

KEEP BALLOTS SECRET: ON THE FUTILITY OF SOCIAL LEARNING IN DECISION MAKING BY VOTING

Joong Bum Rhim and Vivek K Goyal

Research Laboratory of Electronics
Massachusetts Institute of Technology

ABSTRACT

We show that social learning is not useful in a model of team binary decision making by voting, where each vote carries equal weight. Specifically, we consider Bayesian binary hypothesis testing where agents have any conditionally-independent observation distribution and their local decisions are fused by any L -out-of- N fusion rule. The agents make local decisions sequentially, with each allowed to use its own private signal and all precedent local decisions. Though social learning generally occurs in that precedent local decisions affect an agent's belief, optimal team performance is obtained when all precedent local decisions are ignored. Thus, social learning is futile, and secret ballots are optimal. This conclusion contrasts with typical studies of social learning because we include a fusion center rather than concentrating on the performance of the latest-acting agents.

Index Terms—Bayesian hypothesis testing, distributed detection and fusion, sequential decision making, social learning, social networks

1. INTRODUCTION

Consider a set of agents making a decision collectively by voting, with each agent having equal influence. It would seem that the best collective decisions would come from each agent having as much information as possible. For example, if the voting is sequential rather than simultaneous, it would seem that the collective decision would be improved by later-acting agents having knowledge of the votes of the earlier-acting agents. Using the methods of signal processing to analyze a simple model, we show that this is not the case. In a social or political context, this provides a mathematical justification for the use of secret ballots, independent of any other merits, such as avoiding bribery, intimidation, or insincere voting. To be clear, though, our interest is in binary decision making under uncertainty, not voting to express individual preferences.

In contemporary popular culture, the advantage of (properly) aggregating opinions of many (reasonably diverse and independent) agents is known as “the wisdom of crowds” [1]. In statistical signal processing, this is nothing more than a reduction in effective noise level from averaging or aggregating. Suppose *private signals* observed by decision-making agents are not perfectly informative due to noise. If N agents share their private signals, assumed conditionally independent and identically distributed (iid) given the unknown variable of interest, then the sharing effectively increases the signal-to-noise ratio (SNR) by a factor of N . Even if a communication capacity limit restricts the amount of shared information, sharing still helps the detection system work much better than a single-agent system. For example, in systems where agents with Gaussian

noise-corrupted observations share 1-bit signals that are fused with the MAJORITY fusion rule, the effective SNR is asymptotically increased by a factor of $2N/\pi$ [2].

In this paper, we consider a team¹ of agents that together make a binary decision (between 0 and 1) by voting. Binary decisions often made by teams include whether to make a large purchase, whether to hire a job applicant, and whether to convict a defendant; the final example highlights that aggregation by voting, with equal weight to each vote, does not imply the use of majority rule. The agents have their own private signals and make local decisions in some pre-ordered order. Because of this ordering, we refer to the agents as a sequence. The agents form a distributed detection and data fusion system because their local decisions are fused to reach a global decision, and only the global decision determines their performance.

Now assume that the agents can watch other agents' choices as *public signals* so that they can use the predecessors' decisions in decision making. The framework of sequential decision making was independently introduced in [5] and [6]. A key concept in these works is *herding*. A herd of agents takes place when all agents beyond some index have the same local decision. Both works found that a herd of agents making the incorrect decision might arise with positive probability when private signals are bounded; this is a disappointing occurrence since the optimal aggregation of private signals would result in vanishing error probability. Subsequently, [7] showed that agents will asymptotically settle on the correct decision if private signals are unbounded. Recently, [8] extends the result to general network topologies where each agent can observe decisions made by its neighbors instead of all previous agents.

This work contrasts from previous works in two key ways: we consider an arbitrary but finite number of agents N ; and we consider any symmetric fusion of the local decisions (i.e., voting by an L -out-of- N rule²). Because the number of agents is finite, presence or absence of herding is not central to our study; a herd of agents might arise and yet form a small fraction of the N total agents. Also, the fusion by voting makes an early-acting agent important not only for its influence on later-acting agents but also because its vote counts.

As the first to study this scenario, we ask whether it is beneficial for the agents not only to send their local decisions to the fusion center but also to share among themselves. The agents in this work do Bayesian reasoning, and this has been shown to be a good approximation to human behavior [9–13].

