

# BAYESIAN DESIGN OF DECENTRALIZED HYPOTHESIS TESTING UNDER COMMUNICATION CONSTRAINTS

Alla Tarighati, and Joakim Jaldén

ACCESS Linnaeus Centre, Department of Signal Processing,  
KTH Royal Institute of Technology, Stockholm, Sweden

## ABSTRACT

We consider a distributed detection system under communication constraints, where several peripheral nodes observe a common phenomenon and send their observations to a fusion center via error-free but rate-constrained channels. Using the minimum expected error probability as a design criterion, we propose a cyclic procedure for the design of the peripheral nodes using the person-by-person methodology. It is shown that a fine-grained binning idea together with a method for updating the conditional probabilities of the joint index space at the fusion center, decrease the complexity of the algorithm and make it tractable. Also, unlike previous methods which use dissimilarity measures (e.g., the Bhattacharyya distance), a-prior hypothesis probabilities are allowed to contribute to the design in the proposed method. The performance of the proposed method is compared to a method due to Longo *et al.*'s and it is shown that the new method can significantly outperform the previous one at a comparable complexity.

**Index Terms**— Decentralized detection, Bayesian criterion, parallel network, person-by-person optimization.

## 1. INTRODUCTION

We consider a decentralized system where multiple sensors observe the same phenomenon and send their observations to a central processor or fusion center (FC) to make a global decision. If there were no constraints on the communication channels, the sensors could send their complete received data to the fusion center, and the scenario is no different than the centralized case. However, when the channels between the sensors and the fusion center are rate-constrained, the sensors must send a compressed, or quantized, version of the received data to the fusion center. In other words, each sensor acts as a decision maker (DM) and sends its local decision to the FC. A comprehensive survey of early works in decentralized hypothesis testing can be found in [1–3].

The specific problem of quantization for decentralized detection was considered by Longo *et al.* in [4], wherein the channels between the sensors (also called peripheral nodes or DMs) and the fusion center were also considered to be error-free. The peripheral nodes were considered to be scalar quantizers that satisfied the rate constraints, and were to be cooperatively designed according to a system-wide measure of performance. Longo *et al.* argued that the natural criterion of optimization – in their case the power of a Neyman-Pearson test – made the design procedure intractable. They therefore proposed to instead optimize a measure of dissimilarity between the conditional distributions of the joint index space, i.e., the full set of quantizer outputs. Their final design was obtained by a cyclic person-by-person optimization algorithm for maximizing the Bhattacharyya distance [5] between the conditional probability distributions under the two hypotheses, or equivalently for minimizing the

corresponding Bhattacharyya coefficient.

The Bhattacharyya distance, as a member of the Ali-Silvey class of distribution measures [6], has a link to the error probability in a Bayesian formulation of the problem. Namely, by Blackwell's theorem it follows that if one design (say, F) yields a larger distributional distance (or in the present case a smaller Bhattacharyya coefficient) than another design (say, G), there exists a set of a-prior hypothesis probabilities under which F yields a lower (or, more precisely, no greater) probability of error than G, see [4] and also [7]. It is also possible to upper bound the minimum error probability of the fusion center, given a joint design of the DMs that yields a Bhattacharyya coefficient  $b$ , as [8]

$$P_E \leq \sqrt{\pi_0 \pi_1 b}, \quad (1)$$

where  $\pi = (\pi_0, \pi_1)$  denotes the set of a-prior probabilities. This further motivates the use of the Bhattacharyya coefficient as a design measure.

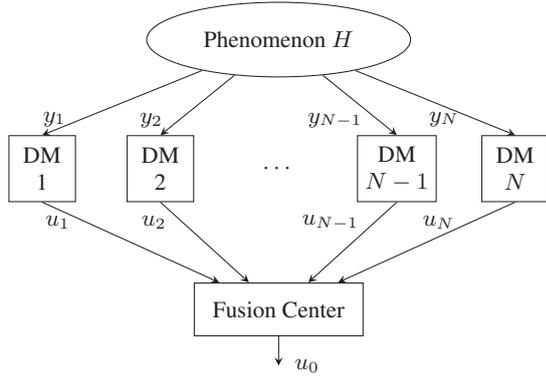
It should however be noted that while Blackwell's theorem asserts the existence of a set of a-prior probabilities for which the design with minimal Bhattacharyya coefficient is also optimal with respect to the probability of error, it does not state that the design is optimal with respect to the probability of error for a *given* set of a-prior probabilities: The Bhattacharyya coefficient, and the design, is independent of the a-prior probabilities. It is further known that (1) is in general a rather loose bound on the probability of error [9], and minimizing the Bhattacharyya coefficient is therefore a rather blunt tool for optimizing the error probability.

