

# Learning without Recall by Random Walks on Directed Graphs

M. A. Rahimian, S. Shahrampour, A. Jadbabaie \*

**Abstract**—We consider a network of agents that aim to learn some unknown state of the world using private observations and exchange of beliefs. At each time, agents observe private signals generated based on the true unknown state. Each agent might not be able to distinguish the true state based only on her private observations. This occurs when some other states are observationally equivalent to the true state from the agent’s perspective. To overcome this shortcoming, agents must communicate with each other to benefit from local observations. We propose a model where each agent selects one of her neighbors randomly at each time. Then, she refines her opinion using her private signal and the prior of that particular neighbor. The proposed rule can be thought of as a Bayesian agent who cannot recall the priors based on which other agents make inferences. This learning without recall approach preserves some aspects of the Bayesian inference while being computationally tractable. By establishing a correspondence with a random walk on the network graph, we prove that under the described protocol, agents learn the truth exponentially fast in the almost sure sense. The asymptotic rate is expressed as the sum of the relative entropies between the signal structures of every agent weighted by the stationary distribution of the random walk.

## I. INTRODUCTION & BACKGROUND

Distributed estimation and detection problems have been interesting subject of study in a variety of disciplines, ranging from control theory to statistics, economics, and signal processing [1]–[8]. In the distributed detection problem, each agent observes a sequence of independent and identically distributed (i.i.d.) private signals generated according to the true (unknown) state. Suppose that each agent forms a belief about the true state, represented by a discrete probability distribution over a *finite* state space, and sequentially performs the Bayes’ rule to her observations at each step. It is well-known [9], [10] that the beliefs formed in the above manner constitute a bounded martingale and converge to a limiting distribution as the number of observations tends to infinity. However, the limiting distribution is not necessarily concentrated on the truth, in which case the agent fails to learn the true state asymptotically. In fact, in many scenarios, the agent faces an identification problem where there are states (other than the true state) that are observationally equivalent to the true state. In other words, these states induce the same distribution on her sequence of privately observed signals. Therefore, rational agents communicate in a social network to distinguish the truth by relying on local observations. This leads to the problem of social learning that is a classical focus of behavioral microeconomic theory [11], [12], also studied in the context of distributed estimation and statistical learning theory [1], [13].

On the other hand, sequentially applying Bayes’ rule in networks can become computationally intractable since the global network structure is not available to individuals. This originates from the fact that agents should use their local data that is increasing with time, and infer about the global signal structure. Therefore, the analysis of rational behavior in networks is an important problem in Bayesian economics, and has attracted a considerable attention [14], [15]. On the other side of the spectrum lie the works such as [16]–[19] which aim to study the problem of learning in networks via iterative applications of non-Bayesian rules. These updates provide the asymptotic properties of learning and consensus under certain conditions. More recently, some works (e.g. see [20], [21]) have also provided the *non-asymptotic* analysis of the problem.

In this paper, we study a distributed learning model where each agent observes a sequence of independent and identically distributed private signals. The structure of the network (which we assume to be strongly connected) is preset in the sense that all agents know their local neighborhood before the learning process. However, they do not necessarily contact all their neighbors every time. At every epoch of time, each agent randomly selects one neighbor, and uses her neighbor’s prior (rather than herself) in the form of the Bayes’ update together with her private signal at that instant of time. This can be seen as a *learning without recall* rule where agents randomly pick their priors from their local neighborhood. Intuitively, asymptotic learning occurs since each agent performs a *random walk* over a strongly connected graph and picks up the privately observed signals of the nodes as they are hit by the random walk. We show that the learning rate for such an agent is exponentially fast with an asymptotic rate that can be expressed as the weighted sum of the relative entropies between the likelihood structures of each agent under various states of the world, and the weights are their probabilities in the stationary distribution of the random walk. In many distributed learning models over random and switching networks, agents must have positive self-reliance at any time. One can observe this condition, for instance, in gossip algorithms [22] and ergodic stationary processes [23]. An interesting and subtle point in our communication structure is the relaxation of this condition, as our agents rely entirely on the beliefs of their neighbors every time that they select a neighbor to gossip with. Moreover, unlike the majority of results that rely on the convergence properties of products of stochastic matrices and are applicable only to irreducible and aperiodic communication matrices, cf. [24, Proposition 1]; our results do not require the transition probability matrix to be aperiodic. This is because our proof of convergence relies on the ergodic theorem for the almost-sure convergence of the long-run fraction of time that is spent in any state of a Markov chain; and it holds true for any irreducible, positive-recurrent

\* The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104-6228 USA (email: jadbabai@seas.upenn.edu). This work was supported by ARO MURI W911NF-12-1-0509.

chain, and in particular any irreducible, finite-state chain [25, Theorem 1.5.6]. It is further true that such a chain has a unique stationary distribution [25, Theorem 1.7.7], which we use to characterize the almost-sure exponentially fast asymptotic rate of convergence under our proposed distributed learning model.

The remainder of this paper is organized as follows. The modeling and formulation are set forth in Section II, where we present the signal and belief structures and their evolution. We end section II by a description of learning without recall updates in sparse structures where the neighborhood of each agent has at most one node. Next in Section III we show how the preceding updates can be used even when the agents' neighborhood are not singletons. This achieved by implementing a gossip-like procedure where a single neighbors is chosen randomly at every time-step and communications are performed with only one neighbor at a time. We study the properties of convergence and learning under this procedure and show a correspondence with random walks on directed graphs that simplifies our analysis. An illustration is provided at the end of Section III, and the paper is concluded by Section IV.

