7-1995

Robustness of Decentralized Tests with ε-Contamination Prior

Chandrakanth H. Gowda
Southern Illinois University Carbondale

R. Viswanathan
Southern Illinois University Carbondale, viswa@engr.siu.edu

Published in Gowda, C.H., & Viswanathan, R. (1995). Robustness of decentralized tests with ε-contamination prior. IEEE Transactions on Information Theory, 41(4), 1164 - 1169. doi: 10.1109/18.391263

Robustness of Decentralized Tests with ϵ-Contamination Prior

Chandrakanth H. Gowda and R. Viswanathan, *Senior Member, IEEE*

***Abstract*—We consider a decentralized detection problem where the prior density is not completely known, but is assumed to belong to an ϵ-contamination class. The expressions for the infimum and the supremum of the posterior probability that the parameter under question is in a given region, as the prior varies over the ϵ-contamination class, are derived. Numerical results are obtained for a specific case of an exponentially distributed observation and an exponentially distributed nominal prior. Asymptotic (as number of sensors tends to a large value) results are also obtained. The results illustrate the degree of robustness achieved with quantized observations as compared to unquantized observations.**

***Index Terms*—Decentralized detection, ϵ-contamination prior, posterior robustness.**

I. Introduction

Research issues in decentralized detection problems have received increased attention in recent years. Tenny and Sandell [1] extended the classical Bayesian decision theory to distributed Bayesian detection problems, in particular to a two-sensor system. Later works include a generalized Bayesian formulation of the distributed detection problem [2], a decentralized version of the sequential Bayesian hypothesis testing problem [3], [4], a survey of results on decentralized detection [5], distributed locally optimal detection [6], and robustness issues in decentralized detection [7], [8]. In [7], Veeravalli, Basar, and Poor have studied the decentralized detection problem in which the distribution of a sensor observation is not completely specified. They have found that under a very general regularity condition on the distribution of the observations, the least favorable density for a decentralized detection problem is exactly the same as the least favorable density for the corresponding centralized detection problem.

We have looked into the posterior robustness in a decentralized binary hypothesis testing problem where the prior distribution is not completely specified. Specifically, we have applied the work of Berger [9] on centralized hypothesis testing to the decentralized case. The posterior distribution of the parameter θ given observation x (denoted by $\pi(\theta|x)$) combines the prior beliefs about θ with the information about θ contained in the sample observation x, to give a composite picture of θ. We consider a system of n sensors and a fusion center. The ith sensor receives an observation $X_i, i = 1, 2, \cdots, n$. When conditioned on $\theta, X_1, X_2, \cdots, X_n$ are a set of independent and identically distributed random variables with a marginal density function $f(x|\theta)$. f is assumed to be known completely. The prior for $\theta, \pi(\theta)$, is not known completely but is known to belong to an ϵ-contamination class Ψ

$$\Psi: \pi(\theta) = (1-\epsilon)\pi_o(\theta) + \epsilon q(\theta). \tag{1}$$

Manuscript received September 14, 1994; revised January 30, 1995. This work was supported by BMDIO/IST and managed by the Office of Naval Research under Contract N00014-94-1-0736. The material in this correspondence was presented at the 28th Annual Conference on Information Sciences and Systems, Princeton University, Princeton, NJ, March 1994.

The authors are with the Department of Electrical Engineering, Southern Illinois University at Carbondale, Carbondale, IL 62901-6603 USA.

IEEE Log Number 9412185.

In (1), ϵ is a positive fraction, π_o is known completely, and q is considered to belong to the set of all distributions. Let C denote a subset of the parameter space. The three different cases considered are as follows.

a) All observations, $X_1, X_2, \cdots, X_n$, are given to the fusion center and no decision is made. In this case we compute the supremum and the infimum of $P(\theta \in C|X_1, X_2, \cdots, X_n)$ as the prior varies over the ϵ-contamination class.

b) A decision U_o based on the observations $X_1, X_2, \cdots, X_n$ is made at the fusion center. In this case we compute the supremum and the infimum of $P(\theta \in C|U_o)$.

c) The fusion center makes decision U_o based on the decisions made at the sensors. We compute the supremum and the infimum of $P(\theta \in C|U_o)$.

Case a), which corresponds to centralized detection, is considered so that the losses associated with quantization in the other two cases can be assessed.

For illustration we take a specific case where the observations X_i and the nominal distribution of θ are exponentially distributed, and the confidence interval $C = (0, \theta_o), \theta_o > 0$. Another case where these variables are normally distributed is considered in [10]. The conclusions drawn from the normal example are similar to those presented here.

This correspondence is organized as follows. For the three scenarios, we state the equations for the infimum and the supremum of the posterior probabilities in Section II and discuss the numerical results obtained in Section III. In Section IV, the convergence of the posterior probabilities for a large number of sensors is discussed. In Section V, we consider an optimization problem in terms of the choice of k in the k-out-of-n fusion rule. Section VI concludes this correspondence.

