

SEQUENTIAL BAYESIAN LEARNING IN LINEAR NETWORKS WITH RANDOM DECISION MAKING

Yunlong Wang and Petar M. Djurić

Department of Electrical and Computer Engineering
Stony Brook University, Stony Brook, NY 11794
Email: {yunlong.wang, petar.djuric@stonybrook.edu}

ABSTRACT

In this paper, we consider the problem of social learning when decisions by agents in a network are made randomly. The agents receive private signals and use them for decision making on binary hypotheses under which the signals are generated. The agents make the decisions sequentially one at a time. All the agents know the decisions of the previous agents. We study a setting where the agents instead of making deterministic decisions by maximizing personal expected utility, they act randomly according to their private beliefs. We propose a method by which the agents learn from the previous agents' random decisions using the Bayesian theory. We define the concept of social belief about the truthfulness of the two hypotheses and analyze its convergence. We provide performance and convergence analysis of the proposed method as well as simulation results that include comparisons with a deterministic decision making system.

Index Terms— social learning, Bayesian learning, information aggregation, multiagent system, decision

1. INTRODUCTION

It is an important issue in social learning how agents make decisions and learn from others' decisions. When individuals make decisions based on private and imperfect information, it is natural that they also learn from others' decisions made in similar situations. The agents can then make their own decisions with the purpose of maximizing some utility function. A large body of literature investigates this problem including non-Bayesian social learning approaches [1, 2, 3, 4], and Bayesian social learning methods [5, 6, 7, 8, 9]. In this paper, we focus on the Bayesian methods. Some work on this issue can be found in [5, 10, 11], where the interest is to study herding behavior and information cascade. In other works [6, 12], the conditions for asymptotic learning are studied. Recently, the concern of the effect on agent's utility function from previous decisions is addressed in [13]. An overview of models and techniques for studying social learning can be found in [5, 8].

In this paper, our contribution lies in that we consider randomness in an agent's decision making. In most of the current literature, the agents are assumed to be "rational", which means that each agent makes decision in order to maximize the expected value of its utility function. Once the agents obtain information and form their beliefs on the true state of nature, they behave deterministically. However, in some scenarios, the behavior of agent may be by design random. In this work, we replace the "rational" assumption of the agent behavior with a model of random decision making that establishes a random mapping from the agent's belief to the action space. We

present a method that implements Bayesian learning by the agents in a sequential decision making scenario, there each agent has the ability to use Bayes' rule to form its belief on the true state of nature based on its private observation and the actions made by previous agents. We analyze the evolution of beliefs in this system theoretically and demonstrate its performance by simulations.

The paper is organized as follows. In the next section we state the problem. In Section 3, we present the proposed Bayesian learning method. We prove the convergence of the social belief and actions in Section 4. Simulation results are provided in Section 5, and concluding remarks in Section 6.

2. PROBLEM STATEMENT

We consider the sequential decision making problem in linear networks of Bayesian agents A_n , $n \in \mathbb{N}^+$, where each agent makes a decision after it gets its private observation and the decisions of all the previous agents. Mathematically, each agent A_n receives an independent private observation y_n that is generated according to one of the following two hypotheses:

$$\begin{aligned} \mathcal{H}_1 : & \quad y_n = \theta + w_n, \\ \mathcal{H}_0 : & \quad y_n = w_n, \end{aligned} \quad (1)$$

where θ is known and w_n is the observation noise modeled as a Gaussian random variable with zero-mean and known variance σ_w^2 . Without loss of generality, we assume that $\theta > 0$. For every agent, its prior probabilities of the hypotheses are noninformative, where we let $p(\mathcal{H}_0) = p(\mathcal{H}_1) = 1/2$.

