

Distributed Hypothesis Testing With Social Learning and Symmetric Fusion

Joong Bum Rhim, *Member, IEEE*, and Vivek K Goyal, *Fellow, IEEE*

Abstract—We study the utility of social learning in a distributed detection model with agents sharing the same goal: a collective decision that optimizes an agreed upon criterion. We show that social learning is helpful in some cases but is provably futile (and thus essentially a distraction) in other cases. Specifically, we consider Bayesian binary hypothesis testing performed by a distributed detection and fusion system, where all decision-making agents have binary votes that carry equal weight. Decision-making agents in the team sequentially make local decisions based on their own private signals and all precedent local decisions. It is shown that the optimal decision rule is not affected by precedent local decisions when all agents observe conditionally independent and identically distributed private signals. Perfect Bayesian reasoning will cancel out all effects of social learning. When the agents observe private signals with different signal-to-noise ratios, social learning is again futile if the team decision is only approved by unanimity. Otherwise, social learning can strictly improve the team performance. Furthermore, the order in which agents make their decisions affects the team decision.

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

I. INTRODUCTION

CONSIDER a team of three physicians, Alexis, Britta, and Carol, who together diagnose a patient and make a decision, say between medication and surgery. Each physician individually examines the patient or runs medical tests within her expertise before the team aggregates the individual opinions. The team would make the best decision if the physicians could share every result and discuss how to treat the patient. This would be as if the agents send their observations in full to a fusion center, which is able to use a sophisticated integration rule, as depicted in Fig. 1(a). However, this level of communication

Manuscript received April 03, 2014; revised July 06, 2014; accepted September 14, 2014. Date of publication October 13, 2014; date of current version November 06, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Pengfei Xia. This work was supported by the National Science Foundation under Grant 1101147. This paper was presented in part in the *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vancouver, BC, Canada, May 2013.

J. B. Rhim was with the Department of Electrical Engineering and Computer Science and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA. He is currently with GroupM Data and Analytics, New York, NY 10001 USA. (e-mail: jbrhim@alum.mit.edu).

V. K. Goyal was with Boston University and the Massachusetts Institute of Technology. He is currently with Google, on leave from the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA (e-mail: v.goyal@ieee.org).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2014.2362885

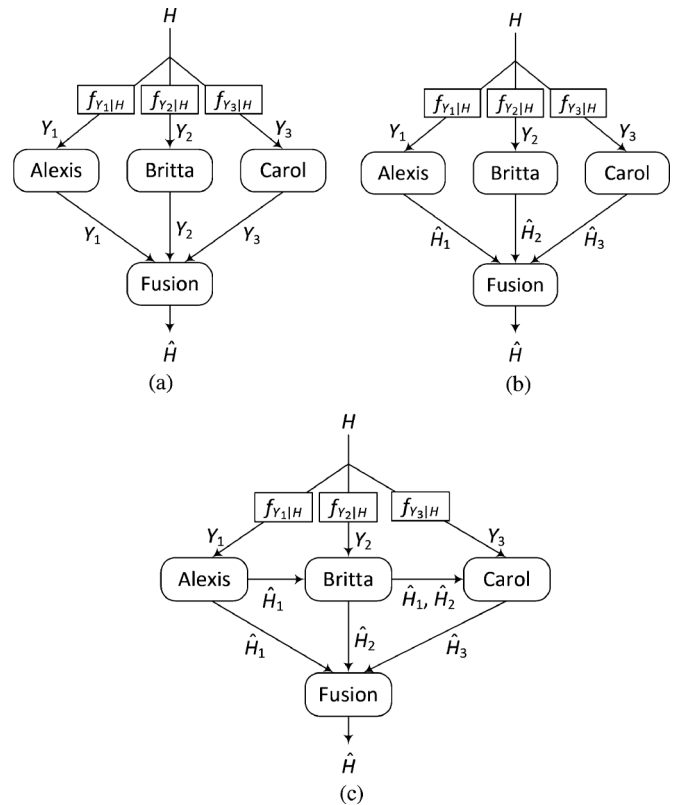


Fig. 1. Examples of group decision making. (a) Agents observe private signals Y_n and send them to the fusion center for centralized detection. (b) Agents observe private signals Y_n , make local decisions \hat{H}_n in a distributed matter, and send their decisions to the fusion center. (c) Agents observe private signals Y_n , make local decisions \hat{H}_n sequentially using all available information, and send their decisions to later-acting agents and to the fusion center. Sending decisions to later-acting agents is termed *public voting*. (a) Centralized detection. (b) Conventional distributed detection. (c) Focus of this paper.

costs a lot of time and resources, and this architecture assumes the ability to jointly interpret the totality of information.

A more common way to aggregate opinions of human decision makers is through some form of voting. In fact, a patient may seek the opinions of three physicians and decide according to the majority, with no interaction among the physicians, as depicted in Fig. 1(b). This is a typical distributed detection model. The optimal local decision-making strategy in such a model has been studied for decades [1], [2].

In some settings, people share their opinions and thus can affect the opinions of others. In the medical example, the patient may see the physicians in sequence, and the recommendation (medication or surgery) of a physician may be known to each later-consulted physician when she is making her own recommendation. A simple example is shown in Fig. 1(c). The effect

of this interaction among decision-making agents—under certain assumptions that are adopted for tractability—is the focus of this paper.

In our model, agents make decisions in a predetermined order and later-acting agents can observe all or some of the decisions made by earlier-acting agents. These local decisions are sent as *public signals*. Thus, each agent has a *private signal* originating from outside the set of agents along with some subset of the public signals. The later-acting agents can perform Bayesian reasoning to learn some information from the public signals and make better local decisions. This process of incorporating information from other agents is called *social learning* or *observational learning*. We will see that social learning provides no improvement when the agents have conditionally independent and identically distributed (iid) private signals; otherwise there is generally an improvement, and the degree of improvement depends on the order in which the agents make their decisions.

A. Relations to Previous Work

Mathematical study of social learning was introduced in [3], [4]. These papers discussed decision-making agents who make decisions in a predetermined order and for individually maximizing their own utilities. They presented models where choices made by all agents are observed by all other agents as public signals. In these models, herding behavior—that agents ignore private signals and make decisions in agreement with the public signals—arises after several earlier-acting agents choose the same alternative. It was later shown that incorrect herding occurs with a nonzero probability if the private signals are boundedly informative.¹ Otherwise, asymptotically a sequence of agents will converge to the correct choice [5].

This work is unique from previous works in that we study social learning in a team decision-making scenario with a symmetric fusion of the local decisions. The fusion by voting makes an earlier-acting agent important not only because she influences later-acting agents but also because her vote counts in the fusion rule. We ask whether it is beneficial for the agents to share their local decisions among themselves. The agents in this work perform Bayesian reasoning, and there is a rich literature that suggests this to be a good approximation to human behavior [6]–[10].

Note that we consider agents who share a common goal. The social learning in previous works occurs among agents that have individual goals and behave selfishly to achieve their own goals. Social learning is helpful in this case if the private signals are unbounded [5] even though the selfishness diminishes the efficiency of social learning [11]. Chamley *et al.* [12] also state that social learning yields a lower rate of learning than collaboration does. They compare social learning by selfish agents to distributed detection by collaborating agents but do not consider social learning by collaborating agents. In contrast, we are considering social learning and distributed detection together in one model.

When the agents' private signals are not conditionally iid, we find that the order in which the agents make their decisions can affect the collective performance. The ordering that maximizes

the amount of information transferred to the fusion center was studied previously in the case when agents observe binary private signals by Ottaviani and Sørensen [13]. They state that for a given group of heterogeneous expert agents, the order in which the agents speak matters. The same statement holds in our model even though the agents in our model have a different motivation than the agents in the previous model, who pursue only the personal goal of improving their reputations. The importance of ordering is also pointed out in [14].

Our preliminary work [15] compares private voting to public voting in which all agents observe the full history of previous decisions and the agents observe conditionally iid private signals. It is shown that public voting does not improve performance. Intuitively, one might think that social learning helps later-acting agents make better local decisions, leading to improvement of the team decision; however, this intuition is shown to be wrong.

