

2012-09

# Quantization of prior probabilities for collaborative distributed hypothesis testing

---

Joong Bum Rhim, Lav R Varshney, Vivek K Goyal. 2012. "Quantization of Prior Probabilities for Collaborative Distributed Hypothesis Testing." IEEE Transactions on Signal Processing, Volume 60, Issue 9, pp. 4537 - 4550. <https://doi.org/10.1109/tsp.2012.2200890>

<https://hdl.handle.net/2144/42552>

*"Downloaded from OpenBU. Boston University's institutional repository."*

# Quantization of Prior Probabilities for Collaborative Distributed Hypothesis Testing

Joong Bum Rhim, Lav R. Varshney, and Vivek K Goyal

**Abstract**—This paper studies the quantization of prior probabilities, drawn from an ensemble, for distributed detection and data fusion. Design and performance equivalences between a team of  $N$  agents tied by a fixed fusion rule and a more powerful single agent are obtained. Effects of identical quantization and diverse quantization are compared. Consideration of perceived common risk enables agents using diverse quantizers to collaborate in hypothesis testing, and it is proven that the minimum mean Bayes risk error is achieved by diverse quantization. The comparison shows that optimal diverse quantization with  $K$  cells per quantizer performs as well as optimal identical quantization with  $N(K-1)+1$  cells per quantizer. Similar results are obtained for maximum Bayes risk error as the distortion criterion.

**Index Terms**—Bayesian hypothesis testing, Bregman divergence, mean Bayes risk minimization, quantization theory, team theory

## I. INTRODUCTION

Consider a team of  $N$  agents that aims to collaboratively choose between hypotheses  $h_0$  and  $h_1$  after each agent obtains a noisy observation. Local decision-making proceeds in parallel, with all agents synchronously observing the object, making hard decisions locally, and sending decisions to the fusion center, without any knowledge of other agents' decisions. The fusion center has some fixed fusion rule known to all agents that generates a global decision based only on local decisions. Fusion rules that are deterministic and symmetric are of the  $L$ -out-of- $N$  form whereby the global choice is  $h_1$  when  $L$  or more agents choose  $h_1$ . Examples are the MAJORITY rule ( $L = \lceil \frac{N+1}{2} \rceil$ ) and the OR rule ( $L = 1$ ). These arise in human affairs, where decision making by juries or committees has been analyzed in economics and political science to understand rational decision rules when the human decision makers possess a common preference for two alternatives [1].

With known prior probabilities for the object state, the local decision rules and the fusion rule can be analyzed and optimized as standard Bayesian hypothesis testing. Here, we consider an ensemble of objects, generally uncountably many. Due to limited memory or limited computational resources,

agents are only able to distinguish at most  $K$  different categories of objects (and therefore categories of prior probabilities). As an example, human decision makers are known to think categorically due to the limitation in their information processing capacity [2]. Due to categorization, the true prior probability of an observed object must be mapped to one of the  $K$  discriminable values before each agent performs Bayesian hypothesis testing. The quality of decisions depends on the individuals' categorization schemes.

In this paper, we study the effect of the categorization and optimization of the categorization. Since we limit our attention to binary hypothesis testing, a problem from the ensemble is specified by a single scalar  $p_0 = \mathbb{P}(H = h_0) = 1 - \mathbb{P}(H = h_1)$ . We model  $p_0$  as a realization of a random variable  $P_0$ , so mapping objects into  $K$  categories is equivalent to  $K$ -level quantization of  $P_0$ . Quantization performance is measured by the quality degradation of decisions made based on the quantized prior probability, a Bregman divergence called Bayes risk error [3] averaged over  $P_0$ , which is called mean Bayes risk error (MBRE). We consider minimum MBRE quantizers in two cases: when all agents use identical quantizers and when they use different quantizers.

A main result of the paper is a large advantage from diversity among agents. The traditional advantage of having multiple agents in hypothesis testing problems is obtaining more observations about objects: agents observe different noise realizations and reduce the effects of noise by information aggregation. New here, we show that diversity among agents' mapping schemes for prior probabilities can decrease the chance that inexact prior probabilities lead to wrong decisions. Beneficial diversity helps each agent in cancelling others' wrong decisions so that the quality of a global decision can be improved on average. Optimal design of the prior-probability quantizers enables  $N$  agents with different  $K$ -level quantizers to perform as well as  $N$  agents with identical  $(N(K-1)+1)$ -level quantizers.

The precise study of quantization of prior probabilities in Bayesian hypothesis testing was recently initiated in [4], which focuses on the minimum MBRE quantizer of a single agent. Quantization of prior probabilities to minimize maximum Bayes risk is considered in [5]. The study of quantization of prior probabilities in distributed hypothesis testing by three agents appears in [6], work generalized herein. We use a single set of Bayes costs as an element of making the agents a *team* in the sense of Marschak and Radner [7], i.e., having a common goal. An alternative is for each agent to have potentially-different Bayes costs. This introduces game-theoretic considerations as described in [8].

This material is based upon work supported by the National Science Foundation under Grants 0729069 and 1101147 and the Korea Foundation for Advanced Studies (KFAS) Fellowship. The material in this paper was presented in part at the IEEE Data Compression Conference, Snowbird, UT, March 2011.

J. B. Rhim (email: jbrhim@mit.edu) and V. K. Goyal (email: vgoyal@mit.edu) are with the Department of Electrical Engineering and Computer Science and the Research Laboratory of Electronics, Massachusetts Institute of Technology.

L. R. Varshney (email: varshney@alum.mit.edu) is with the IBM Thomas J. Watson Research Center.

Most previous work on the effect of quantization in Bayesian distributed detection is focused on the quantization of observations [9]–[11] or the communication topology and rates among agents [12] and/or to the fusion center [13], [14]. We do not consider quantization of observations here, though it may be noted that quantization outside of the system designer's control could be incorporated into the likelihood functions.

The group decision-making model that we are considering is described in Section II. In Section III, we discuss the effect of quantization of prior probabilities on decision making and compare performances of the teams of agents that use identical quantizers and that use different quantizers. Examples of optimal quantizers obtained from our design algorithm are presented in Section IV. In addition to the mean Bayes risk error, the maximum Bayes risk error is considered in Section V. Section VI concludes the paper.

## II. DISTRIBUTED DETECTION AND DATA FUSION MODEL

A binary hypothesis test for a given object is performed by a team of  $N$  agents. The object is in state  $H = h_0$  with probability  $p_0$  and in state  $H = h_1$  with probability  $1 - p_0$ . The agents have a common goal to minimize the cost due to the global decision; the cost for false alarm (misjudgment of  $h_0$  as  $h_1$ ) is  $c_{10}$  and the cost for missed detection (misjudgment of  $h_1$  as  $h_0$ ) is  $c_{01}$  for all agents. For simplicity, we consider zero cost for correct decisions.

The observation by Agent  $i$ ,  $Y_i$ , is governed by the likelihood function  $f_{Y_i|H}(y_i | H = h_m)$ . Each agent makes a hard local decision  $\hat{H}_i \in \{h_0, h_1\}$  based on its observation and the prior probability it believes. The local decision is transferred to a fusion center to be merged with other agents' decisions by some fixed fusion rule. We consider symmetric fusion rules, which are described as  $L$ -out-of- $N$  rules for a specific  $1 \leq L \leq N$ . This fusion rule returns  $\hat{H} = h_1$  if at least  $L$  agents declare  $h_1$ ; otherwise, it returns  $\hat{H} = h_0$ .

Each individual agent makes the two types of errors with the following probabilities:

$$\begin{aligned} P_{e,i}^I &= \mathbb{P}(\hat{H}_i = h_1 | H = h_0), \\ P_{e,i}^{II} &= \mathbb{P}(\hat{H}_i = h_0 | H = h_1). \end{aligned}$$

Then, from the  $L$ -out-of- $N$  fusion rule, the probabilities of the global decision being in error are:

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

$$P_E^{II} = \sum_{n=N-L+1}^N \sum_{\substack{I \subseteq [N] \\ |I|=n}} \prod_{i \in I} P_{e,i}^{II} \prod_{j \in [N]/I} (1 - P_{e,j}^{II}), \quad (2)$$

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

$$R = p_0 c_{10} P_E^I + (1 - p_0) c_{01} P_E^{II}. \quad (3)$$

Agent  $a$  minimizes its Bayes risk by adopting the following likelihood ratio test:

$$\frac{f_{Y_a|H}(y_a | h_1)}{f_{Y_a|H}(y_a | h_0)} \underset{\substack{\hat{H}_a(y_a)=h_1 \\ \gtrsim \\ \hat{H}_a(y_a)=h_0}}{\gtrsim} \frac{p_0 c_{10} \sum_{\substack{I \subseteq [N]/\{a\} \\ |I|=L-1}} \prod_{i \in I} P_{e,i}^I \prod_{j \in [N]/I/\{a\}} (1 - P_{e,j}^I)}{(1 - p_0) c_{01} \sum_{\substack{I \subseteq [N]/\{a\} \\ |I|=N-L}} \prod_{i \in I} P_{e,i}^{II} \prod_{j \in [N]/I/\{a\}} (1 - P_{e,j}^{II})}. \quad (4)$$

Consider the special case of agents making observations through additive Gaussian noise. That is, being in state  $h_m$  sends a signal  $s_m \in \mathbb{R}$  to all agents, but Agent  $i$  receives the corrupted signal  $Y_i = s_m + W_i$ , where the noise  $W_i$  is assumed to be iid Gaussian with zero mean and variance  $\sigma^2$ . Then the likelihood ratio test can be simplified to the decision rule with a decision threshold  $\lambda_a$ :

$$y_a \underset{\substack{\hat{H}_a(y_a)=h_1 \\ \gtrsim \\ \hat{H}_a(y_a)=h_0}}{\gtrsim} \lambda_a.$$

By picking the  $L$ -out-of- $N$  rule as their fusion rule, the agents have symmetry in their decision making under iid additive noise. All agents have the same likelihood ratio test (4), which implies that we can constrain the agents to use identical decision rules.