II. Posterior Robustness of Decentralized Tests

Consider the following hypothesis testing problem:

$$H_o: \quad \theta \in C \equiv (\theta \leq \theta_o) \quad \text{versus} \quad H_1: \quad \theta \in \overline{C} \equiv (\theta > \theta_o).$$

Based on the scheme chosen, either i) each sensor sends its observation directly to the fusion center or ii) each sensor sends its decision to the fusion center. Let $U_o = 1(0)$ represent the action that the fusion center decides the hypothesis $H_1(H_o)$. Similarly, $U_i = 1(0)$ represents the ith sensor decision favoring hypothesis $H_1(H_o)$. In case c), the fusion rules considered are k-out of-n rule (KN), AND ($k = n$), OR ($k = 1$), Majority Logic ($k = n+1/2, n$ odd). The k-out-of-n rule implies $U_o = 1$ if and only if at least k of the U_i's equal 1, where the integer k satisfies $1 \leq k \leq n$.

The calculation of the posterior probabilities for an ϵ-contamination class of priors and a single observation is given in Berger [9]. The infimum and the supremum posterior probabilities for such a case are given by

$$\inf_{\pi \in \Psi} P^{\pi(\theta|x)}(\theta \in C) = \frac{N}{M + \frac{\epsilon}{1-\epsilon} Z} \tag{2}$$

$$\sup_{\pi \in \Psi} P^{\pi(\theta|x)}(\theta \in C) = 1 - \frac{M - N}{M + \frac{\epsilon}{1-\epsilon} Y} \tag{3}$$

where $M = m(x|\pi_o)$ is the marginal density of the observation X, with the nominal prior $\pi_o(\theta)$, P_o is the posterior probability that $\theta \in C$, with the nominal prior

$$N = MP_o, Z = \sup_{\theta \notin C} f(x|\theta)$$

TABLE I
EXPRESSIONS PERTAINING TO POSTERIOR PROBABILITY

Scheme	M	Z	Y
Case (a) $s \in C$	$\int_0^\infty \frac{s^{n-1}}{(\theta/n)^n \Gamma(n)} e^{\frac{-s}{(\theta/n)}} \frac{1}{\lambda} e^{-\theta/\lambda} d\theta$	$(\frac{1}{\theta_o/n})^n \frac{s^{n-1}}{\Gamma(n)} e^{\frac{-s}{(\theta_o/n)}}$	$\frac{n^n}{s\Gamma(n)} e^{-n}$
Case(a) $s \notin C$	Same as in $s \in C$	Interchange Z and Y shown for $s \in C$	
Case (b) $U_0=1$	$\int_0^\infty [\sum_{k=0}^{n-1} \frac{(nh/\theta)^k}{k!} e^{\frac{-h}{\theta/n}}] \frac{1}{\lambda} e^{-\theta/\lambda} d\theta$	1	$\sum_{k=0}^{n-1} \frac{(nh/\theta_o)^k}{k!} e^{\frac{-h}{\theta_o/n}}$
Case (b) $U_0=0$	$\int_0^\infty [1-\sum_{k=0}^{n-1} \frac{(nh/\theta)^k}{k!} e^{\frac{-h}{\theta/n}}] \frac{1}{\lambda} e^{-\theta/\lambda} d\theta$	$1-\sum_{k=0}^{n-1} \frac{(nh/\theta_o)^k}{k!} e^{\frac{-h}{\theta_o/n}}$	1
Case(c) KN $U_0=1$	$\int_0^\infty \{\sum_{p=0}^{n-k} \binom{n}{p} v^{n-p} (1-v)^p\} \frac{1}{\lambda} e^{-\theta/\lambda} d\theta$	1	$\sum_{p=0}^{n-k} \binom{n}{p} w^{n-p} (1-w)^p$
Case(c) KN $U_0=0$	$\int_0^\infty \{1-\sum_{p=0}^{n-k} \binom{n}{p} v^{n-p} (1-v)^p\} \frac{1}{\lambda} e^{-\theta/\lambda} d\theta$	$1-\sum_{p=0}^{n-k} \binom{n}{p} w^{n-p} (1-w)^p$	1

$v = e^{-t/\theta}$ and $w = e^{-t/\theta_o}$

and

$$Y = \sup_{\theta \in C} f(x|\theta).$$

If M is written as $\int_\theta I\, d\theta$, then N is simply $\int_C I\, d\theta$. In (2) and (3) the notation $\pi \in \Psi$ shows that the extremum probabilities are obtained for the Ψ class of priors. For notational simplicity, $\pi \in \Psi$ is omitted from the rest of the material. One can straightforwardly extend (2) and (3) to the three scenarios mentioned earlier. For example, replace the single observation x with the vector $(x_1, x_2, \cdots, x_n)$ or with a sufficient statistic s, if it exists, for a), replace the single observation x with the fusion decision U_o for c), etc. Also, in cases b) and c), where the random variable of interest is discrete, the densities in (2) and (3) are replaced by appropriate probabilities.