This paper builds upon the preliminary results in [15]. We extend our interest to more general models. We allow the likelihoods for the agents' private signals to differ. Also, agents may observe only some of previous decisions, such as a decision made by the agent just before themselves in a sequence [2] or local decisions made by their neighbors in a specific network [16].

No theoretical results in this paper require that private signals are equal to the hypothesis corrupted by additive or Gaussian noises. Agents observe private signals with likelihood functions that are different under state 0 and state 1. All we need is that their likelihood ratios are strictly monotonic, which is a rather natural requirement.

B. Outline

Section II formally describes our decision-making model. Section III provides results for agents observing conditionally iid private signals. Section IV generalizes the model to agents with differing private signal likelihoods and provides examples of helpful social learning. Section V summarizes our results and discusses limitations and extensions of our model.

II. PROBLEM STATEMENT

Consider a distributed detection and decision fusion system with N agents: Alexis, Britta, Carol, Diana, . . . , Norah. We also number the agents $1, 2, \dots, N$; names and numbers are used as convenient. Together, the agents perform a binary hypothesis test to detect $H \in \{0, 1\}$, which has prior probability distribution $p_0 = \mathbb{P}\{H = 0\}$ and $p_1 = \mathbb{P}\{H = 1\} = 1 - p_0$. Agent n observes a *private signal* Y_n about the true state $H = h$, with likelihood function $f_{Y_n|H}(y_n|h)$, as shown in Fig. 2. The private signals $\{Y_n\}_{n=1}^N$ are conditionally independent given H . We assume that the likelihood ratios $f_{Y_n|H}(y_n|1)/f_{Y_n|H}(y_n|0)$ are monotonically increasing.

In our model, Agent n makes *local decision* $\hat{H}_n \in \{0, 1\}$ or, equivalently, quantizes her private signal into 0 or 1, and sends it to a fusion center. The fusion center uses a fixed and symmetric rule to aggregate the local decisions:

$$\hat{H} = \begin{cases} 1, & \text{if } \sum_{n=1}^N \hat{H}_n \geq L; \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

¹A signal Y generated under a state H is called *boundedly informative* if there exists $\kappa > 0$ such that $\kappa < f_{Y|H}(y|h) < 1/\kappa$ for all y and h .

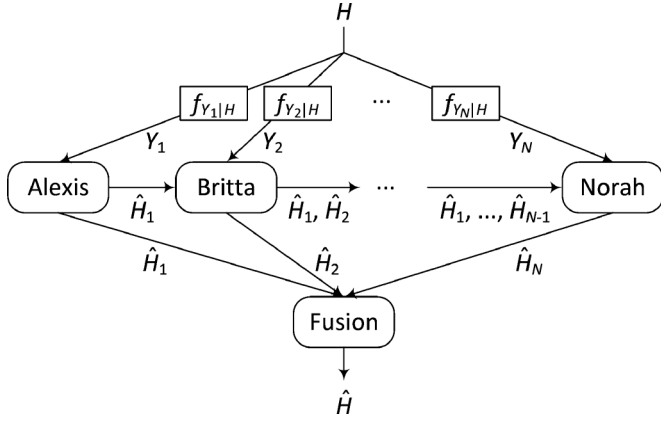


Fig. 2. The distributed detection and decision fusion model considered in this paper. Local decisions are sent not only to the fusion center but also to following agents as public signals.

for some fixed $L \in \{1, 2, \dots, N\}$ that is known to all agents. This rule is symmetric because it is invariant to permutation of the agents; it treats all local decisions with equal weights. We call this the L -out-of- N fusion rule and call the result of the fusion \hat{H} a *team decision*.

The goodness criterion for the decision making is the Bayes risk (mean decision-making cost):

$$R = c_{10}p_0\mathbb{P}\{\hat{H} = 1 | H = 0\} + c_{01}p_1\mathbb{P}\{\hat{H} = 0 | H = 1\}, \quad (2)$$

where c_{10} denotes the cost of a false alarm or Type I error—detecting $H = 0$ as $\hat{H} = 1$ —and c_{01} denotes the cost of a missed detection or Type II error—detecting $H = 1$ as $\hat{H} = 0$. We assume that a correct decision induces no cost. The agents are a team in the sense of Radner [17]: they share the same decision-making costs and the same goal, which is to minimize (2).

Each Bayesian agent performs a likelihood ratio test (LRT) to make optimal local decisions [2], [18]. The test chooses the hypothesis $\hat{H}_n = 1$ if the likelihood ratio is greater than a certain threshold, which is determined by the information the agent has, such as the prior probability, decision-making costs, and the number of agents. Since we are assuming monotonically increasing likelihood ratios, we can express the LRT in a compact form:

$$\begin{aligned} \hat{H}_n(y_n) &= 1 \\ y_n &\gtrsim \lambda_n, \\ \hat{H}_n(y_n) &= 0 \end{aligned} \quad (3)$$

where we call λ_n a decision threshold.

Each agent generates a local decision \hat{H}_n after the LRT. Her decision may be a false alarm or a missed detection, which have probabilities $P_{e,n}^I$ and $P_{e,n}^{II}$, respectively:

$$P_{e,n}^I = \mathbb{P}\{\hat{H}_n = 1 | H = 0\}, \quad (4)$$

$$P_{e,n}^{II} = \mathbb{P}\{\hat{H}_n = 0 | H = 1\}. \quad (5)$$

From the L -out-of- N fusion rule, we can compute the probabilities of false alarm and of missed detection of the team decision, P_E^I and P_E^{II} :

$$\begin{aligned} P_E^I &= \mathbb{P}\{\hat{H} = 1 | H = 0\} \\ &= \mathbb{P}\left\{\sum_{n=1}^N \hat{H}_n \geq L | H = 0\right\} \\ &= \sum_{n=L}^N \sum_{\substack{I \subseteq [N] \\ |I|=n}} \prod_{i \in I} P_{e,i}^I \prod_{j \in [N] \setminus I} (1 - P_{e,j}^I), \\ P_E^{II} &= \mathbb{P}\{\hat{H} = 0 | H = 1\} \\ &= \mathbb{P}\left\{\sum_{n=1}^N \hat{H}_n < L | H = 1\right\} \\ &= \sum_{n=N-L+1}^N \sum_{\substack{I \subseteq [N] \\ |I|=n}} \prod_{i \in I} P_{e,i}^{II} \prod_{j \in [N] \setminus I} (1 - P_{e,j}^{II}), \end{aligned}$$

where $[N]$ denotes the set $\{1, 2, \dots, N\}$. The Bayes risk (2) can be rewritten as

$$R = c_{10}p_0P_E^I + c_{01}p_1P_E^{II}. \quad (6)$$

Note that this is the *team's* Bayes risk, which is the only goodness criterion considered here; agents do not minimize a personal Bayes risk.

In a decision-making task, the agents make decisions sequentially in the preordained order of Alexis, Britta, Carol, through Norah. Through most of the paper, each of these agents can observe all previous local decisions.² For example, Britta can observe Alexis's decision \hat{H}_1 ; Carol can observe Britta's decision \hat{H}_2 as well as \hat{H}_1 ; and so on. We call the local decisions observed by other agents *public signals*.

Later-acting agents have more degrees of freedom in choosing their decision thresholds because the decision thresholds may depend on the observed public signals. Britta has two degrees of freedom: she can apply different thresholds when $\hat{H}_1 = 0$ and when $\hat{H}_1 = 1$. Likewise, Carol has four degrees of freedom because she observes one of four public signal combinations $(\hat{H}_1, \hat{H}_2) \in \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. We use superscripts to specify the public signals that an agent observes. For example, λ_2^0 denotes Britta's decision threshold when she observes $\hat{H}_1 = 0$ and λ_3^{01} denotes Carol's decision threshold when she observes $(\hat{H}_1, \hat{H}_2) = (0, 1)$. Furthermore, $P_{e,2}^I$ denotes Britta's Type I error probability when $\hat{H}_1 = 1$ and $P_{e,3}^{II,0}$ denotes Carol's Type II error probability when $(\hat{H}_1, \hat{H}_2) = (1, 0)$.