## II. THE MODEL

*Notation:* Throughout the paper,  $\mathbb{R}$  is the set of real numbers,  $\mathbb{N}$  denotes the set of all natural numbers, and  $\mathbb{W} = \mathbb{N} \cup \{0\}$ . For  $n \in \mathbb{N}$  a fixed integer the set of integers  $\{1, 2, \dots, n\}$  is denoted by  $[n]$ , while any other set is represented by a calligraphic capital letter. The cardinality of a set  $\mathcal{X}$ , which is the number of its elements, is denoted by  $|\mathcal{X}|$ , and  $\mathcal{P}(\mathcal{X}) = \{\mathcal{M}; \mathcal{M} \subset \mathcal{X}\}$  denotes the power-set of  $\mathcal{X}$ , which is the set of all its subsets. The difference of two sets  $\mathcal{X}$  and  $\mathcal{Y}$  is defined by  $\mathcal{X} \setminus \mathcal{Y} := \{x; x \in \mathcal{X} \text{ and } x \notin \mathcal{Y}\}$ . Boldface letters denote random variables.

We consider a network of  $n$  agents that interact according to a directed graph  $\mathcal{G} = ([n], \mathcal{E})$ , where  $\mathcal{E} \subset [n] \times [n]$  is the set of directed edges. Each agent is labeled with an element of the set  $[n]$ .  $\mathcal{N}(i) = \{j \in [n]; (j, i) \in \mathcal{E}\}$  is the neighborhood of agent  $i$  which is the set of all agents whose beliefs can be observed by agent  $i$ . We let  $\deg(i) = |\mathcal{N}(i)|$  be the degree of node  $i$  corresponding to the number of agent  $i$ 's neighbors.

*The Environment:* We denote by  $\Theta$  the set of states of the world which has a finite cardinality. Also,  $\Delta\Theta$  represents the space of all probability measures on the set  $\Theta$ . Each agent's goal is to decide amongst the finitely many possibilities in the state space  $\Theta$ . A random variable  $\theta$  is chosen randomly from  $\Theta$  by the nature and according to the probability measure  $\nu(\cdot) \in \Delta\Theta$ , which satisfies  $\nu(\hat{\theta}) > 0, \forall \hat{\theta} \in \Theta$  and is referred to as the common prior. For each agent  $i$ , there exists a finite signal space denoted by  $\mathcal{S}_i$ , and given  $\theta$ ,  $\ell_i(\cdot | \theta)$  is a probability measure on  $\mathcal{S}_i$ , which is referred to as the *signal structure* or *likelihood function* of agent  $i$ . Furthermore,  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability triplet, where

$$\Omega = \Theta \times \left( \prod_{i \in [n]} \mathcal{S}_i \right)^{\mathbb{W}},$$

is an infinite product space with a general element  $\omega = (\theta; (s_{1,0}, \dots, s_{n,0}), (s_{1,1}, \dots, s_{n,1}), \dots)$  and the associated sigma field  $\mathcal{F} = \mathcal{P}(\Omega)$ .  $\mathbb{P}(\cdot)$  is the probability measure on  $\Omega$  which assigns probabilities consistently with the common prior  $\nu(\cdot)$  and the likelihood functions  $\ell_i(\cdot | \theta), i \in [n]$ . Conditioned on  $\theta$ , the random vectors  $\{(s_{1,t}, \dots, s_{n,t}), t \in \mathbb{W}\}$  are independent.  $\mathbb{E}\{\cdot\}$  is the expectation operator, which represents integration with respect to  $d\mathbb{P}(\omega)$ .

*Signals:* Let  $t \in \mathbb{W}$  denote the time index and for each agent  $i$ , define  $\{s_{i,t}, t \in \mathbb{W}\}$  to be a sequence of independent and identically distributed random variables with the probability mass function  $\ell_i(\cdot | \theta)$ ; this sequence represents the private observations made by agent  $i$  at each time period  $t$ . The privately observed signals are independent and identically distributed over time, but they could be correlated across the agents.

*Beliefs:* We let  $\mu_{i,t}(\cdot)$  represent the *opinion* or *belief* at time  $t$  of agent  $i$  about the realized value of  $\theta$ . In other words,  $\mu_{i,t}(\cdot)$  is a probability distribution on the set  $\Theta$  at any time  $t$  formed by agent  $i$ . Note the randomness of  $\mu_{i,t}(\cdot)$  due to its dependence on the random observations of the agent. The goal is to study asymptotic learning, i.e. for each agent to learn the true realized value  $\theta \in \Theta$  of  $\theta$  asymptotically. This amounts to having  $\mu_{i,t}(\cdot)$  converge to a point mass centered at  $\theta$ , where the convergence could be in probability or in the stronger almost sure sense that we use in this work.

At  $t = 0$  the value  $\theta = \theta$  is selected by nature. Followed by that,  $s_{i,0}$  for each  $i \in [n]$  is realized and observed by agent  $i$ . Then the agent forms an initial Bayesian opinion  $\mu_{i,0}(\cdot)$  about the value of  $\theta$ . Given  $s_{i,0}$ , and using the Bayes' rule for each agent  $i \in [n]$ , the initial belief in terms of the observed signal  $s_{i,0}$  is given by:

$$\mu_{i,0}(\hat{\theta}) = \frac{\nu(\hat{\theta})\ell_i(s_{i,0} | \hat{\theta})}{\sum_{\tilde{\theta} \in \Theta} \nu(\tilde{\theta})\ell_i(s_{i,0} | \tilde{\theta})}.$$

Afterwards, at any time  $t$  each agent  $i$  observes the realized value of  $s_{i,t}$  as well as the current belief of one of her neighbors  $\mu_{k,t-1}(\cdot)$ , where  $k$  is selected randomly from  $\mathcal{N}(i)$ . She then forms a refined opinion  $\mu_{i,t}(\cdot)$  by incorporating all the data that have been made available to her by the time  $t$ . We elaborate on the update rule in the following.

