Decentralized Detection With Censoring Sensors

Swaroop Appadwedula, *Member, IEEE*, Venugopal V. Veeravalli, *Fellow, IEEE*, and Douglas L. Jones, *Fellow, IEEE*

***Abstract*—In the censoring approach to decentralized detection, sensors transmit real-valued functions of their observations when "informative" and save energy by not transmitting otherwise. We address several practical issues in the design of censoring sensor networks including the joint dependence of sensor decision rules, randomization of decision strategies, and partially known distributions. In canonical decentralized detection problems involving quantization of sensor observations, joint optimization of the sensor quantizers is necessary. We show that under a send/no-send constraint on each sensor and when the fusion center has its own observations, the sensor decision rules can be determined independently. In terms of design, and particularly for adaptive systems, the independence of sensor decision rules implies that minimal communication is required. We address the uncertainty in the distribution of the observations typically encountered in practice by determining the optimal sensor decision rules and fusion rule for three formulations: a robust formulation, generalized likelihood ratio tests, and a locally optimum formulation. Examples are provided to illustrate the independence of sensor decision rules, and to evaluate the partially known formulations.**

***Index Terms*—Distributed detection, least favorable distribution, locally optimum testing, Neyman-Pearson (N-P) testing, robust hypothesis testing.**

I. Introduction

In sensor networks used for detection, geographically separated sensor nodes communicate in order to combine their observations and decide between target and null hypotheses. Decentralized detection problems consider a common fusion center for the sensor nodes where only partial observations are available for global decision-making due to power and bandwidth limitations. For a particular application of a sensor network to a real-world detection problem, an appropriate choice of communication constraints leads to the decentralized detection problem of interest. Its solution describes what partial information should be transmitted by each sensor node and how the fusion center combines the information to make a global decision.

In many detection applications, the target event happens infrequently and the null hypothesis is observed for the majority of time. For sensor networks operating on limited energy resources, an energy-efficient transmission technique would send only when the observations indicate that the target event is likely. In such applications, a communication constraint on the rate of transmission is relevant. Such a "censoring" constraint was introduced by Rago *et al.* [1].

Manuscript received February 27, 2007; revised July 8, 2007. This work was supported in part by the National Science Foundation by Grants EIA-0072043 and CCR 00-49089, through the University of Illinois. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Erchin Serpedin.

S. Appadwedula is with the Massachusetts Institute of Technology Lincoln Laboratory, Lexington, MA 02420 USA (e-mail: swaroop@ll.mit.edu).

V. V. Veeravalli and D. L. Jones are with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: vvv@uiuc.edu; dl-jones@uiuc.edu).

Digital Object Identifier 10.1109/TSP.2007.909355

Under the censoring scheme, sensors either send or do not send some real-valued function of their observations to a fusion node. The censoring scheme differs from the canonical decentralized detection problem [2], where sensors quantize their observations to one of D levels and transmit. The censoring formulation accurately addresses the energy consumption in a scenario where one hypothesis is more likely, so frequent communication is not necessary. In transmission schemes where packets of data are transmitted at a time, it may be possible to accurately represent the local decisions to the precision of the computations with one packet. Quantization may be less relevant in light of the overhead of synchronizing the sensor transmitter and the fusion center receiver, as well as bandwidth needed to send side information (e.g., time, location) that may be necessary to synchronize the sensor data.

The main result of [1] for the censoring formulation is that under a fixed rate of communication, it is optimal [in the Neyman-Pearson (N-P) and Bayesian sense] for the sensors to compute and censor their likelihood ratios in a single interval. In general, the censoring intervals must be determined by joint optimization over all the sensors. In practice, it may be costly or infeasible to perform such a joint optimization. When the fusion center has its own observations, and the rate of transmission under one of the hypotheses is constrained, we show that the censoring intervals can be determined (independently) by the transmit-rate constraint at each sensor. Eliminating the joint optimization of censoring regions has important consequences for adaptive systems and when the distribution of the observations is partially known.

It was shown in [1] and clarified in [3] that the censoring results also apply to Ali-Silvey (A-S) distance metrics [4]. In [5], we identified particular A-S problems where the joint optimization simplifies. A-S distance metrics appear to be a convenient way to simplify the joint optimization problem. In particular, the Kullback-Liebler (K-L) divergence [6, p. 309] and Chernoff distance [7] are relevant for detection with a large number of observations. In [8], asymptotic results for censoring sensor nodes are derived by examining the Chernoff error exponent. Linear programming arguments show that it is optimal for sensor nodes to choose between one of two policies.

In most applications and particularly in decentralized detection applications, the conditional distributions of the observations are neither fixed, nor completely known. For adaptive systems, where the null and target statistics are updated over time, independence of censoring intervals implies that sensors need not communicate to update their censoring intervals as long as they inform the fusion center. While we do not consider adaptive systems in detail here, such approaches have significant promise in outperforming tests based on fixed distributions.

As an alternative to adaptive tests, it may be possible in some settings to define a class of candidate distributions under each hypothesis from which the true distribution is drawn. For example, a database of target statistics may be available, and the

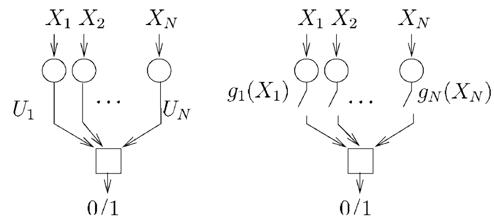


Fig. 1. Setup of: (a) canonical; and (b) censoring decentralized detection problems.

null statistics may be described accurately with standard models (up to a parameter). We consider three formulations of such partially known distribution problems in decentralized detection: a robust formulation, generalized likelihood-ratio tests (GLRTs), and a locally optimum formulation. For the robust formulation and GLRTs, we make use of the independence of the censoring intervals to determine the optimal censoring strategy.

In Section II, we describe the censoring formulation of the decentralized detection problem in the Bayesian setting, give the main result from [1], and introduce randomization. In Section II-B, we show how to eliminate the joint optimization over the sensors. Then, in Section III, we introduce composite formulations of the censoring problem. Finally, in Section IV, we discuss N–P problems.

II. Censoring Sensors

In the target detection problem, the objective is to determine the true state of nature H as being H_0 (null or target-absent hypothesis) or H_1 (target-present hypothesis), given sensor observations and their conditional distributions. In the decentralized setting, partial information from N observations about the state of nature is available for decision making. Consider the parallel topology of the decentralized detection problem in which sensors labeled $i = 1, \ldots, N$, transmit local information to a fusion center for global decision making.

In the canonical decentralized detection problem [Fig. 1(a)], each sensor i maps its real-valued observations X_i to one of D_i levels. The fusion center receives the quantizer outputs $\mathbf{U} = (U_1, \ldots, U_N)$ from the sensors and performs a likelihood ratio test to obtain the global decision.

In the censoring scenario [Fig. 1(b)], the sensors either send some real-valued function $g_i(X_i)$ of their observations X_i when in some informative region R_i or send nothing when in the uninformative region R_i^c, subject to a constraint on the send rate. We define the sensor decision rule as

$$\phi_i(X_i) = \begin{cases} g_i(X_i), & g_i(X_i) \in R_i \\ \rho(R_i^c), & g_i(X_i) \in R_i^c \end{cases} \tag{1}$$

where $\rho(R_i)$ is the real value attributed to the censoring region. The fusion rule ϕ_0 is a binary-valued function of $\phi_i(X_i), i = 1, \ldots, N$. The censoring rules and fusion rule are collectively known as the decision strategy ϕ.

Censoring is an effective communication strategy particularly when one of the hypotheses (say H_0) is significantly more likely. It is appropriate, then, to consider a constraint on transmission under H_0 only.

$$\sum_{i=1}^{N} \mathsf{P}(g_i(X_i) \in R_i \,|\, H_0) \leq \kappa \tag{2}$$

where $\mathsf{P}(\cdot|H_0)$ denotes the conditional probability under H_0, R_i is the send region, and $\kappa < N$ is the communication rate per observation for the N sensors combined. When prior probabilities $\pi_j = \mathsf{P}(H_j), j = 0, 1$ are available, transmission under H_j weighted by π_j can be considered [1].

The following conditional-independence assumption is critical to the development of sensor decision rules based on likelihood ratios. When the observations are dependent, decentralized detection problems are less tractable and complexity of design is an issue [9], [10].

Assumption 2.1 (Conditional Independence): The sensor observations are statistically independent, conditioned on each hypothesis.

We describe our notation for the hypothesis-testing problem. Let $P_j^{(i)}$ be the distribution function of X_i under H_j, $j = 0, 1$ and $p_j^{(i)}(x_i)$ the corresponding probability density function (pdf). We denote the joint distribution function over the observations $\mathbf{X} = (X_1, \ldots, X_N)$ as $P_j = P_j^{(1)} \times P_j^{(2)} \times \ldots \times P_j^{(N)}$. We denote the miss probability as P_M, and the false-alarm probability as P_F, where (by an abuse of notation) $P_M = 1 - E[\phi_0 \,|\, H_1], P_F = E[\phi_0 \,|\, H_0]$, and $E[\cdot|H_j], j = 0, 1$, is the expectation operator under H_j. We consider the error probability $P_E = \pi_0 P_F + \pi_1 P_M$ setting in detail. We use the convenient notation $P_j^{(i)}(R_i)$ and $P_j^{(i)}(R_i^c), j = 0, 1, i = 1, \ldots, N$, to represent the probability of send and no-send under H_j.

