2000

Optimal Distributed Binary Hypothesis testing with Independent Identical Sensors

Q. Zhang
Syracuse University, Department of Electrical Engineering and Computer Science, qizhang@syr.edu

P. K. Varshney
Syracuse University, Department of Electrical Engineering and Computer Science, varshney@syr.edu

R. D. Wesel
University of California Los Angeles, Electrical Engineering Department, wesel@ee.ucla.edu

Optimal Distributed Binary Hypothesis Testing with Independent Identical Sensors

Q. Zhang and P. K. Varshney
EECS Department
Syracuse University
Syracuse, NY 13244
e-mail: qizhang@syr.edu, varshney@syr.edu

R. D. Wesel
Electrical Engineering Department
University of California Los Angeles
Los Angeles, CA 90095-1594
e-mail: wesel@ee.ucla.edu

Abstract - We consider the problem of distributed binary hypothesis testing with independent identical sensors. It is well known that for this problem the optimal sensor rules are a likelihood ratio threshold tests and the optimal fusion rule is a *K*-out-of-*N* rule [1]. Under the Bayesian criterion, we show that for a fixed *K*-out-of-*N* fusion rule, the probability of error is a quasiconvex function of the likelihood ratio threshold used in the sensor decision rule. Therefore, the probability of error has a single minimum and a unique optimal threshold achieves this minimum. We obtain a sufficient and necessary condition on the optimal threshold, except in some trivial situations where one hypothesis is always decided. We present a method for determining whether or not the solution is trivial. Under the Neyman-Pearson criterion, we show that when the Lagrange multiplier method is used for a fixed *K*-out-of-*N* fusion rule, the objective function is quasiconvex and hence has a single minimum point, and the resulting ROC is concave downward. These results are illustrated by means of three examples.

I. INTRODUCTION

The problem of distributed binary hypothesis testing with multiple sensors has been extensively studied in the past decade. A fundamental result is that when the sensors are conditionally independent, the optimal sensor decision rules are necessarily likelihood ratio threshold tests [1]. However, determination of the thresholds is generally difficult. This difficulty is further increased due to the possible existence of multiple local optima [1]. One can obtain a numerical solution via a person-by-person-optimization (PBPO) procedure but the result may be only locally optimal. In this paper, we focus on the problem of distributed binary hypothesis testing with parallel independent identical sensors. By identical sensors, we mean that all the sensor observations follow a common probability distribution and all the sensors use identical decision-making rules. Using identical sensor rules gives a suboptimal result. However, the loss of optimality vanishes when the number of sensors goes to infinity [2,3].

In this paper, we first consider a special case of the Bayesian detection problem. Here we know the prior probabilities and we seek the optimum fusion rule and the optimum sensor decision rules that minimize the average probability of error. We then consider the Neyman-Pearson detection problem. We seek the optimum fusion rule and the optimum sensor decision rules such that the probability of detection is maximized while the probability of false alarm is kept below a prescribed level. For both problems, it is known that the optimum fusion rule is a *K*-out-of-*N* rule and the optimal sensor rule is a likelihood ratio threshold test [1]. A recent study by Shi, Sun and Wesel [4,5] reveals an interesting property of this problem: quasiconvexity. They considered the problem of distributed detection of known signals in additive noise with identical sensors. Each sensor observation was a univariate random variable. Each sensor made a binary decision by comparing its observation to a threshold. For additive generalized Gaussian noise with any priors and for certain other additive noises with equal priors, they showed that the probability of error is a quasiconvex function of the sensor threshold, given a fixed *K*-out-of-*N* fusion rule.

The concept of quasiconvexity [6] can be quite useful since it eliminates the existence of multiple local optima. Hence a local optimum determined by any method is also the global optimum. Furthermore, it will provide justification for the use of the Lagrange multiplier method for solving the Neyman-Pearson detection problem with independent identical sensors. Note that the attempts of using this method to tackle the general Neyman-Pearson detection problem have been questioned because the knowledge of the convexity regarding the problem is lacking.

In this paper, we consider the general situation with sensor observations having an arbitrary probability distribution. We consider the sensor decision rules in the likelihood ratio space, not in the observation space as in [4], since it is well known that an optimal sensor decision rule is a likelihood ratio threshold test [1]. The main contribution of this paper is the following. For a special case of Bayesian detection problem, we show that for any value for the prior probabilities and for a fixed *K*-out-of-*N* fusion rule, the average probability of error is a quasiconvex function of the sensor threshold. For the Neyman-Pearson detection problem, we show that when the Lagrange multiplier method is used for a fixed *K*-out-of-*N* fusion rule, the objective function is a quasiconvex function of the sensor threshold. This quasiconvexity ensures that the objective function has a single minimum point that is achieved by the unique optimal sensor threshold.

In Section II, we present some definitions and notations. In Section III, we consider the Bayesian detection problem. In Section IV, we consider the Neyman-Pearson detection problem. In section V, we illustrate the results of Sections III and IV by means of three examples. In Section VI, we make some concluding remarks.