The aim of this paper is to compare cases in which N agents can observe public signals—called *public voting*, see Fig. 1(c)—to cases in which they cannot—called *secret voting*, see Fig. 1(b). The comparison is done through study of the minimum achievable Bayes risk. In secret voting, minimization of Bayes risk is achieved by optimizing N decision thresholds—one for each agent. In public voting, the optimization

²In Section III-B, we consider incomplete public signals.

is over up to $2^N - 1$ decision thresholds because later-acting agents may have separate decision thresholds for each combination of observed public signals. The minima will depend on L as well ([2], §3.4); since the principal results hold for all L , they additionally hold for optimizing over L .

The comparison takes two forms. We prove that the minima are equal when all agents observe conditionally iid private signals (see Section III) or when the agents require unanimity of one type or the other (see Section IV-A). When neither of these conditions hold, we find through numerical optimization that public voting generally yields lower minimum Bayes risk (see Section IV-B).³ This improvement might be expected from the increased information exchange and greater number of degrees of freedom in optimization; the provable lack of improvement in some cases is perhaps more surprising.

III. IDENTICAL AGENTS

Agents are assumed to observe conditionally iid private signals in this section. We compare the optimal decision-making strategies in two cases: public voting and secret voting. By showing that the optimal decision thresholds are the same in these two cases, we will conclude that public signals are useless.

In secret voting, we restrict the agents to use identical decision thresholds to simplify the problem [15]. Using identical decision thresholds is known to be asymptotically optimum for the binary hypothesis testing problem [1]. Furthermore, it has also been long-known from numerical experiments 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 [19]. We performed numerical experiments for $N \in \{2, 3, \dots, 7\}$, $L \in \{1, 2, \dots, N\}$, and Gaussian and exponential likelihood functions. We found no performance improvement from allowing unequal decision thresholds; see [20, Fig. 3–2], for a depiction that applies to $N = 3$.

A. Complete Public Signals

Suppose that agents can observe all decisions made by previous agents. Specifically, Agent n observes $\hat{H}_1, \dots, \hat{H}_{n-1}$ before making her decision.

First, let us consider the simplest case of $N = 2$. Alexis and Britta make a decision together and their fusion rule is the 1-out-of-2 (OR) rule or the 2-out-of-2 (AND) rule. In secret voting, Alexis and Britta simultaneously make local decisions. They use one decision threshold each.

On the other hand, in public voting Alexis first makes a decision and then Britta makes a decision upon observing Alexis's decision. As always, Alexis has a single decision threshold. Britta also has only one relevant decision threshold, like in secret voting. While Britta seems to be able to choose two decision thresholds, one for $\hat{H}_1 = 0$ and another for $\hat{H}_1 = 1$, for each fusion rule there is one value of the public signal for which her decision becomes irrelevant to the team decision: under the OR rule, when $\hat{H}_1 = 1$ the team decision is $\hat{H} = 1$ regardless of Britta's local decision; and under the AND rule, when $\hat{H}_1 = 0$

the team decision is $\hat{H} = 0$ regardless of Britta's local decision. Thus, both agents equally have one degree of freedom even when the voting is public. This captures the basic idea behind public voting and secret voting yielding the same performance for $N = 2$. We provide a more technical proof below. To draw attention to the distinction between the two cases, we use λ to denote decision thresholds in secret voting and ρ to denote those in public voting.

Theorem 1: The existence of public signals does not affect the optimal local decision rules for $N = 2$ under either of the two L -out-of- N fusion rules.

Proof: The fusion rule will be the 1-out-of-2 (OR) rule or the 2-out-of-2 (AND) rule. Let us compare team Bayes risks in the secret voting and public voting cases under the OR rule. In the secret voting scenario, the Bayes risk is given by

$$R_s = c_{10}p_0 (P_{e,1}^I + (1 - P_{e,1}^I) P_{e,2}^I) + c_{01}p_1 P_{e,1}^{II} P_{e,2}^{II}, \quad (7)$$

where Alexis's decision threshold λ_1 determines local error probabilities $P_{e,1}^I$ and $P_{e,1}^{II}$, and Britta's decision threshold λ_2 determines local error probabilities $P_{e,2}^I$ and $P_{e,2}^{II}$. Their optimal decision thresholds λ_1^* and λ_2^* minimize (7).

In the public voting scenario, the Bayes risk is in the same form

$$R_p = c_{10}p_0 (P_{e,1}^I + (1 - P_{e,1}^I) P_{e,2}^{I_0}) + c_{01}p_1 P_{e,1}^{II} P_{e,2}^{II_0}, \quad (8)$$

except that Britta's error probabilities, $P_{e,2}^{I_0}$ and $P_{e,2}^{II_0}$, are controlled by her decision threshold ρ_2^0 for $\hat{H}_1 = 0$. Britta's decision threshold when $\hat{H}_1 = 1$, ρ_2^1 , is irrelevant; thus, we can assume that $\rho_2^1 = \rho_2^0$ without loss of optimality.

Expressions (7) and (8) are very similar; the only difference is the replacement of $(P_{e,2}^I, P_{e,2}^{II})$ in (7) by $(P_{e,2}^{I_0}, P_{e,2}^{II_0})$ in (8). Now note that the set of achievable values for $(P_{e,2}^I, P_{e,2}^{II})$ and $(P_{e,2}^{I_0}, P_{e,2}^{II_0})$ are identical; they are achieved by varying Britta's decision threshold (λ_2 or ρ_2^0) in precisely the same local decision-making problem. Therefore, minimizing R_s and R_p results in equal Bayes risks, and these are achieved with decision thresholds satisfying the following:

$$\rho_1^* = \lambda_1^*, \quad \rho_2^{0*} = \lambda_2^*. \quad (9)$$

Since $\lambda_1^* = \lambda_2^*$, we also have $\rho_1^* = \rho_2^{0*}$. Therefore, Alexis and Britta should not change their decision thresholds depending on whether or not the voting is public.

The proof for the AND fusion rule is similar. While the precise expressions for R_s and R_p change, again we find that the achievable set of local Type I and Type II error probabilities are identical under secret and public voting, so the minima and optimum decision thresholds are unaffected by the public signal. ■

For larger N , we prove the lack of usefulness of the public signal by mathematical induction.

Theorem 2: Suppose that sharing local decisions does not change Alexis's decision threshold (i.e., $\rho_1^* = \lambda_1^*$). If the optimal decision-making rules are the same in public voting and in secret voting for a specific N and any K -out-of- N fusion rule, then it

³Rather than display minimum Bayes risk, we plot the more evocative receiver operating characteristic (ROC) curves.

is also true for a team of $N + 1$ agents and any L -out-of- $(N + 1)$ fusion rule.

Proof: First, we consider the secret voting scenario with $N + 1$ agents. Since Agent n 's decision is critical only if the other N local decisions are $L - 1$ ones and $N - L + 1$ zeros, the optimal decision threshold λ_n^* is the solution to⁴

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

where we use that $P_{e,1}^I = P_{e,2}^I = \dots = P_e^I$ and $P_{e,1}^{II} = P_{e,2}^{II} = \dots = P_e^{II}$ because the optimal decision thresholds of all agents are identical in secret voting.