Limiting attention to identical decision rules considerably simplifies the problem. The solution with an identical local decision rule constraint is asymptotically optimum for the binary hypothesis testing problem [15]. Furthermore, numerical experience shows that the constraint results in little or no loss of performance for finite  $N$  and the optimal fusion rule has the form of an  $L$ -out-of- $N$  rule [13]. These results have been obtained in decentralized detection models in which the fusion rule is to be optimized in addition to the local decision rules. Our model has a fixed  $L$ -out-of- $N$  fusion rule and only local decision rules are subject to be optimized. Numerical experiments show that the restriction to identical decision rules leads to no loss of performance for  $N \leq 5$  in the Gaussian likelihoods case (see Sections IV-A and IV-B) and the exponential likelihoods case (see Section IV-C). Thus, we constrain the agents to use identical decision rules in the discussion ahead.

## III. QUANTIZATION OF PRIOR PROBABILITIES

We replace the classical formulation that all agents know the true prior probability  $p_0 = \mathbb{P}(H = h_0)$  with the setting where each Agent  $i$  bases its decision rule on  $q_i(p_0)$ , where  $q_i$  a scalar quantizer on  $[0, 1]$ . Such a setting can arise when the team faces an ensemble of binary hypothesis testing problems and the agents lack the ability to have a different decision rule for each problem. The prior probability  $p_0$  is thus modeled as a realization of a random variable  $P_0$  with density function  $f_{P_0}$ . This section discusses computation of Bayes risk averaged over  $P_0$  and the optimization of the quantizers applied to the prior probability.

In order to maximize team performance, all agents must coordinate by sharing their quantized prior probabilities. The following steps describe the collaborative decision making process.

- 1) *Quantizer design*: The  $K$ -level quantizers  $q_1, \dots, q_N$  are designed. These quantizers remain fixed for all time.
- 2) *Coordination*: Agents encounter an object whose prior probability is  $p_0$ . Each agent applies its quantizer to the prior probability and sends the output  $q_i(p_0)$  to all other agents as a part of collaboration.
- 3) *Decision rule design*: Agents design the best common decision rule based on  $q_1(p_0), \dots, q_N(p_0)$ .
- 4) *Signal observation and decision making*: Each agent observes a noisy signal  $Y_i$  and applies the decision rule to make a local decision  $\hat{H}_i$ .
- 5) *Decision fusion*: All local decisions  $\hat{H}_i$  are fused to produce a global decision  $\hat{H}$ .

Quantizer design and decision rule design are detailed in this section. The operations of local decision making and global decision fusion have already been discussed. Coordination simply involves communication.

The Bayes risk incurred when each agent bases its local decision on its quantized prior probability is denoted  $R_M(q_1(p_0), \dots, q_N(p_0))$ .<sup>1</sup> This is called *mismatched Bayes risk* to contrast to the true Bayes risk  $R(p_0)$  if the agents know the true prior probability  $p_0$ . The Bayes risk error is defined as

$$d(p_0, q_1(p_0), \dots, q_N(p_0)) = R_M(q_1(p_0), \dots, q_N(p_0)) - R(p_0).$$

We compute the mean Bayes risk error (MBRE) as the distortion of the quantizers for prior probabilities, which measures the average performance over all  $p_0$ :

$$\begin{aligned} D &= \mathbb{E}[d(P_0, q_1(P_0), \dots, q_N(P_0))] \\ &= \int_0^1 d(p_0, q_1(p_0), \dots, q_N(p_0)) f_{P_0}(p_0) dp_0. \end{aligned}$$

Minimum MBRE quantizers are to be designed within this model.

### A. Single Agent

Let us first review the case of  $N = 1$  [4]. The optimal choice of  $K$  values for quantized prior probabilities is described by the minimum MBRE quantizer

$$q_1^* = \arg \min_{q_1} \mathbb{E}[d(P_0, q_1(P_0))],$$

where the distortion measure of the single quantizer  $q_1$  is the Bayes risk error

$$\begin{aligned} d(p_0, q_1(p_0)) &= R_M(q_1(p_0)) - R(p_0) \\ &= p_0 c_{10} P_E^I(q_1(p_0)) + (1 - p_0) c_{01} P_E^{II}(q_1(p_0)) \\ &\quad - p_0 c_{10} P_E^I(p_0) - (1 - p_0) c_{01} P_E^{II}(p_0). \end{aligned}$$

*Theorem 1*: [4, Thm. 1] The Bayes risk error  $d(p_0, a)$  is nonnegative and only equal to zero when  $p_0 = a$ . As a function of  $p_0 \in (0, 1)$ , it is continuous and strictly convex for all  $a$ .

<sup>1</sup>Note that by coordination, each agent knows other agents' quantized prior probabilities.

*Theorem 2*: [4, Thm. 2] For any deterministic likelihood ratio test  $\hat{H}(\cdot)$ , as a function of  $a \in (0, 1)$  for all  $p_0$ , the Bayes risk error  $d(p_0, a)$  has exactly one stationary point, which is a minimum.

Due to the strict convexity of  $d(p_0, a)$  in  $p_0$  for all  $a$ , quantizers that satisfy necessary conditions for MBRE optimality are regular; quantization cells are subintervals  $\mathcal{R}_1 = [0, b_1], \mathcal{R}_2 = [b_1, b_2), \dots, \mathcal{R}_K = [b_{K-1}, 1]$ , and each representation point  $a_k$  is in  $\mathcal{R}_k$ . The necessary conditions for the optimality of a quantizer for  $f_{P_0}(p_0)$  are now described.

A nearest neighbor condition describes an expression for the cell boundaries  $\{b_k\}$  for fixed representation points  $\{a_k\}$ . Between two consecutive representation points  $a_k$  and  $a_{k+1}$ , the cell boundary  $b_k$  needs to separate  $x_1$  such that  $d(x_1, a_k) < d(x_1, a_{k+1})$  and  $x_2$  such that  $d(x_2, a_k) > d(x_2, a_{k+1})$ . Thus,  $b_k = b^*$  is obtained from the condition

$$d(b^*, a_k) = d(b^*, a_{k+1}). \quad (5)$$

The point is found to be

$$b^* = \frac{c_{01} (P_E^{II}(a_{k+1}) - P_E^{II}(a_k))}{c_{01} (P_E^{II}(a_{k+1}) - P_E^{II}(a_k)) - c_{10} (P_E^I(a_{k+1}) - P_E^I(a_k))}.$$

A centroid condition describes optimal representation points for fixed quantization cells. The MBRE is expressed as the sum of integrals over quantization regions

$$D = \sum_{k=1}^K \int_{\mathcal{R}_k} d(p_0, a_k) f_{P_0}(p_0) dp_0,$$

and the minimization may be performed for each cell separately. The representation point  $a_k = a^*$  of the cell  $\mathcal{R}_k$  is chosen from the optimization problem

$$a^* = \arg \min_a \left\{ \int_{\mathcal{R}_k} d(p_0, a) f_{P_0}(p_0) dp_0 \right\}.$$

Since Bayes risk error is a Bregman divergence [3], the unique minimizer is the centroid of the region  $\mathcal{R}_k$ :

$$a^* = \frac{\int_{\mathcal{R}_k} p_0 f_{P_0}(p_0) dp_0}{\int_{\mathcal{R}_k} f_{P_0}(p_0) dp_0}.$$

The minimum MBRE quantizer  $q_1^*$  can be found by the iterative Lloyd–Max algorithm, which alternates between the nearest neighbor and the centroid conditions.