In the sequel, we assume an exponential distribution with mean θ, for the observation, and an exponential nominal prior with mean λ, for θ. Sensor decision U_i can be written as $U_i = U(X_i - t)$, where

$$U(X_i - t) = \begin{cases} 1, & X_i \geq t \\ 0, & \text{otherwise.} \end{cases}$$

In case b), the fusion center decision based on sufficient statistic

$$S = 1/n \sum_{i=1}^{n} X_i$$

is given by

$$S \underset{U_o=0}{\overset{U_o=1}{\gtrless}} h$$

where h is a threshold. Therefore, the infimum and the supremum posterior probabilities for different cases a), b), and c) can be derived [10]. The derivations are straightforward but care must be exercised in the evaluation of Z and Y. The expressions for M, Z, and Y are given in Table I.

III. NUMERICAL RESULTS

In all of the numerical results, we have assumed $\lambda = 1, \epsilon = 0.1$, and $\theta_o = 1$. Fig. 1 shows the posterior probability for different values of the sufficient statistic s. This describes the situation of a), where all sensor observations are available at the fusion center and the extrema of $P(\theta \in C|s)$ are computed. As seen in Fig. 1, both the posterior probabilities are very high when s is close to zero and are low for sufficiently large values of s. For example, if $n = 5$ and $s = 3$, then the supremum and the infimum probabilities are 0.009799 and 0.00131, respectively. That is, with the observation $s = 3$, small values for the two probabilities indicate that one can confidently decide on hypothesis H_1, in spite of the fact that the prior is not known completely. The data $s = 3$ exhibit strong posterior robustness. Also we can observe that as n tends to infinity,

$$P^{\pi(\theta|s)}(\theta \in C) = \begin{cases} 1, & s \in C \\ 0, & s \notin C. \end{cases}$$

This shows that s correctly estimates the parameter. The prior being not known completely does not entail any loss of information because we have a large number of (unquantized) observations from $f(x|\theta)$. The same inference is validated by the asymptotic analysis in Section IV.

Fig. 2 describes the situation of b), where the fusion center makes a decision U_o based on S. When $U_o = 1$, we would like the posterior probability to be as small as possible (ideally zero) and when $U_o = 0$, we would like the posterior probability to be as large as possible (ideally one). However, as seen from the figure, this is not possible. The performances of case b) and case a) can be compared by taking some specific values for the design parameters. Consider $n = 5$ and $h = 2$. If the sufficient statistic s observed is, say 2.2, the decision in case b) would be $U_o = 1$, and from Fig. 2, with $h = 2$,

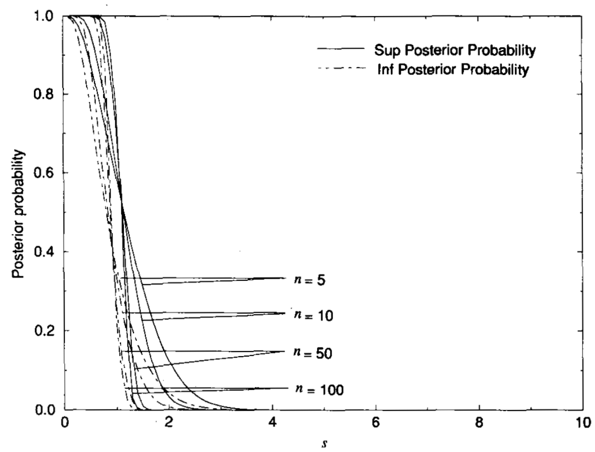

Fig. 1. Posterior probability versus sufficient statistic.

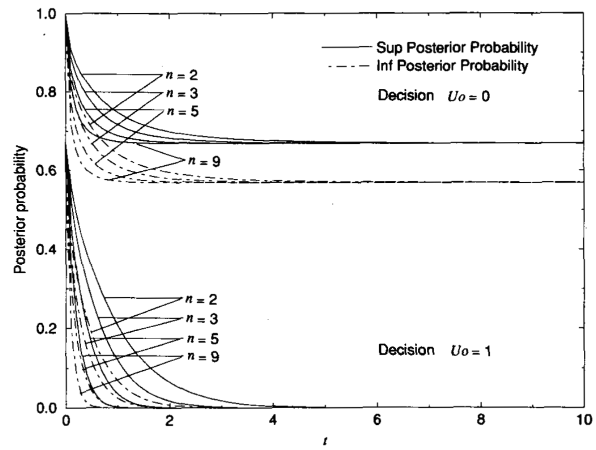

Fig. 3. Posterior probability versus sensor threshold for the AND rule.