## III. COMBINED GOSSIP AND WITHOUT RECALL UPDATES: SIGNALS PICKED UP IN A RANDOM WALK

Consider a digraph  $\mathcal{G}$  satisfying  $\deg(i) \in \{0, 1\}, \forall i \in [n]$ . For this class of networks which include directed circles and rooted trees in [26], [27], the authors propose to use the Bayesian update

$$\mu_{i,t}(\hat{\theta}) = \frac{\mu_{i,t-1}(\hat{\theta})\ell_i(s_{i,t} | \hat{\theta})}{\sum_{\tilde{\theta} \in \Theta} \mu_{i,t-1}(\tilde{\theta})\ell_i(s_{i,t} | \tilde{\theta})}, \forall \hat{\theta} \in \Theta, \quad (1)$$

if  $\deg(i) = 0$ ; and else to use

$$\mu_{i,t}(\hat{\theta}) = \frac{\mu_{j,t-1}(\hat{\theta})\ell_i(\mathbf{s}_{i,t} | \hat{\theta})}{\sum_{\tilde{\theta} \in \Theta} \mu_{j,t-1}(\tilde{\theta})\ell_i(\mathbf{s}_{i,t} | \tilde{\theta})}, \forall \hat{\theta} \in \Theta, \quad (2)$$

where  $j \in [n]$  is the unique vertex  $j \in \mathcal{N}(i)$ . These updates are a special case of the *Learning without Recall* rules that are developed in a companion paper, and they can describe the behavior of Rational but Memoryless agents who share a common prior  $\nu(\cdot)$  and always interpret their current and observed beliefs as having stemmed from this common prior, thus ignoring their entire history of past observations.

Here we propose the application of the Learning without Recall updates that we described in the previous section to general networks, by requiring that at every time step  $t$ , node  $i$  make a random choice from her set of neighbors  $\mathcal{N}(i)$  and uses that choice for the unique  $j$  in (2). To this end, let  $\sigma_t \in \Pi_{i \in [n]} \mathcal{N}(i)$ ,  $t \in \mathbb{N}$  be a sequence of independent and identically distributed random vectors such that  $\forall t \in \mathbb{N}$ ,  $\sigma_{t,i} \in \mathcal{N}(i)$  is that neighbor of  $i$  which she chooses to communicate with at time  $t$ . Hence, for all  $t$  and any  $i$ , (2) becomes

$$\mu_{i,t}(\hat{\theta}) = \frac{\mu_{\sigma_{t,i},t-1}(\hat{\theta})\ell_i(\mathbf{s}_{i,t} | \hat{\theta})}{\sum_{\tilde{\theta} \in \Theta} \mu_{\sigma_{t,i},t-1}(\tilde{\theta})\ell_i(\mathbf{s}_{i,t} | \tilde{\theta})}, \forall \hat{\theta} \in \Theta. \quad (3)$$

To proceed, annex the random choice of neighbors for every node  $i \in [n]$  and all times  $t \in \mathbb{N}$  to the original probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  specified in Section II; and for  $t \in \mathbb{N}$  arbitrary, let  $\mathbb{P}\{\sigma_{t,i} = j\} = p_{i,j} > 0$ . Wherefore,  $\sum_{j \in \mathcal{N}(i)} p_{i,j} = 1 - p_{i,i} \leq 1$ , and  $p_{i,j} = 0$  whenever  $j \notin \mathcal{N}(i) \cup \{i\}$ . Let  $P$  be the row stochastic matrix whose  $(i, j)$ -th entry is equal to  $p_{i,j}$ . Let  $\mathbb{1}_{\{\sigma_{t,i}=j\}} = 1$  if  $\sigma_{t,i} = j$  and  $\mathbb{1}_{\{\sigma_{t,i}=j\}} = 0$  otherwise. Then (3) can be written as

$$\begin{aligned} \mu_{i,t}(\hat{\theta}) &= \sum_{j=1}^n \mathbb{1}_{\{\sigma_{t,i}=j\}} \frac{\mu_{j,t-1}(\hat{\theta})\ell_i(\mathbf{s}_{i,t} | \hat{\theta})}{\sum_{\tilde{\theta} \in \Theta} \mu_{j,t-1}(\tilde{\theta})\ell_i(\mathbf{s}_{i,t} | \tilde{\theta})} \\ &= \ell_i(\mathbf{s}_{i,t} | \hat{\theta}) \prod_{j=1}^n \left( \frac{\mu_{j,t-1}(\hat{\theta})}{\sum_{\tilde{\theta} \in \Theta} \mu_{j,t-1}(\tilde{\theta})\ell_i(\mathbf{s}_{i,t} | \tilde{\theta})} \right)^{\mathbb{1}_{\{\sigma_{t,i}=j\}}} \end{aligned}$$