The public signals raise two changes to the model: the belief update that is standard in social learning and a fusion rule update

¹The term *team* implies that the agents have no conflicts of interest, which here translates to agreement in the relative importance of false alarms and missed detections [3]. Conflicts among agents would make their collective decisions worse [4]; we concentrate on understanding their ideal behavior.

²In the notation of L -out-of- N rule, N denotes the number of agents and L denotes the minimum number of 1 votes for 1 to be the global decision.

This material is based upon work supported by the National Science Foundation under Grant No. 1101147.

that has not appeared previously in the context of social learning.

Each agent will update its belief about the true state based on the public signals (in a Bayesian setting) and design a better decision rule. The previous decisions reflecting the private signals of the precedent agents contain information that the following agents do not have. If an agent observes that most predecessors have chosen 1, then the agent will have a stronger belief on 1 and a weaker belief on 0. Hence, the update of belief is a positive feedback that encourages an agent to follow the precedents.

Each agent will adjust the fusion rule based on the public signals. For example, if the first agent has chosen 0, then the L -out-of- N fusion rule changes to the L -out-of- $(N-1)$ rule for the remaining agents. And if the second agent has subsequently chosen 1, then the fusion rule becomes the $(L-1)$ -out-of- $(N-2)$ rule, etc. Thus, the previous decisions put a following agent in a different position; if most previous agents have chosen 1, then a few more votes for 1 will determine the global decision as 1. This fact makes the agent become more careful to choose 1. The evolution of the fusion rule is a negative feedback that discourages an agent to follow the precedents.

Our mathematical analysis shows that those positive and negative effects are exactly canceled out and that the optimal decision rules with public signals are equal to those without public signals. It implies that observing public signals is practically useless.

Section 2 describes our model formally. Section 3 provides results for two agents and Section 4 extends the results for an arbitrary number of agents. Section 5 concludes the paper.

2. THE FORMAL MODEL

A team of N agents—Alexis, Blake, ..., Norah—performs a binary hypothesis test for an object. The agents are aware of the prior probabilities of the object in state $H = 0$ and in state $H = 1$, which are denoted by $p_0 = \mathbb{P}\{H = 0\}$ and $p_1 = \mathbb{P}\{H = 1\} = 1 - p_0$. The agents individually receive private signals about the state and perform tests to make their own binary decisions \hat{H}_i . They are a team that shares a common cost function and a common goal to minimize the cost of global decisions. The cost for a false alarm is c_{10} and the cost for a missed detection is c_{01} . Thus, for the L -out-of- N fusion rule, the average cost (i.e., Bayes risk) is given as follows:

$$R = c_{10}p_0\mathbb{P}\{\sum_{i=1}^N \hat{H}_i \geq L \mid H = 0\} + c_{01}p_1\mathbb{P}\{\sum_{i=1}^N \hat{H}_i \leq L - 1 \mid H = 1\}. \quad (1)$$

The private signals Y_i are conditionally iid given the state H , with the likelihood function $f_{Y_i|H}$. In other words, all agents observe private signals of equal quality. Bayesian agents use a likelihood ratio test (LRT) as the optimal decision rule, which chooses $\hat{H}_i = 1$ if the ratio $f_{Y_i|H}(y_i | 1)/f_{Y_i|H}(y_i | 0)$ is greater than a certain threshold [14, 15]. We assume that the likelihood ratio is an increasing function of y_i so that the LRT is simplified to a decision rule with a decision threshold λ_i :

$$\begin{array}{c} \hat{H}_i(y_i)=1 \\ y_i \gtrsim \lambda_i \\ \hat{H}_i(y_i)=0 \end{array} \quad (2)$$

We will compare the optimal decision thresholds in the following two scenarios:

- 1) A common distributed detection system: Agents only observe private signals and make local decisions in parallel without knowing other agents' decisions. Then their decisions are fused by a specific L -out-of- N rule to make a global decision. Alternatively, we call this scenario parallel decision making.

- 2) A distributed detection system combined with sequential decision making: Agents observe private signals and sequentially make local decisions. Before each agent makes a decision, the agent observes precedents. Then their decisions are fused by a specific L -out-of- N rule to make a global decision.

Our notations for decision thresholds are distinguished in the two scenarios: λ in the first scenario and ρ in the second scenario.