The main contribution of our work is to show that – contrary to previous claims – it is in fact possible to within the algorithmic framework proposed in [4] minimize the probability of error directly. Furthermore, by proper implementation, the complexity of doing so is not significantly larger than the design based on the Bhattacharyya coefficient; the complexity is of the same order. To obtain our solution, we propose to combine the fine-grained observation-binning already used in [4] together with a method for updating the conditional probability mass functions (PMFs) of the joint index space. Using the proposed method, the complexity does not increase considerably in comparison to the algorithm already proposed in [4], while the receiver operating characteristic (ROC) of the system is much improved.

The outline of this paper is as follows. In Section 2 the system model is defined and the algorithm of Longo *et al.* is briefly described. Section 3 presents the proposed design method and the merits of the proposed design are shown by numerical examples in Section 4. Section 5 concludes the paper.

## 2. PRELIMINARIES

As in [4], we consider a decentralized binary hypothesis testing system with  $N$  peripheral nodes in a parallel network topology as shown



**Fig. 1.** Decentralized hypothesis testing scheme in a parallel network.

in Fig. 1. The peripheral nodes, or decision makers (DMs), observe the same phenomenon and send their observations to the fusion center through rate-constrained channels. The DMs are (scalar) quantizers that map their observations  $y_n \in \mathbb{R}$ ,  $n = 1, \dots, N$ , to an output value (message)  $u_n$  from a discrete set  $\mathcal{M}_n$  using the decision function  $\gamma_n : \mathbb{R} \rightarrow \mathcal{M}_n$ , i.e.,  $u_n = \gamma_n(y_n)$ ,  $n = 1, \dots, N$ . DM  $n$  then sends its message  $u_n$  to the fusion center through a channel of rate  $R_n$  bits and the cardinality of the discrete set  $\mathcal{M}_n$  must therefore satisfy  $0 < \|\mathcal{M}_n\| \leq 2^{R_n}$ .

The observations are in general considered to be correlated with conditional distributions  $p(y_1, y_2, \dots, y_N | H_j)$ , where  $H_j$ ,  $j = 0, 1$ , are the two possible hypotheses. The rate-constrained channels are one-way links from the DMs to the fusion center and there is no communication between the DMs. The fusion center makes the global decision  $u_0 \in \{H_0, H_1\}$  based on the *joint index vector*  $\mathbf{u}^N \triangleq (u_1, u_2, \dots, u_N) \in \mathcal{M}^N$  using the fusion function  $\gamma_0 : \mathcal{M}^N \rightarrow \{H_0, H_1\}$ , where  $\mathcal{M}^N \triangleq \mathcal{M}_1 \times \mathcal{M}_2 \times \dots \times \mathcal{M}_N$ .

Longo *et al.* [4] argued that a cooperative design can outperform a non-cooperative design even if the DMs act independently on their received data. Using the Bhattacharyya coefficient as a design criterion and the person-by-person optimization methodology they proposed a cyclic design algorithm.

In each cycle of the algorithm, the DMs – one by one from DM 1 to DM  $N$  – are individually updated (optimized) while keeping the other DMs fixed. After the optimization of each DM the PMF of the joint index space is updated, and at the end of the cycle a new Bhattacharyya coefficient is calculated. If the improvement in the Bhattacharyya coefficient is greater than a threshold, the algorithm performs another cycle; otherwise the last set of decision functions is the final design.