To analyze the propagation of beliefs under (3) we form the belief ratio

$$\frac{\mu_{i,t}(\check{\theta})}{\mu_{i,t}(\theta)} = \frac{\ell_i(\mathbf{s}_{i,t} | \check{\theta})}{\ell_i(\mathbf{s}_{i,t} | \theta)} \prod_{j=1}^n \left( \frac{\mu_{j,t-1}(\check{\theta})}{\mu_{j,t-1}(\theta)} \right)^{\mathbb{1}_{\{\sigma_{t,i}=j\}}} \quad (4)$$

for any false state  $\check{\theta} \in \Theta \setminus \{\theta\}$  and each agent  $i \in [n]$  at all times  $t \in \mathbb{N}$ . The above has the advantage of removing the normalization factor in the dominator out of the picture; thence, focusing instead on the evolution of belief ratios. To

proceed, we take the logarithms of both sides in (4) to obtain

$$\begin{aligned} \log \left( \frac{\mu_{i,t}(\check{\theta})}{\mu_{i,t}(\theta)} \right) &= \log \left( \frac{\ell_i(\mathbf{s}_{i,t} | \check{\theta})}{\ell_i(\mathbf{s}_{i,t} | \theta)} \right) \\ &\quad + \sum_{j=1}^n \mathbb{1}_{\{\sigma_{t,i}=j\}} \log \left( \frac{\mu_{j,t-1}(\check{\theta})}{\mu_{j,t-1}(\theta)} \right) \end{aligned} \quad (5)$$

Next we can iterate (5) to replace for  $(\mu_{j,t-1}(\check{\theta})/\mu_{j,t-1}(\theta))$  and so on, from which we get (6) at the top of next page. Also note,

$$\sum_{i_1=1}^n \cdots \sum_{i_t=1}^n \mathbb{1}_{\{\sigma_{t,i}=i_1\}} \cdots \mathbb{1}_{\{\sigma_{1,i_{t-1}}=i_t\}} = 1,$$

almost surely, and in fact every where on  $\Omega$ , so that the initial prior belief ratio  $\log(\nu(\check{\theta})/\nu(\theta))$  always appears in the summation (6), and it simplifies as in (7) at the top of next page.

We now claim that whenever  $t \rightarrow \infty$  and the network graph  $\mathcal{G}$  is strongly connected, with  $\mathbb{P}$ -probability one the likelihood ratios of private signals from any node  $m \in [n]$  appears in the summation (7) as  $\ell_m(\mathbf{s}_{m,t-\tau} | \check{\theta})/\ell_m(\mathbf{s}_{m,t-\tau} | \theta)$  for infinitely many values of  $\tau$ . The gist of the proof is in realizing the correspondence between the summation (6) and a random walk on the directed graph  $\mathcal{G}$  that starts at time  $t$  on node  $i$ , proceeds in the reversed time direction, and terminates at time zero. The jumps in this random walk are made from each node  $i$  to one of her in-neighbors  $j \in \mathcal{N}(i)$  and in accordance with the probabilities  $p_{i,j}$  specified by matrix  $P = [p_{i,j}]$ . Indeed, we can denote the random sequence of nodes that are hit by this random walk as  $(i, \mathbf{i}_1, \dots, \mathbf{i}_t)$  where the random variables  $\mathbf{i}_\tau \in [n]$ ,  $\tau \in [t]$  are defined recursively by  $\mathbf{i}_1 := \sigma_{t,i}$ ,  $\mathbf{i}_2 := \sigma_{t-1,\sigma_{t,i}}$ ,  $\mathbf{i}_3 := \sigma_{t-2,\sigma_{t-1,i_2}}$ ,  $\dots$ ,  $\mathbf{i}_t := \sigma_{1,\sigma_{2,i_{t-1}}}$ . Whence (7) is written succinctly as

$$\begin{aligned} \log \left( \frac{\mu_{i,t}(\check{\theta})}{\mu_{i,t}(\theta)} \right) &= \log \left( \frac{\ell_i(\mathbf{s}_{i,t} | \check{\theta})}{\ell_i(\mathbf{s}_{i,t} | \theta)} \right) + \log \left( \frac{\nu(\check{\theta})}{\nu(\theta)} \right) \\ &\quad + \sum_{\tau=1}^t \log \left( \frac{\ell_{\mathbf{i}_\tau}(\mathbf{s}_{\mathbf{i}_\tau,t-\tau} | \check{\theta})}{\ell_{\mathbf{i}_\tau}(\mathbf{s}_{\mathbf{i}_\tau,t-\tau} | \theta)} \right). \end{aligned} \quad (8)$$

As  $t \rightarrow \infty$ , the sequence  $\mathbf{i}_\tau, \tau \in \mathbb{N}$  forms a Markov process with transition matrix  $P$ . Given (8), our claim can be restated as that for every  $m \in [n]$  and as  $t \rightarrow \infty$  there are infinitely many values of  $\tau \in \mathbb{N}$  for which  $\mathbf{i}_\tau = m$ , and it is true because in a finite state Markov chain with transition matrix  $P$  every state is persistent (recurrent) and will be hit infinitely many times provided that the directed graph  $\mathcal{G}$  is strongly connected [25, Theorem 1.5.6], i.e. we have that  $\forall m \in [n]$ ,

$$\mathbb{P}\{\mathbf{i}_\tau = m, \text{ for infinitely many } \tau\} = 1.$$

For any agent  $m \in [n]$  let  $\mathcal{T}_m := \{\tau_{m,j}, j \in \mathbb{N}\}$  be the sequence of stopping times that record the first, second and so on passage times of node  $m$  by the process  $\mathbf{i}_\tau, \tau \in \mathbb{N}$ . That is we have  $\tau_{m,1} = \inf\{\tau \in \mathbb{N} : \mathbf{i}_\tau = m\}$  and for  $j > 1$ ,  $\tau_{m,j} = \inf\{\tau > \tau_{m,j-1} : \mathbf{i}_\tau = m\}$ . Using the above notation,