### B. Identical Quantizers

Consider, for  $N > 1$ , that all agents use the same quantizer  $q$  for prior probabilities; coordination informs all agents of the identity. For an object with prior probability  $p_0$ , all agents use  $p' = q(p_0)$  as the object's prior. The agents will incur the mismatched Bayes risk

$$R_M = p_0 c_{10} P_E^I + (1 - p_0) c_{01} P_E^{II} \quad (6)$$

due to their decisions, but what they minimize is *perceived Bayes risk*:

$$R_P = p' c_{10} P_E^I + (1 - p') c_{01} P_E^{II}. \quad (7)$$

The probabilities  $P_E^I$  and  $P_E^{II}$  in (6) and (7) are identical. In other words,  $P_E^I$  and  $P_E^{II}$  are determined from the decision

rules that minimize the perceived Bayes risk (7) and applied to compute the mismatched Bayes risk (6).

A general way to find local decision rules is to directly optimize (1) and (2), but these are generally complicated functions. One way to simplify the analysis for additive observation noise models is to find a decision-making model with a single agent whose performance is the same as the team of the multiple agents. By the same performance, we mean that the single agent uses the same optimal decision rule and its Type I and Type II error probabilities are respectively equal to  $P_E^I$  and  $P_E^{II}$  when it uses the same quantizer  $q$  for prior probabilities.

*Theorem 3:* Assume that the unquantized prior probability is known. Consider  $N$  agents that perform group decision-making with observations corrupted by additive noises  $W_1, \dots, W_N$ . For convenience, index the agents in descending order of the realizations of the noises:  $W_{(1)} \geq W_{(2)} \geq \dots \geq W_{(N)}$ . When their decisions are fused by the  $L$ -out-of- $N$  rule, their performance is the same as that of a single agent having the same Bayes costs if its observation is corrupted by the  $L$ th largest additive noise  $V = W_{(L)}$ .

*Proof:* The global decision is the same as the decision of Agent  $L$  because all agents adopt the same decision threshold  $\lambda$ . If Agent  $L$  declares  $h_0$ , then Agents  $L+1, \dots, N$ , whose observations are smaller than or the same as that of Agent  $L$ , also declare  $h_0$ . Since at least  $N-L+1$  agents send  $h_0$ , the fusion rule gives  $h_0$  as the global decision. If Agent  $L$  declares  $h_1$ , then Agents  $1, \dots, L-1$ , whose observations are larger than or at least the same as that of Agent  $L$ , also declare  $h_1$  and the global decision is  $h_1$ .

As such, the Bayes risk (3) can be rewritten as

$$\begin{aligned} R &= p_0 c_{10} P_E^I + (1-p_0) c_{01} P_E^{II} \\ &= p_0 c_{10} \mathbb{P}(W_{(L)} + s_0 \geq \lambda | H = h_0) \\ &\quad + (1-p_0) c_{01} \mathbb{P}(W_{(L)} + s_1 < \lambda | H = h_1). \end{aligned} \quad (8)$$

If we consider a new single-agent problem with additive noise  $V = W_{(L)}$ , then the Bayes risk of the single agent is equal to (8). Therefore, the optimal decision rule of the single agent is equal to that of the multiple agents, and the single agent obtains the same performance as the team of multiple agents. ■

When the  $W_i$  are iid continuous random variables with pdf  $f_W$  and cdf  $F_W$ , the random variable  $V = W_{(L)}$  is well understood from the theory of order statistics [16]. The pdf of  $V$  is

$$f_V(v) = \frac{N!}{(N-L)!(L-1)!} F_W^{N-L}(v) [1 - F_W(v)]^{L-1} f_W(v).$$

Thus, we only need to consider a single agent with a different additive noise  $V$ , no matter how many agents there are. The Bayesian decision rule of the single agent that observes  $Y = s_m + V$  is given by

$$\frac{f_{Y|H}(y|h_1)}{f_{Y|H}(y|h_0)} = \frac{f_V(y-s_1)}{f_V(y-s_0)} \frac{\hat{H}(y)=h_1}{\hat{H}(y)=h_0} \frac{p_0 c_{10}}{(1-p_0) c_{01}},$$

which can also be used by the multiple agents.

*Theorem 4:* If multiple agents use identical quantizers for prior probabilities, then their optimal quantizer is equal to the optimal quantizer in the equivalent single-agent model.

*Proof:* Since Theorem 3 is valid for any prior probability, Theorem 3 holds for quantized prior probabilities whenever the multiple agents and the single agent quantize the prior probability to the same value. Thus, their average performances are also the same if they use the same quantizer  $q$  for prior probability. As a result, the single agent can achieve the minimum MBRE by adopting the optimal quantizer of the multiple agents and vice versa. ■

The analysis and design of the minimum MBRE quantizer in a single-agent model in [4] can be applied to the multiple-agent model without any change except the noise model.

### C. Diverse Quantizers

Now consider the setting where each agent may have its own quantizer for prior probabilities. For an object with prior probability  $p_0$ , Agent  $i$  believes that the probability of the object being in state  $h_0$  is  $p^{(i)} = q_i(p_0)$ , which may be different from Agent  $j$ 's quantization  $p^{(j)} = q_j(p_0)$ . This diversity may allow the agents to improve the quality of their decisions as a team.

In order to take advantage of diversity, the agents need to collaborate in their decision making. The coordination phase of the collaborative decision making process is key. When agents use identical quantizers, all agents consider the same perceived Bayes risk (7) even if they do not communicate, but here collaboration must be established.

Agents may try to minimize their own perceived Bayes risks so as to give the best decisions for all agents. When all agents quantize the prior probability to different values, however, they have different perceived Bayes risks:

$$R_P^{(i)} = p^{(i)} c_{10} P_E^I + (1-p^{(i)}) c_{01} P_E^{II} \neq R_P^{(j)}.$$

If all agents individually optimize their decision rules, the resulting performance is not as good as their best [8].

All agents will incur the same cost as a result of their decisions:  $c_{10}$  if their team misreads  $h_0$  as  $h_1$  and  $c_{01}$  if their team misreads  $h_1$  as  $h_0$ . This motivates them to collaborate by sharing a common goal that replaces their individual perceived Bayes risks. As the common goal, we introduce a weighted sum of the perceived Bayes risks called *perceived common risk*:

$$R_{PC} = \sum_{i=1}^N u_i R_P^{(i)},$$

where  $u_i$  are constants that satisfy  $\sum_{i=1}^N u_i = 1$  and  $u_i > 0$  for all  $i \in [N]$ .

The introduction of the perceived common risk allows us to treat identical-quantizer and diverse-quantizer settings on a common footing. The minimum perceived-common-risk decision rule can be used since, if all agents use identical quantizers like in Section III-B, their perceived common risk

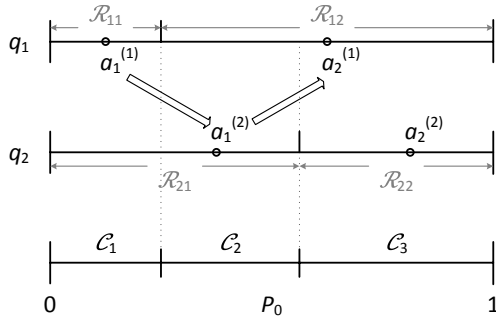


Fig. 1. An example that shows how the dependency among representation points propagates to break the independence between representation points of two different cells  $\mathcal{R}_{11}$  and  $\mathcal{R}_{12}$ .

is equal to their perceived Bayes risks:

$$\begin{aligned} R_{PC} &= \sum_{i=1}^N u_i \left( p^{(i)} c_{10} P_E^I + (1 - p^{(i)}) c_{01} P_E^{II} \right) \\ &= \sum_{i=1}^N u_i \left( p' c_{10} P_E^I + (1 - p') c_{01} P_E^{II} \right) \\ &= p' c_{10} P_E^I + (1 - p') c_{01} P_E^{II} = R_P. \end{aligned}$$

It may seem unintuitive to use the perceived common risk as a criterion of decision making because decision rules, which are determined based on the perceived common risk, may not be optimal for some prior probability  $p_0$ . However, it is the Bayes risk averaged over  $P_0$  that measures the performance of the team. Thus, a set of quantizers that leads to good decisions on average is essential for the agents to use the perceived common risk as a common goal of minimization.

Design of such quantizers has high computational complexity. In Section III-B, a single quantizer is easily designed by the iterative Lloyd–Max algorithm, which utilizes independence among endpoints of quantization cells in the nearest neighbor step (optimizing cell boundaries) and independence among representation points in the centroid step (optimizing representation points). On the other hand, multiple quantizers that define different quantization cells do not have such interdependencies.