Next, in the public voting scenario, we can classify error cases depending on Alexis's detection result, e.g., when the true state is 0 and Alexis's decision is correct ($\hat{H}_1 = 0$), a false alarm occurs if at least L out of the remaining N agents vote for 1. The Bayes risk is given by

$$\begin{aligned} R_p &= c_{10}p_0 (1 - P_{e,1}^I) \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \geq L \mid \hat{H}_1 = H = 0 \right\} \\ &+ c_{10}p_0 P_{e,1}^I \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \geq L - 1 \mid \hat{H}_1 = 1, H = 0 \right\} \\ &+ c_{01}p_1 P_{e,1}^{II} \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \leq L - 1 \mid \hat{H}_1 = 0, H = 1 \right\} \\ &+ c_{01}p_1 (1 - P_{e,1}^{II}) \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \leq L - 2 \mid \hat{H}_1 = H = 1 \right\} \\ &\triangleq R_0 (p_0(1 - P_{e,1}^I) + p_1 P_{e,1}^{II}) + R_1 (p_0 P_{e,1}^I + p_1(1 - P_{e,1}^{II})), \end{aligned} \quad (11)$$

where R_0 and R_1 are specified in (13) and (14) and we define

$$\begin{aligned} q^0 &\triangleq \frac{p_0 (1 - P_{e,1}^I)}{p_0 (1 - P_{e,1}^I) + p_1 P_{e,1}^{II}} = \mathbb{P}\{H = 0 \mid \hat{H}_1 = 0\}, \\ q^1 &\triangleq \frac{p_0 P_{e,1}^I}{p_0 P_{e,1}^I + p_1 (1 - P_{e,1}^{II})} = \mathbb{P}\{H = 0 \mid \hat{H}_1 = 1\}. \end{aligned} \quad (12)$$

When Agents 2, 3, \dots , $N + 1$ observe $\hat{H}_1 = 0$, their optimal decision strategy is to minimize the term R_0 from (11):

$$\begin{aligned} R_0 &= c_{10}q^0 \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \geq L \mid H = 0 \right\} \\ &+ c_{01}(1 - q^0) \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \leq L - 1 \mid H = 1 \right\}, \end{aligned} \quad (13)$$

⁴Please see [21] for a detailed description of how (10) is derived.

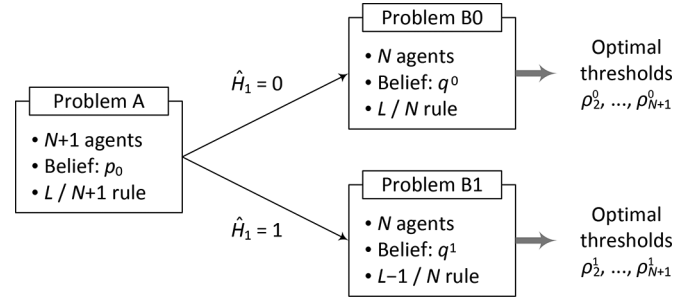


Fig. 3. An $(N + 1)$ -agent problem is divided into two N -agent problems depending on Alexis's decision \hat{H}_1 .

where the condition $\hat{H}_1 = 0$ is embedded in the superscript of \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 the fusion is done by the L -out-of- N rule. It implies that the optimal decision thresholds of Agents 2, 3, \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 Agents 2, 3, \dots , $N + 1$ observe $\hat{H}_1 = 1$, their optimal decision strategy is to minimize the term R_1 from (11):

$$\begin{aligned} R_1 &= c_{10}q^1 \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \geq L - 1 \mid H = 0 \right\} \\ &+ c_{01}(1 - q^1) \mathbb{P} \left\{ \sum_{n=2}^{N+1} \hat{H}_n \leq L - 2 \mid H = 1 \right\}. \end{aligned} \quad (14)$$

Their optimal decision thresholds are the same as those of N agents with prior probability q^1 and the $(L - 1)$ -out-of- N fusion rule. Fig. 3 depicts the evolution of the problem corresponding to Alexis's decision \hat{H}_1 .

Let us find the optimal thresholds $\rho_2^*, \rho_3^*, \dots, \rho_{N+1}^*$ in Problem B0 in Fig. 3. In fact, Problem B0 is also a public voting scenario; agents observe \hat{H}_2, \hat{H}_3 , and so on. However, because of the assumption 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 secret voting. Since an agent's decision is critical only if the other $N - 1$ local decisions consist of $L - 1$ ones and $N - L$ zeros, the optimal decision threshold ρ^{0*} is the solution to

$$\begin{aligned} \frac{f_{Y|H}(\rho^0 | 1)}{f_{Y|H}(\rho^0 | 0)} &= \frac{c_{10}q^0 \binom{N-1}{L-1} (P_e^{I_0})^{L-1} (1 - P_e^{I_0})^{N-L}}{c_{01}(1 - q^0) \binom{N-1}{N-L} (P_e^{II_0})^{N-L} (1 - P_e^{II_0})^{L-1}} \\ &= \frac{c_{10}p_0 (1 - P_e^I) (P_e^{I_0})^{L-1} (1 - P_e^{I_0})^{N-L}}{c_{01}p_1 P_e^{II} (P_e^{II_0})^{N-L} (1 - P_e^{II_0})^{L-1}}, \end{aligned} \quad (15)$$

where q^0 is replaced by (12). Due to the assumption that $\rho_1^* = \lambda_1^*$, P_e^I and P_e^{II} in (15) are the same as P_e^I and P_e^{II} in (10).

Comparing (15) to (10), 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 $\rho_2^*, \dots, \rho_{N+1}^*$ in Problem B1 in Fig. 3 by looking at the N -agent problem without public signals:

$$\begin{aligned} & \frac{f_{Y|H}(\rho^1 | 1)}{f_{Y|H}(\rho^1 | 0)} \\ &= \frac{c_{10} q^1 \binom{N-1}{L-2} (P_e^{I_1})^{L-2} (1 - P_e^{I_1})^{N-L+1}}{c_{01} (1 - q^1) \binom{N-1}{N-L+1} (P_e^{II_1})^{N-L+1} (1 - P_e^{II_1})^{L-2}} \\ &= \frac{c_{10} p_0 P_e^I (P_e^{I_1})^{L-2} (1 - P_e^{I_1})^{N-L+1}}{c_{01} p_1 (1 - P_e^{II}) (P_e^{II_1})^{N-L+1} (1 - P_e^{II_1})^{L-2}}. \quad (16) \end{aligned}$$

Again, due to the assumption that $\rho_1^* = \lambda_1^*$, P_e^I and P_e^{II} in (16) are the same as P_e^I and P_e^{II} in (10). We reach the same conclusion that the two equations have the same solutions, i.e., $\rho_i^* = \lambda_i^*$, by comparing (16) to (10). 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. ■

Corollary 3: Suppose that sharing local 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. ■

This result requires the assumption that Alexis uses the same decision rule in both secret and public voting. This assumption is trivially true for $N = 1$ and also true for $N = 2$ by proof. In addition, our numerical experiments for $N \leq 9$ confirmed that it is true. Thus, this assumption seems heuristically true. In particular, we fail to see how Alexis would choose between increasing or decreasing her decision threshold based on the existence of public signals.

Individual correctness vs. team correctness, and belief update vs. fusion rule evolution: One may have the intuition that, because the public signals are not independent of H , it cannot make sense for agents to ignore the public signals. Indeed, if the goal of Agent n is to make a correct decision herself (ignoring her role in the team) and she observes the public signal of Agent i before making her decision, she should not ignore \hat{H}_i . If $\hat{H}_i = 0$, then Agent n would include the information in her LRT to make a better choice as follows:

$$\frac{P_{Y_n, \hat{H}_i | H}(y_n, 0 | 1) \hat{H}_i \stackrel{>}{=} 1}{P_{Y_n, \hat{H}_i | H}(y_n, 0 | 0) \hat{H}_i \stackrel{<}{=} 0} \stackrel{>}{=} \frac{c_{10} p_0}{c_{01} p_1}. \quad (17)$$

Since Agent n 's private signal Y_n and Agent i 's private signal Y_i are independent given H , so are Y_n and \hat{H}_i . Then (17) is transformed as

$$\begin{aligned} & \frac{f_{Y_n | H}(y_n | 1) \hat{H}_i \stackrel{>}{=} 1}{f_{Y_n | H}(y_n | 0) \hat{H}_i \stackrel{<}{=} 0} \stackrel{>}{=} \frac{c_{10} p_0 \mathbb{P}\{\hat{H}_i = 0 | H = 0\}}{c_{01} p_1 \mathbb{P}\{\hat{H}_i = 0 | H = 1\}} \\ &= \frac{c_{10} p_0 \mathbb{P}\{H = 0 | \hat{H}_i = 0\}}{c_{01} p_1 \mathbb{P}\{H = 1 | \hat{H}_i = 0\}} = \frac{c_{10} q^0}{c_{10} (1 - q^0)}, \quad (18) \end{aligned}$$