II. DEFINITIONS AND NOTATIONS

Let us consider a parallel fusion system that consists of n independent identical sensors and a fusion center. This system is used to determine whether an unknown hypothesis is H_0 or H_1. The sensor observations are conditionally independent given the unknown hypothesis. All the sensors employ the same decision-making rule to make a binary decision regarding the identity of the unknown hypothesis. The sensors transmit their decisions to the fusion center. Based on the received sensor decisions, the fusion center makes the final decision.

Let $\boldsymbol{x}_i$ denote the observation of the ith sensor, i=1,…,n. Let u_i denote the decision of the ith sensor. For these identical sensors, let $p_j(.)$ denote the common probability density function of observations at any sensor when H_j is true, j=0,1. The optimal identical sensor decision rule can be expressed as

$$\frac{p_1(\boldsymbol{x}_i)}{p_0(\boldsymbol{x}_i)} \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \lambda \,, \tag{1}$$

where λ is the common threshold. The quality of a sensor decision u_i can be measured by the probability of false alarm P_F and the probability of detection P_D of the ith sensor. Since the sensors are identical, P_F and P_D are the same for every sensor. The (P_F, P_D) curve, which is generally referred to as the receiver operating characteristic (ROC), is important in the derivation of our result.

Let u_0 denote the final decision made by the fusion center. Let u_0=0 if the fusion center decides H_0. Let u_0=1 if the fusion center decides H_1. Recalling that an optimum fusion rule is a K-out-of-N rule, we express the fusion rule as

$$u_0 = \begin{cases} 1 & \text{if } u_1 + \cdots + u_n \ge k \\ 0 & \text{if } u_1 + \cdots + u_n < k \end{cases}, \tag{2}$$

where k is an integer and $1 \le k \le n$. The quality of the fusion center decision u_0 can be measured by the probability of false alarm Q_F and the probability of detection Q_D of the fusion system

$$Q_F = \sum_{i=k}^{n} \binom{n}{i} P_F^i \left(1 - P_F\right)^{n-i} \,, \tag{3.a}$$

$$Q_D = \sum_{i=k}^{n} \binom{n}{i} P_D^i \left(1 - P_D\right)^{n-i} \,. \tag{3.b}$$

A number of performance criteria can be formulated based on Q_F and Q_D. Our objective is to choose an appropriate performance criterion and find the corresponding optimal k and λ. In the next section, we use the minimum probability of error criterion. It is a specific Bayesian problem. However, we conjecture that the results hold for the general Bayesian problem. In section IV, we use the Neyman-Pearson criterion.

III. THE BAYESIAN DETECTION PROBLEM

In this section, we consider the Bayesian detection problem. Let q_0 and q_1 denote the prior probabilities of H_0 and H_1. Our goal is to find the fusion rule and the sensor decision rule that minimize the probability of error. Using Q_F and Q_D, we express the probability of error P_e as

$$P_e = q_0 Q_F + q_1 \left(1 - Q_D\right). \tag{4}$$

P_e is a function of k, λ, q_0 and q_1. Our goal is to find λ and k such that P_e is minimized. Toward this goal, we minimize P_e for each k, where $1 \le k \le n$. We then choose the smallest of these minima and the corresponding values of λ and k yield the desired solution. This systematic procedure is exhaustive and is guaranteed to result in a globally optimal solution.

Next, we consider the minimization of P_e for each k. We show that for a fixed k, P_e is a quasiconvex function of λ. By a quasiconvex function $f(\lambda)$ of λ, we mean that for some λ^*, $f(\lambda)$ is non-increasing for $\lambda \le \lambda^*$ and $f(\lambda)$ is non-decreasing for $\lambda \ge \lambda^*$ [7]. In this paper, we assume that P_F and P_D have first order derivatives with respect to λ. This assumption is not restrictive in practical situations.

Lemma 1: For a given k, P_e is a quasiconvex function of λ.

Proof:

To prove the lemma, it suffices to show that either $\frac{dP_e}{d\lambda} \le 0$ (or $\frac{dP_e}{d\lambda} \ge 0$) for all λ, or $\frac{dP_e}{d\lambda} \le 0$ when $\lambda \le \lambda^*$ and $\frac{dP_e}{d\lambda} \ge 0$ when $\lambda \ge \lambda^*$ for some λ^*. Some fundamental properties of ROC are used in the proof. Taking the derivative of both sides of equations (3.a), (3.b) and (4) with respect to λ, we obtain

$$\frac{dQ_F}{d\lambda} = n\binom{n-1}{k-1} \cdot \left(\frac{dP_F}{d\lambda}\right) \cdot \left(q_0 P_F^{k-1} \left(1 - P_F\right)^{n-k}\right),$$

$$\frac{dQ_D}{d\lambda} = n\binom{n-1}{k-1} \cdot \left(\frac{dP_F}{d\lambda}\right) \cdot \left(\lambda q_1 P_D^{k-1} \left(1 - P_D\right)^{n-k}\right),$$

$$\frac{dP_e}{d\lambda} = q_0 \frac{dQ_F}{d\lambda} - q_1 \frac{dQ_D}{d\lambda} \,.$$