For the simple example of Fig. 1, the cell  $\mathcal{R}_{21}$  of the quantizer  $q_2$  affects the decisions for objects whose prior probabilities are within the interval  $\mathcal{C}_1$  or  $\mathcal{C}_2$ . The team performance in the interval  $\mathcal{C}_1$  is affected by the representation points  $a_1^{(1)}$  and  $a_1^{(2)}$  and by  $a_2^{(1)}$  and  $a_1^{(2)}$  in the interval  $\mathcal{C}_2$ . Hence, we can observe that dependency is propagated through the representation points: a choice of  $a_2^{(1)}$  depends on a choice of  $a_1^{(2)}$ , which depends on a choice of  $a_1^{(1)}$ . In order to avoid such complexity, we introduce an indirect method to optimize diverse quantizers.

Consider a team of  $N$  agents that respectively use different  $K$ -level quantizers  $q_1, \dots, q_N$  and another team of  $N$  agents that all use an identical  $K_S$ -level quantizer  $q_S$ .

*Theorem 5:* A set of  $N$  different  $K$ -level quantizers  $\{q_1, \dots, q_N\}$  and a  $K_S$ -level quantizer  $q_S$  result in the same

perceived common risk if, for all  $p_0 \in [0, 1]$ ,

$$\sum_{i=1}^N u_i q_i(p_0) = q_S(p_0). \quad (9)$$

*Proof:* For an object with prior probability  $p_0$ , agents respectively using  $\{q_1, \dots, q_N\}$  make decisions based on individually quantized prior probabilities  $p^{(i)} = q_i(p_0)$ . Their perceived common risk is:

$$\begin{aligned} R_{PC,1} &= \sum_{i=1}^N u_i \left( p^{(i)} c_{10} P_E^I + (1 - p^{(i)}) c_{01} P_E^{II} \right) \\ &= \left( \sum_{i=1}^N u_i p^{(i)} \right) c_{10} P_E^I + \left( \sum_{i=1}^N u_i (1 - p^{(i)}) \right) c_{01} P_E^{II}. \end{aligned} \quad (10)$$

For agents using the common quantizer  $q_S$ , with quantized prior probability  $p' = q_S(p_0)$ , their perceived common risk is

$$R_{PC,2} = p' c_{10} P_E^I + (1 - p') c_{01} P_E^{II}. \quad (11)$$

The perceived common risks (10) and (11) only depend on the decision rules used by  $N$  agents. If there exists a constant  $t$  that satisfies

$$R_{PC,1} = t \cdot R_{PC,2}$$

for any  $N$ -tuple of decision rules, then the  $N$ -tuple that minimizes  $R_{PC,1}$  also minimizes  $R_{PC,2}$ . If the proportionality property holds for all  $p_0 \in [0, 1]$ , then the agents in both cases always use the same decision rules. Therefore, the set of different quantizers  $\{q_1, \dots, q_N\}$  and the set of identical quantizers  $\{q_S, \dots, q_S\}$  will cause the same Bayes risk on any problem in the ensemble.

For any  $N$ -tuple and any  $p_0$ ,  $R_{PC,1} = t \cdot R_{PC,2}$  if and only if  $\sum_{i=1}^N u_i p^{(i)} = t p'$  and  $\sum_{i=1}^N u_i (1 - p^{(i)}) = t(1 - p')$ . From the fact that  $\sum_{i=1}^N u_i = 1$ , the constant  $t$  can only be 1, and the condition is simplified to

$$\sum_{i=1}^N u_i q_i(p_0) = q_S(p_0), \quad \text{for all } p_0 \in [0, 1].$$

This simple condition comes from the fact that the perceived common risk is a weighted sum of the perceived Bayes risks. ■

*Theorem 6:* A team of  $N$  agents individually using  $N$  diverse  $K$ -level quantizers can achieve the minimum mean Bayes risk error that they can achieve when they use the same  $(N(K - 1) + 1)$ -level quantizer.

*Proof:* When the agents use an identical  $K_S$ -level quantizer, the optimal quantizer  $q_S$  is always a regular quantizer [4]; each quantization cell  $\mathcal{C}_k$  is an interval and its representation point  $x_k$  is within the interval. For each cell  $\mathcal{C}_k$  of  $q_S$ ,  $k = 1, \dots, K_S$ , the condition (9) is expressed by an equation

$$x_k = \sum_{i=1}^N a^{(i)} u_i,$$

where  $a^{(i)}$  is one of the representation points of  $K$ -level quantizer  $q_i$ ,  $i = 1, \dots, N$ . Overall, we have a total of  $K_S$

equations, which are described in the following matrix form:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_{K_S} \end{bmatrix} = \mathbf{A} \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix}, \quad (12)$$

where  $\mathbf{A}$  is a  $K_S \times N$  matrix that has  $[a_{k_{j_1}}^{(1)} a_{k_{j_2}}^{(2)} \cdots a_{k_{j_N}}^{(N)}]$  as its  $j$ th row and  $a_{k_{j_i}}^{(i)}$  is the representation point of the  $k_{j_i}$ th cell of  $q_i$  for any  $1 \leq k_{j_i} \leq K$ .

There exists an  $\mathbf{A}$  that satisfies (12) if and only if there exists a set of  $N$  different  $K$ -level quantizers  $q_1, \dots, q_N$  that are equivalent to  $q_S$ . In (12), the vector  $[u_1 u_2 \cdots u_N]^T$  is a given parameter and the vector  $[x_1 x_2 \cdots x_{K_S}]^T$  is uniquely determined by the optimization of  $q_S$ ; the matrix  $\mathbf{A}$  consists of unknown parameters (i.e., representation points of  $q_i$ ) to be determined. However, if  $\mathbf{A}$  has any linearly dependent row, the representation points may not exist. On the other hand, if  $\mathbf{A}$  has only linearly independent rows and  $K_S \leq NK$ , then we can solve (12) to find all the representation points of  $q_i$ .

The maximum number of linearly independent rows of  $\mathbf{A}$  is  $N(K-1) + 1$ :

$$\mathbf{A} = \begin{bmatrix} a_1^{(1)} & a_1^{(2)} & \cdots & a_1^{(N)} \\ a_2^{(1)} & a_1^{(2)} & \cdots & a_1^{(N)} \\ a_1^{(1)} & a_2^{(2)} & \cdots & a_1^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{(1)} & a_1^{(2)} & \cdots & a_2^{(N)} \\ a_3^{(1)} & a_1^{(2)} & \cdots & a_1^{(N)} \\ a_1^{(1)} & a_3^{(2)} & \cdots & a_1^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{(1)} & a_1^{(2)} & \cdots & a_3^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ a_N^{(1)} & a_1^{(2)} & \cdots & a_1^{(N)} \\ a_1^{(1)} & a_N^{(2)} & \cdots & a_1^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{(1)} & a_1^{(2)} & \cdots & a_N^{(N)} \end{bmatrix}.$$

Any other row  $[a_{k_1}^{(1)} a_{k_2}^{(2)} \cdots a_{k_N}^{(N)}]$  is a linear combination of the  $N(K-1) + 1$  rows. Thus, the existence of  $\{q_1, \dots, q_N\}$  that is equivalent to  $q_S$  is only guaranteed when  $K \leq K_S \leq N(K-1) + 1$ .

A quantizer with greater  $K_S$  does not increase mean Bayes risk error [4]. Therefore, the minimum mean Bayes risk error of agents using diverse  $K$ -level quantizers is upper bounded by that of agents using the identical  $(N(K-1) + 1)$ -level quantizer. ■

The optimal set of  $K$ -level quantizers for  $N$  agents is always a set of diverse quantizers if  $K > 1$ . The optimal set of quantizers can be designed by a two-step algorithm. The first step is to design the optimal  $(N(K-1) + 1)$ -level quantizer  $q_S$  to be commonly used by  $N$  agents, e.g. with the Lloyd–Max algorithm [4].

The second step is to disassemble the quantizer  $q_S$  into  $N$  different  $K$ -level quantizers  $q_1, \dots, q_N$  that lead to the same

mean Bayes risk. For convenience, we consider the  $K$ -level quantizers whose quantization cells are intervals, i.e., the cell  $\mathcal{R}_{ik}$  is defined to be  $[b_{k-1}, b_k)$  for any Agent  $i$  and  $1 \leq k \leq K$ .

In order to satisfy (9), each cell boundary of  $q_S$  needs to be a cell boundary of at least one of  $q_1, \dots, q_N$  because, for any cell boundary  $y$  of  $q_S$  and for any  $x_1 < y$  and  $x_2 > y$ ,  $\sum_{i=1}^N u_i q_i(x_1) = q_S(x_1) \neq q_S(x_2) = \sum_{i=1}^N u_i q_i(x_2)$ . Also, any two or more quantizers out of  $q_1, \dots, q_N$  cannot have a common cell boundary since  $q_S$  has  $N(K-1) + 1$  cells. As a result, each cell boundary of  $q_S$ , except 0 and 1, needs to be a cell boundary of exactly one of  $q_1, \dots, q_N$ .

