000.

[20] L. Epstein, J. Noor, and A. Sandroni, "Non-Bayesian updating: A theoretical framework," *Theoretical Economics*, vol. 3, no. 2, 2008.