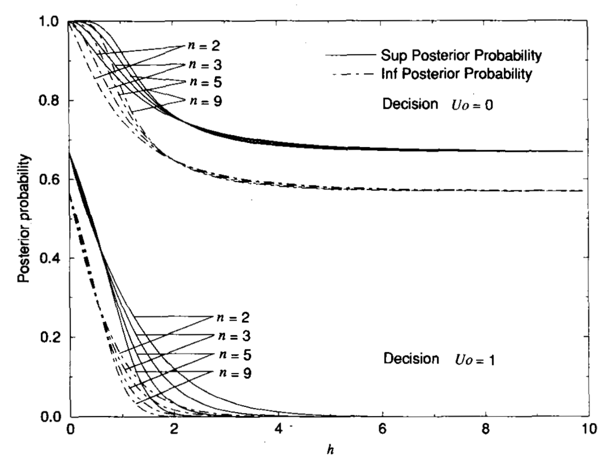

Fig. 2. Posterior probability versus threshold for case b).

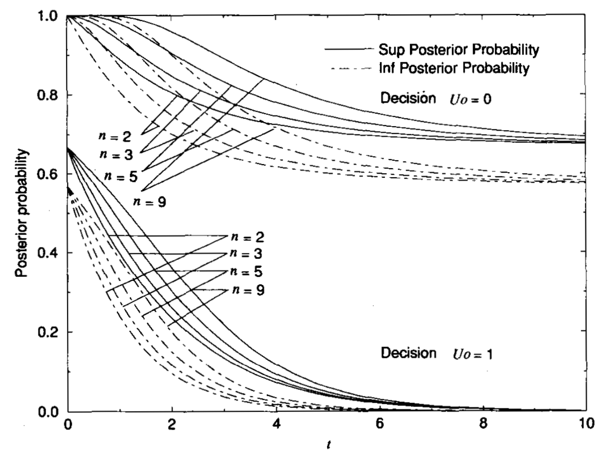

Fig. 4. Posterior probability versus sensor threshold for the OR rule.

the corresponding supremum and infimum probabilities are 0.03306 and 0.005585, respectively. For case a), with $s = 2.2$ in Fig. 1, the two corresponding probabilities are 0.08052 and 0.01785. Therefore, for a relatively weak data close to the threshold, case b) exhibits a slightly better robustness than case a). On the other hand, if $s = 4$, the two probabilities corresponding to case b) are unchanged, but the two corresponding probabilities for the central case are 0.00044 and 0.000034. Certainly with a strong data, the central scheme exhibits superior posterior robustness. The loss due to quantization is evident.

Figs. 3–5 describe the situation of c), where the fusion center makes a decision U_o based on the sensor decisions $(U_1, U_2, \cdots, U_n)$. Figs. 3–5 correspond to the AND, OR, and ML rules, respectively. For the AND rule, for both $U_o = 1$ and 0, the posterior probabilities decrease as the number of sensors increases as shown in Fig. 3. This monotonicity is proved theoretically in the Appendix . Even though, with an increasing number of sensors, the decrease of the posterior probability for $U_o = 1$ is desirable, this is achieved at the cost of decreasing probability for the $U_o = 0$ case. Recall that a higher posterior probability is desirable when the decision $U_o = 0$ is made. The effect of the loss of information due to 1-bit quantization of sensor observations on the robustness of posterior probability is evident from the figure. Also, even though a large t would achieve a desirable vanishing probability when $U_o = 1$, this would also produce a posterior probability of lowest possible value for $U_o = 0$. Therefore, as a compromise, a not too large or a not too small t value is needed. For the OR rule, the posterior probabilities increase as the number of sensors increases (Fig. 4). This behavior is opposite to that of AND. At any given t, the behavior of posterior probabilities with increasing n is desirable for $U_o = 0$ and is undesirable for $U_o = 1$. As in the AND case, a compromise t value is required for OR as well as for ML (Fig. 5) rules. A quantitative comparison of cases b) and c) has to take into account many possible choices for the values of t and h. Because of greater quantization of data in c) as compared to b), the former is, in general, less robust than the latter. In Section V we consider the choice of k in the k-out-of-n rule, from a worst case viewpoint.

IV. Asymptotic Performance

In this section, the convergence of posterior probabilities as n become large is considered. We consider only situations a) and b). The analysis of case c) is in general more involved when either the sensor threshold changes with n and the fusion rule is fixed or when the sensor threshold is fixed and the fusion rule is dependent on n [11].