Therefore, we determine cell boundaries of  $q_1, \dots, q_N$  by splitting the cell boundaries of  $q_S$  into  $N$  sets by the following conditions:

$$\begin{aligned} \bigcup_{i=1}^N B_i &= B_S, \\ B_i \cap B_j &= \{0, 1\}, \quad \text{for all } i \text{ and all } j \neq i, \end{aligned} \quad (13)$$

where  $B_S$  is the set of cell boundaries of  $q_S$  and  $B_i$  are the sets of cell boundaries of  $q_i$ ;  $|B_S| = N(K-1) + 2$  and  $|B_i| = K + 1$ . The sets  $B_1, \dots, B_N$  that satisfy (13) are not unique.

Representation points of the cells of  $q_1, \dots, q_N$  are determined after cell boundaries of the quantizers are fixed. We have  $K_S = N(K-1) + 1$  equations of representation points of  $q_1, \dots, q_N$  that describe the condition (9). Once we find the representation points that satisfy all the equations, then the quantizers are minimum MBRE diverse quantizers for the  $N$  agents. The optimal identical quantizer  $q_S$  is always regular because the optimal quantizer for single agent is regular, but the optimal diverse quantizers  $q_1, \dots, q_N$  need not be regular.

The total number of the representation points is  $NK$ , but we have  $N-1$  less equations than what we need to uniquely determine the representation points. Furthermore, a different arrangement of cell boundaries changes the equations and, consequently, proper representation points. Therefore, optimal diverse  $K$ -level quantizers are not unique; any choice leads to the same result because Theorem 5 shows that any set of the optimal quantizers causes the same mean Bayes risk error as  $q_S$  does.

#### D. Comparison to Team-Oblivious Agents

We have discussed teams of agents that are aware of the existence of the other  $N-1$  agents and the  $L$ -out-of- $N$  fusion rule. Let us now consider the case when the agents do not know  $N$  and  $L$ ; individual agents maximize their own probability of being correct. Agent  $i$  considers the Bayes risk

$$R = p_0 c_{10} P_{e,i}^I + (1 - p_0) c_{01} P_{e,i}^{II}.$$

All agents choose their quantizers as if they are single agents with additive noise  $W$  drawn from  $f_W$ . Consequently, they have identical quantizers  $q$  because they have the same Bayes costs  $c_{10}$  and  $c_{01}$ .

Their quantizers are obviously not optimal. In Section III-B, it has been shown that their optimal identical quantizers need to be designed based on the additive noise  $V = W_{(L)}$ ,

whose density function is different from  $f_W$ , in a single-agent model. Furthermore, any identical quantizers can be improved by transformation of diverse quantizers, as in Section III-C. Therefore, for each agent, minimization of the MBRE of its own decision is not the best strategy as an optimal team member and its optimal behavior depends on  $N$  and  $L$ . The effect of  $N$  and  $L$  on the design of optimal quantizers is reflected by the equivalent single-agent model in Theorem 3.

#### IV. EXAMPLES

Throughout this section, assume that the prior probability  $p_0$  of an object being in state  $h_0$  is drawn from the uniform distribution  $f_{P_0}(p_0) = 1$  for all  $p_0 \in [0, 1]$ .

##### A. Gaussian Likelihoods, MAJORITY Rule

Suppose the object sends the signal  $s_0 = 0$  in state  $h_0$  and  $s_1 = 1$  in state  $h_1$ , and each agent receives the signal corrupted by additive iid noise  $W_i$  with cdf  $F_W$  and pdf  $f_W$  where

$$f_W(w) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-w^2/2\sigma^2}.$$

Each likelihood  $f_{Y_i|H}$  is thus Gaussian with mean  $H$ . The MAJORITY rule is used to fuse local decisions made by  $N$  agents. We assume that  $N$  is an odd number so that voting never results in a tie.

From Theorem 3, the equivalent single-agent model has an additive noise  $V_N^{\text{MAJ}}$  with pdf given by

$$f_{V_N^{\text{MAJ}}}(v) = \frac{N!}{\left(\frac{N-1}{2}\right)!^2} F_W^{(N-1)/2}(v) [1 - F_W(v)]^{(N-1)/2} f_W(v).$$

The mean of  $V_N^{\text{MAJ}}$  is zero and the variance of  $V_N^{\text{MAJ}}$  is proportional to that of  $W_i$ :

$$\begin{aligned} \text{var}(V_N^{\text{MAJ}}) &= \int_{-\infty}^{\infty} v^2 \frac{N!}{\left(\frac{N-1}{2}\right)!^2} F_W^{\frac{N-1}{2}}(v) [1 - F_W(v)]^{\frac{N-1}{2}} f_W(v) dv \\ &= \int_{-\infty}^{\infty} v^2 \frac{N!}{\left(\frac{N-1}{2}\right)!^2} F_{\mathcal{N}}^{\frac{N-1}{2}}\left(\frac{v}{\sigma}\right) [1 - F_{\mathcal{N}}\left(\frac{v}{\sigma}\right)]^{\frac{N-1}{2}} f_W(v) dv \\ &= \sigma^2 \int_{-\infty}^{\infty} z^2 \frac{N!}{\left(\frac{N-1}{2}\right)!^2} F_{\mathcal{N}}^{\frac{N-1}{2}}(z) [1 - F_{\mathcal{N}}(z)]^{\frac{N-1}{2}} f_{\mathcal{N}}(z) dz \\ &\triangleq \sigma^2 \zeta_N^{\text{MAJ}}, \end{aligned} \quad (14)$$

where  $f_{\mathcal{N}}$  and  $F_{\mathcal{N}}$  denote the pdf and cdf of a standard normal random variable, and  $z = v/\sigma$ . The factor  $\zeta_N^{\text{MAJ}}$  is the variance of the median of  $N$  iid standard normal random variables. The means and variances of Gaussian order statistics have been studied extensively [17]–[19], and the value of  $\zeta_N^{\text{MAJ}}$  for any  $1 \leq N \leq 20$  can be found in [18]. For example,  $\zeta_3^{\text{MAJ}} = 0.4487$  and  $\zeta_5^{\text{MAJ}} = 0.2863$ .

This equivalent single-agent model is considered for optimization of identical  $(N(K-1)+1)$ -level quantizers. An optimal set of diverse  $K$ -level quantizers for the agents can be designed from the optimal  $(N(K-1)+1)$ -level quantizer.

Fig. 2 shows an example of a team of five agents<sup>2</sup>; it depicts an optimal set of diverse  $K$ -level quantizers and the resulting

<sup>2</sup>An example of a team of three agents is given in [6].

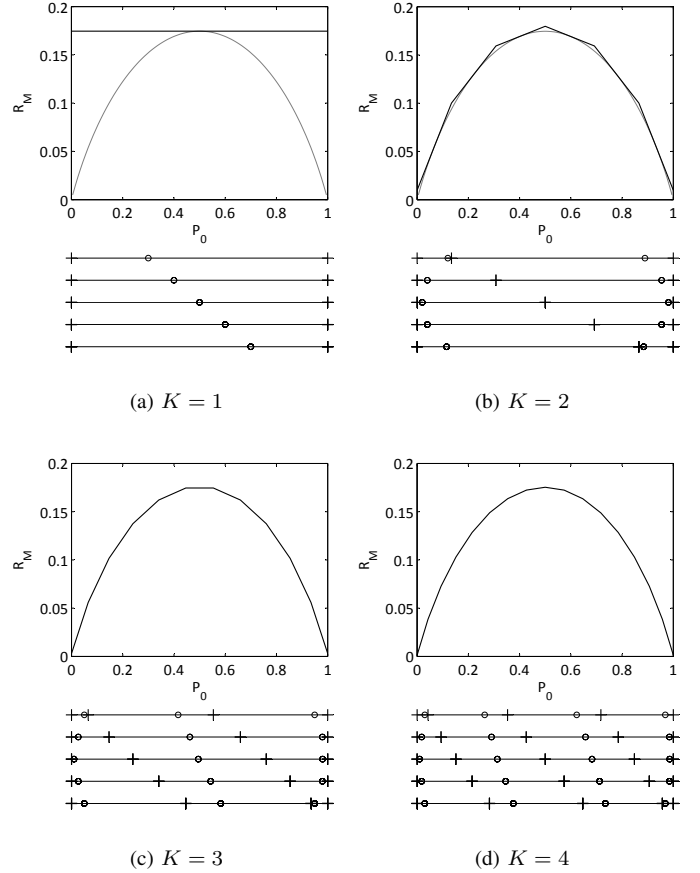


Fig. 2. Optimal diverse  $K$ -level quantizers (cell boundaries as +’s and representation points as o’s) for uniformly distributed  $P_0$  and the resulting mismatched Bayes risk when  $N = 5$  agents perform distributed hypothesis testing fused by the MAJORITY rule. The parameters are defined as  $c_{10} = c_{01} = 1$ ,  $\sigma = 1$ , and  $u_i = 1/5$ , for  $i = 1, \dots, 5$ . For comparison, the unquantized Bayes risk curve is depicted in gray in (a) and (b).