Here we note that $\frac{dP_D}{dP_F} = \lambda$, a property of ROC. Define

$$g(\lambda,k) = n\binom{n-1}{k-1} \cdot \left(-\frac{dP_F}{d\lambda}\right) \cdot \left(q_0 P_F^{k-1} \cdot (1 - P_F)^{n-k}\right),$$

$$r(\lambda,k) = ln \frac{q_1}{q_0} + ln\lambda + (k-1) \cdot ln \frac{P_D}{P_F} + (n-k) \cdot ln \frac{1-P_D}{1-P_F} \,, \tag{5}$$

and we obtain

$$\frac{dP_e}{d\lambda} = g(\lambda,k) \cdot \left(e^{r(\lambda,k)} - 1\right). \tag{6}$$

Since P_F decreases as λ increases, we have $g(\lambda,k) \ge 0$. Therefore, the sign of $\frac{dP_e}{d\lambda}$ is determined by $r(\lambda,k)$. According to equation (6), it suffices to show that $r(\lambda,k)$ is either always negative (positive), or $r(\lambda,k) \le 0$ for $\lambda \le \lambda^*$ and $r(\lambda,k) \ge 0$ for $\lambda \ge \lambda^*$ for some λ^*.

Taking the derivative of both sides of equation (5) with respect to λ, we obtain

$$\frac{dr(\lambda,k)}{d\lambda} = \frac{1}{\lambda} + \left(-\frac{dP_F}{d\lambda}\right) \cdot \left[\frac{k-1}{P_D} \cdot \left(\frac{P_D}{P_F} - \lambda\right) + \frac{n-k}{1-P_D} \cdot \left(\lambda - \frac{1-P_D}{1-P_F}\right)\right]. \tag{7}$$

Since (P_F, P_D) is a point on a ROC curve, as shown in Figure 1, we have $\frac{P_D}{P_F} \ge \lambda \ge \frac{1-P_D}{1-P_F}$. Based on this result and that $\frac{dP_F}{d\lambda} \le 0$ and $1 \le k \le n$, we have $\frac{dr(\lambda,k)}{d\lambda} \ge 0$, i.e. $r(\lambda,k)$ is a monotone non-decreasing function of λ. Hence, $r(\lambda,k)$ either intersects the λ-axis at some λ^*, or it does not intersect the λ-axis at all.

If $r(\lambda,k)$ does not intersect the λ-axis, then $r(\lambda,k)$ is always negative (positive). From equation (6), $\frac{dP_e}{d\lambda}$ is always negative (positive). Thus P_e monotonically decreases (or increases) with λ.

If $r(\lambda,k)$ intersects the λ axis at some λ^*, then $r(\lambda,k) \le 0$ holds for $\lambda \le \lambda^*$ and $r(\lambda,k) \ge 0$ holds for $\lambda \ge \lambda^*$. ■

In many situations, use of the log likelihood ratio is preferred. Let $\tau = ln\lambda$. Because τ is a monotone increasing function of λ, Lemma 1 holds when λ is replaced by τ.

Corollary 1: For a given k, P_e is a quasiconvex function of τ.

Remark 1:

The proof of Lemma 1 shows that $\frac{dr(\lambda,k)}{d\lambda} \ge 0$. A careful study of equation (7) shows that equality holds only if $\lambda = +\infty$. Hence, $r(\lambda,k)$ is a monotone increasing function of λ and $r(\lambda,k)=0$ has at most one root λ^*, where λ^* can take the value of $+\infty$. For $\lambda < \lambda^*$, $r(\lambda,k) < 0$ and P_e is a monotone decreasing function of λ. For $\lambda > \lambda^*$, $r(\lambda,k) > 0$ and P_e is a monotone increasing function of λ.

Remark 2:

The proof of Lemma 1 and Remark 1 imply that P_e has a single minimum that is achieved by a unique λ. This result is stated in the following theorem.

Theorem 1: For a given k, P_e has a single minimum, which is achieved by a unique λ.

The proof of Lemma 1 suggests that if λ satisfies $r(\lambda,k)=0$, then the corresponding value of λ minimizes P_e. This is a sufficient condition stated in the following theorem.

Theorem 2: For a given k, λ minimizes P_e if it satisfies

$$ln\frac{q_1}{q_0}+ln\lambda+(k-1)\cdot ln\frac{P_D}{P_F}+(n-k)\cdot ln\frac{1-P_D}{1-P_F}=0\,. \quad (8)$$

It is shown later in Remarks 4 and 5 that for a non-trivial solution, i.e. when the fusion system does not always decide one hypothesis, equation (8) is also a necessary condition on the optimal λ.

To use Theorems 1 and 2 to find the optimal λ, $r(\lambda,k)=0$ must have a positive root. This condition is satisfied for a class of sensors. On the sensor ROC, let $\lambda_{0,0}$ denote the slope at the point (0,0), and $\lambda_{1,1}$ the slope at the point (1,1). The following theorem shows that if $\lambda_{0,0}=\infty$ and $\lambda_{1,1}=0$, then $\lambda_{1,1}=0$ has a positive root for any non-zero prior probabilities.

Theorem 3: For a given k, $1\le k\le n$, if $\lambda_{0,0}=\infty$, $\lambda_{1,1}=0$, and q_0, $q_1>0$, then $r(\lambda,k)=0$ has a unique positive root.

Proof:

The uniqueness is shown in Remark 1. To prove the existence, it suffices to show that $\lim_{\lambda\to 0} r(\lambda,k)=-\infty$ and $\lim_{\lambda\to +\infty} r(\lambda,k)=+\infty$.