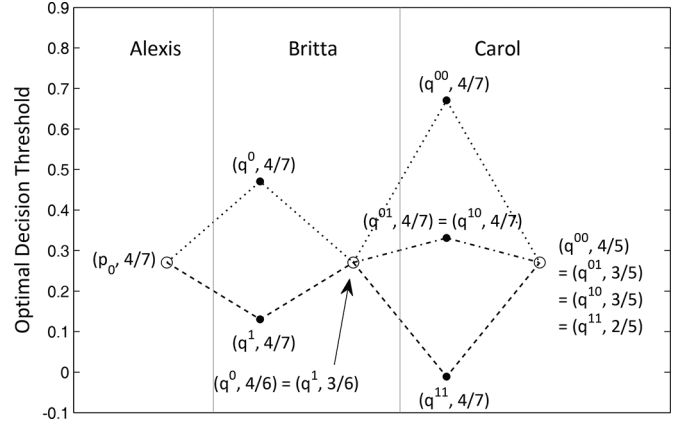


Fig. 4. Illustration of the separate effects of belief updates and fusion rule updates. The diagram depicts decision thresholds for the first three agents in an example with $p_0 = 0.25$ and the 4-out-of-7 fusion rule. Agents observe $H \in \{0, 1\}$ corrupted by iid Gaussian additive noises with zero mean and unit variance, and $c_{10} = c_{01} = 1$. The $(p, L/N)$ label on a marker (\bullet or \circ) indicates that its height represents the optimal decision threshold for prior probability p and the L -out-of- N fusion rule. Alexis has an initial decision threshold depending only on the prior probability p_0 and the 4-out-of-7 fusion rule (leftmost \circ). If Britta considers belief updates only, the optimal decision threshold is changed from Alexis's decision threshold to a new value that depends on \hat{H}_1 (two leftmost \bullet 's). However, after adopting the fusion rule evolution as well, Britta's optimal decision threshold returns to equal Alexis's decision threshold (center \circ). Similarly, if Carol considers belief updates only, the four values for (\hat{H}_1, \hat{H}_2) lead to three distinct decision thresholds (three rightmost \bullet 's). After accounting for the fusion rule evolution, Carol's optimal decision threshold returns to equal Alexis's decision threshold (rightmost \circ).

where $q^0 = \mathbb{P}\{H = 0 | \hat{H}_i = 0\}$. Agent n 's LRT with two observations Y_n and \hat{H}_i becomes an LRT with one observation Y_n as if she updates the prior probability of $H = 0$ from p_0 to q^0 upon observing $\hat{H}_i = 0$. More generally, the updated belief of $H = h$ has a form of $\mathbb{P}\{H = h | \text{observed public signals}\}$, assuming the agents observe conditionally independent private signals. The updated belief of $H = 0$ will be higher than the prior probability if previous agents have chosen 0 and lower if they have chosen 1.

Return now to the optimization of team performance. In the proof of Theorem 2, (13) and (14) reveal that the belief is updated in the same manner when the agents perform group decision making. The difference is that these formulas contain *fusion rule evolution* (depicted in Fig. 3) along with the *belief update*. While the belief update is common in the social learning literature—in fact, central to it—the evolution of the fusion rule hardly appears in the literature.

The fusion rule evolution arises in a public voting scenario because an agent should hesitate to invalidate the votes of later-acting agents. Each agent can keep a running tally of the numbers of 0 and 1 votes. A large number of 1 votes implies that only a few more 1 votes will determine the team decision to be 1 even before all agents vote; thus, in order to vote 1, an agent should need a strong private signal in support of 1 to take the risk of making later-acting agents' votes null. By Corollary 3, the effects of the belief update and fusion rule evolution cancel exactly. A numerical example is detailed in Fig. 4. Reading that figure from left to right, for each agent after Alexis, the belief update is done first and then the fusion rule evolution brings the

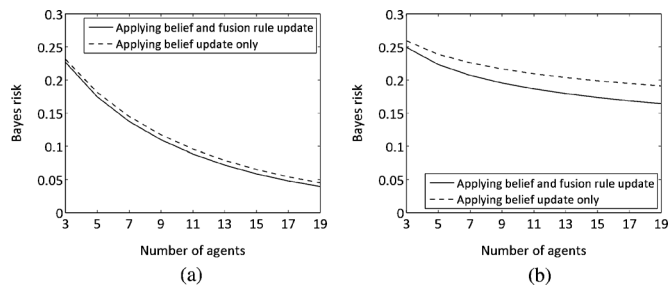


Fig. 5. Comparisons of Bayes risks when belief updates are or are not combined with fusion rule updates. Agents observe $H \in \{0, 1\}$ corrupted by iid Gaussian additive noises with zero mean and unit variance, $c_{10} = c_{01} = 1$, and $p_0 = 0.5$. (a) MAJORITY fusion rule. (b) OR fusion rule.

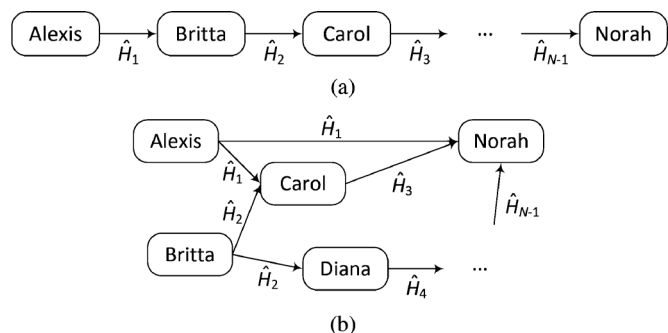


Fig. 6. Examples of partial public voting. (a) Each agent sequentially observes the decision made by the agent right before her. (b) Each agent observes the decisions made by her neighbors.

optimal decision threshold back exactly to the optimal decision threshold of Alexis.

To corroborate the theoretical results above, we made numerical computations of Bayes risks obtained with belief updates alone and those obtained with the combination of belief updates and fusion rule updates. Fig. 5 shows comparisons as the numbers of agents are varied for the MAJORITY and OR fusion rules. Consistent with the theoretical results, belief updates must be combined with fusion rule updates to minimize the Bayes risk.

In conclusion, while belief updates encourage later-acting agents to agree with the earlier-acting agents (the phenomenon that classically causes herding), decision-making by voting causes fusion rule evolution as well. This discourages agreement of later-acting agents with the earlier-acting agents. By proving that the optimal thresholds of the last N agents are the same in the cases when $\hat{H}_1 = 0$, when $\hat{H}_1 = 1$, and when they do not know \hat{H}_1 , it is proved that the effects of the former and the latter are exactly canceled out.

B. Incomplete Public Signals

We now extend our analysis to scenarios in which each agent is aware of an arbitrary subset of the previously made decisions. For example, each agent may observe the public signal only from its neighbors in a sequence (see Fig. 6(a)) or the communication topology may be more arbitrary (see Fig. 6(b)). Let us say that agents perform *partial public voting* when agents observe proper subsets of precedent local decisions.

Corollary 4: Observing a subset of public signals does not affect optimal decision rules and performance of a team.

Proof: First, we will show that a team of agents observing only a subset of public signals (Team A) cannot outperform a team of the same size that consists of those who observe full public signals (Team B). The proof is by contradiction. Let us assume that Team A can outperform Team B using their optimal decision strategy. Then, since what each agent in Team A observes is also observed by the corresponding agents in Team B, Team B can mimic the optimal strategy of Team A. For Agent n in Team B, all she has to do is ignore the public signals that Agent n in Team A cannot observe. After mimicking the strategy of Team A, the performance of Team B becomes the same as that of Team A. This contradicts our assumption. Hence Team A cannot outperform Team B.

Next, let us consider Team A and a team of agents not observing any public signals (Team C). We can prove that Team C cannot outperform Team A through similar logic.