Bayes risk for  $K = 1, \dots, 4$ . The Bayes risk error due to quantization of prior probabilities is depicted in Fig. 3, which shows improvement of decision making as  $K$  increases. Also, the mean Bayes risk error is given in Fig. 4; it shows the advantage of using optimal diverse quantizers against using optimal identical quantizers.

Asymptotically,  $\zeta_N^{\text{MAJ}}$  decays inversely with  $N$ ; more precisely [20],

$$\lim_{N \rightarrow \infty} N \cdot \zeta_N^{\text{MAJ}} = \frac{\pi}{2}. \quad (15)$$

If the agents could share their observations rather than their local hard decisions, the team would base its decision on the sample mean of the observations. The resulting performance is governed by the variance of the sample mean of the noise variables  $W_1, \dots, W_N$ . For  $\sigma = 1$ ,

$$\lim_{N \rightarrow \infty} N \cdot \text{var} \left( N^{-1} \sum_{i=1}^N W_i \right) = 1.$$

The ratio of  $\pi/2$  between the two asymptotic variances quantifies the loss in using majority vote rather than fusing full measurements.

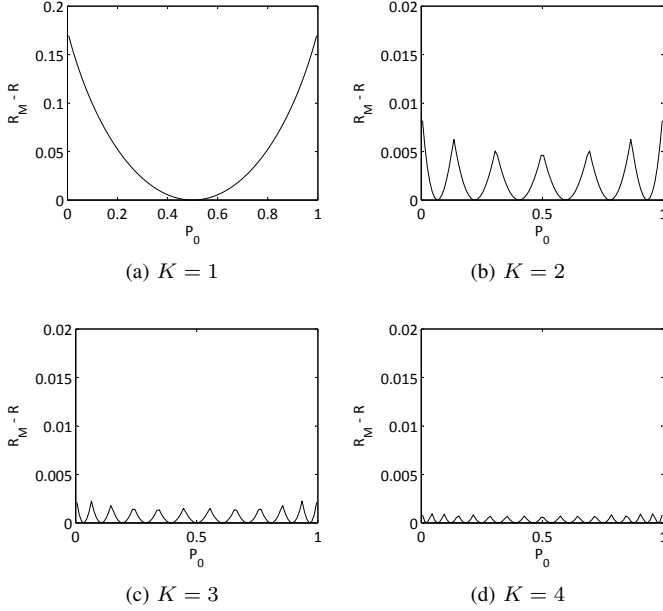


Fig. 3. Bayes risk error of the quantizers in Fig. 2.

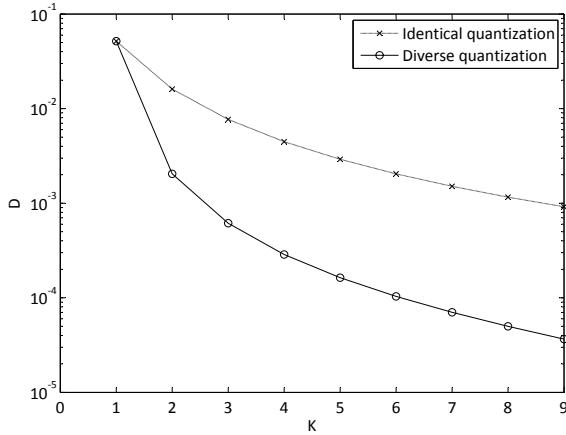


Fig. 4. Mean Bayes risk error for uniformly distributed  $P_0$  in the example of Fig. 2.

### B. Gaussian Likelihoods, OR Rule

Maintaining the Gaussian observation model from Section IV-A, now consider fusion using the OR rule: the global decision is  $h_0$  only when all agents declare  $h_0$ . The equivalent single-agent model has an additive noise  $V_N^{\text{OR}}$ , which is the maximum order statistic of noises  $W_1, \dots, W_N$ . The pdf of  $V_N^{\text{OR}}$  is given by

$$f_{V_N^{\text{OR}}}(v) = NF_W^{N-1}(v)f_W(v).$$

The mean of  $V_N^{\text{OR}}$  is proportional to  $\sigma$ ,

$$\mathbb{E}[V_N^{\text{OR}}] = \mu_N^{\text{OR}} \cdot \sigma,$$

and the variance of  $V_N^{\text{OR}}$  is proportional to  $\sigma^2$ ,

$$\text{var}(V_N^{\text{OR}}) = \zeta_N^{\text{OR}} \cdot \sigma^2;$$

these can be proven analogously to (14). The factor  $\mu_N^{\text{OR}}$  is increasing in  $N$  but  $\zeta_N^{\text{OR}}$  is decreasing in  $N$ . For example,

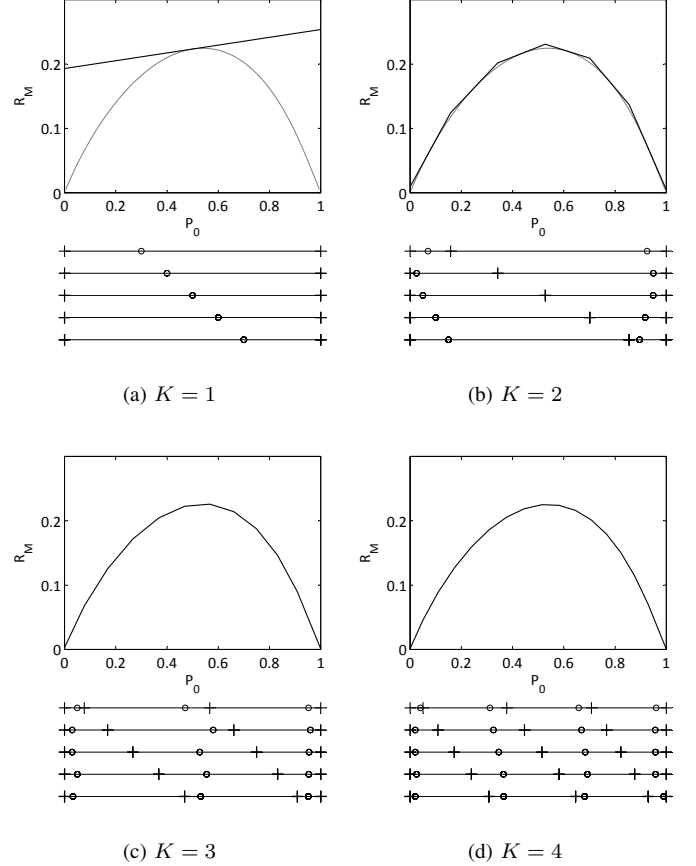


Fig. 5. Optimal diverse  $K$ -level quantizers (cell boundaries as '+'s and representation points as 'o's) for uniformly distributed  $P_0$  and the resulting mismatched Bayes risk when  $N = 5$  agents perform distributed hypothesis testing fused by the OR rule. The parameters are defined as  $c_{10} = c_{01} = 1$ ,  $\sigma = 1$ , and  $u_i = 1/5$ , for  $i = 1, \dots, 5$ . For comparison, the unquantized Bayes risk curve is depicted in gray in (a) and (b).

$\mu_3^{\text{OR}} = 0.8463$  and  $\mu_5^{\text{OR}} = 1.1630$ ;  $\zeta_3^{\text{OR}} = 0.5595$  and  $\zeta_5^{\text{OR}} = 0.4475$ . For other  $N$  between 1 and 20, the values of  $\mu_N^{\text{OR}}$  and  $\zeta_N^{\text{OR}}$  are listed in [18].

An optimal set of diverse  $K$ -level quantizers for a team of five agents is given in Fig. 5 for  $K = 1, \dots, 4$ . Fig. 6 and Fig. 7 show the Bayes risk error and the mean Bayes risk error due to the quantization. The trends are similar to those in Section IV-A, which considers the same model except for the fusion rule. The difference in the fusion rule changes the equivalent single-agent model. However, the same optimization algorithm of quantizers for prior probabilities can be applied in any case: designing the optimal identical  $(N(K-1)+1)$ -level quantizer of the equivalent single agent and disassembling the quantizer into  $N$  diverse  $K$ -level quantizers. The algorithm comes from the relationship between the perceived Bayes risk and the common risk, which is defined as the weighted sum of the perceived Bayes risk. Thus, the algorithm does not depend on how the team of agents make decisions and the error probabilities  $P_E^I$  and  $P_E^{II}$  are computed.