Let $\alpha_n \in \mathcal{A} = \{0, 1\}$ be the decision of agent A_n and $\alpha_{1:n}$ denote the decision sequence from agent A_1 to agent A_n . Then the agent A_n can formulate its private belief in \mathcal{H}_1 , $p(\mathcal{H}_1 | \alpha_{1:n-1}, y_n)$ by using the Bayes rule given by the following equation, $\forall n \in \mathbb{N}^+$,

$$p(\mathcal{H}_1 | \alpha_{1:n-1}, y_n) = \frac{1}{1 + \frac{(1 - \pi_{n-1})p(y_n | \mathcal{H}_0)}{\pi_{n-1}p(y_n | \mathcal{H}_1)}}, \quad (2)$$

where we define π_n to be the social belief as a posterior on \mathcal{H}_1 conditioned on the action sequence until agent A_n , i.e.,

$$\pi_n = p(\mathcal{H}_1 | \alpha_{1:n}), \forall n \in \mathbb{N}^+, \quad (3)$$

with π_0 being defined by $1/2$. Here we remark that the social belief π_{n-1} serves as the prior knowledge of agent A_n before it has its private observation y_n .

For any agent A_n , after obtaining the private belief by the Bayes rule in (2), we assume that it makes its decision by drawing it from

This work was supported by NSF under Award CCF-1320626.

the following probability mass function:

$$p(\alpha_n = k | \alpha_{1:n-1}, y_n) = p(\mathcal{H}_k | \alpha_{1:n-1}, y_n), \quad \forall k \in \{0, 1\}. \quad (4)$$

In other word, agent A_n makes its decision by generating a Bernoulli random variable with probability $p(\mathcal{H}_1 | \alpha_{1:n-1}, y_n)$ to be one as its decision. By contrast, in most of the existing literature, “rational” agents make decision 1 whenever $p(\mathcal{H}_1 | \alpha_{1:n-1}, y_n) > 0.5$ and 0 otherwise. When the utility is 1 if an agent makes the right decision and 0 otherwise, the expected utility of agent A_n is maximized by this deterministic policy. After making the decision α_n , agent A_n broadcasts its decision to all the subsequent agents, and they all update their social beliefs from π_{n-1} to π_n . For clarity, we present a diagram that depicts the random system in Fig. 1.

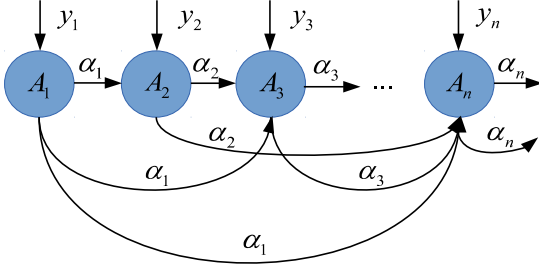


Fig. 1. Sequential decision making based on private signals and all the decisions made by previous agents.

In the following sections, we propose a Bayesian learning method for updating the beliefs of the agents. We show that according to the method, the expected value of social belief asymptotically converges to the true state of nature, or mathematically, we claim that

$$\lim_{n \rightarrow \infty} \mathbb{E}\pi_n = k, \quad \forall k \in \{0, 1\}, \quad (5)$$

where k is the index of the true hypothesis.

3. THE PROPOSED BAYESIAN LEARNING METHOD

We propose a method where the agents update their social beliefs by Bayes' rule from π_{n-1} to π_n . For all the agents A_n , $n \in \mathbb{N}^+$, the social belief is updated according to,

$$\pi_n = \left(\frac{\pi_{n-1}(1 - l_n^{(1)})}{\pi_{n-1}(1 - l_n^{(1)}) + (1 - \pi_{n-1})(1 - l_n^{(0)})} \right)^{1-\alpha_n} \times \left(\frac{\pi_{n-1}l_n^{(1)}}{\pi_{n-1}l_n^{(1)} + (1 - \pi_{n-1})l_n^{(0)}} \right)^{\alpha_n}, \quad (6)$$

where $l_n^{(0)} = p(\alpha_n = 1 | \alpha_{1:n-1}, \mathcal{H}_0)$ and $l_n^{(1)} = p(\alpha_n = 1 | \alpha_{1:n-1}, \mathcal{H}_1)$ denote the probability of agent A_n making decision $\alpha_n = 1$ given the decision sequence up to α_{n-1} and the true state of nature being \mathcal{H}_0 or \mathcal{H}_1 , respectively. For the case where $n = 1$, we define that $l_1^{(0)} = p(\alpha_1 = 1 | \mathcal{H}_0)$ and $l_1^{(1)} = p(\alpha_1 = 1 | \mathcal{H}_1)$.