Corollary 3 implies that Team C in fact performs as well as Team B. Therefore Team A is also as good as Teams B and C with respect to their optimal performances. ■

The convenience of secret voting emerges especially when agents cannot observe all public signals. Even though partial public voting cannot outperform public voting, the former requires more computations for Bayesian learning. When the agents observe a subset of public signals, they need to consider all possible realizations of the public signals that they cannot observe in order to perform Bayesian learning. For example, in Fig. 6(a), Carol observes Britta's decision but not Alexis's decision. Her updated belief when $\hat{H}_2 = 0$ will be computed as follows:

$$\begin{aligned}
 q_3^0 &= \mathbb{P}\{H = 0 \mid \hat{H}_2 = 0\} \\
 &= \sum_{\hat{h}_1} \mathbb{P}\{H = 0 \mid \hat{H}_2 = 0, \hat{H}_1 = \hat{h}_1\} \mathbb{P}\{\hat{H}_1 = \hat{h}_1 \mid \hat{H}_2 = 0\} \\
 &= \sum_{\hat{h}_1} \frac{\mathbb{P}\{H = 0 \mid \hat{H}_2 = 0, \hat{H}_1 = \hat{h}_1\} \mathbb{P}\{\hat{H}_1 = \hat{h}_1, \hat{H}_2 = 0\}}{\mathbb{P}\{\hat{H}_1 = 0, \hat{H}_2 = 0\} + \mathbb{P}\{\hat{H}_1 = 1, \hat{H}_2 = 0\}}.
 \end{aligned} \tag{19}$$

This process is more complicated than belief update with the knowledge of both $\hat{H}_1 = 0$ and $\hat{H}_2 = 0$, which is just to compute $\mathbb{P}\{H = 0 \mid \hat{H}_2 = 0, \hat{H}_1 = 0\}$.

Instead of accepting this complexity, the agents should ignore the public signals. Since the optimal secret voting strategy performs equally to the optimal public voting strategy, it is economical for them to not share any public signals at all.

IV. AGENTS WITH DIFFERENT LIKELIHOODS

Agents may have private signals that relate differently to the hypothesis. As one example, this occurs when we model each agent's private signal as the hypothesis observed over an additive noise channel, and the channels have different SNRs. We could think of the agents with relatively high SNRs as *experts* and those with low SNRs as *novices*. Their decisions are not equally informative, unlike in the identical-agent case of Section III. We will show that the public signals are futile in cases where the fusion rule requires unanimity but useful in other cases.

Now that the agents' private signals do not have the same distributions, the order in which the agents act matters to their team performance. To distinguish the agents, we name them in descending order of SNRs: Amy has the highest SNR, Beth has the second-highest SNR, and so on. However, we keep the notation that the numbering of agents (and hence the subscript indices) indicate the order of decision making, i.e., Agent 1 acts first, but may or may not be Amy.

A. AND and OR Fusion Rules

The AND (N -out-of- N) and OR (1-out-of- N) rules have a common feature. The team decision requires unanimity of the agents of one type or the other: for a team decision of 1 under the AND rule, the agents must unanimously decide 1; and for a team decision of 0 under the OR rule, they must unanimously decide 0. We thus call these *unanimity fusion rules*. This characteristic gives a special result for these fusion rules.

Let us consider a team of two agents with the OR rule: Amy with a higher SNR and Beth with a lower SNR. From the discussion of the two-agent case in Section III, both Amy and Beth have one degree of freedom. Suppose Amy makes her decision first with decision threshold λ_A , and Beth then makes her decision with decision threshold λ_B^0 , regardless of Amy's decision. Their minimum Bayes risk is

$$\min_{\lambda_A, \lambda_B^0} c_{10}p_0 \left(P_{e,A}^I + (1 - P_{e,A}^I) P_{e,B}^{I_0} \right) + c_{01}p_1 P_{e,A}^{II} P_{e,B}^{II_0}. \quad (20)$$

Now suppose that they switch their positions: Beth first makes her decision with decision threshold ρ_B , and Amy then makes her decision with decision threshold ρ_A^0 , regardless of Beth's decision. Their minimum Bayes risk is now given by

$$\min_{\rho_B, \rho_A^0} c_{10}p_0 \left(P_{e,B}^I + (1 - P_{e,B}^I) P_{e,A}^{I_0} \right) + c_{01}p_1 P_{e,B}^{II} P_{e,A}^{II_0}. \quad (21)$$

The situation is very similar to the proof of Theorem 1. We have two nearly identical expressions, (20) and (21), and each agent has a single decision threshold to vary. This leads to identical minima achieved with identical decision thresholds. Specifically, (21) can be rearranged to

$$\min_{\rho_B, \rho_A^0} c_{10}p_0 \left(P_{e,A}^{I_0} + (1 - P_{e,A}^{I_0}) P_{e,B}^I \right) + c_{01}p_1 P_{e,A}^{II_0} P_{e,B}^{II}. \quad (22)$$

Now the only difference between (20) and (22) is the replacement of $(P_{e,A}^I, P_{e,A}^{II}, P_{e,B}^{I_0}, P_{e,B}^{II_0})$ in (20) by $(P_{e,A}^{I_0}, P_{e,A}^{II_0}, P_{e,B}^I, P_{e,B}^{II})$ in (22). Each of these 4-tuples is the combinations of Type I and Type II error probabilities for Amy and Beth as each agent varies a single threshold; thus, the achievable 4-tuples are identical. Therefore, the optimal decision thresholds are also the same:

$$\lambda_A^* = \rho_A^{0*} \text{ and } \lambda_B^* = \rho_B^*. \quad (23)$$

In conclusion, not only is the public signal useless but also the agents' decision-making order does not affect the optimal

team decision. Their optimal strategy is just to adopt the decision thresholds λ_1^* and λ_2^* that minimize the Bayes risk:

$$(\lambda_1^*, \lambda_2^*) = \arg \min_{(\lambda_1, \lambda_2)} \{ c_{10}p_0 (1 - (1 - P_{e,1}^I) (1 - P_{e,2}^I)) + c_{01}p_1 P_{e,1}^{II} P_{e,2}^{II} \}. \quad (24)$$

We can extend this result to N agents as long as the fusion is performed by the OR rule or the AND rule.

Theorem 5: For a unanimity fusion rule, secret voting is the optimal strategy even when agents observe private signals with different SNRs. Specifically, public signals are useless and the ordering of agents does not affect their optimal decision rules nor the resulting performance.

Proof: For the OR rule, where the Bayes risk is given by

$$c_{10}p_0 \left(1 - \prod_{n=1}^N (1 - P_{e,n}^I) \right) + c_{01}p_1 \prod_{n=1}^N P_{e,n}^{II}, \quad (25)$$

each agent has a meaningful decision threshold only if all previous agents declare 0. Otherwise, decisions of the remaining agents are irrelevant. Thus, without loss of optimality, we can constrain that the agents optimize their decision thresholds λ_n for the case when $\hat{H}_1 = \hat{H}_2 = \dots = \hat{H}_{n-1} = 0$ and use the same decision threshold for any public signals. In fact, they need not know the public signals; they just need to perform decision making as if all public signals are 0.

Furthermore, the symmetry in (25) implies that indices of Agents m and n are interchangeable. Therefore, the ordering of agents does not affect the optimal values of decision thresholds and, consequently, the minimum Bayes risk.

For the AND rule, the Bayes risk is given by

$$c_{10}p_0 \prod_{n=1}^N P_{e,n}^I + c_{01}p_1 \left(1 - \prod_{n=1}^N (1 - P_{e,n}^{II}) \right), \quad (26)$$

and we can prove the statement in a similar way. ■

In general, the solution to minimizing (25) or (26) does not satisfy $\lambda_1^* = \lambda_2^* = \dots = \lambda_N^*$, unlike in Section III. Agents will use different optimal decision thresholds corresponding to their SNRs. We are arguing not that the agents should have the same decision thresholds but that they should keep their decision thresholds the same in any decision-making position. Furthermore, each agent should make a decision sincerely based on her private signal and should not care about public signals because her vote will count only when all the other agents vote in a particular way (0 for the OR rule or 1 for the AND rule).