Average performance with the OR rule (Fig. 7) is slightly worse than with the MAJORITY rule (Fig. 4). A qualitative

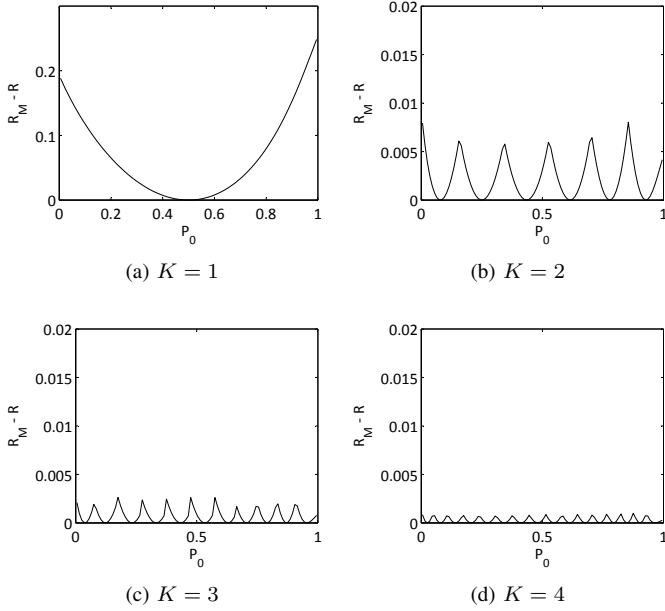
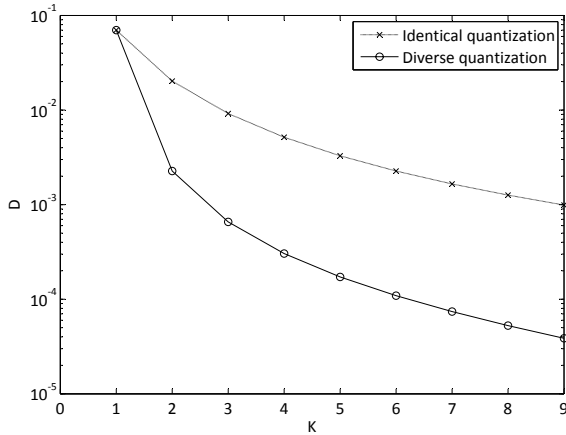


Fig. 6. Bayes risk error of the quantizers in Fig. 5.

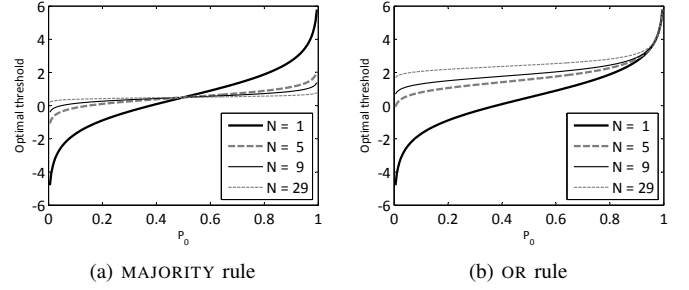
Fig. 7. Mean Bayes risk error for uniformly distributed  $P_0$  in the example of Fig. 5.

understanding can be obtained through study of  $V_N^{\text{OR}}$  and  $V_N^{\text{MAJ}}$ , and asymptotic behavior of the variances of these random variables suggests that the performance gap increases with  $N$ .

While the mean of  $V_N^{\text{MAJ}}$  is zero for all  $N$ ,  $\mu_N^{\text{OR}}$  (the mean of  $V_N^{\text{OR}}$ ) is positive for  $N \geq 2$  and increasing with  $N$ . In the equivalent single-agent model under the OR rule, the single agent tends to observe something larger than a true signal by about  $\mu_N^{\text{OR}}$  because of the noise  $V_N^{\text{OR}}$ . Hence, the optimal decision threshold of the single agent is larger than that in the model of Section IV-A. This can be interpreted as any individual agent optimally requiring “stronger evidence” to declare  $h_1$  as the number of agents increases. The optimal decision thresholds for  $N = 5$  are shown in Fig. 8. The asymptotic growth of  $\mu_N^{\text{OR}}$  is given by (see [16, Ex. 10.5.3])

$$\lim_{N \rightarrow \infty} \mu_N^{\text{OR}} / (2 \log N)^{1/2} = 1.$$

If all the equivalent noise variables had the same standard-

Fig. 8. Optimal decision threshold for  $N$  agents for  $c_{10} = c_{01} = 1$  and  $\sigma = 1$ .

ized distribution, then MAJORITY and OR fusion rules for fixed  $N$  could be compared through the variances  $\zeta_N^{\text{MAJ}}$  and  $\zeta_N^{\text{OR}}$ . We saw in (15) that  $\zeta_N^{\text{MAJ}} = \Theta(1/N)$ . The decay of  $\zeta_N^{\text{OR}}$  is much slower; specifically (see [16, Ex. 10.5.3]),

$$\lim_{N \rightarrow \infty} \left( \frac{12}{\pi^2} \log N \right) \cdot \zeta_N^{\text{OR}} = 1.$$

This suggests that for large  $N$  (and the Gaussian likelihood case), the MAJORITY rule is more effective than the OR rule.

### C. Exponential Likelihoods

Consider a particle that disappears with rate  $s_0$  in state  $h_0$  and with rate  $s_1$  in state  $h_1$ , where  $s_0 > s_1$ . Conditioned on  $H$ , the particle has an exponentially-distributed lifetime  $Y$ :

$$f_{Y|H}(y | h_m) = s_m e^{-s_m y}.$$

An agent observes that the particle disappears at time  $Y = y$  and performs Bayesian hypothesis testing by the likelihood ratio test:

$$\frac{f_{Y|H}(y | h_1)}{f_{Y|H}(y | h_0)} \underset{\hat{H}=h_0}{\overset{\hat{H}=h_1}{>}} \frac{p_0 c_{10}}{(1-p_0)c_{01}}.$$

The likelihood ratio test can be simplified to the decision rule

$$y \underset{\hat{H}=h_0}{\overset{\hat{H}=h_1}{>}} \frac{1}{s_0 - s_1} \log \left( \frac{s_0}{s_1} \frac{p_0 c_{10}}{(1-p_0)c_{01}} \right) = \lambda,$$

which yields errors with probabilities

$$P_e^I = e^{-s_0 \lambda} \quad \text{and} \quad P_e^{II} = 1 - e^{-s_1 \lambda}.$$

Now suppose that  $N$  agents perform Bayesian hypothesis testing with  $N$  particles in the same state; Agent  $i$  observes that the  $i$ th particle disappears at time  $Y_i = y_i$  and applies a common decision threshold  $\lambda$  to its observation to make its local decision. All agents' decisions are fused by  $L$ -out-of- $N$  rule. Then the equivalent single-agent model is to consider a particle with lifetime  $Y_{(L)}$ , which is the  $L$ th longest lifetime among  $\{Y_1, \dots, Y_N\}$ ,

$$\begin{aligned} & f_{Y_{(L)}|H}(y | h_m) \\ &= \frac{N!}{(N-L)!(L-1)!} (1 - e^{-s_m y})^{N-L} (e^{-s_m y})^{L-1} s_m e^{-s_m y}. \end{aligned}$$

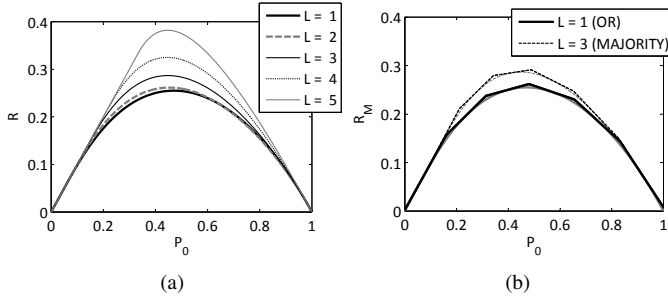


Fig. 9. Bayes risk (a) without quantization of prior probabilities and (b) with diverse 2-level minimum MBRE quantizers for  $N = 5$ ,  $s_0 = 2$ ,  $s_1 = 1$ ,  $c_{10} = c_{01} = 1$ , and  $u_i = 1/5$  for  $i = 1, \dots, 5$ .

The probabilities of global errors are

$$P_E^I = \sum_{n=M}^N \binom{N}{n} (e^{-s_0 \lambda})^n (1 - e^{-s_0 \lambda})^{N-n},$$

$$P_E^{II} = \sum_{n=N-M+1}^N \binom{N}{n} (1 - e^{-s_1 \lambda n})^n (e^{-s_1 \lambda})^{N-n}.$$

In this scenario, a small  $L$  is a good choice for the fusion rule. Fig. 9a shows that the smallest mean Bayes risk without quantization of prior probabilities is achieved by  $L = 1$ . Fig. 9b depicts an example of Bayes risk when five agents use diverse minimum MBRE quantizers and their fusion rule is the OR rule or the MAJORITY rule.