Since $\lambda_{0,0}=\infty$ and $\lambda_{1,1}=0$, we have

$$\lim_{\lambda\to 0}\frac{P_D}{P_F}=1\,,\ \lim_{\lambda\to 0}\frac{1}{\lambda}\cdot\frac{1-P_D}{1-P_F}=1\,,\ \lim_{\lambda\to +\infty}\frac{1-P_D}{1-P_F}=1 \text{ and } \lim_{\lambda\to +\infty}\frac{1}{\lambda}\cdot\frac{P_D}{P_F}=1\,.$$

Since $0<q_0$, $q_1<1$, $ln\frac{q_1}{q_0}$ is finite. Using these limits, the fact that $1\le k\le n$ and the definition of $r(\lambda,k)$ in equation (5), we have

$$\lim_{\lambda\to 0} r(\lambda,k)=-\infty \text{ and } \lim_{\lambda\to +\infty} r(\lambda,k)=+\infty\,.$$

We recall that $\frac{dr(\lambda,k)}{d\lambda}>0$ from Remark 1 and we conclude that $r(\lambda,k)=0$ has a positive root. ■

Remark 3:

Putting the four finite limits that are used in the proof of Theorem 3 into the definition of $r(\lambda,k)$ in equation (5) and using $\tau=ln\lambda$, we have

$$\lim_{\tau\to -\infty}\frac{r(\tau,k)}{(n-k+1)\cdot\tau}=1 \text{ and } \lim_{\tau\to +\infty}\frac{r(\tau,k)}{k\cdot\tau}=1\,.$$

Therefore, $r(\tau,k)$ is approximately a linear function of τ at $\pm\infty$.

Because of this property, the $r(\tau,k)$ curve can be approximated by two straight lines and the SECANT algorithm [7] can be used to find the root of $r(\tau,k)=0$. This algorithm is stated as follows

1. Arbitrarily choose τ_1, τ_2 and a positive ε. Compute $r_1=r(\tau_1,k)$ and $r_2=r(\tau_2,k)$. Set $i=3$.
2. Let $\tau_i=\frac{r_{i-1}\cdot\tau_{i-2}-r_{i-2}\cdot\tau_{i-1}}{r_{i-1}-r_{i-2}}$. Compute $r_i=r(\tau_i,k)$.
3. If $|r_i|\le\varepsilon$, stop*;* otherwise, let $i=i+1$, go to step *2*.

When $\lambda_{0,0}=\infty$ and $\lambda_{1,1}=0$, the algorithm converges quickly because $r(\tau,k)$ is well approximated by two straight lines. This algorithm is illustrated in Figure 2.

Remark 4:

If a sensor does not satisfy $\lambda_{0,0}=\infty$ and $\lambda_{1,1}=0$, it is possible that $r(\lambda,k)=0$ has no root. Actually, it can be determined whether or not $r(\lambda,k)=0$ has a root. Since $r(\lambda,k)$ is a monotone increasing function of λ, $r(\lambda,k)=0$ has a root if and only if $r(\lambda_{0,0},k)\ge 0$ and $r(\lambda_{1,1},k)\le 0$. When these conditions are satisfied, the uniqueness of the optimal λ (Theorem 1) implies that $r(\lambda,k)=0$ is a necessary condition on the optimal λ.

Remark 5:

When the conditions given in Remark 4 are not satisfied, one can easily obtain the minimum P_e and the corresponding optimal values of λ. If $r(\lambda_{0,0},k)<0$, P_e is a monotone decreasing function of λ and its minimum occurs at $\lambda=\lambda_{0,0}$. In this case, H_0 is always decided and the minimum P_e is equal to q_1. On the other hand, if $r(\lambda_{1,1},k)>0$, P_e is a monotone increasing function of λ and its minimum occurs at $\lambda=\lambda_{1,1}$. In this case, H_1 is always decided and the minimum P_e is equal to q_0. These are trivial solutions.

From Theorem 2, we observe that the optimal value of λ is intimately related to k, q_0 and q_1. Here we present some results on these relationships. The proofs are not included due to limited space. They can be found in [11]. Suppose $r(\lambda,k)=0$ has a positive root for each k, where $1\le k\le n$. Let λ_k denote this root.

Lemma 2: λ_k is a decreasing function of k for fixed q_1 and q_0, except when $P_D=P_F$, or when the sensors always decide H_1.

Lemma 3: λ_k is a decreasing function of q_1/q_0 for fixed k, except when the sensors always decide H_1.

IV. The Neyman-Pearson Detection Problem

In this section, we consider the Neyman-Pearson detection problem. We show that this problem can be formulated as a minimum probability of error problem that we considered in the Section III.

In a Neyman-Pearson detection problem, Q_D is maximized while Q_F is kept below a prescribed level α. Since the sensor decisions are discrete random variables, “dependent randomization” [8] or “scheduling” [9] may be necessary. Basically, these schemes employ synchronized randomization of two sets of rules namely the fusion rule and sensor rules. These schemes introduce undesired degree of freedom and require additional computational resource for synchronization. In this paper, we do not consider these schemes that are undesirable in most practical situations. Without “dependent randomization” or “scheduling”, Warren and Willett prove that randomized fusion rules are suboptimal when the sensor observations contain no point-mass of probability [10]. Since this condition is assumed satisfied in this paper, we only need to consider *K*-out-of-*N* fusion rules.