B. Other Fusion Rules

Optimal decision making is more complex for other fusion rules due to the increase of degrees of freedom. Even in the simplest case when three agents make a decision with the MAJORITY (2-out-of-3) rule, the last two agents have two meaningful degrees of freedom each. The second agent can have different decision thresholds for $\hat{H}_1 = 0(\lambda_2^0)$ and for $\hat{H}_1 = 1(\lambda_2^1)$, and the last agent can also have different ones for $(\hat{H}_1, \hat{H}_2) = (0, 1)(\lambda_3^{01})$ and for $(\hat{H}_1, \hat{H}_2) = (1, 0)(\lambda_3^{10})$. (The third agent is irrelevant for $(\hat{H}_1, \hat{H}_2) = (0, 0)$ or $(\hat{H}_1, \hat{H}_2) = (1, 1)$ because the team decision has been made without her decision.)

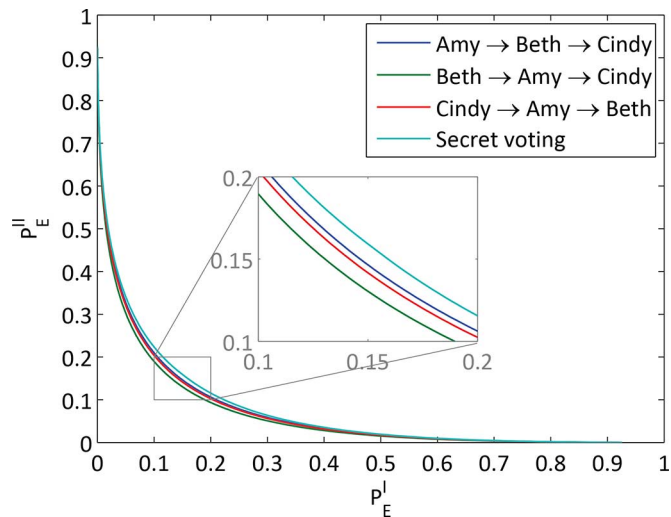


Fig. 7. Lower bounds of operating regions for different orderings of three agents, Amy ($\sigma_A^2 = 0.25$), Beth ($\sigma_B^2 = 1$), and Cindy ($\sigma_C^2 = 2.25$), and the MAJORITY fusion rule.

Learning from public signals can be helpful in decision making due to these extra degrees of freedom, unlike in Section IV-A.

Our symmetric fusion rule always treats all local decisions with equal weights even though they are made by agents that experience different SNRs. Thus, the team decision can be improved by social learning, which inevitably unbalances weights of local decisions.

Figs. 7 and 8 present quantitative evidence of helpful social learning. In these examples, we consider additive Gaussian noises with zero mean and variances σ_n^2 . The noises are mutually independent but not identically distributed.

In the case of Fig. 7, there are three agents using the MAJORITY fusion rule. Amy has the highest SNR ($\sigma_A^2 = 0.25$), Beth has the median SNR ($\sigma_B^2 = 1$), and Cindy has the lowest SNR ($\sigma_C^2 = 2.25$). Fig. 7 depicts the optimal reversed receiver operating characteristic (ROC) curves for all possible orderings of the actions of the three agents.⁵ We only need to consider three orderings because the order of the last two agents does not matter as we discussed in Section IV-A. The team's updated fusion rule after Amy makes a decision will be the AND or the OR rule, depending on Amy's decision.

There are two notable things in Fig. 7. First, the reversed ROC curve of secret voting is above that of public voting for any ordering. This implies that public voting strictly outperforms secret voting, regardless of the order in which the agents make decisions. Second, among the public voting scenarios, the best is when Beth makes her decision first. Further numerical experiments for $N = 3$ with several different noise variances also show that the team performance is the best when the agent with the median SNR acts first.

Fig. 8 shows the optimal reversed ROC curves for four agents and the 2-out-of-4 fusion rule. Again, we need not compare all 16 orderings. After the first agent makes a decision, the updated fusion rule will be the 1-out-of-3 rule if the decision is 1, in

⁵As decision threshold parameters are varied with the order of agent actions fixed, some set of (P_E^1, P_E^{11}) pairs is achievable. We call the lower boundary of this set the reversed ROC curve.

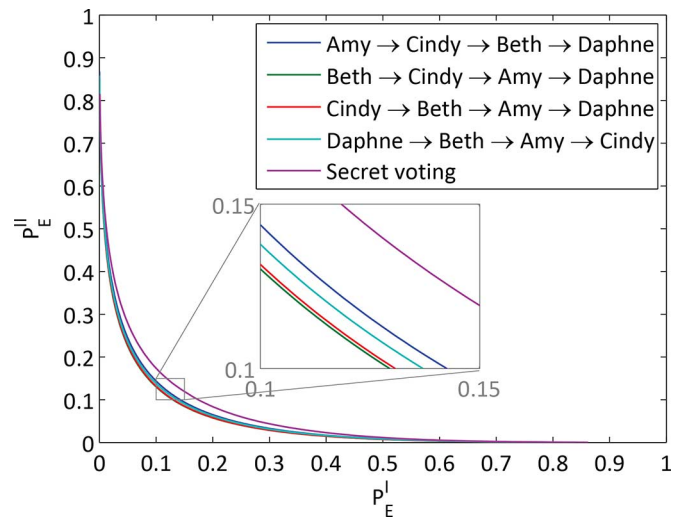


Fig. 8. Lower bounds of operating regions for different orderings of four agents, Amy ($\sigma^2 = 0.25$), Beth ($\sigma^2 = 0.5$), Cindy ($\sigma^2 = 1$), and Daphne ($\sigma^2 = 2.25$), and the 2-out-of-4 fusion rule.

which case the ordering of the rest of the agents does not matter (1-out-of-3 is a unanimity rule), and the 2-out-of-3 rule if the decision is 0, in which case we infer that the agent with median SNR should act next. Thus, we only need to compare four orderings, where the first agent is varying and the median of the remaining three becomes the second agent. Fig. 8 shows that the agent with the second-highest SNR should act first as well as that public voting always outperforms secret voting.

In this section, we have provided evidence that agents can exploit social learning to improve their team decision when the qualities of the private signals vary. The sequence of the agents also needs to be carefully chosen to achieve the best team performance.

V. CONCLUSION

A. Summary and Interpretation of Results

We have discussed the utility of social learning in a distributed detection and fusion system by a team of agents who have the same decision-making costs. We showed that it is not a trivial question whether intra-team social learning is useful. The model is intentionally designed to be simple so that we can understand fundamentals: agents have no conflicts of interests and their fusion rule is fixed and symmetric.

When Agent i casts a vote \hat{H}_i as an opinion on the value of the hypothesis H , except in degenerate situations, this vote is not independent of H . Thus, when the vote is available to other agents as a public signal, it is natural for them to update their beliefs about H ; this is the basis of the social learning concept. This paper highlights a countervailing phenomenon that arises when opinions are aggregated by voting: the running tally of previous votes changes the effective fusion rule for the agents who have yet to act. This fusion rule update is a less frequently explored topic.

When the agents observe conditionally iid private signals, the effects of the belief update and the evolution of the fusion rule cancel exactly. Consequently, the optimal performance with or

without public signals is the same; internally flowing information does not improve the team performance.

When the agents observe signals that are conditionally independent but not identically distributed, the existence of public signals may improve team performance, and when it does, the degree of improvement depends on the order in which the agents act. We first showed that the 1-out-of- N and N -out-of- N fusion rules are peculiar in that they essentially require unanimity among agents: under the 1-out-of- N rule, all agents must vote 0 for the team decision to be 0; under the N -out-of- N rule, all agents must vote 1 for the team decision to be 1. The unanimity rules do not allow a chance for the agents to specialize their decision thresholds for various public signals. Social learning can play a role to improve team decisions as long as the fusion rule is not one of these unanimity rules.

In examples with Gaussian likelihoods, we showed that team performance improves when agents with differing observation SNRs do social learning. With three agents making a team decision by the MAJORITY fusion rule, it is best for the agent with median SNR to act first. With four agents using the 2-out-of-4 fusion rule, it is best for the agent with second-best SNR to act first.

None of our results show that the existence of public signals makes team performance worse; logically, if being influenced by public signals would be detrimental, optimal agents would simply ignore them. We speculate, however, that public signals may be detrimental because, in human agents, belief updates overwhelm the decision rule update. In fact, without it being explained to them, people may not have intuition for the idea that a good agent does not take lightly the disenfranchisement of later-acting agents. Experiments with human subjects support the view that the public signals are detrimental, though the interpretations are merely preliminary [22], [23].