We show how the assumption of conditional independence leads to censored likelihood-ratio tests. For simplicity, consider combinations of send/no-send for each sensor when there are $N = 2$ sensors. Extension of the result to a larger number of sensors is straightforward. The error probability can be written as

$$\begin{aligned} P_E &= \pi_1 + \pi_0 P(\text{say } H_1 \,|\, H_0) - \pi_1 P(\text{say } H_1 \,|\, H_1) \\ &= \pi_1 + \int \phi_0(\phi_1(x_1), \phi_2(x_2)) \\ &\quad \times \left[\pi_0 p_0^{(1)}(x_1) p_0^{(2)}(x_2) - \pi_1 p_1^{(1)}(x_1) p_1^{(2)}(x_2)\right] dx_1 dx_2. \end{aligned} \tag{3}$$

Breaking up the integral into send and no send regions for each sensor node, we obtain for the four terms

$$\begin{aligned} P_E = \pi_1 &+ \int_{g_1(x_1) \in R_1, g_2(x_2) \in R_2} \phi_0(\phi_1(x_1), \phi_2(x_2)) \\ &\times \left[\pi_0\, p_0^{(1)}(x_1) p_0^{(2)}(x_2) - \pi_1 p_1^{(1)}(x_1) p_1^{(2)}(x_2)\right] \, dx_1 \, dx_2 \\ &+ \int_{g_1(x_1) \in R_1} \phi_0\left(\phi_1(x_1), R_2^c\right) \\ &\times \left[\pi_0\, p_0^{(1)}(x_1) P_0^{(2)}\left(R_2^c\right) - \pi_1\, p_1^{(1)}(x_1) P_1^{(2)}\left(R_2^c\right)\right] \, dx_1 \\ &+ \int_{g_2(x_2) \in R_2} \phi_0\left(R_1^c, \phi_2(x_2)\right) \\ &\times \left[\pi_0\, P_0^{(1)}\left(R_1^c\right) p_0^{(2)}(x_2) - \pi_1\, P_1^{(1)}\left(R_1^c\right) p_1^{(2)}(x_2)\right] \, dx_2 \\ &+ \phi_0\left(R_1^c, R_2^c\right) \Big[\pi_0\, P_0^{(1)}\left(R_1^c\right) P_0^{(2)}\left(R_2^c\right) \\ &\qquad - \pi_1\, P_1^{(1)}\left(R_1^c\right) P_1^{(2)}\left(R_2^c\right)\Big] \end{aligned} \tag{4}$$

where the dependency of ϕ_0 on the censoring operation is expressed explicitly. Examining the bracketed term in each of the four integrals of (4), we see that a likelihood rule in the send region and a fixed ratio of probabilities in the no-send region minimizes the error probability.

Without loss of generality, we choose $g_i(X_i)$ at every sensor node to be the likelihood ratio defined as $l(X_i) = (p_1^i(X_i))/(p_0^i(X_i))$. In Section II-B, the choice of g_i as the likelihood ratio leads to a simple description of R_i.

We consolidate the description of the optimal decision strategy by defining the censored likelihood ratio at the sensor nodes as

$$\phi_i(X_i) = \begin{cases} l(X_i), & l(X_i) \in R_i \\ \rho_i, & l(X_i) \in R_i^c \end{cases} \tag{5}$$

where

$$\rho_i = \frac{P_1^{(i)}\left(l(X_i) \in R_i^c\right)}{P_0^{(i)}\left(l(X_i) \in R_i^c\right)}. \tag{6}$$

Given send regions R_i and the choice $g_i(X_i) = l(X_i)$, the optimal fusion rule for the Bayesian problem $(\min P_E)$ under the communication constraint (2) is

$$\phi_0\left(\phi_i(X_i)|_{i=1}^N\right) = \begin{cases} 1, & \prod_{i=1}^N \phi_i(X_i) \geq \tau \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

It is optimal for sensor nodes to send the likelihood ratio between the distributions $P_1^{(i)}$ and $P_0^{(i)}$ in the send region and for the fusion center to evaluate the likelihood of no-send region (from knowledge of R_i and $P_j^{(i)}(X_i), j = 0, 1$).

In the Bayesian problem, randomization of the test (7) for the event $\phi_0(\phi_i(X_i)|_{i=1}^N) = \tau$ (as in a N-P problem [11]) can be eliminated by introducing the following assumption.

Assumption 2.2 (No Point Mass): For each sensor i, the likelihood $l(X_i)$ has no point mass under either hypothesis

$$P_j^{(i)}(l(X_i) = t) = 0 \quad \text{for } t \in [0, \infty), \quad j = 0, 1. \tag{8}$$

The censoring regions or alternatively the censoring functions $g_i(x_i)$ may need randomization due to the communication-rate constraint.

A. Randomization

In centralized detection problems with constraints, it is well known that randomization over decision rules can improve performance over deterministic rules (e.g., N-P Lemma [11, pp. 23-25]). Under a censoring constraint, randomization over the choice of censoring regions can improve performance not only in the N-P sense, but also in the Bayesian sense.

We define the set of deterministic, independently randomized and dependently randomized decision strategies for the censoring problem [10, p. 301]. Let Φ_i be the set of all deterministic sensor rules for sensor i, where a particular deterministic sensor rule is the fixed choice of censoring function g_i and censoring region R_i^c for all time. Let Φ_0 be the set of all deterministic fusion rules. Then, we can describe the set of all decision strategies as $\Phi = \Phi_0 \times \Phi_1 \times \Phi_2 \times \cdots \Phi_N$. Let $\bar{\Phi}$ be the set of independently randomized decision strategies in the sense that each sensor has a finite set of candidate sensor rules $\{\phi_i^{(k)}\}_{k=1}^{K_i}$ and chooses to use rule $\phi_i^{(k)}$ for a fraction of the time $\gamma_i^{(k)}$, where $\gamma_i^{(k)}$ are statistically independent across i. Let $\hat{\Phi}$ be the set of dependently randomized decision strategies in the same sense as for $\bar{\Phi}$, except that $\gamma_i^{(k)}$ may be dependent across sensors. It is easy to see that $\Phi \subset \bar{\Phi} \subset \hat{\Phi}$, so performance can not degrade when going from the optimal deterministic to the optimal independently randomized to the optimal dependently randomized strategy. For the randomized decision strategies, the false-alarm, detection, and transmit probabilities are determined by the expected value over the candidate rules.

Over the set of randomized decision strategies, the maximum number of deterministic sensor rules or decision strategies over which to randomize can be determined by simple linear programming arguments, as shown in [12]. (Results given in [13, pp. 65-67] can also be applied.)

Theorem 2.1: Given a communication-rate constraint to be satisfied,

a) Over the set $\bar{\Phi}$, the decision strategy minimizing P_E randomizes between at most two deterministic sensor decision rules $\phi_i^{(k)}, k = 1, 2$, at each sensor where $\phi_i^{(k)}$ is given by (5).

b) Over the set $\hat{\Phi}$, the decision strategy which minimizes P_E randomizes between at most two deterministic decision strategies $\phi^{(k)}, k = 1, 2$, where $\phi^{(k)}$ are as given in (7) and (5).

The optimality of randomizing between two candidate strategies in the censoring problem is analogous to [10, Proposition 3.7] for the canonical decentralized detection problem.

B. Determining the Censoring Regions

Conveniently, the choice of censoring regions simplifies considerably for the communication rate constraint (2). The first part of the simplification, which applies to deterministic censoring regions, is due to Rago *et al.*, and requires the assumption of no point mass.

Under Assumption 2.2, and given a communication-rate constraint to be satisfied, at each sensor, consolidation of no-send intervals into a single no-send interval does not increase the error probability, Ali-Silvey distance, or miss probability (subject to a false-alarm rate constraint) [1]. The no-send interval can then be described as

$$R_i^c = \{l(X_i) : t_{1,i} < l(X_i) \leq t_{2,i}\}. \tag{9}$$

The approach for consolidating no-send regions into a single interval is as follows. Consider a no-send region consisting of two intervals. Under Assumption 2.2, it is always possible to choose a single interval which lies within the outer endpoints of the original two intervals while preserving $P_j^{(i)}(l(X_i) \in R_i^c), j = 0, 1$. In [1], it is shown that the error probability does not decrease, and the Ali-Silvey distance does not increase for such a consolidation.

Determining the optimal deterministic decision strategy is a $(2N+1)$-dimensional optimization problem over two thresholds at each sensor and the fusion-center threshold. In [1], it was found that the lower threshold of the censoring region at each sensor goes to zero for sufficiently small communication rates (and false-alarm rates in the N-P problem). Here, we derive a stronger result, the second part of the simplification, which does not involve asymptotics.

In many applications, it is appropriate to consider a fusion center that has its own observations. One can imagine the situation where sensor nodes are organized into clusters (geographically, perhaps) with a fusion node located within each cluster. For robustness and to share the generally costly role of fusion [14], the role of fusion may be rotated among identical sensor nodes.

Definition 2.1 (Zero Lower Thresholds): Define the set of deterministic decision strategies $\Phi^\star \subset \Phi$ as those strategies with fusion rule (7) and sensor decision rules (5) where the censoring regions have zero lower thresholds

$$R_i^c = \{0 \le l(X_i) \le t_{2,i}\}, \quad \forall i. \tag{10}$$

Consider a communication-rate constraint on each sensor

$$P_0^{(i)}(l(X_i) \in R_i) = \kappa_i \tag{11}$$

where $\kappa_i < 1$ is the send rate for sensor i. Then we have the following result.

Lemma 2.1: Suppose the fusion node has its own observations. Under Assumption 2.2, and given the communication-rate constraint (11) to be satisfied, for any decision strategy $\phi \in \Phi, \phi \notin \Phi^\star$, there exists a decision strategy $\phi^\star \in \Phi^\star$ with lower error probability (i.e., $P_E(\phi^\star) < P_E(\phi)$).

Proof Outline: A fusion node which has its own observations can be considered as an uncensored sensor node, so that the censoring formulation requires no modification. Based on (4), we concluded that the censoring rule and fusion rule which minimize P_E are given by (5) and (7), respectively. Rago's result further simplifies the choice of censoring regions (9). Now, consider the change in P_E as the censoring regions of type (9) are modified to censoring regions of type (10) for a particular sensor k.