$$\begin{aligned}
\log \left( \frac{\boldsymbol{\mu}_{i,t}(\check{\theta})}{\boldsymbol{\mu}_{i,t}(\theta)} \right) &= \log \left( \frac{\ell_i(\mathbf{s}_{i,t} | \check{\theta})}{\ell_i(\mathbf{s}_{i,t} | \theta)} \right) + \sum_{i_1=1}^n \mathbb{1}_{\{\sigma_{t,i}=i_1\}} \log \left( \frac{\ell_{i_1}(\mathbf{s}_{i_1,t-1} | \check{\theta})}{\ell_{i_1}(\mathbf{s}_{i_1,t-1} | \theta)} \right) \\
&+ \mathbb{1}_{\{\sigma_{t,i}=i_1\}} \sum_{i_2=1}^n \mathbb{1}_{\{\sigma_{t-1,i_1}=i_2\}} \log \left( \frac{\ell_{i_2}(\mathbf{s}_{i_2,t-2} | \check{\theta})}{\ell_{i_2}(\mathbf{s}_{i_2,t-2} | \theta)} \right) \\
&+ \mathbb{1}_{\{\sigma_{t,i}=i_1\}} \mathbb{1}_{\{\sigma_{t-1,i_1}=i_2\}} \sum_{i_3=1}^n \mathbb{1}_{\{\sigma_{t-2,i_2}=i_3\}} \log \left( \frac{\ell_{i_3}(\mathbf{s}_{i_3,t-3} | \check{\theta})}{\ell_{i_3}(\mathbf{s}_{i_3,t-3} | \theta)} \right) + \dots \\
&+ \mathbb{1}_{\{\sigma_{t,i}=i_1\}} \mathbb{1}_{\{\sigma_{t-1,i_1}=i_2\}} \dots \mathbb{1}_{\{\sigma_{1,i_{t-2}}=i_{t-1}\}} \sum_{i_t=1}^n \mathbb{1}_{\{\sigma_{1,i_{t-1}}=i_t\}} \left\{ \log \left( \frac{\ell_{i_t}(\mathbf{s}_{i_t,0} | \check{\theta})}{\ell_{i_t}(\mathbf{s}_{i_t,0} | \theta)} \right) + \log \left( \frac{\nu(\check{\theta})}{\nu(\theta)} \right) \right\} \quad (6)
\end{aligned}$$

$$\begin{aligned}
\log \left( \frac{\boldsymbol{\mu}_{i,t}(\check{\theta})}{\boldsymbol{\mu}_{i,t}(\theta)} \right) &= \log \left( \frac{\ell_i(\mathbf{s}_{i,t} | \check{\theta})}{\ell_i(\mathbf{s}_{i,t} | \theta)} \right) + \log \left( \frac{\nu(\check{\theta})}{\nu(\theta)} \right) + \sum_{i_1=1}^n \mathbb{1}_{\{\sigma_{t,i}=i_1\}} \left\{ \log \left( \frac{\ell_{i_1}(\mathbf{s}_{i_1,t-1} | \check{\theta})}{\ell_{i_1}(\mathbf{s}_{i_1,t-1} | \theta)} \right) \right. \\
&+ \sum_{i_2=1}^n \mathbb{1}_{\{\sigma_{t,i_1}=i_2\}} \left\{ \log \left( \frac{\ell_{i_2}(\mathbf{s}_{i_2,t-2} | \check{\theta})}{\ell_{i_2}(\mathbf{s}_{i_2,t-2} | \theta)} \right) + \dots + \sum_{i_\tau=1}^n \mathbb{1}_{\{\sigma_{t-\tau+1,i_{\tau-1}}=i_\tau\}} \left\{ \log \left( \frac{\ell_{i_\tau}(\mathbf{s}_{i_\tau,t-\tau} | \check{\theta})}{\ell_{i_\tau}(\mathbf{s}_{i_\tau,t-\tau} | \theta)} \right) \right. \\
&+ \dots + \sum_{i_{t-1}=1}^n \mathbb{1}_{\{\sigma_{1,i_{t-2}}=i_{t-1}\}} \left\{ \log \left( \frac{\ell_{i_{t-1}}(\mathbf{s}_{i_{t-1},1} | \check{\theta})}{\ell_{i_{t-1}}(\mathbf{s}_{i_{t-1},1} | \theta)} \right) + \sum_{i_t=1}^n \mathbb{1}_{\{\sigma_{1,i_{t-1}}=i_t\}} \log \left( \frac{\ell_{i_t}(\mathbf{s}_{i_t,0} | \check{\theta})}{\ell_{i_t}(\mathbf{s}_{i_t,0} | \theta)} \right) \right\} \dots \left. \right\} \quad (7)
\end{aligned}$$

(8) can be rewritten as

$$\begin{aligned}
\log \left( \frac{\boldsymbol{\mu}_{i,t}(\check{\theta})}{\boldsymbol{\mu}_{i,t}(\theta)} \right) &= \log \left( \frac{\ell_i(\mathbf{s}_{i,t} | \check{\theta})}{\ell_i(\mathbf{s}_{i,t} | \theta)} \right) + \log \left( \frac{\nu(\check{\theta})}{\nu(\theta)} \right) \\
&+ \sum_{m=1}^n \sum_{\substack{\tau \in \mathcal{T}_m, \\ \tau \leq t}} \log \left( \frac{\ell_m(\mathbf{s}_{m,t-\tau} | \check{\theta})}{\ell_m(\mathbf{s}_{m,t-\tau} | \theta)} \right). \quad (9)
\end{aligned}$$