In Scenario 1, we will say that the optimal decision thresholds are identical, i.e., $\lambda_1^* = \dots = \lambda_N^*$ for the following reasons: Using identical decision thresholds is asymptotically optimum for the binary hypothesis testing problem [16]. Furthermore, by numerical experiments, it turns out that constraining to identical decision rules causes little or no loss of performance for finite N and the corresponding optimal fusion rule has the L -out-of- N form [17]. Our numerical experiments (not reported here) for fixed fusion rules and conditionally iid private signals show that the optimal decision thresholds are in fact identical at least for any $N \leq 7$ in multiple cases with such as Gaussian and exponential likelihood functions.

We use P_e^I and P_e^{II} as notations for probabilities of Type I (false alarm) and Type II (missed detection) errors. We use a subscript index to indicate a specific agent. For example, error probabilities of Alexis are written as

$$P_{e,1}^I = \mathbb{P}\{\hat{H}_1 = 1 \mid H = 0\}, \quad P_{e,1}^{II} = \mathbb{P}\{\hat{H}_1 = 0 \mid H = 1\}. \quad (3)$$

Since all agents have equally good private signals, their error probabilities are the same as (3) in the parallel decision-making scenario.

In Scenario 2, the agents sequentially make hard decisions: Alexis first makes a decision \hat{H}_1 , Blake next makes a decision \hat{H}_2 , and so on. Predecessors' decisions, which we call public signals, may cause different error probabilities by following agents. In this case, we use superscript notation to specify the public signals. For example, $P_{e,2}^{I_0}$ denotes the probability of Type I error of Blake upon observing $\hat{H}_1 = 0$ and $P_{e,3}^{II_{01}}$ denotes the probability of Type II error of Chuck (the third agent) upon observing $\hat{H}_1 = 0$ and $\hat{H}_2 = 1$. We sometimes use a notation like $P_{e,3}^{I_0}$ that specifies Alexis's decision as $\hat{H}_1 = 0$ but does not specify Blake's decision \hat{H}_2 .

Once the error probabilities of all agents are computed, the conditional probabilities of the global decision being in error are given from the L -out-of- N fusion rule:

$$P_E^I = p_{\hat{H}|H}(1|0) = \sum_{n=L}^N \sum_{\substack{I \subseteq [N] \\ |I|=n}} \prod_{i \in I} P_{e,i}^I \prod_{j \in [N] \setminus I} (1 - P_{e,j}^I),$$

$$P_E^{II} = p_{\hat{H}|H}(0|1) = \sum_{n=N-L+1}^N \sum_{\substack{I \subseteq [N] \\ |I|=n}} \prod_{i \in I} P_{e,i}^{II} \prod_{j \in [N] \setminus I} (1 - P_{e,j}^{II}),$$

where $[N]$ denotes the set $\{1, 2, \dots, N\}$. The Bayes risk is based on the team decision:

$$R = c_{10}p_0P_E^I + c_{01}p_1P_E^{II}. \quad (4)$$

3. TWO AGENTS

Let us consider the simplest case for distributed detection, which is $N = 2$. We will compare two cases, which are depicted in Fig. 1. We start from an assumption that Alexis does not change her strategy whether the agents observe public signals or not. The assumption comes from an intuition that she is the first agent to make a decision and does not observe any previous decisions in both scenarios.

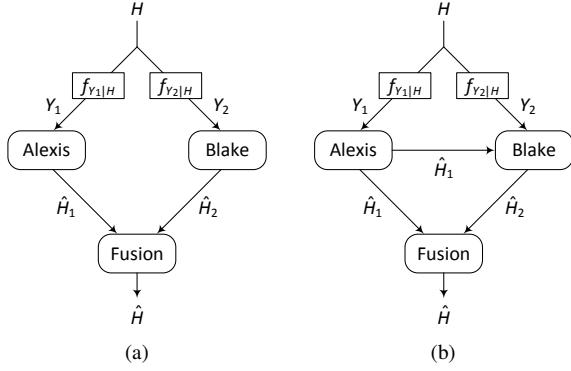


Fig. 1. Two agents and a fusion center. (a) Parallel decision making. (b) Sequential decision making.

Theorem 1. *Suppose that sharing previous decisions does not change Alexis's decision rule. Then, for $N = 2$ and any L -out-of- N fusion rule, the existence of public signals does not affect the decision strategies of agents.*

Proof. There are only two meaningful fusion rules when $N = 2$: 1-out-of-2 (OR) rule and 2-out-of-2 (AND) rule. Let's consider the OR rule first.