In order to obtain $l_n^{(0)}$ and $l_n^{(1)}$, we first need to get the probability that $\alpha_n = 1$ with a known observation y_n , and then marginalize y_n .

Given the observation y_n , the agent A_n makes its decision according to its private belief by the following equation:

$$\begin{aligned} p(\alpha_n = 1 | \alpha_{1:n-1}, y_n) &= p(\mathcal{H}_1 | \alpha_{1:n-1}, y_n) \\ &= \frac{1}{1 + e^{-\log \left(\frac{p(\mathcal{H}_1 | \alpha_{1:n-1}, y_n)}{p(\mathcal{H}_0 | \alpha_{1:n-1}, y_n)} \right)}} \\ &= \frac{1}{1 + e^{-\beta(y_n - \gamma_n)}}, \end{aligned} \quad (7)$$

where $\gamma_n = \frac{\theta}{2} - \frac{\sigma_w^2}{\theta} \log \frac{\pi_{n-1}}{1 - \pi_{n-1}}$, and $\beta = \frac{\theta}{\sigma_w^2}$, and where the last sign of equality is due to the data model in (1). Thus, we can write

$$\begin{aligned} \log \left(\frac{p(\mathcal{H}_1 | \alpha_{1:n-1}, y_n)}{p(\mathcal{H}_0 | \alpha_{1:n-1}, y_n)} \right) &= \log \left(\frac{p(\mathcal{H}_1 | \alpha_{1:n-1})p(y_n | \mathcal{H}_1)}{p(\mathcal{H}_0 | \alpha_{1:n-1})p(y_n | \mathcal{H}_0)} \right) \\ &= \log \left(\frac{\pi_{n-1}}{1 - \pi_{n-1}} \right) + \frac{\theta}{\sigma_w^2} y_n - \frac{\theta^2}{2\sigma_w^2} \\ &= \frac{\theta}{\sigma_w^2} (y_n - \gamma_n). \end{aligned} \quad (8)$$

With this we have shown that with π_{n-1} , the agent A_n 's behavior can be modeled by a logistic function where the argument is the observation y_n and the dependent variable is the probability of making a decision of 1. Here we remark that the shaping parameter β is defined by the signal to noise ratio of the data model, and the shift parameter γ_n is just the threshold given by the Bayesian hypothesis testing method. When $y_n = \gamma_n$, we have $p(\mathcal{H}_1 | \alpha_{1:n-1}, y_n) = p(\mathcal{H}_0 | \alpha_{1:n-1}, y_n) = 1/2$.

The agent A_n needs to implement the marginalization of y_n to get $l_n^{(k)}$, $\forall k \in \{0, 1\}$, according to

$$\begin{aligned} l_n^{(k)} &= p(\alpha_n = 1 | \alpha_{1:n-1}, \mathcal{H}_k) \\ &= \int_{-\infty}^{\infty} p(\alpha_n = 1 | \alpha_{1:n-1}, y_n) p(y_n | \mathcal{H}_k) dy_n \\ &= \int_{-\infty}^{\infty} \frac{1}{1 + e^{-\beta(y_n - \gamma_n)}} \frac{1}{\sqrt{2\pi}\sigma_w} \exp \left(-\frac{(y_n - \theta_k)^2}{2\sigma_w^2} \right) dy_n, \end{aligned} \quad (9)$$

where $\theta_0 = 0$ and $\theta_1 = \theta$.

We summarize the behavior of agent A_n at time slot $t \leq n$ by the following steps:

Step 1: If $t = n$, after observing y_n , agent A_n calculates the private belief by (2) and makes a random decision by (4). Otherwise, agent A_n implements step 2.