The Bhattacharyya coefficient is defined as

$$b = \sum_{\mathbf{u}^N \in \mathcal{M}^N} \sqrt{P_0(\mathbf{u}^N)P_1(\mathbf{u}^N)}, \quad (2)$$

where  $P_j(\mathbf{u}^N) \triangleq P(u_1, \dots, u_N | H_j)$ ,  $j = 0, 1$ , are the conditional PMFs. The updating rule for the  $N$ th<sup>1</sup> DM used in [4] is given by (cf. [4, Eq. (13)])

$$\gamma_N^+(y_N) = \arg \min_{u_N} \sum_{\mathbf{u}^{N-1} \in \mathcal{M}^{N-1}} f_0(\mathbf{u}^{N-1}, y_N) L^{1/2}(\mathbf{u}^N) + f_1(\mathbf{u}^{N-1}, y_N) L^{-1/2}(\mathbf{u}^N), \quad (3)$$

<sup>1</sup>The updating rule for the  $n$ th DM is analogous, but focussing on the  $N$ th DM parallels the exposition in [4] and simplifies notation.

where  $\mathcal{M}^{N-1} \triangleq \mathcal{M}_1 \times \mathcal{M}_2 \times \dots \times \mathcal{M}_{N-1}$ , and where the optimization is over all possible indices  $u_N \in \mathcal{M}_N$  for every input value  $y_N \in \mathbb{R}$  of the  $N$ th DM. In (3),  $L$  is the likelihood ratio defined as  $L(\mathbf{u}^N) \triangleq P_1(\mathbf{u}^N)/P_0(\mathbf{u}^N)$ , and the functions  $f_0$  and  $f_1$  are auxiliary functions defined as

$$f_j(\mathbf{u}^{N-1}, y_N) \triangleq \int_{\gamma_1^{-1}(u_1)} \dots \int_{\gamma_{N-1}^{-1}(u_{N-1})} p(\mathbf{y}^N | H_j) dy_1 \dots dy_{N-1}, \quad (4)$$

where  $\gamma_n^{-1}(u_n)$  is the quantizer partitions for  $u_n$ , i.e., the set of  $y_n$ s that satisfies  $\gamma_n(y_n) = u_n$ .

The practical implementation of Longo *et al.*'s method must store a representation of the quantizer partitions  $\gamma_n^{-1}$  and the function  $f_j$ . Since it is impossible to store an exact representation of an arbitrary partition or an arbitrary function in computer memory, Longo *et al.* suggested in the examples section of [4] the idea of dividing a portion of the observation space of DM  $n$ , i.e.,  $\mathbb{R}$ , containing most of the probability mass into the equally sized fine-grained *bins*, and required  $f_j$  to be constant across each bin. The optimization problem in (3) was then (approximately) solved bin-by-bin: The assigned index to each bin was chosen to satisfy (3), with  $y_N$  chosen as the *midpoint* of the bin. In short, the algorithm finally implemented in [4] can be viewed as an algorithm that assigns an index from  $\mathcal{M}_N$  to each bin. The same index assignment problem is solved by the algorithm proposed herein, albeit with the (Bayesian) probability of error metric in place of the Bhattacharyya coefficient.

### 3. PROPOSED METHOD

In the Bayesian formulation of the decentralized hypothesis testing problem, the expected error probability when using the maximum a-posteriori (MAP) criterion at the fusion center is [10]

$$P_E = 1 - \sum_{\mathbf{u}^N \in \mathcal{M}^N} \max_{j=0,1} \{P(H_j | \mathbf{u}^N) P(\mathbf{u}^N)\} = 1 - \sum_{\mathbf{u}^N \in \mathcal{M}^N} \max_{j=0,1} \{P_j(\mathbf{u}^N) \pi_j\}, \quad (5)$$

where, again,  $P_j(\mathbf{u}^N) \triangleq P(\mathbf{u}^N | H_j)$ , and where  $\pi_j \triangleq P(H_j)$  is the a-prior probability of hypothesis  $H_j$ ,  $j = 0, 1$ . It can further be shown that the conditional PMFs that appear in (5) can be expressed in terms of the auxiliary functions  $f_j(\mathbf{u}^{N-1}, y_N)$ , defined in (4), and the decision function  $\gamma_N$  of the  $N$ th DM as

$$P_j(\mathbf{u}^N) = \int_{\gamma_N^{-1}(u_N)} f_j(\mathbf{u}^{N-1}, y_N) dy_N. \quad (6)$$