In the parallel decision-making scenario, the Bayes risk is given by

$$R_p = c_{10}p_0 \left(P_{e,1}^I + P_{e,2}^I - P_{e,1}^I P_{e,2}^I \right) + c_{01}p_1 P_{e,1}^{II} P_{e,2}^{II}, \quad (5)$$

where Alexis's error probabilities $P_{e,1}^I$ and $P_{e,1}^{II}$ are governed by her decision threshold λ_1 and Blake's $P_{e,2}^I$ and $P_{e,2}^{II}$ are governed by his decision threshold λ_2 . Their optimal decision thresholds λ_1^* and λ_2^* are the minimizer of (5).

In the sequential decision-making scenario, Blake observes Alexis's decision \hat{H}_1 . Thus, we take care of Alexis's decision threshold, ρ_1 , and Blake's decision threshold upon observing $\hat{H}_1 = 0$, which is denoted by ρ_2^0 , only because his decision does not matter when $\hat{H}_1 = 1$. The Bayes risk is given by

$$R_s = c_{10}p_0 \left(P_{e,1}^I + (1 - P_{e,1}^I) P_{e,2}^{I_0} \right) + c_{01}p_1 P_{e,1}^{II} P_{e,2}^{II_0}, \quad (6)$$

which is exactly the same formula as (5). Therefore, the optimal decision thresholds would be the same as those in the first case:

$$\rho_1^* = \lambda_1^*, \quad \rho_2^{0*} = \lambda_2^*, \quad (7)$$

and $\rho_1^* = \rho_2^{0*}$ because $\lambda_1^* = \lambda_2^*$. Since we do not care about Blake's decision threshold when $\hat{H}_1 = 1$, ρ_2^1 , we can set $\rho_2^1 = \rho_2^{0*}$, too. Therefore, Alexis and Blake's optimal decision thresholds are not affected by the existence and the value of the public signal \hat{H}_1 .

The same statement for the AND fusion rule can also be proven in a similar way. \square

4. N AGENTS

In Section 3, we have shown that knowing Alexis's decision practically does not change Blake's optimal strategies for $N = 2$ and any L -out-of- N fusion rule. Now let us expand the problem to a general N -agent problem by mathematical induction.

Theorem 2. *Suppose that sharing previous decisions does not change Alexis's decision rule (i.e., $\rho_1^* = \lambda_1^*$). If the existence of the public signals does not affect optimal decision thresholds of a team of N agents for a specific N and any K -out-of- N fusion rule, then the existence of the public signals also does not affect optimal decision thresholds of a team of $N+1$ agents and any L -out-of- $(N+1)$ fusion rule.*

Proof. First, consider the parallel decision-making scenario with $N+1$ agents. Since a decision of an agent is critical only if the other N local decisions are $L-1$ ones and $N-L+1$ zeros, the optimal decision threshold λ^* is the solution to³

$$\begin{aligned} \frac{f_{Y|H}(\lambda|1)}{f_{Y|H}(\lambda|0)} &= \frac{c_{10}p_0 \binom{N}{L-1} (P_e^I)^{L-1} (1 - P_e^I)^{N-L+1}}{c_{01}p_1 \binom{N}{N-L+1} (P_e^{II})^{N-L+1} (1 - P_e^{II})^{L-1}} \\ &= \frac{c_{10}p_0 (P_e^I)^{L-1} (1 - P_e^I)^{N-L+1}}{c_{01}p_1 (P_e^{II})^{N-L+1} (1 - P_e^{II})^{L-1}}, \end{aligned} \quad (8)$$

where we embed the idea that the optimal decision thresholds of all agents are identical.