Step 2: The agent A_n calculates γ_t from the current social belief π_{t-1} by (8) and the two likelihoods $l_t^{(0)}$ and $l_t^{(1)}$ by integration in (9).

Step 3: After observing the action of agent A_t , the agent A_n updates its social belief from π_{t-1} to π_t by (6).

We also point out that the above sequential learning algorithm belongs to the “social learning filter” in [5].

4. ANALYSIS

In this section, we analyze the asymptotic property of the proposed sequential method. As pointed out in (3), the social belief is a deterministic function of the action sequence, whereas the action sequence $\alpha_{1:n}$ is random due to the data model. We first examine the

expected value of the social belief, $\mathbb{E}\pi_n$, which is given by,

$$\begin{aligned}\mathbb{E}\pi_n &= \sum_{\alpha_{1:n} \in \mathcal{A}^n} p(\alpha_{1:n} | \mathcal{H}_1) p(\mathcal{H}_1 | \alpha_{1:n}) \\ &= \sum_{\alpha_{1:n-1} \in \mathcal{A}^{n-1}} p(\alpha_{1:n-1} | \mathcal{H}_1) \\ &\quad \times \sum_{\alpha_n=0}^1 p(\alpha_n | \alpha_{1:n-1}, \mathcal{H}_1) p(\mathcal{H}_1 | \alpha_{1:n}).\end{aligned}\quad (10)$$

Now we present our main analytical result by the following theorem:

Theorem 1 *In the proposed random decision making system, if the agents update their social beliefs using Bayes' rule according to (6), the expected value of social belief asymptotically converge to 1 when the true state of nature is $\mathcal{H}_t = \mathcal{H}_1$, and 0 otherwise, i.e.,*

$$\lim_{t \rightarrow \infty} \mathbb{E}\pi_n[t] = \begin{cases} 1, & \text{if } \mathcal{H}_t = \mathcal{H}_1, \\ 0, & \text{if } \mathcal{H}_t = \mathcal{H}_0. \end{cases}\quad (11)$$

Proof: Considering the symmetric structure of the proposed system, we just prove the convergence of social belief to 1 in the case where the true state of nature is \mathcal{H}_1 . The proof of convergence to 0 can be repeated with just notational changes. Without loss of generality, we assume that the true state of nature is \mathcal{H}_1 .

We first show that $\mathbb{E}\pi_n[t]$ increases in terms of n . By (10), we can write,

$$\mathbb{E}\pi_n - \mathbb{E}\pi_{n-1} = \sum_{\alpha_{1:n-1} \in \mathcal{A}^{n-1}} p(\alpha_{1:n-1} | \mathcal{H}_1) \Delta(\alpha_{1:n-1}), \quad (12)$$

where $\Delta(\alpha_{1:n-1})$ is a function of the action sequence $\alpha_{1:n-1}$, which is given by,

$$\Delta(\alpha_{1:n-1}) = \sum_{\alpha_n=0}^1 p(\alpha_n | \alpha_{1:n-1}, \mathcal{H}_1) p(\mathcal{H}_1 | \alpha_{1:n}) - p(\mathcal{H}_1 | \alpha_{1:n-1}). \quad (13)$$

Therefore, to prove that $\mathbb{E}\pi_n > \mathbb{E}\pi_{n-1}$, it is sufficient to show,

$$\Delta(\alpha_{1:n-1}) > 0, \quad \forall \alpha_{1:n-1} \in \mathcal{A}^{n-1}. \quad (14)$$

By equation (6), we can write,

$$\begin{aligned}\Delta(\alpha_{1:n-1}) &= \frac{\pi_{n-1}(1 - l_n^{(1)})^2}{\pi_{n-1}(1 - l_n^{(1)}) + (1 - \pi_{n-1})(1 - l_n^{(0)})} \\ &\quad + \frac{\pi_{n-1}(l_n^{(1)})^2}{\pi_{n-1}l_n^{(1)} + (1 - \pi_{n-1})l_n^{(0)}} - \pi_{n-1}.\end{aligned}\quad (15)$$

From the above equation, one can derive that,

$$\Delta(\alpha_{1:n-1}) = \frac{\pi_{n-1}}{Z} (1 - \pi_{n-1})^2 (l_n^{(1)} - l_n^{(0)})^2, \quad (16)$$

where Z is the multiplication of the two denominators and is given by $Z = (\pi_{n-1}(1 - l_n^{(1)}) + (1 - \pi_{n-1})(1 - l_n^{(0)}))(\pi_{n-1}l_n^{(1)} + (1 - \pi_{n-1})l_n^{(0)})$.