The fusion threshold is fixed $\tau = (\pi_0/\pi_1)$ for any choice of censoring regions, and the upper threshold of the censoring region $t_{2,k}$ is completely determined by $t_{1,k}$ and κ_k (11), so we compute $(\partial P_E)/(\partial t_{1,k})$. Due to Assumption 2.2 and the fusion center taking observations, the $(\partial P_E)/(\partial t_{1,k})$ is defined for all $t_{1,k}$. We find that $(\partial P_E)/(\partial t_{1,k})$ is strictly positive for all $t_{1,k}$, so P_E is monotone increasing in $t_{1,k}$. Therefore, it is optimal to choose $t_{1,i} = 0, \forall i$. The details of the proof are given in the Appendix. ∎

Lemma 2.1 is particularly useful in terms of design. A communication-rate constraint on each sensor is easily motivated from the point of view of energy consumption. In a distributed sensor network, each sensor will have its own battery or solar-cell and may be forced to maintain its own resources. Under the constraint (11), it is clear that $t_{2,i}$ can be chosen directly from κ_i. This implies that the joint optimization of censoring regions is eliminated when a communication-rate constraint at each sensor is considered. Each sensor can then adjust its threshold independently based on its resource constraint without affecting the other sensors, as long as the fusion center is informed of the change.

For the communication-rate constraint across all sensor nodes (2), Lemma 2.1 can be used to derive the result.

Theorem 2.2: Suppose the fusion node has its own observations. Under Assumption 2.2, and given the communication-rate constraint (2) to be satisfied, for any decision strategy $\phi \in \Phi, \phi \notin \Phi^\star$, there exists a decision strategy $\phi^\star \in \Phi^\star$ with lower error probability (i.e., $P_E(\phi^\star) < P_E(\phi)$).

Proof: The optimal decision strategy under (2) chooses some value $\kappa_i^\star$ for $P_0^{(i)}(R_i), i = 1, \ldots, N$. Since Lemma 2.1 holds for any choice of $\kappa_i, i = 1, \ldots, N$, it holds for the choice $\kappa_i^\star, \sum_{i=1}^N \kappa_i^\star = \kappa$. ∎

For the problem of minimizing P_E over the set $\bar{\Phi}$ or $\hat{\Phi}$, we can restrict attention to randomization over decision strategies from $\Phi^\star$. Under the communication-rate constraint (2), independent randomization between a transmit rate of $\kappa_i^{(1)}$ and $\kappa_i^{(2)}$ at each sensor is optimal, and dependent randomization between the set of transmit rates $[\kappa_1, \ldots, \kappa_N]^{(k)}, k = 1, 2$ is optimal. Under the communication-rate constraint (11), randomization over N sensor rules or N decision strategies is optimal.

C. Gaussian Example

In this example, we compare decision strategies with censoring intervals that have nonzero lower thresholds against those that have lower thresholds equal to zero. Both the cases of the fusion center taking observations and not taking observations are considered.

Consider the problem of detecting a mean-shift in Gaussian noise with two sensors

$$H_0 : X_i \sim \mathcal{N}(0, \sigma^2) \text{ versus } H_1 : X_i \sim \mathcal{N}(\theta, \sigma^2) \tag{12}$$

where $\theta > 0$ is the mean shift and σ^2 is the variance of the noise. Since the likelihood ratio is monotone in the observation, censoring the observations X_i in a single interval is equivalent to censoring the likelihood ratio.

Fig. 2 shows the error probability P_E as a function of π_0 for identical communication constraints ($\kappa_1 = \kappa_2 = 0.1, \theta = 1$, and $\sigma = 1$). The observations at the fusion center have the same distribution as at the other sensor nodes. We determined the minimum number of trials to estimate P_F (or P_M) by requiring that the estimate be within a factor ϵ of the true value with probability β at $P_F = \alpha$. Treating P_F as a binomial random variable to be estimates, and using the Gaussian approximation, which is valid for a large number of trials, we find that at least $[(1-\alpha)/(\alpha\epsilon^2)](Q^{-1}((1-\beta)/2))^2$ trials are needed, where Q is the complementary cdf of a $\mathcal{N}(0,1)$ random variable. Five-hundred-thousand Monte Carlo trials were simulated to ensure that P_F (or P_M) is within 10% of its true value with 95% probability at $P_F = 0.001$ (or $P_M = 0.001$). For $P_F < 0.001$ (or $P_M < 0.001$), the accuracy of the estimate degrades.

When π_0 is large, the performance of the censoring scheme is relatively close to the performance of centralized detection, where complete observations are available. Censoring is an effective approach for reducing communication when H_0 is very likely. From Fig. 2(a), we observe that censoring intervals with lower threshold equal to zero are optimal when $\pi_0 > 0.37$, a meaningful regime of operation. When the fusion center has its own observations, Fig. 2(b) shows that intervals with lower threshold equal to zero are optimal over the entire range of π_0.

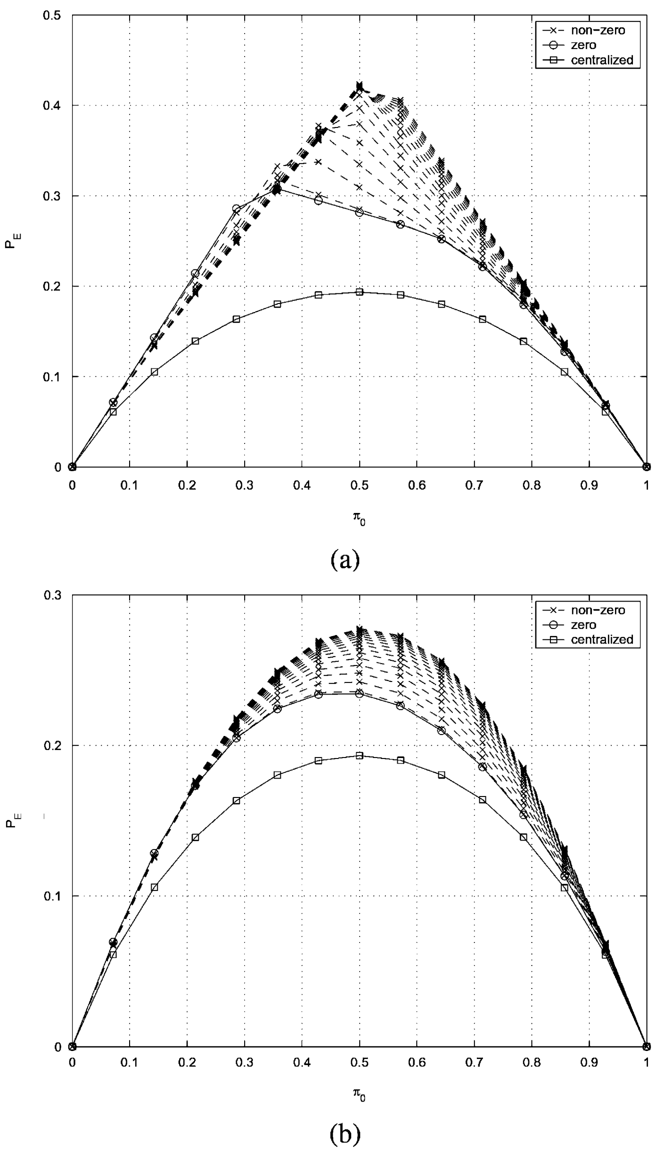


Fig. 2. Error probability as a function of the prior for the problem of mean shift in identically distributed (i.i.d.) Gaussian noise without (a) and with (b) fusion center observations. Without fusion center observations, nonzero lower thresholds at each sensor node achieve minimum error probability. With fusion center observations, using nonzero thresholds at each sensor node does not decrease the error probability over zero thresholds.

III. Composite Problems

We consider three formulations of composite testing problems in decentralized detection. In the robust formulation, the class of distributions under each hypothesis has an associated stochastic ordering. In the locally optimum formulation and for generalized likelihood-ratio tests, the conditional distributions are parametrized by a single parameter which belongs to some range of values. The locally optimum formulation is particularly suited to constant-false alarm rate (CFAR) adaptive approaches, where the parameter of the null distribution is estimated on-line.

A. Robust Formulation

Huber considered composite problems by defining a stochastic ordering of the distributions within each class [15]. He found that in a minimax sense, it is optimal to design for a certain pair of least-favorable distributions (LFDs). We seek to address the uncertainty in the distributions at each sensor by following Huber's approach. In the censoring problem, the uncertainty in the distributions leads to the consideration of a worst-case communication-rate constraint.

Consider the robust hypothesis-testing problem [15]

$$\begin{aligned} H_0 &: P_0 \in \mathcal{P}_0 \\ H_1 &: P_1 \in \mathcal{P}_1 \end{aligned}$$

where $\mathcal{P}_0$ and $\mathcal{P}_1$ are disjoint classes of distributions under H_0 and H_1, respectively. The distribution of each observation may come from a different family (i.e., $X_i \sim P_j^{(i)}, P_j^{(i)} \in \mathcal{P}_j^{(i)}$). Under hypothesis $H_j, j = 0, 1$ the joint distribution of the observations then comes from the class of distributions $\mathcal{P}_j \triangleq \mathcal{P}_j^{(1)} \times \ldots \times \mathcal{P}_j^{(N)}$.

In order to formulate the robust decentralized detection problem for censoring sensors, the communication constraint needs to be considered over the uncertainty classes.

$$\begin{aligned} \text{RB1} : \inf_{\phi} \sup_{P_0 \in \mathcal{P}_0, P_1 \in \mathcal{P}_1} & P_E(P_0, P_1, \phi) \\ \text{s.t.} \sup_{P_0 \in \mathcal{P}_0} & P_T(P_0, \phi) \le \kappa \end{aligned}$$

where $P_E(P_0, P_1, \phi) = \pi_0 P_F(P_0, \phi) + \pi_1 P_M(P_1, \phi)$ and $P_T(P_0, \phi) = \sum_{i=1}^N P_0(g_i(X_i) \in R_i)$ is the communication-rate for decision strategy ϕ with censoring regions R_i and censoring function g_i.