We break the original Neyman-Pearson detection problem into a set of Neyman-Pearson detection problems for each value of k. We solve these problems for each value of k, then choose the solution that yields the maximal Q_D. For a given k, our goal is

Maximize Q_D,
Subject to $Q_F\le\alpha$,
where α is a prescribed positive constant.

We are particularly interested in solving this problem via the Lagrange multiplier method. This method has seen limited use in solving the general distributed Neyman-Pearson detection problem with non-identical sensors [1]. The reason is that the objective function may not be convex, even with a fixed fusion rule. This makes it difficult to determine whether or not there are multiple local extreme points, and if so, the number of local extreme points and their locations. The same difficulty arises when the fusion rule is not fixed. Fortunately for our problem, we show that for a fixed K-out-of-N fusion rule, the objective function is a quasi-convex function and therefore the Lagrange multiplier method can be employed to obtain the unique global minimum. Furthermore, we show that for a fixed K-out-of-N fusion rule, the ROC is concave downward.

Define $L_k(\lambda,s)=s(Q_F-\alpha)-Q_D$, where s is the Lagrange multiplier and $s\geq 0$. Our goal is to minimize $L_k(\lambda,s)$ with respect to s and λ.

Lemma 4: $L_k(\lambda,s)$ is a quasiconvex function of λ.
Proof:
Since $s\geq 0$, we define prior probabilities $q_0=s/(1+s)$, $q_1=1/(1+s)$. Putting q_0 and q_1 into equation (4), we have $P_e=q_0Q_F+q_1(1-Q_D)$. Putting P_e into the definition of $L_k(\lambda,s)$, we obtain

$$L_k(\lambda,s)=\frac{1}{q_1}P_e-\left(\frac{q_0}{q_1}\alpha+1\right).$$

Since P_e is a quasiconvex function of λ, so is $L_k(\lambda,s)$. ■

Now $L_k(\lambda,s)$ has a single minimum that is uniquely achieved by the optimal threshold. Noting that $s=q_0/q_1$, we obtain the solution by solving the following equations

$$Q_F-\alpha=0\ ,$$

$$\ln\frac{1}{s}+\ln\lambda+(k-1)\cdot\ln\frac{P_D}{P_F}+(n-k)\cdot\ln\frac{1-P_D}{1-P_F}=0\ .$$

To find the overall optimal solution, we repeat the above procedure for all possible values of k, and then choose the pair of k and λ that gives the largest Q_D.

In general, the resulting ROC is not concave downward as in centralized detection. However, we show that for a fixed K-out-of-N fusion rule, the ROC is concave.

Lemma 5: For a given k, Q_D is a concave function of Q_F.
Proof:
It suffices to show that $\frac{d^2Q_D}{dQ_F^2}\leq 0$. From the proof of Lemma 1, we obtain $\frac{dQ_D}{dQ_F}=e^{r(\lambda,k)}$. Thus, we have

$$\frac{d^2Q_D}{dQ_F^2}=\frac{de^{r(\lambda,k)}}{dQ_F}=\frac{de^{r(\lambda,k)}}{d\lambda}\Big/\frac{dQ_F}{d\lambda}=e^{r(\lambda,k)}\cdot\frac{dr(\lambda,k)}{d\lambda}\Big/\frac{dQ_F}{d\lambda}\ .$$

From Remark 1, we have $\frac{dr(\lambda,k)}{d\lambda}\geq 0$. Since $\frac{dP_F}{d\lambda}\leq 0$, we have $\frac{dQ_F}{d\lambda}\leq 0$. Hence we have $\frac{d^2Q_D}{dQ_F^2}\leq 0$. ■

This concavity ensures that the Lagrange multiplier method can be used to uniquely determine the optimal threshold in the case considered here [1].

V. NUMERICAL RESULTS

In this section, we illustrate the results of the previous sections by means of three examples.

Example 1:
In this example, we consider the detection of known signals in Gaussian noise. The sensor observation x is $x=s_x+n_x$, where $s_x=\pm d$ is the transmitted signal and n_x is a Gaussian random variable with zero mean and unit variance. Define $H_0\equiv\{s_x=-d\}$ and $H_1\equiv\{s_x=+d\}$. The log likelihood ratio τ_x for this problem is given by $\tau_x=2dx$. The sensor false alarm and detection probabilities can be computed as

$$P_F=Q\left(\tfrac{\tau}{2d}+d\right),$$

$$P_D=Q\left(\tfrac{\tau}{2d}-d\right),$$

where $Q(z)=\int_z^\infty\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$ and τ is the log likelihood ratio threshold.

With fixed n=11 and d=0.1, we consider various combinations of q_0 and q_1, and observe the relationship between P_e, τ, k, q_0 and q_1.

In Figure 3, with $q_0=q_1=0.5$, P_e is plotted against τ for each value of k. We can see that for any given k, P_e is a quasiconvex function of τ and has a single minimum achieved by a unique value of τ. These results agree with the main results of Section III. We also notice that the optimal value of τ decreases with k, as suggested in Lemma 2.