B. Limitations and Possible Extensions

Within the scope of our model, an intriguing open problem is to explain the observed patterns on the optimal order for agents with differing observation SNRs to act. The optimal ordering is nontrivial. Our experimentation with Gaussian likelihoods and up to four agents was extensive, but we do not know if the results extend to other likelihoods. Are there systematic ordering rules for larger N ?

Our model has several limitations, such as conditional independence of private signals, knowledge of the prior probability, and the agents having no motivation beyond minimization of the Bayes risk (a team performance criterion). Each of these merits further study. Perhaps the most interesting extensions are to the set of fusion rules.

We have assigned an equal weight to each vote in (1) and hence called the fusion symmetric. Suppose the fusion rule is subject to optimization. Then one could use the Chair-Varshney fusion rule [24], which is an algorithm that weights each individual vote according to its quality and accumulates them to compare to an optimally determined threshold. The effect of social learning in this scenario is an open question.

Our focus on equal weighting was motivated by its prevalence in most situations in which human opinions are aggregated by voting, ranging from government elections to juries

to decision-making by committees. In human affairs, unequal weighting of votes is unusual, except for the extreme cases of veto power and a single ultimate decision maker.

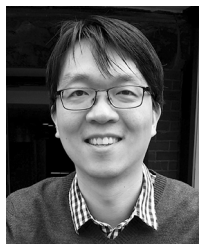
Having a single ultimate decision maker changes the situation rather dramatically. With respect to Fig. 2, one may consider Alexis, Britta, . . . , Mary to be *advisers* to the *decider* Norah, and simplify the fusion rule to $\hat{H} = \hat{H}_N$.⁶ Here the availability of public signals does improve team performance. This is similar to the situation studied in [11], where it is shown that an adviser with an inaccurate prior probability for H may provide a more helpful public signal to the decider.

REFERENCES

- [1] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Control, Signals, Syst.*, vol. 1, no. 2, pp. 167–182, 1988.
- [2] P. K. Varshney, *Distributed Detection and Data Fusion*. Berlin, Germany: Springer-Verlag, 1997.
- [3] A. V. Banerjee, "A simple model of herd behavior," *Quart. J. Econ.*, vol. 107, no. 3, pp. 797–817, Aug. 1992.
- [4] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *J. Polit. Econ.*, vol. 100, no. 5, pp. 992–1026, Oct. 1992.
- [5] L. Smith and P. Sørensen, "Pathological outcomes of observational learning," *Econometrica*, vol. 68, no. 2, pp. 371–398, Mar. 2000.
- [6] 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.
- [7] W. K. Viscusi, "Are individuals Bayesian decision makers?," *Amer. Econ. Rev.*, vol. 75, no. 2, pp. 381–385, May 1985.
- [8] 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.
- [9] C. J. Gill, L. Sabin, and C. H. Schmid, "Why clinicians are natural Bayesians," *Brit. Med. J.*, vol. 330, pp. 1080–1083, May 2005.
- [10] M. Glanzer, A. Hilford, and L. T. Maloney, "Likelihood ratio decisions in memory: Three implied regularities," *Psychonom. Bull. Rev.*, vol. 16, no. 3, pp. 431–455, June 2009.
- [11] J. B. Rhim and V. K. Goyal, "Social teaching: Being informative vs. being right in sequential decision making," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, July 2013, pp. 2602–2606.
- [12] C. Chamley, A. Scaglione, and L. Li, "Models for the diffusion of beliefs in social networks," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 16–29, May 2013.
- [13] M. Ottaviani and P. Sørensen, "Information aggregation in debate: Who should speak first?," *J. Pub. Econ.*, vol. 81, no. 3, pp. 393–421, Sept. 2001.
- [14] 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 Process. Mag.*, vol. 30, no. 3, pp. 43–57, May 2013.
- [15] J. B. Rhim and V. K. Goyal, "Keep ballots secret: On the futility of social learning in decision making by voting," in *Proc. IEEE Int. Conf. on Acoustics, Speech, Signal Process.*, Vancouver, Canada, May 2013, pp. 4231–4235.
- [16] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *Rev. Econ. Stud.*, vol. 78, no. 4, pp. 1201–1236, Mar. 2011.
- [17] R. Radner, "Team decision problems," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 857–881, Sept. 1962.
- [18] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Phil. Trans. Roy. Soc. London Ser. A*, vol. 231, pp. 289–337, Jan. 1933.
- [19] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, H. V. Poor and J. B. Thomas, Eds. Greenwich, CT, USA: JAI Press, 1993, pp. 297–344.

⁶This scenario gives the decider a private signal; if only advisers have private signals, then Norah could be another adviser, and the fusion center could be considered the decider.

- [20] J. B. Rhim, "Quantization of prior probabilities in bayesian group decision-making," Master's thesis, Massachusetts Inst. of Tech., Cambridge, MA, USA, Sep. 2010.
- [21] 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, Sep. 2012.
- [22] J. B. Rhim, "Aggregation and influence in teams of imperfect decision makers," Ph.D. dissertation, Massachusetts Inst. of Tech., Cambridge, MA, Jun. 2014.
- [23] J. B. Rhim and V. K. Goyal, "Social learning in team decision making," in *Proc. Collective Intell.*, Cambridge, MA, June 2014, paper 113.
- [24] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. ASE-22, no. 1, pp. 98–101, Jan. 1986.



Joong Bum Rhim (S'11–M'14) received the B.S. degree with honors in electrical engineering (summa cum laude) from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2008. He received the S.M. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, in 2010 and 2014, respectively.

He is a research scientist at GroupM Data and Analytics, New York. He was a Korea Foundation for

Advanced Studies Fellow and a research assistant at MIT.

Dr. Rhim received the student best paper award at the IEEE Sensor Array and Multichannel Signal Processing Workshop in 2012. His research interests include human perception, decision making, machine learning, and game theory.



Vivek K Goyal (S'92–M'98–SM'03–F'14) received the B.S. degree in mathematics and the B.S.E. degree in electrical engineering from the University of Iowa, where he received the John Briggs Memorial Award for the top undergraduate across all colleges. He received the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, where he received the Eliahu Jury Award for outstanding achievement in systems, communications, control, or signal processing.

He was a Member of Technical Staff in the Mathematics of Communications Research Department of Bell Laboratories, Lucent Technologies, 1998–2001; and a Senior Research Engineer for Digital Fountain, Inc., 2001–2003. He was with the Massachusetts Institute of Technology 2004–2013, where he was the Esther and Harold E. Edgerton Associate Professor of Electrical Engineering. He is now on leave from the Department of Electrical and Computer Engineering of Boston University. He was an adviser to 3dim Tech, Inc., and is now with Google. His research interests include computational imaging, human perception, decision making, sampling, quantization, and source coding theory.

Dr. Goyal is a member of Phi Beta Kappa, Tau Beta Pi, Sigma Xi, Eta Kappa Nu, SIAM, and AAAS. He was awarded the 2002 IEEE Signal Processing Society Magazine Award and an NSF CAREER Award. Work he supervised won student best paper awards at the IEEE Data Compression Conference in 2006 and 2011 and the IEEE Sensor Array and Multichannel Signal Processing Workshop in 2012 as well as four MIT thesis awards. He served on the IEEE Signal Processing Society (SPS) Image and Multiple Dimensional Signal Processing Technical Committee (TC) and the Steering Committee of the IEEE TRANSACTIONS ON MULTIMEDIA. He currently serves on the IEEE SPS Image, Video, and Multidimensional Signal Processing TC, the Editorial Board of *Foundations and Trends and Signal Processing*, and the Scientific Advisory Board of the Banff International Research Station for Mathematical Innovation and Discovery. He is a Technical Program Committee Co-chair of IEEE ICIP 2016 and Sampling Theory and Applications 2015, and a permanent Conference Co-chair of the SPIE Wavelets and Sparsity conference series. He is a co-author of *Foundations of Signal Processing* (Cambridge University Press, 2014).