1) Background: To provide background for the solution to the robust problem RB1, we describe the results of [15] and [16]. Huber considered three different minimax detection criteria for the centralized problem and showed in each case that designing for LFDs is optimal for conditionally independent identically distributed observations. In [16], it was shown that designing for LFDs is optimal in the minimax sense for the canonical decentralized detection problem (Fig. 1) with conditionally independent but not necessarily identically distributed observations. We follow the approach in [16].

We begin by defining the property of joint stochastic boundedness (JSB), a fundamental property of certain uncertainty classes that enables the robust problem to be simplified.

Definition 3.1 (JSB [16]): A pair $(\mathcal{P}_0^{(i)}, \mathcal{P}_1^{(i)})$ of classes is said to be jointly stochastically bounded by $(Q_0^{(i)}, Q_1^{(i)})$ if there exists $(Q_0^{(i)}, Q_1^{(i)}) \in \mathcal{P}_0^{(i)} \times \mathcal{P}_1^{(i)}$ such that for any $(P_0^{(i)}, P_1^{(i)}) \in \mathcal{P}_0^{(i)} \times \mathcal{P}_1^{(i)}$ and all $t \ge 0$

$$\begin{aligned} &P_0^{(i)}(l_{q,i}(X_i) > t) \le Q_0^{(i)}(l_{q,i}(X_i) > t), \quad \text{and} \\ &P_1^{(i)}(l_{q,i}(X_i) \le t) \le Q_1^{(i)}(l_{q,i}(X_i) \le t) \end{aligned}$$

where $l_{q,i}$ is the likelihood ratio between $Q_1^{(i)}$ and $Q_0^{(i)}$.

In terms of a binary hypothesis-testing problem, the JSB property implies that the miss probability under $P_1^{(i)}$ and the false-alarm probability under $P_0^{(i)}$ are smaller than the corresponding probabilities under $Q_0^{(i)}$ and $Q_1^{(i)}$ when the test is $l_{q,i}(X_i)$.

The following lemma describes JSB for the set of observations $\mathbf{X}$ in terms of JSB for each observation X_i. For a proof of the lemma, refer to [16].

Lemma 3.1 ([16]): For each $i, i = 1, \ldots, N$, let the pair $(\mathcal{P}_0^{(i)}, \mathcal{P}_1^{(i)})$ be jointly stochastically bounded by $(Q_0^{(i)}, Q_1^{(i)})$. Then the pair $(\mathcal{P}_0, \mathcal{P}_1)$ is jointly stochastically bounded by (Q_0, Q_1).

In order to apply Lemma 3.1 to the decentralized problem, the JSB property needs to be extended to censored tests. To this end, consider the censoring function at each sensor

$$\phi_{q,i}(X_i) = \begin{cases} l_{q,i}(X_i), & l_{q,i}(X_i) \in R_i \\ \rho_i, & \text{otherwise} \end{cases} \tag{13}$$

Since ρ_i is the average of $l_{q,i}(X_i)$ under $Q_0^{(i)}$ in the interval R_i, we can see that $t_{1,i} < \rho_i < t_{2,i}$. So, $\phi_{q,i}(X_i)$ is nondecreasing in $l_{q,i}(X_i)$ and the event $\{\phi_{q,i}(X_i) > t\}, t \geq 0$ is equivalent to $\{l_{q,i}(X_i) > t'\}$ for an appropriate choice of t'. Therefore, the JSB property applies to the censoring test $l_{q,FC} = \prod_{i=1}^N \phi_{q,i}(X_i)$ as well.

$$P_0(l_{q,FC} > t) \leq Q_0(l_{q,FC} > t) \tag{14}$$

$$P_1(l_{q,FC} \leq t) \leq Q_1(l_{q,FC} \leq t) \tag{15}$$

2) Censoring Sensors: As in [15] and [16], our goal is to show that the robust problem RB1 has the same solution as the simpler problem

$$\begin{aligned} \text{SB1}: &\inf_{\phi} P_E(Q_0, Q_1, \phi) \\ &\text{s.t. } P_T(Q_0, \phi) \leq \kappa, \end{aligned}$$

where $Q_0 \in \mathcal{P}_0$ and $Q_1 \in \mathcal{P}_1$ are the least-favorable distributions in the sense of Defn. 3.1. Problem SB1 is just the Bayesian censoring problem for testing Q_1 versus Q_0 described in Section II.

By applying Theorem 2.2, it is possible to establish the equivalence of problems RB1 and SB1.

Theorem 3.1: Suppose the fusion node has its own observations. Under Assumption 2.2, the robust Bayesian problem of minimizing the worst-case error-probability subject to a worst-case transmission constraint (i.e., problem RB1) can be solved by designing for the least-favorable distributions in the sense of Defn. 3.1 (i.e., problem SB1).

Proof: Let $\phi^\dagger$ be the optimal decision strategy for SB1, then it must be a decision strategy with a fusion test based on $l_{q,FC}$. Applying (14) and (15), we obtain

$$\begin{aligned} \sup_{P_0 \in \mathcal{P}_0} P_F(P_0, \phi^\dagger) &= P_F(Q_0, \phi^\dagger) \\ \sup_{P_1 \in \mathcal{P}_1} P_M(P_1, \phi^\dagger) &= P_M(Q_1, \phi^\dagger) \end{aligned} \tag{16}$$

On the other hand, using simple arguments about supremum over the uncertainty class and infimum over the decision strategies, we can see that

$$\begin{aligned} P_E(Q_0, Q_1, \phi^\dagger) &\leq P_E(Q_0, Q_1, \phi), \forall\, \phi, \text{s.t.} P_T(Q_0, \phi) \leq \kappa \\ &\leq \sup_{P_0 \in \mathcal{P}_0, P_1 \in \mathcal{P}_1} P_E(P_0, P_1, \phi), \forall\, \phi, \\ &\quad \text{s.t. } \sup_{P_0 \in \mathcal{P}_0} P_T(P_0, \phi) \leq \kappa \end{aligned} \tag{17}$$

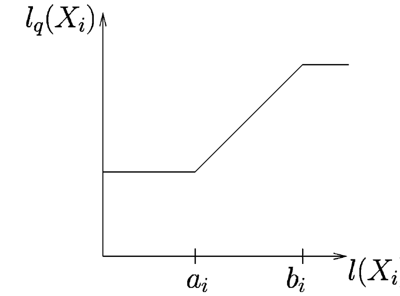


Fig. 3. Robust test for mixture distributions shows clipping at a small level a_i and large level b_i.

where in the last inequality the JSB property (14) can be applied to the communication-rate constraint $\sum_{i=1}^N Q_0(t_{1,i} < l_{q,i}(X_i) \leq t_{2,i}) \leq \kappa$, since the censoring intervals have a lower threshold equal to zero. From (16), we can argue

$$\begin{aligned} \sup_{P_0 \in \mathcal{P}_0, P_1 \in \mathcal{P}_1} P_E(P_0, P_1, \phi^\dagger) &\geq \inf_{\phi} \sup_{P_0 \in \mathcal{P}_0, P_1 \in \mathcal{P}_1} \\ &\times P_E(P_0, P_1, \phi), \forall\, \phi, \text{ s.t. } \sup_{P_0 \in \mathcal{P}_0} P_T(P_0, \phi) \leq \kappa. \end{aligned} \tag{18}$$

By taking the infimum of the right-hand side of (17) we see that $P_E(Q_0, Q_1, \phi^\dagger)$ is sandwiched between the same term on the right-hand side of (18). So, $\phi^\dagger$ is the saddle-point solution to RB1. ■

3) Example: ϵ-Contamination Class: We apply the results of the robust Bayesian problem to the problem of testing between two mixture classes, as considered in [15]. Consider the mixture distribution $P_j^{(i)}$ at sensor i under hypothesis H_j

$$\mathcal{P}_j^{(i)} = \left\{ P_j^{(i)} = (1 - \epsilon_{j,i}) G_j^{(i)} + \epsilon_{j,i} H_j^{(i)},\ H_j^{(i)} \in \mathcal{H}^{(i)} \right\} \tag{19}$$

where $G_j^{(i)}$ is the nominal distribution, $H_j^{(i)}$ is the contaminating distribution from an arbitrary class of distributions $\mathcal{H}^{(i)}$, and $\epsilon_{j,i}$ is the contaminating factor.

It was found in [15] that the optimal robust test for the centralized problem clips the likelihood ratio $l_i(X_i) = (dG_1^{(i)}(X_i))/(dG_0^{(i)}(X_i))$ produced by the nominal distributions at a small value a_i due to uncertainty in H_1 and a large value b_i due to uncertainty in H_0. The clipping levels are related to the contamination levels according to

$$a_i\, G_0(l_i(X_i) \leq a_i) - G_1(l_i(X_i) \leq a_i) = \frac{\epsilon_{1,i}}{1 - \epsilon_{1,i}} \tag{20}$$

$$G_0(l_i(X_i) < b_i) + \frac{1}{b_i} G_1(l_i(X_i) \geq b_i) = \frac{1}{1 - \epsilon_{0,i}}. \tag{21}$$

Fig. 3 shows the likelihood ratio $l_{q,i}(X_i)$ of the LFD as a function of the likelihood ratio of the nominal distribution $l_i(X_i)$.

Recognizing that censoring tests and robust tests are both clipping tests, we make the important observation that in some cases, there is no loss in performance due to censoring. For the censoring problem, the censored robust test clips small values of the likelihood ratio based on the communication-rate constraint (i.e., $\sum_{i=1}^{N} G_0(l_i(X_i) \leq t_{2,i}) \leq \kappa$). When the communication-rate constraint is related to the contamination in H_0 so that $t_{2_i} \leq a_i$, the censored robust test is identical to the robust test.