We next need to show $\forall n \in \mathbb{N}, 0 < \pi_n < 1$. This claim can be proved by mathematical induction. Since we define $\pi_0 = 0.5$, we have $0 < \pi_0 < 1$. Assuming that $0 < \pi_{n-1} < 1$ is true,

then by (6), it holds that $0 < \pi_n < 1$ if and only if $0 < l_n^{(0)} < 1$ and $0 < l_n^{(1)} < 1$. By (9), $\forall k \in \{0, 1\}$, $0 < l_n^{(k)} < 1$ when $\gamma_n = \frac{\theta}{2} - \frac{\sigma_w^2}{\theta} \log \frac{\pi_{n-1}}{1 - \pi_{n-1}}$ is finite, which is guaranteed by the assumption in the induction step that $0 < \pi_{n-1} < 1$. Thus, we have shown that

$$0 < \pi_n < 1, \quad \forall n \in \mathbb{N}. \quad (17)$$

Following the result in (17) and (9), we have $(l_n^{(1)} - l_n^{(0)})^2 > 0$. Then (16) shows that $\Delta(\alpha_{1:n-1}) > 0$, which is a sufficient condition for $\mathbb{E}\pi_n$ to be increasing in terms of n . Considering the fact that $\mathbb{E}\pi_n \leq 1 \quad \forall n \in \mathbb{N}^+$, we have shown that given the true state of nature is \mathcal{H}_1 , the limit of $\mathbb{E}\pi_n$ exists and it is 1. With this, the proof of (11) is completed. \square

From theorem 1, one can immediately state a corollary that both, the social belief π_n and the action of agent A_n , α_n , converge to $k \in \{0, 1\}$ in probability when \mathcal{H}_k is the true state of nature. Then it holds that

$$\lim_{n \rightarrow \infty} p(\pi_n = k) = 1, \quad (18)$$

$$\lim_{n \rightarrow \infty} p(\alpha_n = k) = 1. \quad (19)$$

The statement (18) can be shown by contradiction. Without loss of generality, we continue to assume that $k = 1$. Then if we assume that $p(\pi_n = 1) < 1$ when n goes to infinity, noting that π_n can be no larger than 1, we have that $\mathbb{E}\pi_n < 1$, which is in contradiction with theorem 1. Therefore (18) holds.

For the statement (19), we also assume that $k = 1$. From (2), it holds that when $\pi_n = 1$, the private belief $p(\mathcal{H}_1 | \alpha_{1:n}, y_{n+1}) = 1$ regardless of the observation y_{n+1} . Hence, the event $\alpha_{n+1} = 1$ is a subset of the event $\pi_n = 1$. Therefore we get that $p(\alpha_{n+1} = 1) > p(\pi_n = 1)$. Noting that $p(\alpha_n = 1) < 1$, (19) can be proved by (18).

5. SIMULATION RESULTS

In this section, we present simulation results on the evolution of social belief in our sequential system along with some numerical comparisons with a deterministic decision making method from [8]. In the deterministic method, given the same observation model in (1), after calculating the private belief $p(\mathcal{H}_1 | \alpha_{1:n-1}, y_n)$, the agent A_n made a decision by the following rule:

$$\alpha_n = \begin{cases} 1, & \text{if } p(\mathcal{H}_1 | \alpha_{1:n-1}, y_n) > 1/2, \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

If we set the reward to one when the agent made a decision identical to the true hypothesis, and zero otherwise, by the above rule, the expected utility of the agent A_n was maximized.