Using the fine-grained binning idea, the observation space of DM  $n$  is divided into  $K_n$  equally sized (small) bins

$$C_k^n \triangleq (c_k^n, d_k^n], \quad 1 \leq k \leq K_n, 1 \leq n \leq N, \quad (7)$$

and an index is assigned to the observations in each bin, i.e.,

$$\gamma_n(y_n) = v_k^n \in \mathcal{M}_n, \quad \text{if } y_n \in C_k^n. \quad (8)$$

With a slight abuse of notation we write (8) in an abbreviated form as  $\gamma_n(C_k^n) = v_k^n$ . Each DM is then designed in a bin-by-bin manner where an index is assigned to a bin while the other bins are kept fixed. The index assigned to a bin is chosen to minimize the probability of error given in (5). In other words, focusing again on the

$N$ th DM for notational simplicity, the optimization formulation for the  $k$ th bin of the  $N$ th DM  $\mathcal{C}_k^N$  is given as (cf. (5))

$$\gamma_N^+(\mathcal{C}_k^N) = \arg \max_{v_k^N \in \mathcal{M}_N} \sum_{\mathbf{u}^N \in \mathcal{M}_N} \max_{j=0,1} \{P_j(\mathbf{u}^N)\pi_j\}, \quad (9)$$

where the index assignment  $v_k^N$  influences the PMFs  $P_j(\mathbf{u}^N)$  through the full decision function  $\gamma_N$  according to (6). The optimizer of (9) is found over all possible indices  $v_k^N \in \mathcal{M}_N$ , in turn for every bin  $\mathcal{C}_k^N$ ,  $1 \leq k \leq K_N$  and then for every DM  $n$  where  $1 \leq n \leq N$ .

As seen by (9), in order to update the index assignments of a bin, the set of all conditional PMFs for every possible joint index vector need to be computed. This has a complexity order of  $\mathcal{O}(\|\mathcal{M}^N\|)$  and the potential to make the algorithm difficult to implement for larger rates and number of DMs. That said, the complexity of the proposed design procedure based on (9) is not significantly higher than the design procedure proposed in [4] and based on (3).

To see this, assume that we want to update the index of  $\mathcal{C}_k^N$  (which is defined by (7) for  $n = N$ ). Assume also that the index of  $\mathcal{C}_k^N$  before updating is  $\alpha \in \mathcal{M}_N$  and the corresponding conditional PMFs are  $P_j(\mathbf{u}^N)$ ,  $j = 0, 1$ . Then  $\gamma_N(\mathcal{C}_k^N) = \alpha$  and the conditional PMFs where  $u_N = \alpha$  are

$$\begin{aligned} P_j(\mathbf{u}^{N-1}, \alpha) &= \sum_{i \in \mathcal{I}_N(\alpha)} \int_{\mathcal{C}_i^N} f_j(\mathbf{u}^{N-1}, y_N) dy_N \\ &= \sum_{\substack{i \in \mathcal{I}_N(\alpha) \\ i \neq k}} \int_{\mathcal{C}_i^N} f_j(\mathbf{u}^{N-1}, y_N) dy_N + \int_{\mathcal{C}_k^N} f_j(\mathbf{u}^{N-1}, y_N) dy_N, \end{aligned} \quad (10)$$

where  $\mathcal{I}_N(\alpha)$  is the set of bins of DM  $N$  with index  $\alpha \in \mathcal{M}_N$  (including bin  $k$ ).

Assume that the index of the  $k$ th bin  $\mathcal{C}_k^N$  changes to  $\alpha^+ \neq \alpha$  as a result of (9). Then this bin does not belong to  $\mathcal{I}_N(\alpha)$  anymore and the new conditional PMFs  $P^+(\mathbf{u}^{N-1}, \alpha | H_j)$  for which  $u_N = \alpha$  are

$$\begin{aligned} P_j^+(\mathbf{u}^{N-1}, \alpha) &= \sum_{\substack{i \in \mathcal{I}_N(\alpha) \\ i \neq k}} \int_{\mathcal{C}_i^N} f_j(\mathbf{u}^{N-1}, y_N) dy_N \\ &= P_j(\mathbf{u}^{N-1}, \alpha) - \Delta_j(\mathbf{u}^{N-1}, \mathcal{C}_k^N) \end{aligned} \quad (11)$$

where we define  $\Delta_j(\mathbf{u}^{N-1}, \mathcal{C}_k^N)$  in terms of the auxiliary function  $f_j(\mathbf{u}^{N-1}, y_N)$  in (4) and (6) as

$$\Delta_j(\mathbf{u}^{N-1}, \mathcal{C}_k^N) \triangleq \int_{\mathcal{C}_k^N} f_j(\mathbf{u}^{N-1}, y_N) dy_N. \quad (12)$$