Next, consider the sequential decision-making scenario when the agents observe previous decisions. The Bayes risk is given by

$$\begin{aligned} R_s &= c_{10}p_0 \left(1 - P_{e,1}^I \right) \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \geq L \mid \hat{H}_1 = H = 0 \right\} \\ &\quad + c_{10}p_0 P_{e,1}^I \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \geq L - 1 \mid \hat{H}_1 = 1, H = 0 \right\} \\ &\quad + c_{01}p_1 P_{e,1}^{II} \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \leq L - 1 \mid \hat{H}_1 = 0, H = 1 \right\} \\ &\quad + c_{01}p_1 \left(1 - P_{e,1}^{II} \right) \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \leq L - 2 \mid \hat{H}_1 = H = 1 \right\} \\ &\triangleq R_0 \left(p_0(1 - P_{e,1}^I) + p_1 P_{e,1}^{II} \right) + R_1 \left(p_0 P_{e,1}^I + p_1(1 - P_{e,1}^{II}) \right), \end{aligned} \quad (9)$$

where R_0 and R_1 are specified in (11) and (12) and we define

$$\begin{aligned} q^0 &\triangleq \frac{p_0 (1 - P_{e,1}^I)}{p_0 (1 - P_{e,1}^I) + p_1 P_{e,1}^{II}} = \mathbb{P}\{H = 0 \mid \hat{H}_1 = 0\}, \\ q^1 &\triangleq \frac{p_0 P_{e,1}^I}{p_0 P_{e,1}^I + p_1 (1 - P_{e,1}^{II})} = \mathbb{P}\{H = 0 \mid \hat{H}_1 = 1\}. \end{aligned} \quad (10)$$

When the agents $2, \dots, N+1$ observe that $\hat{H}_1 = 0$, their optimal decision strategy is to minimize the term R_0 from (9):

$$\begin{aligned} R_0 &= c_{10}q^0 \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n^0 \geq L \mid H = 0 \right\} \\ &\quad + c_{01}(1 - q^0) \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n^0 \leq L - 1 \mid H = 1 \right\}, \end{aligned} \quad (11)$$

where the condition $\hat{H}_1 = 0$ is intended in the term \hat{H}_n^0 . Please note that R_0 is the same as the Bayes risk of N agents when the prior probability is q^0 and fusion is by the L -out-of- N rule. It implies that the optimal decision thresholds of Agents $2, \dots, N+1$ are the same as those of N agents with prior probability q^0 and the L -out-of- N fusion rule.

Likewise, when the agents $2, \dots, N+1$ observe that $\hat{H}_1 = 1$, their optimal decision strategy is to minimize the term R_1 from (9):

$$\begin{aligned} R_1 &= c_{10}q^1 \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n^1 \geq L - 1 \mid H = 0 \right\} \\ &\quad + c_{01}(1 - q^1) \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n^1 \leq L - 2 \mid H = 1 \right\}. \end{aligned} \quad (12)$$

³Please see [18] for a detailed description of how (8) is derived.

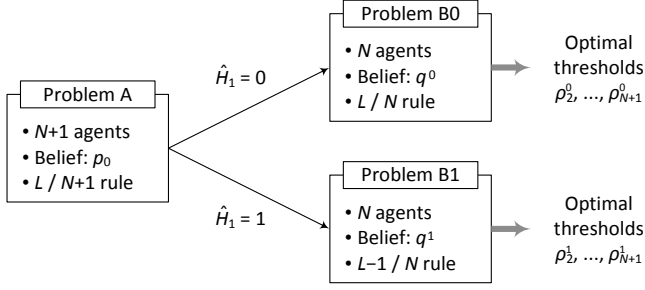


Fig. 2. An $(N+1)$ -agent problem is divided into two N -agent problems depending on Alexis's decision \hat{H}_1 .

Their optimal decision thresholds are the same as those of N agents with prior probability q^1 and $(L-1)$ -out-of- N fusion rule. Fig. 2 depicts the evolution of the problem corresponding to Alexis's decision \hat{H}_1 .

Let us find the optimal thresholds $\rho_2^{0*}, \dots, \rho_{N+1}^{0*}$ in Problem B0 in Fig. 2. In fact, the N agents in Problem B0 also observe public signals. However, because of the condition that the existence of the public signals does not affect optimal decision thresholds of a team of N agents for any K -out-of- N fusion rule, we can find the optimal thresholds as if the agents do not observe the public signals. Since a decision of an agent is critical only if the other $N-1$ local decisions are $L-1$ ones and $N-L$ zeros, the optimal decision threshold ρ^{0*} is the solution to

$$\begin{aligned} \frac{f_{Y|H}(\rho^0 | 1)}{f_{Y|H}(\rho^0 | 0)} &= \frac{c_{10} q^0 \binom{N-1}{L-1} (P_e^{I_0})^{L-1} (1-P_e^{I_0})^{N-L}}{c_{01} (1-q^0) \binom{N-1}{N-L} (P_e^{II_0})^{N-L} (1-P_e^{II_0})^{L-1}} \\ &= \frac{c_{10} p_0 (1-P_e^I) (P_e^{I_0})^{L-1} (1-P_e^{I_0})^{N-L}}{c_{01} p_1 P_e^{II} (P_e^{II_0})^{N-L} (1-P_e^{II_0})^{L-1}}, \end{aligned} \quad (13)$$

where q^0 is replaced by (10). Due to the assumption that $\rho_1^* = \lambda_1^*$, P_e^I and P_e^{II} in (13) are the same as P_e^I and P_e^{II} in (8).