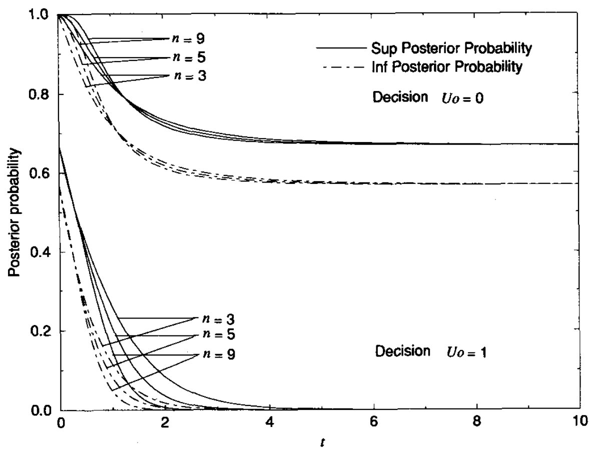

Fig. 5. Posterior probability versus sensor threshold for majority logic rule.

Case a): In this case, where the sufficient statistic is available at the fusion center and no decision is made, we see from the expression for M in Table I, that as n increases without bound

$$\lim_{n\to\infty} \frac{N}{M} = \begin{cases} 1, & s \in C \\ 0, & s \notin C. \end{cases}$$

Using this value of N/M, (2) as well as (3), and Table I, we get

$$\inf P^{\pi(\theta|s)}(\theta \in C) = \begin{cases} 1, & s \in C \\ 0, & s \notin C \end{cases} \tag{4}$$

$$\sup P^{\pi(\theta|s)}(\theta \in C) = \begin{cases} 1, & s \in C \\ 0. & s \notin C. \end{cases} \tag{5}$$

Hence, a perfect decision is possible.

Case b): Asymptotically, as $n \to \infty$, the sufficient statistic

$$S = \frac{1}{n}\sum_{i=1}^{n} X_i$$

is degenerate at $S = \theta$. As n becomes unbounded, we get

$$P(U_o = 1|\theta) = P(S > h|\theta) = \begin{cases} 1, & h < \theta \\ 0, & h > \theta \end{cases} \tag{6}$$

$$P(U_o = 0|\theta) = P(S < h|\theta) = \begin{cases} 1, & h > \theta \\ 0, & h < \theta. \end{cases} \tag{7}$$

Using (2), (3), (6), (7), and Table I, we obtain the following results [10].

Case i): Let $U_o = 1$ and $h < \theta_o$.

$$\inf P^{\pi(\theta|U_o=1)}(\theta \in C) = \frac{e^{-h/\lambda} - e^{-\theta_o/\lambda}}{e^{-h/\lambda} + \dfrac{\epsilon}{1-\epsilon}} \tag{8}$$

$$\sup P^{\pi(\theta|U_o=1)}(\theta \in C) = 1 - \frac{e^{-\theta_o/\lambda}}{e^{-h/\lambda} + \dfrac{\epsilon}{1-\epsilon}}. \tag{9}$$

If we let

$$\theta_o = 1, \quad \lambda = 1, \quad \text{and} \quad \epsilon = 0.1 \tag{10}$$

then we get

$$\inf P^{\pi(\theta|U_o=1)}(\theta \in C) = 0.5689, \quad \text{at } h = 0 \tag{11}$$

$$\sup P^{\pi(\theta|U_o=1)}(\theta \in C) = 0.6689, \quad \text{at } h = 0. \tag{12}$$

Case ii): Let $U_o = 1$ and $h > \theta_o$

$$\inf P^{\pi(\theta|U_o=1)}(\theta \in C) = 0 \tag{13}$$

$$\sup P^{\pi(\theta|U_o=1)}(\theta \in C) = 0. \tag{14}$$

Case iii): Let $U_o = 0$ and $h < \theta_o$

$$\inf P^{\pi(\theta|U_o=0)}(\theta \in C) = 1 \tag{15}$$

$$\sup P^{\pi(\theta|U_o=0)}(\theta \in C) = 1. \tag{16}$$

Case iv): Let $U_o = 0$ and $h > \theta_o$

$$\inf P^{\pi(\theta|U_o=0)}(\theta \in C) = \frac{1 - e^{-\theta_o/\lambda}}{1 - e^{-h/\lambda} + \dfrac{\epsilon}{1-\epsilon}} \tag{17}$$

$$\sup P^{\pi(\theta|U_o=0)}(\theta \in C) = 1 - \frac{e^{-\theta_o/\lambda} - e^{-h/\lambda}}{1 - e^{-h/\lambda} + \dfrac{\epsilon}{1-\epsilon}}. \tag{18}$$

For the parameters values as in (10)

$$\inf P^{\pi(\theta/U_o=0)}(\theta \in C) = 0.5689, \quad \text{at } h = \infty \tag{19}$$

$$\sup P^{\pi(\theta/U_o=0)}(\theta \in C) = 0.6689, \quad \text{at } h = \infty \tag{20}$$