The new PMFs where  $u_N = \alpha^+$  are

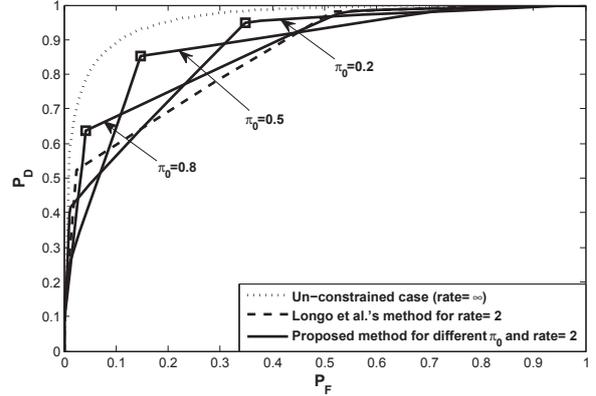
$$P_j^+(\mathbf{u}^{N-1}, \alpha^+) = P_j(\mathbf{u}^{N-1}, \alpha^+) + \Delta_j(\mathbf{u}^{N-1}, \mathcal{C}_k^N), \quad (13)$$

while all other PMFs remain fixed, i.e.,

$$P_j^+(\mathbf{u}^{N-1}, u_N \neq \alpha, \alpha^+) = P_j(\mathbf{u}^{N-1}, u_N \neq \alpha, \alpha^+). \quad (14)$$

Thus, only a subset of the conditional probabilities need to be updated for each change of the bin index  $v_k^N$ .

From the above, it can be seen that for the index assignment of each bin, the changes in PMFs corresponding to (11) and (13) should be found which requires calculating the function  $\Delta_j(\mathbf{u}^{N-1}, \mathcal{C}_k^N)$  for all joint index vectors  $\mathbf{u}^{N-1} \in \mathcal{M}^{N-1}$ . This is essentially the same computations required in the updating rule of [4] in the sense that  $f_j$  must be evaluated for all different joint index vectors  $\mathbf{u}^{N-1}$  in order



**Fig. 2.** ROC curves for different values of the a-prior probabilities for the example in Section 4. Squares correspond to the points which give minimum error probability in each case.

to evaluate (3). There is a small discrepancy in that  $\Delta_j(\mathbf{u}^{N-1}, \mathcal{C}_k^N)$  involves the integral of  $f_j$  over  $\mathcal{C}_k^N$ , but given that  $f_j$  in (4) is already defined in terms of a high dimensional integral, the added complexity of evaluating (12) is minor. In a practical implementation of either approach, we believe that the most reasonable general strategy is to choose the bins so small that the functions involved varies slowly over the bins, and to approximate (4) and (12) using only the midpoint of the integration interval; in which case the complexity is of the same order. For the (Gaussian) examples studied in Section 4, the involved quantities can also be obtained in closed form using differences of  $Q$ -functions and the complexities per evaluation are also then of the same order. In the end, evaluating (or approximating) the required quantities for all  $\mathbf{u}^{N-1}$  has a complexity of  $\mathcal{O}(\|\mathcal{M}^{N-1}\|)$ , and solving either (3) or (9) by searching over  $\mathcal{M}_N$  yields an overall complexity of  $\mathcal{O}(\|\mathcal{M}^N\|)$ .

In closing, we should stress that both Longo *et al.*'s method [4] and the proposed method (which use person-by-person optimization) only lead to locally optimal solutions and the results depend on the initialization of DMs. A straightforward way to further improve upon the method of [4] is to initialize the method proposed herein with the solution obtained as prescribed in [4].