Comparing (13) to (8), we can find that they have the same solutions, i.e., $\rho_i^{0*} = \lambda_i^*$. Therefore, the agents should not change their decision thresholds after observing $\hat{H}_1 = 0$.

We can also find the optimal thresholds $\rho_2^{1*}, \dots, \rho_{N+1}^{1*}$ in Problem B1 in Fig. 2 by looking at the N -agent problem without public signals:

$$\begin{aligned} \frac{f_{Y|H}(\rho^1 | 1)}{f_{Y|H}(\rho^1 | 0)} &= \frac{c_{10} q^1 \binom{N-1}{L-2} (P_e^{I_1})^{L-2} (1-P_e^{I_1})^{N-L+1}}{c_{01} (1-q^1) \binom{N-1}{N-L+1} (P_e^{II_1})^{N-L+1} (1-P_e^{II_1})^{L-2}} \\ &= \frac{c_{10} p_0 P_e^I (P_e^{I_1})^{L-2} (1-P_e^{I_1})^{N-L+1}}{c_{01} p_1 (1-P_e^{II}) (P_e^{II_1})^{N-L+1} (1-P_e^{II_1})^{L-2}}. \end{aligned} \quad (14)$$

Again, due to the assumption that $\rho_1^* = \lambda_1^*$, P_e^I and P_e^{II} in (14) are the same as P_e^I and P_e^{II} in (8). We reach to the same statement that the two equations have the same solutions, i.e., $\rho_i^{1*} = \lambda_i^*$, by comparing (14) to (8). Thus, the agents should not change their decision thresholds after observing $\hat{H}_1 = 1$.

Consequently, for a team of $N+1$ agents and any L -out-of- $(N+1)$ rule, their optimal decision thresholds are the same whether they observe previous decisions or not. \square

We want to clarify the term *belief* in Fig. 2, which is distinguished from prior probability. The prior probability is fixed as p_0 and known to all agents. The belief is how probable the agents think $H = 0$ is. The belief changes as agents observe previous decisions. The above proof does not mean that the prior probability is changed after the agents observe Alexis's decision. It is saying that the problem of determining optimal thresholds upon observing Alexis's decision is equivalent to Problem B0 or B1 corresponding to \hat{H}_1 .

The essence of this proof is to show that the effect of new fusion rules, L -out-of- N and $(L-1)$ -out-of- N , exactly cancels out the effect of new beliefs q^0 and q^1 . It is observed when we compare (13) and (14) to (8). From an overall standpoint, the public signals neither harm nor help the decision-making task.

Corollary 3. *Suppose that sharing previous decisions does not change Alexis's decision rule (i.e., $\rho_1^* = \lambda_1^*$). For any N and L -out-of- N fusion rule, the existence of the public signals does not affect optimal decision thresholds of a team of N agents.*

Proof. Use mathematical induction with Theorems 1 and 2. \square

Corollary 3 is trivial for $N = 1$ and is proven by Theorem 1 for $N = 2$. For $N \geq 3$, we can iteratively break down the problem until we have 2^{N-2} two-agent problems like in Fig. 2. The backward process from the leaves (two-agent problems) to the root (N -agent problem) will demonstrate Corollary 3.

Corollary 3 only refers to the optimal decision thresholds. However, it implies that team's performance (i.e. Bayes risk) is not affected by the public signals. This is because decision thresholds determine the probabilities of errors and, thus, the Bayes risk.

5. CONCLUSION

We have discussed sequential decision making in a distributed detection and fusion system by a team of agents. The system is intentionally simple for the purpose of understanding fundamentals; agents share a cost function and know the prior probability, and their observations are conditionally iid. It would be intuitively desired to have as much information as possible when performing a hypothesis test. However, our study has revealed that it is useless to observe other agents' decisions in this system. The agents just need to individually make the best decisions.