In the first experiment, we verified the analytical result of the expected social belief by Monte Carlo methods, which were conducted with 2,000 trials. In each trial, we set the number of agents to be 1000. The parameters were $\sigma_w = 1$, with different $\theta = 1$, $\theta = 0.5$ and $\theta = 0.2$. The results are shown in Fig. 2. On the abscissa, we plotted the agent index and on the ordinate the estimate of the expected social belief, which was given by the average social belief from all the 2,000 trials. The private signals of the agents were generated according to \mathcal{H}_1 .

In Fig. 2, it was shown that the expected social belief was increasing in both methods, and we observed crossovers of the evolution of expected social belief in cases where $\theta = 1$ and $\theta = 0.5$. If we extended the abscissa to $n = 2,000$, we also see the crossover

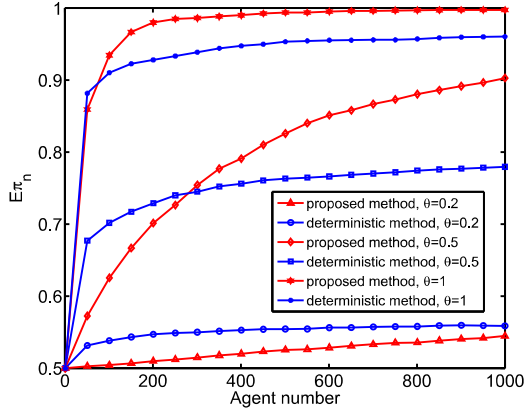


Fig. 2. The convergence of social beliefs with two methods, where the red and blue lines correspond to the random and deterministic decision making methods, respectively.

showing in case where $\theta = 0.2$. It can be seen that the higher signal-to-noise ratio in the data model yielded faster convergence.

In order to show the herding behavior of the agents, we plotted the histogram of social belief in Fig. 3, where the parameters were $\sigma_w = 1$ and $\theta = 0.5$.

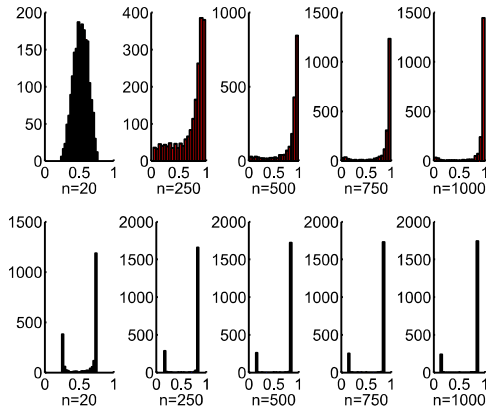


Fig. 3. The histograms of social beliefs with random and deterministic decision making.

In Fig. 3, the upper plot shows the histogram of social belief in the random system for agents A_{20} , A_{250} , A_{500} , A_{750} , A_{1000} , and the lower plot displays the social belief in the deterministic system for the same agents. It can be seen that in the deterministic system, the agents were more likely to show herding behavior entailing that the social belief became almost unchanged. In [10] and [11], more analysis is provided about the herding behavior in deterministic system. It was shown that in these systems, once several consecutive agents made a decision in favor of \mathcal{H}_0 , it was very hard for the subsequent agents to make the opposite decision. In the random system, the agents' behaviors allow for choosing the correct hypothesis even if several consecutive agents made identical decisions, the social belief kept evolving.

In Fig. 4, we plotted the proportion of agents that made the right decision to verify (19). As shown in the figure, the agents with both methods had a trend to make the right decision asymptotically,

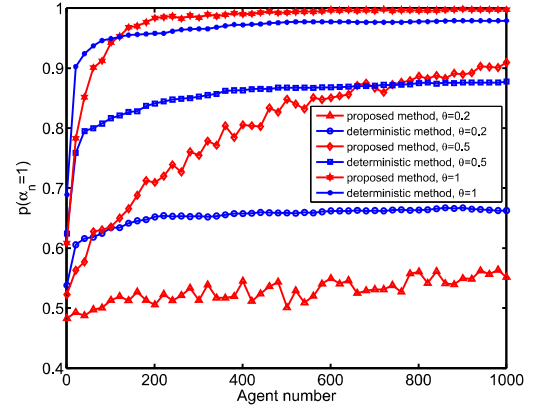


Fig. 4. The proportion of agents making the right decision.

whereas in the random system the agents' actions had some fluctuations.