B. Generalized Likelihood-Ratio Tests

The GLRT is a suboptimal technique often employed when the value of the parameter which characterizes the distribution of the observations is unknown. Given the composite hypothesis-testing problem

$$H_0 : \theta \in \Theta_0 \text{ vs. } H_1 : \theta \in \Theta_1$$

where $\Theta_j, j = 0, 1$, is some (disjoint) partitioning of the parameter space, the generalized likelihood ratio is

$$\frac{\max_{\theta \in \Theta_1} p_\theta(\mathbf{X})}{\max_{\theta \in \Theta_0} p_\theta(\mathbf{X})} \tag{22}$$

which is compared to some threshold.

For decentralized detection problems, it may be reasonable to perform a GLRT at each individual sensor and some associated fusion test. However, it is unclear how to determine the censoring regions in such a proposed scheme since the choice of censoring regions and likelihood depends on the true value of θ under $H_j,\ j = 0, 1$. Furthermore, at the fusion center, a reliable estimate of the censored likelihood is necessary for the test to perform well.

For some distribution classes, it is possible to find censoring regions at the sensor nodes which are uniformly optimal over the distribution parameter. In this respect, consider the following definition [17, p. 78].

Definition 3.2 (Monotone Likelihood Ratio): The real-parameter family of densities $p_\theta(x)$ is said to have monotone likelihood ratio if there exists a real-valued function $T(x)$ such that for any $\theta < \theta'$

a) the distributions P_θ and $P_{\theta'}$ are distinct;
b) the ratio $(p_{\theta'}(x))/(p_\theta(x))$ is a nondecreasing function of $T(x)$.

The one-parameter exponential family [17, p. 80] and certain location-parameter families [17, p. 509] are important examples that have monotone likelihood ratios.

Given that the observation at each sensor node has a monotone likelihood ratio, and by applying Lemma 2.1, the censoring regions can be expressed as $R_i^c = \{T(X_i) \leq t_i\}, i = 1, 2, \ldots, N$. The choice of t_i to meet the communication-rate constraint $P_0(R_i^c) = \kappa_i$ depends on the parameter value θ_0 of the H_0 distribution. When θ_0 is fixed, estimated on-line, or has a worst-case value in terms of communication, t_i can be determined at each sensor. At the fusion center, the censored likelihood ratio ρ_i would be determined by periodic updates from the sensor nodes. Analysis of the proposed GLRT procedure is best done through simulation and experimentation.

C. Locally Optimum Formulation

In many detection applications, the distribution under the null hypothesis H_0 may be known (or can be estimated on-line) whereas the distribution under the target hypothesis H_1 is not completely known. When the distribution under each hypothesis is parameterized by a common parameter, a test which is optimal when the parameter value under H_1 is in the neighborhood of the parameter value under H_0 is known as a locally optimum test [17, p. 527].

Consider the following hypothesis-testing problem, where θ parameterizes the distribution under each hypothesis

$$H_0 : \theta = \theta_0 \text{ vs. } H_1 : \theta > \theta_0.$$

In the locally optimum setting, the Taylor's expansion of the detection probability around θ_0 is maximized subject to a false-alarm constraint.

$$\begin{aligned} &\sup_\delta P_D' \triangleq \left[\frac{d^m P_D(\delta(\mathbf{X}), \theta)}{d\theta^m}\right]_{\theta_0} \\ &\text{s.t} P_F(\delta) \leq \alpha \end{aligned}$$

where $\mathbf{X}$ is the vector of observations, δ is the decision rule, and $m > 0$ is the smallest integer for which P_D' is nonzero and bounded.

For the censoring scenario, consider a Bayesian formulation of the locally optimum problem where censored observations are available at a fusion center and the following local risk is to be minimized

$$\tilde{r} \triangleq \pi_0 P_F - (1 - \pi_0) P_D'. \tag{23}$$

We denote the probability distribution of the observations parametrized by θ as P_θ and the corresponding pdf as p_θ. We assume that $p_\theta(\mathbf{X})$ is sufficiently smooth [11, p. 38] so that the order of integration and differentiation can be interchanged in the computation of P_D'.

By using an approach identical to the Bayesian censoring problem, it is possible to obtain the optimal fusion rule as

$$\psi_0\left(\psi_i(X_i)\,|_{i=1}^N\right) = \begin{cases} 1, & \sum_{i=1}^N \psi_i(X_i) \geq \tau \\ 0, & \text{otherwise} \end{cases} \tag{24}$$

where the sensor decision rules are given by

$$\psi_i(X_i) = \begin{cases} \tilde{l}(X_i), & \tilde{l}(X_i) \in R_i \\ \tilde{\rho}_i, & \tilde{l}(X_i) \in R_i^c \end{cases} \tag{25}$$

where $\tilde{l}(X_i) = [(d^m/d\theta^m)(p_\theta(X_i))/(p_{\theta_0}(X_i))]_{\theta_0}$ is the local likelihood ratio, and $\tilde{\rho}_i = (\mathsf{P}(\tilde{l}(X_i) \in R_i^c \,|\, H_1))/(\mathsf{P}(\tilde{l}(X_i) \in R_i^c \,|\, H_0))$ is the local likelihood in the no-send region.

An appropriate constraint on the communication rate can then be written as

$$\sum_{i=1}^N \mathsf{P}(\tilde{l}(X_i) \in R_i \,|\, H_0) \leq \kappa. \tag{26}$$

Just as in the Bayesian censoring problem of Section II, our goal is to limit the search for the optimal censoring regions as much as possible. For centralized problems, the LO formulation can be directly related to the N-P problem by considering [11, p. 38]

$$H_0 : p_{\theta_0}(\mathbf{X}) \text{ vs. } H_1 : \left[\frac{d^m}{d\theta^m}\frac{p_\theta(\mathbf{X})}{p_{\theta_0}(\mathbf{X})}\right]_{\theta_0}.$$

For the LO censoring problem, such an approach is not possible since the result that a single censoring interval is optimal applies under the assumption that there is a pdf under each hypothesis.

Fortunately, we do have the result that censoring in a single interval of the local likelihood ratio is optimal

$$R_i^c = \{t_{1,i} < \tilde{l}(X_i) \leq t_{2,i}\} \tag{27}$$

As in Section II-A, define the set of deterministic Ψ, independently $\bar{\Psi}$ and dependently $\hat{\Psi}$ randomized decision strategies for the locally optimum censoring problem. We then have the following result. See the Appendix for a proof.

Theorem 3.2: Under Assumption 2.2 and given a communication-rate constraint κ to be satisfied, then for any censoring strategy in Ψ, there exists another in Ψ with lower local risk, where in the latter the fusion rule is given by (24), and the sensor rules are given by (25) and (27).

As in the Bayesian censoring problem of Section II, a second simplification in the censoring regions is possible for the LO censoring problem. Define the set of deterministic decision strategies $\Psi^\star \subset \Psi$ with fusion rule (24) and sensor decision rules (25) having censoring regions with lower thresholds at $-\infty$

$$R_i^c = \{\tilde{l}(X_i) \leq t_{2,i}\},\ \forall i. \tag{28}$$

Lemma 3.2: Suppose the fusion node has its own observations. Under Assumption 2.2 and given a communication-rate constraint κ to be satisfied, then for any censoring strategy $\psi \in \Psi$, there exists a decision strategy $\psi^\star \in \Psi^\star$ with lower local risk (i.e., $P_E(\psi^\star) < P_E(\psi)$).

Proof: The proof is very similar to the proof of Lemma 2.1 and Theorem 2.2. ■

IV. N-P Problems

In this section, we consider N-P problems in decentralized detection where the rate of *global* false-alarms is constrained. Whereas in the Bayesian problems which have been addressed, the communication-rate constraint can be interpreted as a "local" false-alarm constraint, in N-P problems to be addressed, the global false-alarm rate is also constrained. Such a constraint increases the role of randomization in the detection problem.

The N-P censoring problem of minimizing the miss probability subject to false-alarm and communication-rate constraints is

$$\begin{aligned} \min \quad & P_M \\ \text{s.t.} \quad & P_F \leq \alpha \\ \text{and} \quad & \sum_{i=1}^{N} \mathsf{P}(g_i(X_i) \in R_i \,|\, H_0) \leq \kappa. \end{aligned} \tag{29}$$

Consider the simplest nontrivial scenario of $N = 2$ sensor nodes, and apply the N-P Lemma [11, pp. 23-25] for a given choice of censoring regions to each of the four cases of transmit/no transmit separately.

When both sensor nodes transmit, we simply apply the N-P Lemma for the region $g(x_i) \in R_i, i = 1, 2$ to find that a likelihood ratio test (with randomization when the likelihood equals the threshold) is optimal. When both sensor nodes do not transmit, the miss and false-alarm probabilities are

$$\begin{aligned} P_M &= 1 - \int_{g_1(X_1)\in R_1^c, g_2(X_2)\in R_2^c} \\ &\quad \times \phi_0\left(R_1^c, R_2^c\right)\, p_1^{(1)}(x_1)\, p_1^{(2)}(x_2)\, dx_1\, dx_2 \\ &= 1 - \phi_0(R_1^c, R_2^c) P_1^{(1)}\left(R_1^c\right) P_1^{(2)}\left(R_2^c\right) \\ P_F &= \int_{g_1(X_1)\in R_1^c, g_2(X_2)\in R_2^c} \\ &\quad \times \phi_0\left(R_1^c, R_2^c\right)\, p_0^{(1)}(x_1)\, p_0^{(2)}(x_2)\, dx_1\, dx_2 \\ &= \phi_0\left(R_1^c, R_2^c\right) P_0^{(1)}\left(R_1^c\right) P_0^{(2)}\left(R_2^c\right) \end{aligned} \tag{30}$$

where ϕ_0 is the fusion center rule to be determined. Since P_M and P_F are independent of the sensor observations due to censoring, a given false-alarm rate can only be achieved by randomizing ϕ_0.