This result is justified by symmetries throughout our model. The first symmetry is the equal quality of private signals due to the identical likelihood functions. The second symmetry is the equal (1-bit) votes that agents have. The third symmetry is the equal weights of the votes implied by the L -out-of- N fusion rule.

One interpretation of the main result is as follows: Each agent sends 1 bit of information about its private signal to the fusion center. If Blake refers to Alexis's decision, then the fusion center effectively receives less than 1 bit of information about Blake's private signal from him. Therefore, in order to prevent such an efficiency loss, Blake should make a reasonable decision only based on his own private signal. Hence, a secret ballot encourages optimal decision making, since it is not good for the agents to be affected by public signals.

This result is obtained under an assumption that Alexis uses the same decision thresholds in both scenarios, and the assumption is not unreasonable. The assumption is trivially true for $N = 1$. It is also true for $N = 2$; our proof of Theorem 1 even does not use this assumption. In addition, we have confirmed that it is true at least for any $N \leq 9$ by numerical experiments. This assumption heuristically seems true, and our future work is to verify it for arbitrary N .

6. REFERENCES

- [1] J. Surowiecki, *The Wisdom of Crowds: Why the Many are Smarter Than the Few*, Little, Brown & Co., New York, NY, 2004.
- [2] L. R. Varshney, J. B. Rhim, K. R. Varshney, and V. K. Goyal, "Categorical decision making by people, committees, and crowds," in *Proc. Information Theory and Applications Workshop*, La Jolla, CA, Feb. 2011.
- [3] R. Radner, "Team decision problems," *Ann. Math. Stat.*, vol. 33, no. 3, pp. 857–881, Sept. 1962.
- [4] J. B. Rhim, L. R. Varshney, and V. K. Goyal, "Conflict in distributed hypothesis testing with quantized prior probabilities," in *Proc. IEEE Data Compression Conf.*, Snowbird, UT, Mar. 2011, pp. 313–322.
- [5] A. V. Banerjee, "A simple model of herd behavior," *Quart. J. Econ.*, vol. 107, no. 3, pp. 797–817, Aug. 1992.
- [6] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *J. Political Economy*, vol. 100, no. 5, pp. 992–1026, Oct. 1992.
- [7] L. Smith and P. Sørensen, "Pathological outcomes of observational learning," *Econometrica*, vol. 68, no. 2, pp. 371–398, Mar. 2000.
- [8] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *Rev. Economic Studies*, vol. 78, no. 4, pp. 1201–1236, Mar. 2011.
- [9] J. A. Swets, W. P. Tanner, Jr., and T. G. Birdsall, "Decision processes in perception," *Psychol. Rev.*, vol. 68, no. 5, pp. 301–340, Sept. 1961.
- [10] W. K. Viscusi, "Are individuals Bayesian decision makers?," *Am. Econ. Rev.*, vol. 75, no. 2, pp. 381–385, May 1985.
- [11] G. L. Brase, L. Cosmides, and J. Tooby, "Individuation, counting, and statistical inference: The role of frequency and whole-object representations in judgment under uncertainty," *J. Exp. Psychol. Gen.*, vol. 127, no. 1, pp. 3–21, Mar. 1998.
- [12] C. J. Gill, L. Sabin, and C. H. Schmid, "Why clinicians are natural Bayesians," *British Medical J.*, vol. 330, pp. 1080–1083, May 2005.
- [13] M. Glanzer, A. Hilford, and L. T. Maloney, "Likelihood ratio decisions in memory: Three implied regularities," *Psychon. Bull. Rev.*, vol. 16, no. 3, pp. 431–455, June 2009.
- [14] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Phil. Trans. Royal Soc. London. Ser. A*, vol. 231, pp. 289–337, Jan. 1933.
- [15] P. K. Varshney, *Distributed Detection and Data Fusion*, Springer-Verlag, 1997.
- [16] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems*, vol. 1, no. 2, pp. 167–182, 1988.
- [17] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, H. V. Poor and J. B. Thomas, Eds., pp. 297–344. JAI Press, Greenwich, CT, 1993.
- [18] J. B. Rhim, L. R. Varshney, and V. K. Goyal, "Quantization of prior probabilities for collaborative distributed hypothesis testing," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4537–4550, Sept. 2012.