## V. MINIMAX BAYES RISK ERROR QUANTIZERS

Rather than mean Bayes risk error, let us consider maximum Bayes risk error as the criterion for optimizing quantizer design [5]. Such quantizers maximize worst case performance, whereas minimum MBRE quantizers maximize average performance. The minimax Bayes risk error quantizer is defined by the following optimization problem:

$$(q_1^*, \dots, q_N^*) = \arg \min_{(q_1, \dots, q_N)} \max_{p_0} d(p_0, q_1(p_0), \dots, q_N(p_0)).$$

The minimax Bayes risk error quantizer has the same nearest neighbor condition as the minimum MBRE quantizer for  $N = 1$ . On the other hand, a centroid condition for optimality of a regular quantizer for  $N = 1$  is different from that of the minimum MBRE quantizer [5]. For any quantization point  $a$ , the Bayes risk error  $d(p_0, a) = R_M(a) - R(p_0)$  is nonnegative and strictly convex in  $p_0$ , with minimum value of zero attained only at  $p_0 = a$ . Thus, its maximum point within its  $k$ th cell  $\mathcal{R}_k = [b_{k-1}, b_k]$  is a cell boundary:  $b_{k-1}$  or  $b_k$ . The point  $a^*$  that satisfies

$$d(b_{k-1}, a^*) = d(b_k, a^*) \quad (16)$$

minimizes the maximum Bayes risk error within  $\mathcal{R}_k$ . This is the centroid condition for the representation point  $a_k = a^*$  of cell  $\mathcal{R}_k$ . The minimax Bayes risk error quantizer can be found by alternatively applying the nearest neighbor condition (5) and the centroid condition (16) through the iterative Lloyd–Max algorithm.

For  $N > 1$ , agents can take advantage of diversity in the same way as in Section III-C. A team of  $N$  agents bonded by

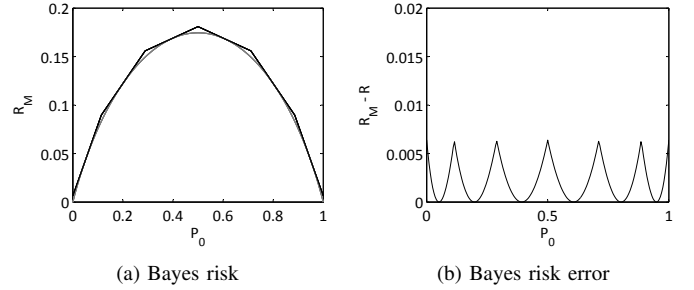


Fig. 10. (a) Bayes risk and (b) Bayes risk error for  $N = 5$  agents performing distributed hypothesis testing fused by the MAJORITY rule. They observe signals corrupted by iid additive Gaussian noise  $\mathcal{N}(0, \sigma^2)$ . The parameters are defined as  $s_0 = 0$ ,  $s_1 = 1$ ,  $c_{10} = c_{01} = 1$ ,  $\sigma = 1$ , and  $u_i = 1/5$  for  $i = 1, \dots, 5$ .

the  $L$ -out-of- $N$  fusion rule is equivalent to a single agent with noise equal to their  $L$ th largest noise  $W_{(L)}$  (cf. Theorem 3). Thus, identical quantizers that minimize maximum Bayes risk error of the  $N$  agents also minimize that of the single agent and vice versa (cf. Theorem 4).

Furthermore, if the  $N$  agents collaborate by sharing the perceived common risk, there exists a set of  $N$  diverse  $K$ -level quantizer that leads to the same Bayes risk for any  $p_0$  as a set of identical  $(N(K-1)+1)$ -level quantizers does (cf. Theorem 5). Therefore, the identical  $(N(K-1)+1)$ -level minimax Bayes risk error quantizers can be transformed into  $N$  diverse  $K$ -level minimax Bayes risk error quantizers (cf. Theorem 6).

Fig. 10 shows an example of Bayes risk when five agents use diverse 2-level minimax Bayes risk error quantizers and their fusion rule is MAJORITY. A property of minimax Bayes risk error quantizers is that the maximum Bayes risk errors at all cell boundaries are the same, which is shown in Fig. 10b. In addition, the minimax Bayes risk error quantizers are not dependent on the distribution of  $P_0$  as long as  $f_{P_0}(p_0) > 0$  for all  $p_0 \in [0, 1]$  because they minimize the worst case error, not the average.

## VI. CONCLUSION

We have discussed distributed detection and data fusion performed by a team of agents when there is a distribution of prior probabilities and the agents only know quantized versions of prior probabilities. We have focused on how to take advantage of diversity in quantization. When all agents use identical quantizers, they are affected by the same perceived Bayes risk and the distributed hypothesis testing problem can be analyzed by existing theorems of decision theory. On the contrary, when they do not use identical quantizers, then they consider different perceived Bayes risks, which prevents them from collaborating in hypothesis testing. We let the agents use the perceived common risk as a new distortion measure of hypothesis testing so as to unite them as a team to perform distributed hypothesis testing in any case.

We have defined mean Bayes risk error as the optimization criterion for prior-probability quantizers. We have presented theorems to show that diverse quantizers are better than identical quantizers. The equivalence between multiple-agent

decision making and single-agent decision making simplifies a team of agents that use identical quantizers to a single agent. By combining the equivalence theorem with the equivalence between diverse quantizers and identical quantizers used by a team of agents, we can take advantage of the study of the minimum MBRE quantizer of a single agent to analyze optimal diverse quantization for multiple agents. It is shown that if the agents use diverse  $K$ -level quantizers, it has the same effect of using identical  $(N(K-1)+1)$ -level quantizers.

The equivalence theorems hold under the condition that all agents collaborate to perform Bayesian hypothesis testing. Hence, for any given distortion function of quantizers, diverse quantizers that minimize the distortion can be easily found. While the case of minimum MBRE was covered in detail, we also discussed the minimax Bayes risk error criterion. In the minimax case, one can again first design an  $(N(K-1)+1)$ -level quantizer for a single agent and then disassemble the quantizer into  $N$  diverse  $K$ -level quantizers.

#### REFERENCES

- [1] D. Austen-Smith and J. S. Banks, "Information aggregation, rationality, and the Condorcet jury theorem," *Am. Polit. Sci. Rev.*, vol. 90, no. 1, pp. 34–45, Mar. 1996.
- [2] C. N. Macrae and G. V. Bodenhausen, "Social cognition: Thinking categorically about others," *Annual Review of Psychology*, vol. 51, pp. 93–120, Feb. 2000.
- [3] K. R. Varshney, "Bayes risk error is a Bregman divergence," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4470–4472, Sep. 2011.
- [4] K. R. Varshney and L. R. Varshney, "Quantization of prior probabilities for hypothesis testing," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4553–4562, Oct. 2008.
- [5] —, "Multilevel minimax hypothesis testing," in *Proc. IEEE/SP Workshop Stat. Signal Process.*, Nice, France, Jun. 2011, pp. 109–112.
- [6] J. B. Rhim, L. R. Varshney, and V. K. Goyal, "Collaboration in distributed hypothesis testing with quantized prior probabilities," in *Proc. IEEE Data Compression Conf.*, Snowbird, UT, Mar. 2011, pp. 303–312.
- [7] J. Marschak and R. Radner, *Economic Theory of Teams*. New Haven: Yale University Press, 1972.
- [8] 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.
- [9] S. Kassam, "Optimum quantization for signal detection," *IEEE Trans. Commun.*, vol. COM-25, no. 5, pp. 479–484, May 1977.
- [10] H. V. Poor and J. B. Thomas, "Applications of Ali–Silvey distance measures in the design of generalized quantizers for binary decision systems," *IEEE Trans. Commun.*, vol. COM-25, no. 9, pp. 893–900, Sep. 1977.
- [11] R. Gupta and A. O. Hero, III, "High-rate vector quantization for detection," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1951–1969, Aug. 2003.
- [12] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks: Quantized data and random link failures," *IEEE Trans. Signal Process.*, vol. 58, pp. 1383–1400, Mar. 2010.
- [13] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, H. V. Poor and J. B. Thomas, Eds. Greenwich, CT: JAI Press, 1993, pp. 297–344.
- [14] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I—fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.
- [15] 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.
- [16] H. A. David and H. N. Nagaraja, *Order Statistics*, 3rd ed. Hoboken, NJ: John Wiley & Sons, 2003.
- [17] L. H. C. Tippett, "On the extreme individuals and the range of samples taken from a normal population," *Biometrika*, vol. 17, no. 3/4, pp. 364–387, Dec. 1925.
- [18] D. Teichroew, "Tables of expected values of order statistics and products of order statistics for samples of size twenty and less from the normal distribution," *Ann. Math. Stat.*, vol. 27, no. 2, pp. 410–426, Jun. 1956.
- [19] R. C. Bose and S. S. Gupta, "Moments of order statistics from a normal population," *Biometrika*, vol. 46, no. 3/4, pp. 433–440, Dec. 1959.
- [20] 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.