In Figure 4, with $q_0=q_1=0.5$, $r(\tau,k)$ is plotted against τ for each value of k. We can see that $r(\tau,k)$ is a monotonically increasing function of τ for any given k. We also notice that an $r(\tau,k)$ curve can be well approximated by two concatenated straight lines.

In Figure 5, with q_0=0.75 and q_1=0.25, P_e is plotted against τ for each value of k. Comparing this figure to Figure 3, we find that the optimal τ for each value of k has increased. Such increases are due to the decrease in q_1/q_0, as implied by Lemma 3. In this figure, the global minimum P_e occurs at k=7. This provides a counter example to the conjecture that the best k is $0.5(n+1)$ [4]. However, the conjecture may still hold for the equal priors case.

Example 2:
In this example, we consider the non-coherent detection of a quadrature signal in Gaussian noise. The received signal has random phase and is subject to Rayleigh fading. When the signal is present in the environment, the sensor observation is

$$x_I=s_I+n_I$$

$$x_Q=s_Q+n_Q$$

where subscripts I and Q denote the in-phase and quadrature components. The signal components s_I and s_Q have random phase and are subject to Rayleigh fading. The average power of the signal is equal to $0.5\sigma^2$. n_I and n_Q are independent Gaussian random variables with zero mean and unit variance. When the signal is absent, the observation is just Gaussian noise.

Define $H_0\equiv$\{signal is absent\} and $H_1\equiv$\{signal is present\}. Because no phase information is available, non-coherent detection is employed. In this case, we use the statistic $r^2=x_I^2+x_Q^2$. It can be shown that r has Rayleigh density function under either H_0 or H_1.
The log likelihood ratio is $\tau_r=0.5r^2\sigma^2/(1+\sigma^2)-ln(1+\sigma^2)$. We note that the minimum value of τ_r is $-ln(1+\sigma^2)$. Similarly, the minimum value of the joint log likelihood ratio for n sensors is $-nln(1+\sigma^2)$. By the minimum value, we mean that any smaller value has no physical meaning.

The sensor false alarm and detection probabilities are given by

$$P_F=\left(\frac{1}{1+\sigma^2}\right)^{1+1/\sigma^2}e^{-\tau(1+1/\sigma^2)},$$

$$P_D = (P_F)^{\frac{1}{1+\sigma^2}},$$

where τ is the log likelihood ratio threshold and $\tau \geq -ln(1+\sigma^2)$.

In Figure 6, with n=7, σ^2=0.5, q_0=0.55 and q_1=0.45, P_e is plotted against τ for each value of k. In this case, P_e is a non-symmetric function of τ and the minimum τ is equal to $-ln(1+\sigma^2) \approx$-0.4055.

Since the minimum value of the joint log likelihood ratio for n sensors is $-nln(1+\sigma^2)$, it is possible that for some values of q_0 and q_1, P_e monotonically increases with τ. In fact, this happens when $q_1/q_0 \geq (1+\sigma^2)$. In this case, the minimum P_e occurs at $\tau = -ln(1+\sigma^2)$ and H_1 is always decided. This phenomenon is shown in Figure 7, with n=5, σ^2=0.5, q_0=0.1164 and q_1=0.8836.

Example 3:

In this example, we consider the previous two examples under the Neyman-Pearson criterion. We solve the Neyman-Pearson detection problem for each value of k and then select the solution that yields the maximum probability of detection. We recall that Q_F is the probability of false alarm of the fusion system and Q_D is the probability of detection of the fusion system.

First, we consider the detection of known signals in Gaussian noise as defined in Example 1. In Figure 8, with n=3, d=0.5, the ROCs are plotted for each value of k. We find that the K-out-of-N fusion rule with k=2 always gives the best solution. When $Q_F \leq 0.25$, k=1 fusion rule is better than the k=3 fusion rule. When $Q_F \geq 0.25$, k=3 fusion rule is better than the k=1 fusion rule.

Next, we consider the non-coherent detection of a quadrature signal in Gaussian noise as defined in Example 2. In Figure 9, with n=4, σ^2=1, the ROCs are plotted for each value of k. We find that for $Q_F \geq 0.0162$, the k=1 fusion rule is the best. For $Q_F \leq 0.0162$, the k=2 fusion rule is best.

VI. SUMMARY

We considered the Bayesian and Neyman-Pearson detection problems with distributed independent identical sensors. The goal was to find the optimal K-out-of-N fusion rule and the optimal sensor likelihood ratio threshold test. We showed that for a given K-out-of-N fusion rule, the corresponding objective function exhibits the property of quasiconvexity. This property ensures that the objective function has a single minimum that is uniquely achieved by the optimal sensor likelihood ratio threshold.

ACKNOWLEDGEMENT

Research sponsored by the Space and Naval Warfare Systems Center San Diego, under Grant No. N66001-99-1-8922.

REFERENCES

[1] Pramod K. Varshney, *Distributed Detection and Data Fusion*, Springer-Verlag New York, Inc., 1997.

[2] J.N. Tsitsiklis, "Decentralized detection by a large number of sensors," Mathematics of control, Signals, and Systems, vol. 1, pp. 167-182, 1988.