On the other hand, note that  $\log(\ell_m(\mathbf{s}_{m,t-\tau_{m,j}} | \check{\theta}) / \ell_m(\mathbf{s}_{m,t-\tau_{m,j}} | \theta))$ ,  $j \in \mathbb{N}$  is a sequence of independent and identically distributed signals, so that by the strong of large numbers we obtain that with  $\mathbb{P}$ -probability one,

$$\begin{aligned}
&\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \log \left( \frac{\ell_m(\mathbf{s}_{m,t-\tau_{m,j}} | \check{\theta})}{\ell_m(\mathbf{s}_{m,t-\tau_{m,j}} | \theta)} \right) \\
&= \mathbb{E} \log \left( \frac{\ell_m(\mathbf{s}_{m,0} | \check{\theta})}{\ell_m(\mathbf{s}_{m,0} | \theta)} \right) := -D_{KL}(\ell_m(\cdot | \check{\theta}) \| \ell_m(\cdot | \theta)) \leq 0, \quad (10)
\end{aligned}$$

where the non-positivity follows from the information inequality for the Kullback-Leibler divergence  $D_{KL}(\cdot \| \cdot)$  and is strict whenever  $\ell_m(\cdot | \check{\theta}) \not\equiv \ell_m(\cdot | \theta)$ , i.e.  $\exists s \in \mathcal{S}_i$  such that  $\ell_i(s | \check{\theta}) \neq \ell_i(s | \theta)$  [28, Theorem 2.6.3]. Note that whenever  $\ell_i(\cdot | \hat{\theta}) \equiv \ell_i(\cdot | \theta)$  or equivalently  $D_{KL}(\ell_i(\cdot | \hat{\theta}) \| \ell_i(\cdot | \theta)) = 0$ , then the two states  $\hat{\theta}$  and  $\theta$  are statically indistinguishable to agent  $i$ . In other words, there is no way for agent  $i$  to differentiate  $\hat{\theta}$  from  $\theta$  based only on her private signals. This follows from the fact that both  $\theta$  and  $\hat{\theta}$  induce the same probability distribution on her sequence of observed i.i.d. signals. On the other hand, having  $D_{KL}(\ell_m(\cdot | \theta) \| \ell_m(\cdot | \check{\theta})) < 0$  for

some agent  $m \in [n]$  would ensure per (10) and persistence of state  $m$  that with  $\mathbb{P}$ -probability one,

$$\sum_{\substack{\tau \in \mathcal{T}_m, \\ \tau \leq t}} \log \left( \frac{\ell_m(\mathbf{s}_{m,t-\tau} | \check{\theta})}{\ell_m(\mathbf{s}_{m,t-\tau} | \theta)} \right) \rightarrow -\infty$$

as  $t \rightarrow \infty$  in (9); consequently,  $\log(\boldsymbol{\mu}_{i,t}(\check{\theta}) / \boldsymbol{\mu}_{i,t}(\theta)) \rightarrow -\infty$  for all agent  $i \in [n]$  and any such  $\check{\theta} \in \Theta$ ,  $\check{\theta} \neq \theta$ . Indeed, having  $\log(\boldsymbol{\mu}_{i,t}(\check{\theta}) / \boldsymbol{\mu}_{i,t}(\theta)) \rightarrow -\infty$  for all  $\check{\theta} \neq \theta$  is necessary and sufficient for learning, and we therefore, have the following characterization.

**Definition 1** (Global Identifiability). *In a strongly connected topology, the true state  $\theta$  is globally identifiable, if for all  $\check{\theta} \neq \theta$  there exists some agent  $m \in [n]$  such that  $D_{KL}(\ell_m(\cdot | \theta) \| \ell_m(\cdot | \check{\theta})) < 0$ , i.e.  $m$  can distinguish between  $\check{\theta}$  and  $\theta$  based only on her private signals.*

We have thus established the conditions for learning under the without recall updates in (1) and (2), where the neighbor  $j$  is chosen randomly with strictly positive probabilities specified in transition matrix  $P$ . We dub this procedure “gossips without recall” and summarize our findings as follows:

**Theorem 1** (Almost-Sure Learning). *Under the gossips without recall updates in a strongly connected network where the truth is globally identifiable, all agents learn the truth asymptotically almost surely.*

We can extend the above analysis to derive an asymptotic rate of learning for the agents that is exponentially fast and is expressed as  $\sum_{m=1}^m \pi_m D_{KL}(\ell_m(\cdot | \theta) \| \ell_m(\cdot | \check{\theta})) < 0$ , where  $\bar{\pi} := (\pi_1, \dots, \pi_n)$  is the stationary distribution of the

transition matrix  $P$ , which for a strongly connected  $\mathcal{G}$  is the unique probability distribution on  $[n]$  satisfying  $\bar{\pi}P = \bar{\pi}$ . To see how, for each agent  $m \in [n]$  and all time  $t$ , define  $\mathcal{T}_m(t) := \{\tau_{m,j}, j \in \mathbb{N} : \tau_{m,j} \leq t\}$  and divide both sides of (9) by  $t$  to obtain