Similarly, when only one of the two sensor node transmits, let's consider the argument for optimality [11, p. 24] in detail for the region $g(x_1) \in R_1$ and $g(x_2) \in R_2^c$. Let $\phi_i(X_i)$ be defined as in (5), and define

$$\phi_0\left(\phi_i(X_i)|_{i=1}^N\right) = \begin{cases} 1, & \prod_{i=1}^N \phi_i(X_i) > \tau \\ \mu, & \prod_{i=1}^N \phi_i(X_i) = \tau \\ 0, & \prod_{i=1}^N \phi_i(X_i) < \tau \end{cases} \tag{31}$$

where μ is some randomization parameter independent of $\phi_i(X_i)|_{i=1}^N$.

A fusion rule ϕ_0' that satisfies the conditions of (29) would have miss and false-alarm probability

$$\begin{aligned} P_M &= 1 - P_1^{(2)}\left(R_2^c\right) \int_{g(x_1)\in R_1} \\ &\quad \times \phi_0'\left(\phi_1(x_1), R_2^c\right) p_1^{(1)}(x_1)\, dx_1 \\ P_F &= P_0^{(2)}\left(R_2^c\right) \int_{g(x_1)\in R_1} \\ &\quad \times \phi_0'\left(\phi_1(x_1), R_2^c\right) p_0^{(1)}(x_1)\, dx_1 \leq \alpha \end{aligned} \tag{32}$$

since ϕ_0' does not depend on x_2. Based on the definition of the fusion rule (31), we have

$$\begin{aligned} &\left[\phi_0\left(\phi_1(x_1), R_2^c\right) - \phi_0'\left(\phi_1(x_1), R_2^c\right)\right] \Big(p_1^{(1)}(x_1) P_1 \\ &\quad \times \left(R_2^c\right) - \tau\, p_0^{(1)}(x_1) P_0\left(R_2^c\right)\Big) > 0, \forall x_1 \in R_1. \end{aligned} \tag{33}$$

Integrating over R_1, we find that $P_D(\phi_0) \geq P_D(\phi_0')$, so ϕ_0 is the optimal fusion rule for the N-P censoring problem (29). In general, for the N-P problem, randomization at the fusion center may be necessary to achieve a desired false-alarm rate.

When the fusion center has its own observations, randomization at the fusion center is unnecessary since the likelihood ratio at the fusion center has no point mass under either hypothesis. It is straightforward to extend the Bayesian results (Lemma 2.1 and Theorem 2.2) to the N-P problem.

Lemma 4.1: Suppose the fusion node takes observations. Under Assumption 2.2, and given communication-rate (11) and false-alarm rate constraints to be satisfied, for any decision strategy $\phi \in \Phi$, there exists a decision strategy $\phi^\star \in \Phi^\star$ with lower miss probability (i.e., $P_M(\phi^\star) < P_M(\phi)$).

Proof: In [1], it was shown that consolidation of no-send intervals does not decrease P_D for a given false-alarm rate. Let $\phi \in \Phi$ be a decision strategy employing sensor rules (5) with censoring intervals (9) and fusion rule (31). Let $\phi^\star \in \Phi^\star$ be a decision strategy with fusion rule (31).

We further restrict attention to $\phi \in \Phi$ and $\phi^\star \in \Phi^\star$ which meet the false-alarm rate and communication-rate with equality (such strategies exist). For a fixed decision strategy, P_E is linear in π_0 since P_F and P_M are fixed. Then, we must have $P_M(\phi^\star) \leq P_M(\phi)$; otherwise $P_E(\phi^\star) > P_E(\phi)$, which contradicts Lemma 2.2. ■

We end this subsection by describing several implications of Lemma 4.1. As in the Bayesian problem, the N-P result for a communication-rate constraint across sensor nodes (2) is

Theorem 4.1: Suppose the fusion node takes observations. Under Assumption 2.2, and given communication-rate (2) and false-alarm rate constraints to be satisfied, for any decision strategy $\phi \in \Phi$, there exists a decision strategy $\phi^\star \in \Phi^\star$ with lower miss probability (i.e., $P_M(\phi^\star) < P_M(\phi)$). A proof of the theorem follows from arguments identical to the proof of Theorem 2.2.

Under the communication-rate constraint 2, for the N-P censoring problem over the set $\bar{\Phi}$ or $\hat{\Phi}$, it is sufficient to consider randomization over decision strategies from $\Phi^\star$. By a simple extension of Theorem 2.1 [12], randomization over at most three decision strategies (three censoring regions at each sensor) is optimal for dependently (independently) randomized decision strategies.

Finally, for false-alarm rates close to one, switching the role of H_0 and H_1 would lower the rate of communication for at least the same level of performance.

A. Locally Optimum Censoring

In Section III-C, minimizing the local risk in the locally optimum formulation of the censoring problem was considered. The local-risk formulation is somewhat artificial in that the rate of change of miss probability is combined with the false-alarm probability. Of greater interest is the N-P formulation where given the communication-rate constraint, an α-level LO decision strategy is to be determined [i.e., $\max P'_D$ s.t. $P_F \leq \alpha$, and (26)].

The α-level LO formulation can be related to the local risk formulation in the same way that the N-P formulation was related to the Bayesian risk formulation. Applying the N-P Lemma to the LO censoring problem, we obtain the fusion rule

$$\psi_0\left(\psi_i(X_i)\,|_{i=1}^N\right) = \begin{cases} 1, & \sum_{i=1}^N \psi_i(X_i) > \tau \\ \mu, & \sum_{i=1}^N \psi_i(X_i) = \tau \\ 0, & \sum_{i=1}^N \psi_i(X_i) < \tau \end{cases} \quad (34)$$

where μ is a randomization parameter independent of $\psi_i(X_i)|_{i=1}^N$.

The relationship between the α-level LO censoring problem and the local-risk censoring problem parallels the relationship between the Bayesian and N-P censoring problem.

Consider the individual sensor communication-rate constraint

$$\mathsf{P}(\tilde{l}(X_i) \in R_i \,|\, H_0) = \kappa_i \quad (35)$$

then, we have the following result.

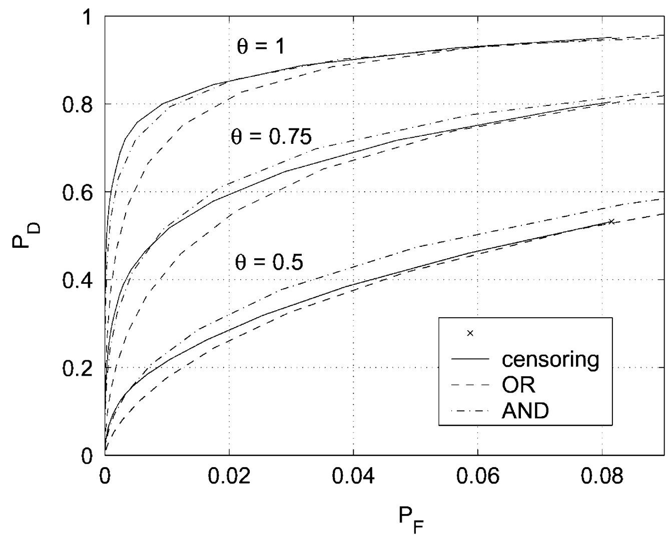


Fig. 4. ROC for the LO decentralized detection problem of mean-shift in i.i.d. Gaussian noise.

Theorem 4.2: Suppose the fusion center takes observations. Under Assumption 2.2, and given communication-rate (35) and false-alarm rate constraints to be satisfied, for any decision strategy $\psi \in \Psi$, there exists a decision strategy $\psi^\star \in \Psi^\star$ with larger rate of change of detection probability (i.e., $P'_D(\psi^\star) > P'_D(\psi)$).

Results for the communication-rate constraint (26) and randomized strategies follow similarly.

1) Example: To illustrate the LO censoring approach, we consider an LO formulation of the mean-shift example from Section II

$$H_0 : X_i \sim \mathcal{N}(0, \sigma^2) \text{ vs. } H_1 : X_i \sim \mathcal{N}(\theta, \sigma^2)$$

where X_i are identically distributed (i.i.d.) observations, σ^2 is the variance, and $\theta > 0$ is the distribution parameter which represents the signal strength. The local likelihood ratio for the mean-shift problem is monotone in the sufficient statistic X_i, so we censor X_i. It can be shown that for distributions from the one-parameter exponential family [17, p. 80], the LO and N-P censoring schemes are identical except for the likelihood in the no-send region.

We compare the LO censoring scheme with the binary LO decentralized detection scheme in which sensor observations are quantized to one bit ($D = 2$). Assuming that the censored transmission requires 24 bits to transmit the likelihood ratio with enough precision, a communication-rate constraint $\kappa = (1/24)$ for each sensor would transmit at the same average bit-rate as the binary scheme. Although this comparison is fair in terms of bandwidth, the energy consumption due to transmission would depend on the transmission scenario. It is more likely that wireless sensor networks are designed to transmit packets of information at a time rather than a single bit at a time.

Fig. 4 shows the optimal ROC for LO detection of a mean-shift in i.i.d. Gaussian noise with $N = 2$ sensors, and $\sigma = 0.4$. For the censoring scheme, the ROC was obtained by Monte Carlo simulation using 500 000 trials, whereas for the binary scheme, analytical expressions were derived. For simplicity, we consider censoring rules $\psi \in \Psi^\star$ and show their performance for false-alarm rates where it is optimal to have $t_{1,i} = -\infty$. We do not consider the fusion center to take its own observations in this example.

For the binary scheme, we limit our consideration to the AND and OR fusion rules. The performance of the censoring scheme is similar to that of the binary scheme at large θ and degrades as θ decreases. In the censoring scheme, the communication-rate constraint becomes more severe as θ decreases. For comparison in a typical setting, the mean shifts required to achieve $P_F = 0.001$, and $P_D = 0.9$ in the censoring, the AND, and the OR rules are $\theta = 1.31, 1.4$, and 1.5, respectively.

B. Robust N-P Censoring

In the robust N-P formulation of the censoring problem we are interested in maximizing the worst-case detection probability subject to worst-case false-alarm and worst-case communication-rate constraints.

To describe our formulation of the robust N-P problem, we need the following definition of vector ordering.