Ideally we would like

$$P^{\pi(\theta|U_o=1)}(\theta \in C) = 0$$

and

$$P^{\pi(\theta|U_o=0)}(\theta \in C) = 1.$$

But this is not achieved, whatever may be the value of h. Because a decision (1-bit quantization) is made, there is a loss of information. Therefore, even with a large number of observations, one cannot compensate for the quantization loss. However, the best performance is obtained when h is chosen as a number arbitrarily close to θ_o. This makes sense because θ_o is the boundary between the hypotheses, the sufficient statistic S asymptotically estimates θ correctly, and best inference regarding the hypotheses is achieved when S is compared to θ_o.

V. Optimal k in k-out-of-n Counting Fusion Rule

It is desirable to have a high posterior probability when $U_o = 0$ and at the same time have a low posterior probability when $U_o = 1$. Keeping this in mind, let us use the following criterion. For a specific value of k in the k-out-of-n fusion rule, we place a lower bound on the infimum posterior probability when $U_o = 0$ and find the sensor decision threshold t_k that attains the lower bound. We then find the supremum posterior probability given $U_o = 1$ when the threshold equals t_k. This is repeated for all values of k. The value of k that gives the lowest of the supremum posterior probabilities given $U_o = 1$ is found and we call this optimal k among the k-out-of-n fusion rules.

The values of ϵ, λ, and θ_o are the same as in (10). The lower bound for the infimum posterior probability when $U_o = 0$ is 0.9 in Table II and is 0.8 in Table III. As we can see from Tables II and III, for the case of an exponentially distributed observation and an exponentially distributed nominal prior, the optimal k is close to or equal to 1. That is, for small values of n, the optimal rule is nearly the OR rule for this exponential example. The result in [10] for the normal example shows that the optimal rule is nearly the ML rule. In any case, given the observation density, the contamination proportion, and the nominal prior, an optimal k in the above sense can be obtained.

TABLE II
OPTIMAL k FOR DIFFERENT n

# of Sensors n	Optimal k	Threshold t_k	Supremum $P^{\pi(\theta\|U_o=1)}(\theta \in C)$
2	1	0.46999	0.52156
3	1	0.84592	0.46638
4	1	1.14025	0.43265
5	1	1.37798	0.40940
6	1	1.57619	0.39215
7	2	1.09324	0.37579
8	2	1.23436	0.36071
9	2	1.35985	0.34853

$\inf P^{\pi_o(\theta|U_o=0)}(\theta \in C)=0.9$

TABLE III
OPTIMAL k FOR DIFFERENT n

# of Sensors n	Optimal k	Threshold t_k	Supremum $P^{\pi(\theta\|U_o=1)}(\theta \in C)$
2	1	1.05653	0.38764
3	1	1.53815	0.33300
4	1	1.89395	0.29910
5	1	2.17437	0.27711
6	2	1.43915	0.25821
7	2	1.62924	0.23906
8	2	1.79415	0.22412
9	2	1.93960	0.21206

$\inf P^{\pi_o(\theta|U_o=0)}(\theta \in C)=0.8$

VI. CONCLUSIONS

A binary hypothesis decentralized detection problem with an ϵ-contamination prior is considered. The performance figures of interest are the supremum and the infimum of the posterior probability that the parameter θ is in a given set C, as the prior varies over the contaminant class. Numerical results are obtained for a specific case of exponentially distributed observation and an exponentially distributed nominal prior. The results illustrate the degree of robustness achievable with quantized observations as compared to unquantized observations. It can be expected that similar general conclusions would result from a study of other examples. It is also shown that given the observation density, the contamination proportion, and the nominal prior, the value of k in the k-out-of-n fusion rule that gives the minimum

$$\sup P^{\pi(\theta|U_o=1)}(\theta \in C)$$

when

$$\inf P^{\pi(\theta|U_o=0)}(\theta \in C)$$

is lower-bounded, can be found.

APPENDIX
VARIATION OF POSTERIOR PROBABILITY AS A FUNCTION OF THE NUMBER OF SENSORS

In this Appendix, we establish the monotonicity of the posterior probabilities with respect to n, for the AND and OR rules, for any arbitrary observation density and nominal prior density with the restriction that the family $f(x|\theta)$ is stochastically increasing. That is, the CDF's satisfy $F_{\theta_1}(x) \leq F_{\theta_2}(x)$ for all x, when $\theta_1 > \theta_2$. Examples of stochastically increasing families include i) simple exponential density, and ii) a family of $f(x|\theta)$ with θ as the location parameter. Rewriting (3) for case c) yields,

$$\sup P^{\pi(\theta|U_o)}(\theta \in C) = 1 - (1 - P_o) \cdot \left[1 + \frac{\epsilon \sup_{\theta \in C} P(U_o|\theta)}{(1-\epsilon) m(U_o|\pi_o)}\right]^{-1}. \tag{A1}$$