$$\begin{aligned} \frac{1}{t} \log \left( \frac{\boldsymbol{\mu}_{i,t}(\check{\theta})}{\boldsymbol{\mu}_{i,t}(\theta)} \right) &= \frac{1}{t} \log \left( \frac{\ell_i(\mathbf{s}_{i,t} | \check{\theta})}{\ell_i(\mathbf{s}_{i,t} | \theta)} \right) + \frac{1}{t} \log \left( \frac{\nu(\check{\theta})}{\nu(\theta)} \right) \\ &\quad + \frac{1}{t} \sum_{m=1}^n \sum_{\tau \in \mathcal{T}_m(t)} \log \left( \frac{\ell_m(\mathbf{s}_{m,t-\tau} | \check{\theta})}{\ell_m(\mathbf{s}_{m,t-\tau} | \theta)} \right). \end{aligned}$$

Upon invoking (10) we obtain

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{1}{t} \log \left( \frac{\boldsymbol{\mu}_{i,t}(\check{\theta})}{\boldsymbol{\mu}_{i,t}(\theta)} \right) &= \\ - \sum_{m=1}^n \lim_{t \rightarrow \infty} \frac{|\mathcal{T}_m(t)|}{t} D_{KL}(\ell_m(\cdot | \theta) \| \ell_m(\cdot | \check{\theta})). \quad (11) \end{aligned}$$

Finally the ergodic theorem ensures that the average time spent in any state  $m \in [n]$  converges almost surely to its stationary probability  $\pi_m$ , i.e. with probability one  $\lim_{t \rightarrow \infty} |\mathcal{T}_m(t)|/t = \pi_m$ , [25, Theorem 1.10.2]. Hence, (11) becomes

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \left( \frac{\boldsymbol{\mu}_{i,t}(\check{\theta})}{\boldsymbol{\mu}_{i,t}(\theta)} \right) = - \sum_{m=1}^n \pi_m D_{KL}(\ell_m(\cdot | \theta) \| \ell_m(\cdot | \check{\theta})),$$

completing the proof for the claimed asymptotically exponentially fast rate.

**Example 1.** *Eight Agents with Binary Signals in a Tri-State World.*

As an illustration consider the network of agents in Fig. 1 with the true state of the world being 1, the first of the tree possible states  $\Theta = \{1, 2, 3\}$ . The likelihood structure for the first three agents is given in the table and note that none of them can learn the truth on their own; indeed, agent 3 does not receive any informative signals and her beliefs shall never depart from their initial priors following (1). We further set  $l_j(\cdot | \cdot) \equiv l_3(\cdot | \cdot)$  for all  $j \in [8] \setminus \{3\}$ , so that all the remaining agents are also unable to infer anything about the true state of the world from their own private signals.

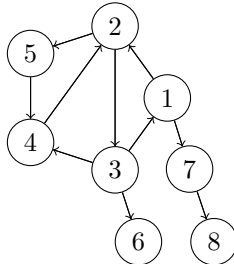


Fig. 1: Network Structure for Example 1

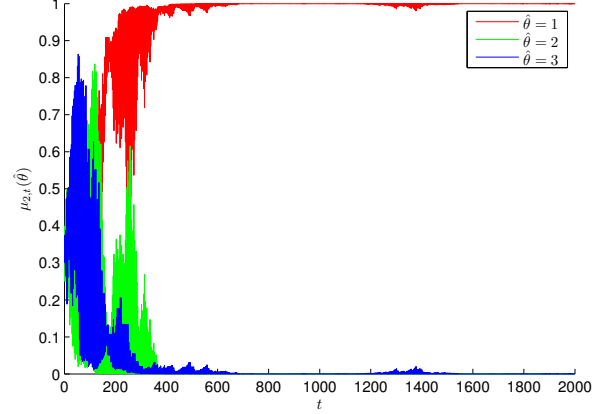


Fig. 2: Evolution of the second agent's beliefs over time

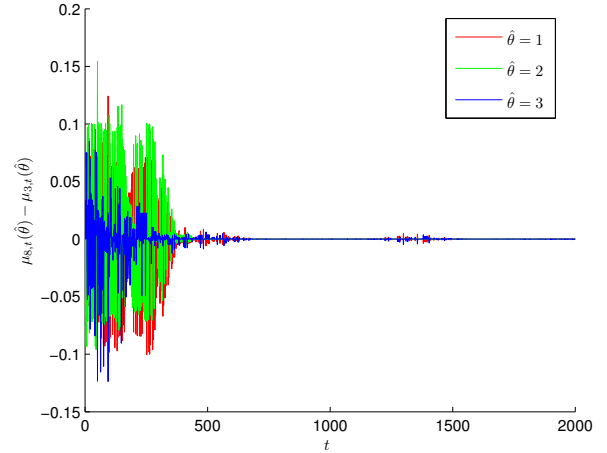


Fig. 3: The difference between the third and eighth agent's beliefs over time

likelihoods	$\hat{\theta} = 1$	$\hat{\theta} = 2$	$\hat{\theta} = 3$
$l_1(\mathbf{s}_{1,t} = 0   \hat{\theta})$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{5}$
$l_2(\mathbf{s}_{2,t} = 0   \hat{\theta})$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{2}$
$l_3(\mathbf{s}_{3,t} = 0   \hat{\theta})$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

Starting from a uniform common prior and following the proposed gossip without recall scheme with neighbors chosen uniformly at random, all agents asymptotically learn the true state, even though none of them can learn the true state on their own. The plots in Figs. 2 and 3 depict the belief evolution for the second agent, as well as the difference between the beliefs for the third and eighth agents. It is further observable that all agents learn the true state at the same exponentially fast asymptotic rate of learning.