#### 4. EXAMPLES

In this section we illustrate the benefit of the proposed approach over that of [4] by numerical examples. Like in [4] a decentralized hypothesis testing problem with two peripheral nodes ( $N = 2$ ) is considered. The observation model is as

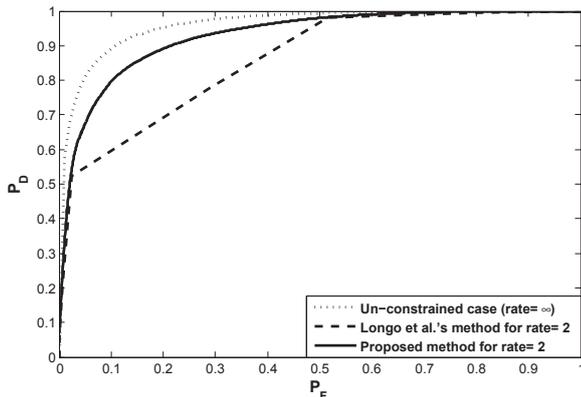
$$\begin{aligned} H_0 : \mathbf{y} &= \mathbf{n} \\ H_1 : \mathbf{y} &= \mathbf{a} + \mathbf{n}, \end{aligned}$$

where  $\mathbf{a}$  is a constant known signal and  $\mathbf{n}$  is a zero mean Gaussian vector of covariance matrix

$$\Sigma = \begin{pmatrix} \sigma^2 & r\sigma^2 \\ r\sigma^2 & \sigma^2 \end{pmatrix}$$

where  $r$  ( $|r| < 1$ ) is the spatial correlation coefficient. As in [4], we assume that the

- signals have equal energy  $|a_1|^2 = |a_2|^2 = \mathcal{E}$  with  $a_1 = -a_2$ ,
- per channel signal-to-noise ratio ( $\mathcal{E}/\sigma^2$ ) is  $-5$  dB,
- spatial correlation coefficient is  $r = 0.9$ , and
- channel rates are the same, i.e.,  $R_n = R$  for  $n = 1, 2$ .



**Fig. 3.** The complete ROC curve achievable by tuning the a-prior probabilities of the proposed design. The design of [4], which is independent of the a-prior probabilities, and the optimum centralized performance are included for comparison.

The interval  $[-4, +4]$  (containing 0.9997 of the total probability mass for each DM) is divided into 256 bins, and the DMs are designed for various values of a-prior probabilities and channel rates (in bits/sample).

Although the proposed design is based on the (Bayesian) error probability, the close relation between the error probability and the ROC makes it possible to compare the results of proposed method with those of [4] in terms of the ROC curve as well. The error probability is given by

$$P_E = \pi_0 P(u_0 = H_1 | H_0) + \pi_1 P(u_0 = H_0 | H_1) \quad (15)$$

$$= \pi_0 P_F + \pi_1 (1 - P_D),$$

where  $P_F$  and  $P_D$  are the probability of false alarm and the probability of detection, respectively [11]. The relation in (15) allows us to trace out an ensemble of designs for different  $P_F$ s and  $P_D$ s by tuning the a-prior probabilities  $\pi$ .

Once the peripheral nodes are designed<sup>2</sup>, the ROC curve [11] of the system can be found (as is shown in Fig. 2 for different values of a-prior probabilities). The fusion function  $\gamma_0$  is allowed to be stochastic and is for each method chosen optimally for each  $P_F$ . The (deterministic) MAP fusion rule that achieves the minimum error probability  $P_{E,\min}$  for the proposed design is shown by a square in Fig. 2, and the corresponding pair  $(P_F^*, P_D^*)$  [cf. (15)] is a Pareto optimal point on the ROC curve [11]. For the proposed design method, there is a unique design of the DMs (corresponding to a unique ROC curve) for every set of a-prior probabilities. However, in the method proposed by Longo *et al.* there is only one design; it is independent of any a-prior probabilities. This gives the proposed design methodology a distinct advantage.

Fig. 2 depicts the resulting ROC curves for the proposed design for rate  $R = 2$  channels and for different values of a-prior probabilities along with corresponding  $(P_F^*, P_D^*)$  pairs for each design. The resulting curve from Longo *et al.*'s method (coinciding with that of [4]) and the optimum achievable ROC curve for a centralized hypothesis test (corresponding to  $R = \infty$ ) are shown for comparison. The envelope of achievable ROC curves resulting from our proposed design can be found from the resulting  $(P_F^*, P_D^*)$  pairs obtained for

<sup>2</sup>For rate  $R = 1$  we initialize the DMs in all methods with the optimal local (threshold) decision rules, and for rate  $R > 1$  we uniformly quantize the two decision regions of the  $R = 1$  initialization.