Let

$$f(n) = P^{\pi(\theta|U_o,n)}(\theta \in C).$$

Then

$$\sup f(n) = 1 - (1 - P_o(n))\left[1 + \frac{\epsilon H(n)}{(1-\epsilon)M(n)}\right]^{-1}$$

where

$$N(n) = \int_{-\infty}^{\theta_o} P(U_o|\theta, n)\pi_o(\theta)\, d\theta$$
$$M(n) = \int_{-\infty}^{\infty} P(U_o|\theta, n)\pi_o(\theta)\, d\theta$$
$$P_o(n) = \frac{N(n)}{M(n)}$$
$$H(n) = \sup_{\theta \in C} P(U_o|\theta, n).$$

Below we are interested in the infimum and the supremum of the posterior probability that θ is in C. For convenience, the explicit notation $\theta \in C$ is suppressed.

We observe that the following relations are equivalent:

$$\sup f(n) \gtreqless \sup f(n+1)$$

$$1 - (1 - P_o(n))\left[1 + \frac{\epsilon H(n)}{(1-\epsilon)M(n)}\right]^{-1} \gtreqless 1 - (1 - P_o(n+1))\left[1 + \frac{\epsilon H(n+1)}{(1-\epsilon)M(n+1)}\right]^{-1}$$

$$(1-\epsilon)[M(n)X(n+1) - M(n+1)X(n)] + \epsilon[H(n)X(n+1) - H(n+1)X(n)] \gtreqless 0 \tag{A2}$$

where $X(n) = M(n) - N(n)$.

Consider the specific case where the fusion rule is the AND rule and the decision is $U_o = 1$.

$$M(n) = \int_{-\infty}^{\infty} [P(X_i > t|\theta]^n \pi_o(\theta)\, d\theta,$$
$$X(n) = \int_{\overline{C}} [P(X_i > t|\theta]^n \pi_o(\theta)\, d\theta$$
$$H(n) = [P(X_i > t|\theta_1]^n$$

where $\theta_1 \in C$ is the value of θ for which the $\sup P(U_o = 1|\theta)$ is attained. The factor multiplying $(1-\epsilon)$ in (A2) is

$$\begin{aligned} M(n)X(n+1) - M(n+1)X(n) &= \int_C \int_{\overline{C}} [P(X_i > t|a]^n \pi_o(a) \\ &\quad \cdot [P(X_i > t|b)]^n \pi_o(b)[P(X_i > t|b) \\ &\quad - P(x_i > t|a)]\, db\, da \end{aligned} \tag{A3}$$

Because of the stochastically larger property of

$$f(x|\theta), P(X_i > t|b) - P(X_i > t|a)$$

and (A3) are positive. Similarly, the factor multiplying ϵ in (A2) is positive. Hence the left-hand side of (A2) is positive. That is, the sup $P^{\pi(\theta|U_o=1)}$ for the AND rule decreases as n increases. Considering $U_o = 0$

$$M(n) = \int_{-\infty}^{\infty} \{1 - [P(X_i > t|\theta)]^n\}\pi_o(\theta)\, d\theta$$

$$X(n) = \int_{\overline{C}} \{1 - [P(X_i > t|\theta)]^n\}\pi_o(\theta)\, d\theta$$

$$H(n) = 1 - [P(X_i > t|\theta_2)]^n$$

where $\theta_2 \in C$ is the value of θ for which the sup $P(U_o = 0|\theta)$ is attained. The factor multiplying $(1 - \epsilon)$ in (A2) is positive, since

$$\begin{aligned}(M(n)&X(n+1) - M(n+1)X(n)) \\ &= \int_C \int_{\overline{C}} [1 - P(X_i > t|a)]\pi_o(a) \\ &\cdot [1 - P(X_i > t|b)]\pi_o(b) \\ &\cdot \left\{ \sum_{j=0}^{n-1} [P(X_i > t|a)]^j \sum_{k=0}^{n} [P(X_i > t|b)]^k \right. \\ &\left. - \sum_{l=0}^{n} [P(X_i > t|a)]^l \sum_{m=0}^{n-1} [P(X_i > t|b)]^m \right\} db\, da \end{aligned} \tag{A4}$$

and the quantity inside the curly bracket is positive. Similarly, the factor multiplying ϵ in (A2) is positive. Therefore, the left-hand side of (A2) is positive. That is, the sup $P^{\pi(\theta|U_o=0)}$ for the AND rule decreases as n increases. Based on similar steps, we can prove that both inf $P^{\pi(\theta|U_o=1)}$ and inf $P^{\pi(\theta|U_o=0)}$ decrease with increasing n.