[3] P. Chen and A. Papamarcou, "New asymptotic results in parallel distributed detection," *IEEE Trans. Info. Theory*, vol. 39, no. 6, pp. 1847-1863, Nov. 1993.

[4] W. Shi, T. W. Sun and R. D. Wesel, "Optimal binary distributed detection," *Proceedings of the 33rd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA. October 24-27, 1999.

[5] W. Shi, T. W. Sun and R. D. Wesel, "Optimal binary distributed detection," *Proceedings of the2000 IEEE International Symposium on Information Theory*, Sorrento, Italy June 2000.

[6] Lars Hörmander, *Notions of Convexity*, Birkhäuser Boston, 1994.

[7] W.H. Press et al., *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, New York, 1986.

[8] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, vol.2 (Signal Detection), H. V. Poor and J. B Thomas, Eds. Greenwich, CT: Jai Press, 1991.

[9] P. Willett and D. Warren, "Randomization and scheduling in decentralized detection," *Proc. 27th Allerton Conf. Commun., Contr., and Comput.*, pp649-650, Oct. 1989.

[10] P. Willett and D. Warren, "The suboptimality of randomized tests in distributed and quantized detection systems," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, March 1992.

[11] Q. Zhang, *Ph.D. dissertation in progress*.

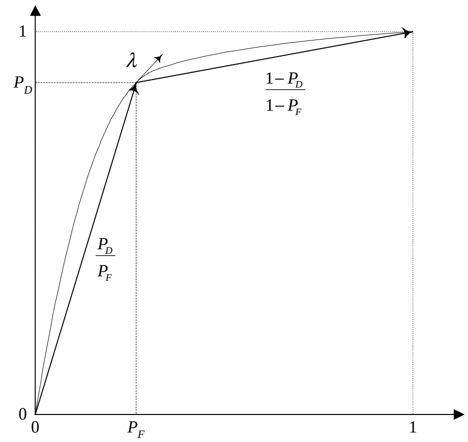

Figure 1: Sensor ROC

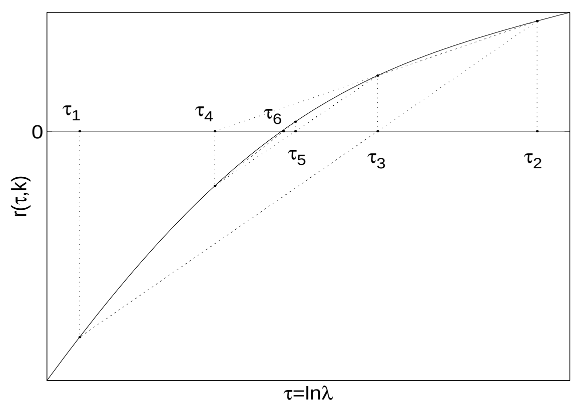

Figure 2: The SECANT algorithm

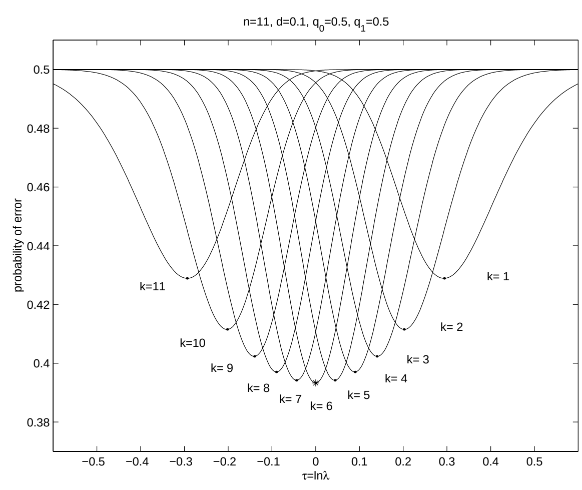

Figure 3: P_e vs. τ curves with equal priors for Example 1

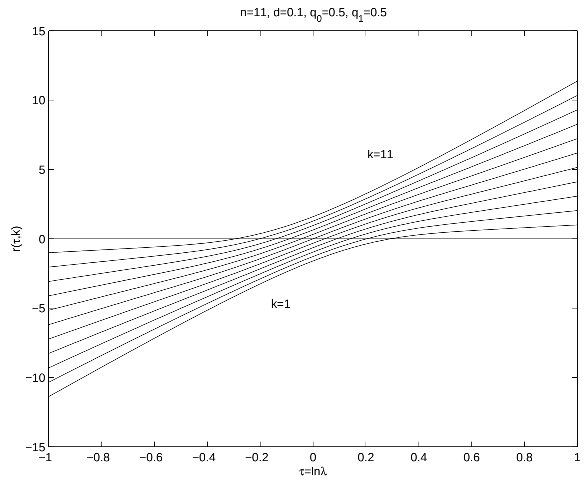

Figure 4: $r(\tau,k)$ curves for Example 1

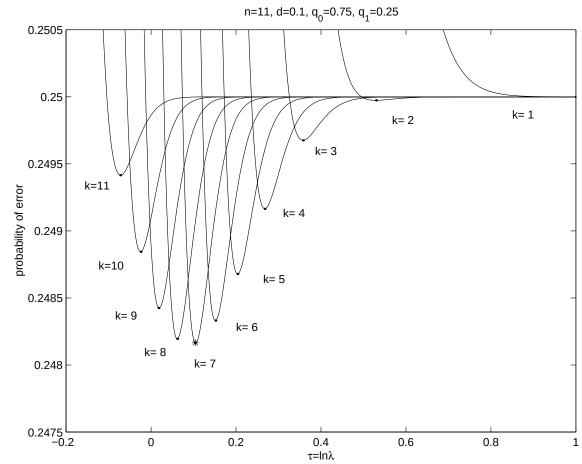

Figure 5: The optimal k with unequal priors for Example 1

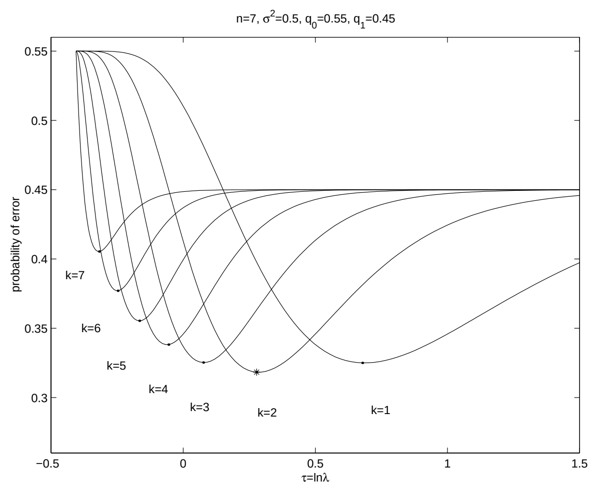

Figure 6: Non-symmetric P_e vs. τ curves for Example 2

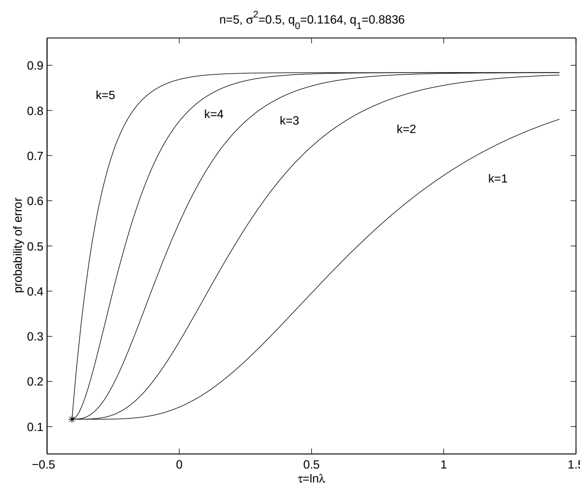

Figure 7: Monotonic P_e vs. τ curves for Example 2

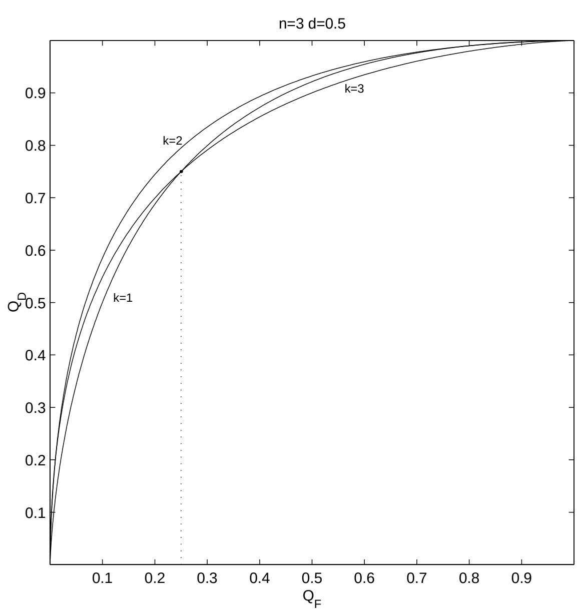

Figure 8: ROC curves for the detection of known signals in Gaussian noise for Example 3

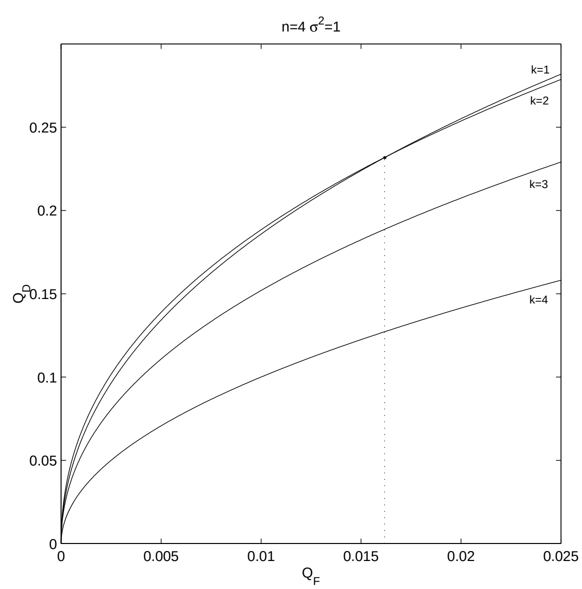

Figure 9: ROC curves for non-coherent signal detection in Rayleigh fading channel for Example 3