**Table 1.** Comparison between the minimum error probability of the method from [4] and the proposed method.

$\pi_0$	$R = 1$			$R = 2$		
	$P_{E,\min}^B$	$P_{E,\text{ub}}^B$	$P_{E,\min}^*$	$P_{E,\min}^B$	$P_{E,\text{ub}}^B$	$P_{E,\min}^*$
0.3	0.237	0.390	0.220	0.168	0.309	0.135
0.4	0.313	0.417	0.249	0.217	0.331	0.145
0.5	0.389	0.426	0.311	0.251	0.338	0.148
0.6	0.313	0.417	0.249	0.205	0.331	0.145
0.7	0.237	0.390	0.220	0.160	0.309	0.133

many different a-priors. Fig. 3 illustrates the achievable ROC curves of the proposed method, along with Longo *et al.*'s method in addition to the optimum centralized performance. It can be observed that the design of [4] achieves the minimum probability of error and the same performance as the proposed design at two points on the ROC curve, as predicted by Blackwell's theorem, but for many other operational points the proposed design significantly outperforms the design of [4] and is at rate  $R = 2$  already quite close to the optimal centralized performance.

The minimum error probability resulting from [4] and the proposed method are shown in Table 1 for various values of a-prior probabilities and channel rates. In the table,  $P_{E,\min}^B$  is the minimum error probability resulting from Longo *et al.*'s method,  $P_{E,\text{ub}}^B$  is the corresponding upper bound of the error probability resulting from (1), and  $P_{E,\min}^*$  is the minimum error probability of the proposed method. The results in Table 1 show that for rates  $R = 1$  and  $R = 2$ , there is a significant gain in performance achieved by the proposed method, and increasing the rate improves performance. Comparing the resulting error probabilities to the upper bound in (1) confirms that the Bhattacharyya coefficient provide a rather loose bound, and optimizing the probability error directly is therefore well justified.

## 5. CONCLUSION

As in [4], we have considered a distributed detection scenario with binary hypotheses, distributed sensors or local decision makers, and rate constrained channels from the DMs to the fusion center. By explicitly using the Bayesian measure of probability of error in place of the dissimilarity measure, i.e., the Bhattacharyya distance, used in [4], we have obtained a design that not only with similar complexity outperforms the prior design in terms of the probability of error in the Bayesian setting, but which also yields superior ROC characteristics: For a given probability of false alarm  $P_F$ , we achieve a significant improvement in the probability of detection  $P_D$  over that obtained by the approach of [4]. The key to the performance improvement of the proposed design lies in the parameterization of the design offered by tuning the asserted prior probabilities.

## 6. REFERENCES

- [1] John N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, 1993, pp. 297–344.
- [2] R. Viswanathan and P.K. Varshney, "Distributed detection with multiple sensors I. Fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, 1997.
- [3] R.S. Blum, S.A. Kassam, and H.V. Poor, "Distributed detection with multiple sensors II. Advanced topics," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, 1997.

- [4] M. Longo, T.D. Lookabaugh, and R.M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. Inf. Theory*, vol. 36, no. 2, pp. 241–255, Mar. 1990.
- [5] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Trans. Commun.*, vol. 15, no. 1, pp. 52–60, Feb. 1967.
- [6] H.V. Poor and John B. Thomas, "Applications of Ali-Silvey distance measures in the design of generalized quantizers for binary decision systems," *IEEE Trans. Commun.*, vol. 25, no. 9, pp. 893–900, 1977.
- [7] W. A. Hashlamoun, *Application of distance measures and probability of error bounds to distributed detection systems*, Ph.D. thesis, Syracuse University, May 1991.
- [8] D.E. Boekke and J.C.A. van der Lubbe, "Some aspects of error bounds in feature selection," *Pattern Recognition*, vol. 11, no. 5-6, pp. 353–360, 1979.
- [9] W. A. Hashlamoun, P.K. Varshney, and V. N S Samarasooriya, "A tight upper bound on the Bayesian probability of error," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 2, pp. 220–224, 1994.
- [10] M. Feder and N. Merhav, "Relations between entropy and error probability," *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp. 259–266, 1994.
- [11] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, Wiley-Interscience, 2001.