Similarly, for the OR rule, we can prove that

sup $P^{\pi(\theta|U_o=1)}$, sup $P^{\pi(\theta|U_o=0)}$, inf $P^{\pi(\theta|U_o=1)}$, and inf $P^{\pi(\theta|U_o=0)}$

all increase with n.

REFERENCES

[1] R. R. Tenney and N. R. Sandell, Jr., "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, pp. 501–510, July 1981.

[2] I. Y. Hoballah and P. K. Varshney, "Distributed Bayesian signal detection," *IEEE Trans. Inform. Theory*, vol. 35, pp. 995–1000, Sept. 1989.

[3] H. R. Hashemi and I. B. Rhodes, "Decentralized sequential detection," *IEEE Trans. Inform. Theory*, vol. 35, pp. 509–520, May 1989.

[4] V. V. Veeravalli, "Comments on decentralized sequential detection," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1428–1429, July 1992.

[5] J. N. Tsitsiklis, " Decentralized detection," in *Advances in Statistical Processing*, vol. 2, *Signal Detection*, H. V. Poor and J. B. Thomas, Eds. Greenwhich, CT: JAI Press, 1990.

[6] R. S. Blum and S. A. Kassam, "Optimum distributed detection of weak signals in dependent sensors," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1066–1079, May 1992.

[7] V. V. Veeravalli, T. Basar, and H. V. Poor, "Minimax robust decentralized detection," *IEEE Trans. Inform. Theory*, vol. 40, pp. 35–40, Jan. 1994.

[8] E. Geraniotis and Y. A. Chau, "Robust data fusion for multi-sensor detection systems," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1265–1279, Nov. 1990.

[9] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag, 1985.

[10] C. H. Gowda, "Robustness of decentralized tests with ϵ-contamination prior," M.S. Thesis, Dept. elec. Eng.,, Southern Illinois Univ., Carbondale, IL 62901, Aug. 1994.

[11] R. Viswanathan and V. Aalo, "On counting rules in distributed detection," *IEEE Transa. Acoust., Speech, Signal Process.*, pp. 772–775, May 1989.

Asymptotically Optimum Detection of a Weak Signal Sequence with Random Time Delays

Igor M. Arbekov

***Abstract*—The problem of designing asymptotically optimum detectors for a weak signal sequence with random time delays in the presence of a white Gaussian noise is considered. The multidimensional probability distribution of the time delays is assumed to be known. As a result of asymptotic analysis of the log-likelihood ratio, the asymptotically optimum linear or quadratic detectors and their probability distributions and efficiencies are found.**

***Index Terms*— Detection of dependent random weak signals, log-likelihood ratio, limiting probability distribution.**

"The cognitive essence of the theory of probability is opened only by the limiting theorems."

B. V. Gnedenko and A. N. Kolmogorov [7]

I. INTRODUCTION

The problem of detecting a signal with unknown time delay is one of the most important problems of statistical radioengineering and has applications in the construction of broadcasting systems, radiolocation, etc. When the signal power is much less than the noise power, the signal can be repeated a few times to improve signal reception. At the receiver, the detection of the signal can be complicated by several factors. One of them is the random time delay of each transmitted signal.

In this correspondence, we will be concerned with detection of a weak signal sequence with random dependent time delays in additive white Gaussian noise. In Section II, we will prove the theorems establishing the limiting distribution of the log-likelihood ratio for the corresponding statistical hypotheses when m (the "depth" of the dependence of the time delays) is much less than n (the length of the signal sequence, $m \ll n$). Here we will assume that the signal-to-noise ratio μ decreases to zero and the length of the signal sequence n grows to infinity, so that the following asymptotic representation of the log-likelihood ratio $\Lambda(\boldsymbol{X})$ is true:

$$\Lambda(\boldsymbol{X}) = L_n(\boldsymbol{X}) + \epsilon_n$$

where the variance of $L_n(X)$ is constant and $\epsilon_n \to 0$ in probability under both hypotheses. In this case the contiguity of the sequences of the probability measures corresponding to the statistical hypotheses takes place [5].

In addition it will be proved that the limiting distribution of $L_n(X)$ is Gaussian under both hypotheses. These results give the opportunity to determine the minimum value n under which confident (efficient) detection of a weak signal sequence with given detection errors takes place. This is considered in Section III.

The observation $x(t)$ of a deterministic signal $s(t)$ with random time delay η in additive noise $n(t)$ can be written as

$$X(t) = s(t - \eta) + n(t).$$

Manuscript received November 10, 1992; revised November 5, 1994.

The author is with the Laboratory for Control Systems Theory, Moscow University, Bldg. G, no. 101, Moscow, Russian Federation B-234.

IEEE Log Number 9411670.