Definition 4.1 (Vector Ordering): If every component of the vector difference $\mathbf{x}-\mathbf{y}$ is less than or equal to zero, then $\mathbf{x} \preceq \mathbf{y}$.

Consider the robust N-P problem

$$\begin{aligned}\text{RNP1}: \quad & \inf_{\phi} \sup_{P_1 \in \mathcal{P}_1} P_M(P_1, \phi) \\ & \text{s.t.} \sup_{P_0 \in \mathcal{P}_0} (P_F(P_0, \phi), P_T(P_0, \phi)) \preceq (\alpha, \kappa)\end{aligned}$$

where $\phi = (\phi_0, g_1, R_1, \ldots, g_N, R_N)$ is some decision strategy with fusion rule ϕ_0, sensor functions g_i and censoring regions $\{R_i^c\}$, and $P_T(P_0, \phi) = \sum_{i=1}^N P_0^{(i)}(g_i(X_i) \in R_i)$ is the communication rate which is to be below κ. The false-alarm and communication rate are jointly constrained over the uncertainty class $\mathcal{P}_0$.

Consider the simple N-P censoring problem

$$\begin{aligned}\text{SNP1}: \quad & \inf_{\phi} P_M(Q_1, \phi) \\ & \text{s.t.} (P_F(Q_0, \phi), P_T(Q_0, \phi)) \preceq (\alpha, \kappa)\end{aligned}$$

where (Q_0, Q_1) is the pair of LFDs defined in Section III-A.

Using exactly the same arguments as in Theorem 3.1, we obtain

Theorem 4.3: Suppose the fusion node takes observations. Under Assumption 2.2, the robust N-P problem of minimizing the worst-case miss probability subject to worst-case transmission and worst-case false-alarm constraints (i.e., problem RNP1) can be solved by designing for the least-favorable distributions in the sense of Defn. 3.1 (i.e., problem SNP1).

V. Conclusion

Under a censoring scheme for transmission, we have shown that the design of sensor networks for detection simplifies considerably, even when the distribution of the observations is only partially known. For a censoring rate at each sensor node and the case where the fusion center has its own observations, we proved that the censoring regions at each sensor node can be chosen independently, since the lower threshold goes to zero. For adaptive systems, where the null and target statistics are updated over time, the independence of censoring regions simplifies redesign. Updating across sensor nodes in unnecessary; only updates at individual sensor nodes and the fusion center are required. Alternatively, we considered formulations where the distribution of the observations is only partially known. In the robust formulation, we found that designing for a pair of least-favorable distributions is optimal in the sense of worst-case detection performance and communication rate. In ϵ-contamination classes, it is significant to find that robustness and censoring are achieved by the same clipping test. For problems where the target is characterized by a shift in the distribution parameter, we determined the locally optimum censoring strategy. We found that it involves the fusion of censored local likelihoods using a sum rule. The simplicity of the resulting tests for partially known formulations makes them quite useful in practice. We demonstrated the detection performance of the various censoring approaches for the problem of detecting a mean shift in Gaussian noise.

Appendix

Proof Lemma 2.1: In finding $(\partial P_E)/(\partial t_{1,k})$, the derivative of the dependent variable $t_{2,k}$ with respect to $t_{1,k}$ must be considered. Since P_E can be written as an integral whose limits are functions of $t_{1,k}$ and $t_{2,k}$, we can apply the fundamental theorem of integral calculus. The derivative $(\partial P_E)/(\partial t_{1,k})$ exists since we assume that the pdfs of the observations (the integrand) have no point mass. Taking the derivative of (11) with respect to $t_{1,k}$, we obtain

$$\frac{\partial t_{2,k}}{\partial t_{1,k}} = \frac{f_0(t_{1,k})}{f_0(t_{2,k})} \tag{36}$$

where $f_j(l_k), j = 0, 1$, is the pdf of the likelihood ratio at the k-th sensor under H_j.

Define the likelihood ratio at the fusion center of all sensors but sensor k as

$$U = \prod_{i \neq k} \phi_i(X_i) \tag{37}$$

where $\phi_i(X_i) = l(X_i)\, \mathcal{I}_{\{l(X_i) \in R_i\}} + \rho_i\, \mathcal{I}_{\{l(X_i) \in R_i^c\}}$, and $\mathcal{I}_{\{\cdot\}}$ is the indicator function. Due to the fact that U is a likelihood ratio, it has the following property (which comes from finding the slope of P_D with respect to P_F ([18], pp. 13–14))

$$u = \frac{dG_1(u)}{dG_0(u)} \tag{38}$$

where $G_j(u)$ is the cumulative distribution function (cdf) of U under H_j. We will write the error probability in terms of the cdf of U instead of the pdf of U, since U is a censored likelihood ratio with point mass under both hypotheses.

The error probability for the censoring scheme is then given by $P_E = \pi_1 + \pi_0 P_0[U\phi_k(X_k) \geq \tau] - \pi_1 P_1[U\phi_k(X_k) \geq \tau]$, or

$$\begin{aligned}\frac{P_E}{\pi_1} = 1 - & \iint_{(u\phi_k(x_k)) \geq \tau} p_1(x_k)\, dG_1(u)\, dx_k \\ & + \tau \iint_{(u\phi_k(x_k)) \geq \tau} p_0(x_k)\, dG_0(u)\, dx_k.\end{aligned}$$

Applying (38) and separating the integral over x_k into send and no-send regions, we obtain

$$\frac{P_E}{\pi_1} = 1 - \int\limits_{l_k \in R_k} f_0(l_k) \int\limits_{u=\frac{\tau}{l_k}}^{\infty} (l_k u - \tau)\, dG_0(u)\, dx_k - \int\limits_{l_k \in R_k^c} f_0(l_k) \int\limits_{u=\frac{\tau}{\rho_k}}^{\infty} (\rho_k u - \tau)\, dG_0(u)\, dx_k$$

where ρ_k is the censored value defined in (6).

Under Assumption 2.2, and using the fact that the no-send region for each sensor is a single interval of the likelihood ratio, we can write

$$\begin{aligned}\frac{P_E}{\pi_1} = 1 &- \int_0^{t_{1,k}} f_0(l_k) \int_{\frac{\tau}{l_k}}^{\infty} (l_k u - \tau)\, dG_0(u)\, dl_k \\ &- \int_{t_{1,k}}^{t_{2,k}} f_0(l_k)\, dl_k \int_{\frac{\tau}{\rho_k}}^{\infty} (\rho_k u - \tau) dG_0(u) \\ &- \int_{t_{2,k}}^{\infty} f_0(l_k) \int_{\frac{\tau}{l_k}}^{\infty} (l_k u - \tau)\, dG_0(u)\, dl_k.\end{aligned}$$

Taking the derivative of P_E with respect to $t_{1,k}$, we obtain

$$\frac{1}{\pi_1}\frac{\partial P_E}{\partial t_{1,k}} = -f_0(t_{1,k}) \int_{\frac{\tau}{t_{1,k}}}^{\infty} (t_{1,k} u - \tau)\, dG_0(u) \tag{39}$$

$$+ f_0(t_{2,k}) \frac{\partial t_{2,k}}{\partial t_{1,k}} \int_{\frac{\tau}{t_{2,k}}}^{\infty} (t_{2,k} u - \tau)\, dG_0(u) \tag{40}$$

$$- \kappa_k \frac{\partial}{\partial t_{1,k}} \left\{ \int_{u=\frac{\tau}{\rho_k}}^{\infty} (\rho_k u - \tau)\, dG_0(u) \right\}. \tag{41}$$

Given that the fusion center takes observations, the conditional distribution of U has no point mass under either hypothesis and the last term (41) becomes zero.

Substituting for $(\partial t_{2,k})/(\partial t_{1,k})$, we obtain the lower bound

$$\frac{\partial P_E}{\partial t_{1,k}} > \pi_1 f_0(t_{1,k}) \int_{\frac{\tau}{t_{2,k}}}^{\frac{\tau}{t_{1,k}}} (t_{2,k} u - \tau)\, dG_0(u). \tag{42}$$

Since the integrand of (42) is nonnegative in the region of integration, $(\partial P_E)/(\partial t_{1,k}) > 0$ as required. ∎

When the fusion center does not take observations, we do not have the simple result about optimality of $\Phi^\star$. However, we can obtain some conditions about "local" optimality as the censoring intervals are shifted slightly. Examining the term (41) due to the censored value, we find that $G_j(u)$ has a point mass at $\rho' = \prod_{i \neq k} \rho_i$. For $\rho' \geq \tau/\rho_k$, (41) can then be expanded as

$$\begin{aligned}- \kappa_k \frac{\partial}{\partial t_{1,k}} \Bigg\{ &\int_{u=\frac{\tau}{\rho_k}}^{\rho'} (\rho_k u - \tau)\, dG_0(u) \\ &+ \int_{u=\rho'}^{\infty} (\rho_k u - \tau)\, dG_0(u) \\ &+ (\rho_k \rho' - \tau)\, \mathsf{P}(U = \rho' \mid H_0) \Bigg\}.\end{aligned} \tag{43}$$

Taking the derivative of (6), we obtain

$$\frac{\partial \rho_k}{\partial t_{1,k}} = \frac{1}{\kappa_k}(t_{2,k}\, f_0(t_{1,k}) - t_{1,k}\, f_0(t_{1,k})). \tag{44}$$

Substituting for $(\partial \rho_k)/(\partial t_{1,k})$ and simplifying (43), we obtain for the term (41)

$$-\kappa_k \frac{\partial}{\partial t_{1,k}}\{\cdot\} = \begin{cases} (t_{1,k} - t_{2,k}) \prod_{i \neq k} P_1(R_i^c) & \tau \leq \prod_{i=1}^N \rho_i \\ 0, & \text{otherwise.} \end{cases}$$

Consider starting with a strategy $\phi^\star \in \Phi^\star$. Let $\rho_{\text{tot}} = \prod_{i=1}^N \rho_i$. For $\pi_0 > (\rho_{\text{tot}})/(1 + \rho_{\text{tot}})$, it is clear that $(\partial P_E)/(\partial t_{1,i})$ is positive as $t_{1,i}, i = 1, 2, \ldots, N$ is increased, until $\pi_0 \geq (\rho_{\text{tot}})/(1 + \rho_{\text{tot}})$ since the term (41) is zero. In this sense, ϕ is locally optimal but not globally optimal (since $(\partial P_E)/(\partial t_{1,i})$ is not positive for all π_0, a strategy $\phi \notin \Phi^\star$ cannot be compared with $\phi^\star$ simply by considering $(\partial P_E)/(\partial t_{1,i})$).