#### IV. CONCLUDING REMARKS

This work addressed a social and observational learning model in multi-agent networks. Agents attempt to learn some

unknown state of the world which belongs to a finite state space. Conditioned on the true state, a sequence of i.i.d. private signals are generated and observed by each agent of the network. The private signals do not provide each agent with adequate information to identify the truth. Hence, agents contact their neighbors to augment their imperfect observations with those of their neighbors. In our model, every time, each agent picks a neighbor randomly and updates her belief using the prior of that particular neighbor but using the likelihood for her own private signal. The communication protocol is an instance of a *learning without recall* and is implemented in such a way that signals likelihoods that comprise an agent's belief are picked up by a random walk on the network graph. We proved that agents learn the truth exponentially fast and in the almost sure sense, provided that the network is strongly and the truth is globally identifiable. The asymptotic rate is expressed as a weighted sum of the relative entropies between the signal structures of each agent, where the weights come from the stationary distribution of the transition probability matrix according to which neighbors are chosen at every time instant.

#### REFERENCES

- [1] V. Borkar and P. Varaiya, "Asymptotic agreement in distributed estimation," *Automatic Control, IEEE Transactions on*, vol. 27, no. 3, pp. 650–655, Jun 1982.
- [2] J. N. Tsitsiklis *et al.*, "Decentralized detection," *Advances in Statistical Signal Processing*, vol. 2, pp. 297–344, 1993.
- [3] C. Lopes and A. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4064–4077, 2007.
- [4] F. Bullo, J. Cortés, and S. Martínez, *Distributed control of robotic networks: a mathematical approach to motion coordination algorithms*. Princeton Univ Pr, 2009.
- [5] S. Kar and J. Moura, "Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 355–369, 2009.
- [6] K. Drakopoulos, A. Ozdaglar, and J. N. Tsitsiklis, "On learning with finite memory," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6859–6872, 2013.
- [7] S. Shahrampour, S. Rakhlin, and A. Jadbabaie, "Online learning of dynamic parameters in social networks," in *Advances in Neural Information Processing Systems*, 2013.
- [8] M. Alanyali, S. Venkatesh, O. Savas, and S. Aeron, "Distributed bayesian hypothesis testing in sensor networks," in *IEEE American Control Conference (ACC)*, vol. 6, 2004, pp. 5369–5374.
- [9] D. Blackwell and L. Dubins, "Merging of opinions with increasing information," *The Annals of Mathematical Statistics*, vol. 33, pp. 882 – 886, 1962.
- [10] E. Lehrer and R. Smorodinsky, "Merging and learning," *Lecture Notes-Monograph Series*, pp. 147–168, 1996.
- [11] C. P. Chamley, *Rational Herds: Economic Models of Social Learning*. Cambridge University Press, 2004.
- [12] M. O. Jackson, *Social and Economic Networks*. Princeton, NJ, USA: Princeton University Press, 2008.
- [13] K. Rahnama Rad and A. Tahbaz-Salehi, "Distributed parameter estimation in networks," in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, Dec. 2010, pp. 5050–5055.
- [14] M. Mueller-Frank, "A general framework for rational learning in social networks," *Theoretical Economics*, vol. 8, no. 1, pp. 1–40, 2013.
- [15] E. Mossel, A. Sly, and O. Tamuz, "Asymptotic learning on bayesian social networks," *Probability Theory and Related Fields*, vol. 158, no. 1-2, pp. 127–157, 2014.
- [16] 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.
- [17] S. Shahrampour and A. Jadbabaie, "Exponentially fast parameter estimation in networks using distributed dual averaging," in *52nd IEEE Conference on Decision and Control (CDC)*, 2013, pp. 6196–6201.
- [18] A. Lalitha, A. Sarwate, and T. Javidi, "Social learning and distributed hypothesis testing," in *2014 IEEE International Symposium on Information Theory (ISIT)*, June 2014, pp. 551–555.
- [19] S. Shahrampour, M. A. Rahimian, and A. Jadbabaie, "Switching to learn," in *American Control Conference (ACC)*, July 2015, pp. 2918–2923.
- [20] S. Shahrampour, A. Rakhlin, and A. Jadbabaie, "Distributed detection: Finite-time analysis and impact of network topology," *arXiv preprint arXiv:1409.8606*, 2014.
- [21] A. Nedic, A. Olshevsky, and C. Uribe, "Nonasymptotic convergence rates for cooperative learning over time-varying directed graphs," in *American Control Conference (ACC)*, July 2015, pp. 5884–5889.
- [22] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.
- [23] A. Tahbaz-Salehi and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 225–230, 2010.
- [24] B. Golub and M. O. Jackson, "Naïve Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, vol. 2, no. 1, pp. 112–149, Feb. 2010.
- [25] J. R. Norris, *Markov Chains*. Cambridge University Press, 1999.
- [26] M. A. Rahimian and A. Jadbabaie, "Learning without recall in directed circles and rooted trees," in *American Control Conference*, 2015.
- [27] M. A. Rahimian, P. Molavi, and A. Jadbabaie, "(Non-) bayesian learning without recall," in *IEEE Conference on Decision and Control (CDC)*, 2014, pp. 5730–5735.
- [28] T. Cover and J. Thomas, *Elements of Information Theory*, ser. A Wiley-Interscience publication. Wiley, 2006.