Finally, we showed the social belief evolution with one outlier. In the second experiment, using the same data model with $\sigma_w = 1$ and $\theta = 1$, we set the sequence of decisions to be $\alpha_{1:40}$ all equal to ones except that $\alpha_{10} = 0$. Figure 4 showed the evolution of social beliefs of the deterministic and proposed methods conditioned on this decision sequence. It can be seen that the social belief in the deterministic method decreased from 0.9 to 0.4 just because one decision α_{10} is in favor of \mathcal{H}_0 . On the contrary, the proposed random method was not so sensitive to the outlier decision.

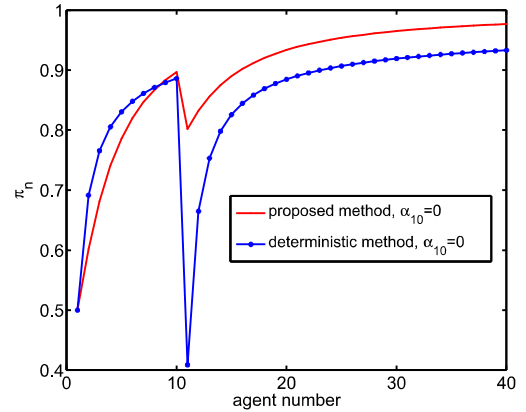


Fig. 5. The evolution of social belief with one outlier decision.

6. CONCLUSION

In this paper we proposed a Bayesian learning method for agents in a sequential random decision making system. We modeled the observations of agents as Gaussian random variables with either zero mean or a known value. The agents drew their decisions randomly from their beliefs about the true state of nature. We proved that by the proposed Bayesian learning method, both the social belief and the action converged to the state of nature in probability. We demonstrated the performance of the proposed method by Monte Carlo simulations and compared it with that of a deterministic method. We showed that the system using random decision policy has a faster convergence rate than the one with deterministic policy.

7. REFERENCES

- [1] A. V. Banerjee, "A simple model of herd behavior," *The Quarterly Journal of Economics*, vol. 107, no. 3, pp. 797–817, 1992.
- [2] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, pp. 992–1026, 1992.
- [3] 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.
- [4] B. Golub and M. O. Jackson, "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, vol. 2, no. 1, pp. 112–149, 2010.
- [5] V. Krishnamurthy and H. V. Poor, "Social learning and Bayesian games in multiagent signal processing: How do local and global decision makers interact?," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 43–57, 2013.
- [6] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *The Review of Economic Studies*, vol. 78, no. 4, pp. 1201–1236, 2011.
- [7] D. Gale and S. Kariv, "Bayesian learning in social networks," *Games and Economic Behavior*, vol. 45, no. 2, pp. 329–346, 2003.
- [8] P. M. Djurić and Y. Wang, "Distributed Bayesian learning in multiagent systems: Improving our understanding of its capabilities and limitations," *IEEE Signal Processing Magazine*, vol. 29, no. 2, pp. 65–76, 2012.
- [9] Y. Wang and P. M. Djurić, "A gossip method for optimal consensus on a binary state from binary actions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 274–283, 2013.
- [10] V. Krishnamurthy, "Bayesian sequential detection with phase-distributed change time and nonlinear penalty pomdp lattice programming approach," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 7096–7124, 2011.
- [11] C. Chamley, *Rational Herds: Economic Models of Social Learning*, Cambridge University Press, 2004.
- [12] E. Mossel, A. Sly, and O. Tamuz, "Asymptotic learning on Bayesian social networks," *Probability Theory and Related Fields*, pp. 1–31, 2012.
- [13] C. Wang, Y. Chen, and K. J. R. Liu, "Sequential Chinese restaurant game," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 571–584, 2012.