Proof of Theorem 3.2: We follow the same approach as in [1] to show that censoring in a single region of the local likelihood ratio minimizes the local risk. Consider the censoring region for a particular sensor k. It is possible to show that the consolidation of a censoring region (for sensor k) consisting of two intervals into a single interval does not increase the risk.

For conditionally independent observations, the local risk can be written as

$$\begin{aligned}\frac{\tilde{r}}{\pi_1} &= -P_D' + \tau P_F \\ &= -\frac{d}{d\theta} \Bigg[\int_{\mathcal{D}_1} \prod_{i=1}^N \Big(p_\theta(x_i)\, \mathcal{I}_{\{\tilde{l}(x_i) \in R_i\}} \\ &\qquad + P_\theta\left(\tilde{l}(x_i) \in R_i^c\right) \mathcal{I}_{\{\tilde{l}(x_i) \in R_i^c\}} \Big)\, dx_i \Bigg]_{\theta_0} \\ &\quad + \tau \int_{\mathcal{D}_1} \prod_{i=1}^N p_{\theta_0}(x_i)\, dx_i\end{aligned} \tag{45}$$

where $\mathcal{D}_1 = \{\sum_{i=1}^N \psi_i(X_i) \geq \tau\}$ is the decision region for H_1 defined by (24), and $\tau = (\pi_0/\pi_1)$. Interchanging the order of integration and differentiation, and applying the product rule to the first integral term, we obtain

$$\begin{aligned}\tilde{r} = \pi_1 \int_{\mathcal{D}_1} \left[\tau - \sum_{i=1}^N \psi_i(x_i) \right] \prod_{i=1}^N \Big(&p_{\theta_0}(x_i)\, \mathcal{I}_{\{\tilde{l}(x_i) \in R_i\}} \\ &+ P_\theta\left(\tilde{l}(x_i) \in R_i^c\right) \mathcal{I}_{\{\tilde{l}(x_i) \in R_i^c\}} \Big)\, dx_i.\end{aligned} \tag{46}$$

Define the random variables $U = \exp(\sum_{i \neq k} \tilde{l}(X_i)), V = \exp(\tilde{l}(X_k))$, and the threshold $\tau' = \exp(\tau)$, then the decision region becomes $\mathcal{D}_1 = \{UV > \tau'\}$.

Using Assumption 2.2, it is possible to separate the N-dimensional integral for the local risk into the product of an integral over sensor k and an integral over all other sensors

$$\tilde{r} = \pi_1 \int_{u=0}^{\infty} \int_{v=\frac{\tau'}{u}}^{\infty} \left[\log\left(\frac{\tau'}{u}\right) - \log(v) \right] p_{\theta_0}(v)\, dv\, p_{\theta_0}(u)\, du. \tag{47}$$

Once in this form, many of the arguments given in [1] can be used directly. As in [1], we choose to introduce the additional constraint that the communication under H_1

$$\int_{\tilde{l}(X_i)\in R_i} \left[\frac{d}{d\theta} p_\theta(x_j)\right]_{\theta_0} dx_j \tag{48}$$

remain the same under the consolidation of the no-send interval. The introduction of this constraint may eliminate feasible solutions, but enables us to show that single intervals are optimal.

To complete the proof, one step requires that the integral over V in (47) is nonincreasing in (τ'/u). Since V has no-point mass under either hypothesis, we can obtain

$$\frac{d}{d\sigma}\left\{\int_\sigma^\infty [\log(v) - \log(\sigma)]\, p_{\theta_0}(v)\, dv\right\} = -\sigma^{-1}\mathsf{P}(V > \sigma \,|\, H_0). \tag{49}$$

The rate of change is increasing and negative, which implies that the integral is nonincreasing. ■

REFERENCES

[1] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.

[2] R. R. Tenney and N. R. Sandell, Jr., "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, no. 4, pp. 501–510, Jul. 1981.

[3] S. Appadwedula, "Energy-efficient sensor networks for detection applications," Ph.D. dissertation, Univ. Illinois at Urbana-Champaign, Urbana, 2003.

[4] S. M. Ali and S. D. Silvey, "A general class of coefficients of divergence of one distribution from another," *J. R. Statist. Soc., Series B.*, vol. 28, no. 1, pp. 131–142, 1966.

[5] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Energy-efficient detection in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 693–702, Apr. 2005.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[7] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Ann. Math. Statist.*, vol. 23, pp. 493–507, Dec. 1952.

[8] W. P. Tay, J. N. Tsitsiklis, and M. Z. Win, "Asymptotic performance of a censoring sensor network," *IEEE Trans. Inf. Theory*, vol. 53, pp. 4191–4209, Nov. 2007.

[9] J. N. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Trans. Autom. Control*, vol. AC-30, no. 5, pp. 440–446, May 1985.

[10] J. N. Tsitsiklis, , H. V. Poor and J. B. Thomas, Eds., "Decentralized detection," in *Advances in Statistical Signal Processing*. Greenwich, CT: JAI, 1991, vol. 2, (Signal Detection), pp. 297–344.

[11] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.

[12] S. Appadwedula, D. L. Jones, and V. V. Veeravalli, "Energy-efficient sensor networks for detection applications," in *Proc. 6th Int. Conf. Inf. Fusion*, Cairns, Australia, Jul. 2003, pp. 56–63.

[13] E. Ertin, "Polarimetric processing and sequential detection for automatic target recognition systems," Ph.D. dissertation, Ohio State Univ., Columbus, 1999.

[14] W. B. Heinzelman, "Application-specific protocol architectures for wireless networks," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, 2000.

[15] P. J. Huber, "A robust version of the probability ratio test," *Ann. Math. Statist.*, vol. 36, pp. 1753–1758, Dec. 1965.

[16] V. V. Veeravalli, T. Basar, and H. V. Poor, "Minimax robust decentralized detection," *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp. 35–40, Jan. 1994.

[17] E. L. Lehmann, *Testing Statistical Hypotheses*. New York: Wiley, 1986.

[18] C. W. Helstrom, *Elements of Signal Detection and Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

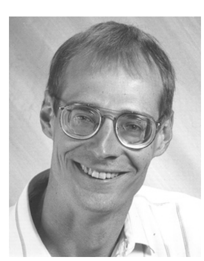
Swaroop Appadwedula (M'05) received the B.S. degree from Cornell University, Ithaca, NY, in 1996, and the M.S. and Ph.D. degrees from the University of Illinois at Urbana-Champaign (UIUC) in 1998 and 2003, respectively.

As a graduate student at UIUC, he spent a considerable amount of time interacting with students and developing materials for the Digital Signal Processing Laboratory. Since 2003, he has been with the Massachusetts Institute of Technology (MIT) Lincoln Laboratory, Cambridge, where he currently works on a array processing algorithms. His research interests are in digital signal processing, communications and detection theory, particularly sensor networks, system-level design, and array processing.

Venugopal V. Veeravalli (S'86–M'92–SM'98–F'06) received the B.Tech. degree in 1985 from the Indian Institute of Technology, Bombay, (Silver Medal Honors), the M.S. degree in 1987 from Carnegie-Mellon University, Pittsburgh, PA, and the Ph.D. degree in 1992 from the University of Illinois at Urbana-Champaign (UIUC), all in electrical engineering.

He joined the UIUC in 2000, where he is currently an Associate Professor with the Department of Electrical and Computer Engineering, and a Research Associate Professor with the Coordinated Science Laboratory. He is also currently serving as Program Director for Communications Research at the U.S. National Science Foundation, Arlington, VA. He was an Assistant Professor with Cornell University, Ithaca, NY, during 1996–2000. His research interests include detection and estimation theory, information theory, wireless communications, and sensor networks.

Dr. Veeravalli is an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY, and on the editorial board for *Communications in Information and Systems* (CIS), *Sensor Letters*, and the *Journal of Advances of Information Fusion*. Among the awards he has received for research and teaching are the IEEE Browder J. Thompson Best Paper Award in 1996, the National Science Foundation CAREER Award in 1998, the Presidential Early Career Award for Scientists and Engineers (PECASE) in 1999, the Michael Tien Excellence in Teaching Award from the College of Engineering, Cornell University in 1999, the Xerox Award for faculty research from the College of Engineering, UIUC, in 2003. He was a Beckman Associate at the Center for Advanced Study, University of Illinois in 2002.



Douglas L. Jones (S'82–M'83–SM'97–F'02) received the B.S.E.E., M.S.E.E., and Ph.D. degrees from Rice University, Houston, TX, in 1983, 1985, and 1987, respectively.

During the 1987–1988 academic year, he was with the University of Erlangen-Nuremberg, Germany, on a Fulbright Postdoctoral fellowship. Since 1988, he has been with the University of Illinois at Urbana-Champaign, where he is currently a Professor of electrical and computer engineering with the Coordinated Science Laboratory, and the Beckman Institute. He was on sabbatical leave at the University of Washington during spring 1995 and at the University of California at Berkeley in spring 2002. In spring semester 1999, he served as the Texas Instruments Visiting Professor at Rice University. He is an author of two DSP laboratory textbooks, and was selected as the 2003 Connexions Author of the Year.

Dr. Jones is currently serving on the Board of Governors of the IEEE Signal Processing Society. His research interests are in digital signal processing and communications, including nonstationary signal analysis, adaptive processing, multisensor data processing, OFDM, and various applications such as advanced hearing aids.