

# Energy-Efficient Detection in Sensor Networks

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

**Abstract**—There is significant interest in battery-powered sensor networks to be used for detection in a wide variety of applications, from surveillance and security to health and environmental monitoring. Severe energy and bandwidth constraints at each sensor node demand system-level approaches to design that consider detection performance jointly with system-resource constraints. Our approach is to formulate detection problems with constraints on the expected cost arising from transmission (sensor nodes to a fusion node) and measurement (at each sensor node) to address some of the system-level costs in a sensor network. For a given resource constraint, we find that randomization over the choice of measurement and over the choice of when to transmit achieves the best performance (in a Bayesian, Neyman–Pearson, and Ali–Silvey sense). To facilitate design, we describe performance criteria in the send/no-send transmission scenario, where the joint optimization over the sensor nodes decouples into optimization at each sensor node.

**Index Terms**—Decentralized detection, distributed detection, energy constraints, resource constraints, sensor networks.

## I. INTRODUCTION

**I**NTRUDER detection, traffic monitoring, wildlife observation, and data collection are just some of the wide variety of applications envisioned for battery-powered sensor networks. In many of these applications, detection of a particular “trigger” event is the initial (or primary) step executed before any other processing. When trigger events occur rarely, sensor nodes will spend a vast majority of their lifetime in the detection loop. The efficient use of system resources during detection then plays a key role in the longevity of the sensor nodes.

Traditionally, decentralized detection problem formulations have addressed detection in sensor networks by focusing on measures of performance such as error probability and have neglected system costs. The energy dissipation in the system including transmit energy and the energy due to processing has not been considered directly in the detection problem. We address the problem of detection in sensor networks with energy constraints as a system-level problem that requires the consideration of detection performance jointly with system resources. We formulate detection problems with constraints on the expected cost arising from transmission (sensor nodes to a fusion node)

and measurement (at each sensor node) to address some of the system-level costs in a sensor network.

The following example illustrates one realistic problem addressed by our formulation. Consider sensor nodes having two different sensing modalities (e.g., any two of the following: acoustic, seismic, and image). Suppose a primary detector can make use of both modalities, while a secondary detector uses only one of the modalities; the cost of measurement and operation for the primary detector is, thus, greater than the costs in the secondary detector. A lower transmission rate using the primary detector may achieve the same performance in terms of error probability as a higher transmission rate using the secondary detector. Given a constraint on the expected cost to be satisfied, the relevant problem is to decide how often, and at what transmit rate, to use a particular detector.

In typical application scenarios, target arrivals do not follow any particular distribution. For example, the occurrence of intruders in a home is typically a rare and unpredictable event. In such scenarios, our main result is that under a constraint on the expected cost, the sensor nodes must randomize their choice of measurement and how frequently they transmit to achieve the best performance (in a Bayesian, Neyman–Pearson (N-P), or Ali–Silvey (A-S) sense). In particular, for sensor nodes that are either awake or asleep, the result implies that sensor nodes should sleep for periods of time to save energy. In practice, the issue of how to determine the optimal decision strategy, including the randomization parameters, is a major issue, particularly in time-varying noise and target environments. Therefore, to facilitate design, we describe several cases in which the joint optimization over the sensor nodes is decoupled.

There has been some important work on incorporating system cost into detection problems. Fu *et al.* [1] consider a pattern-recognition problem in which a set of features with associated costs of computation may be used for discriminating various patterns. Various assumptions about the dependence between feature measurements were considered to reduce the complexity of the optimization, posed as a dynamic programming problem. Ertin [2] uses a decision-theoretic formulation to formalize the work of Fu *et al.* Ertin finds that under resource constraints, the optimal decision strategy involves randomization over deterministic decision strategies. The randomization results are applied to a sequential detection problem with a constraint on the expected cost. Wald’s approximations from sequential analysis and a receiver operating characteristic (ROC)-based technique are used to determine the optimal design parameters.

The introduction of measurement costs to detection problems has been considered in detail in sequential detection problems, where an interpretation of cost as delay is typical. In sequential problems, the test statistic is the sequential probability ratio test

Manuscript received December 1, 2003; revised July 30, 2004. This work was supported in part by the National Science Foundation under Grant EIA-0072043 and in part by CAREER/PECASE under Award CCR 00-49089. This paper was presented in part at the Sixth International Conference on Information Fusion, Cairns, Queensland, Australia, July 2003.

S. Appadwedula is with the Advanced Sensor Techniques Group, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02420-9185 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/JSAC.2005.843536

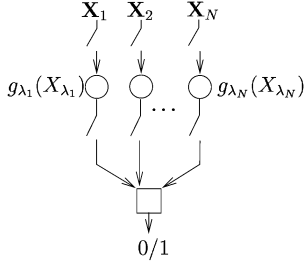


Fig. 1. Topology of “censoring” decentralized detection problem showing choice of measurement and choice of transmit.

and stopping thresholds determine when to stop taking observations and declare a result [3, p. 101]. Our problem formulation differs from the sequential problem formulation in that we consider sensor nodes that can collect, compute and transmit in each observation interval. Randomization of how often nodes collect observations and transmit plays the important role in the problems we consider.

The novelty of our work is in the particular (convenient) formulation of the decentralized detection problem with associated system costs due to transmission and observation. By making use of existing results [4]–[6] which consider only the cost of transmission and known results on randomization [2], [7], we are able to make statements regarding the design of sensor nodes.

The remainder of the paper is organized as follows. In Section II, the send/no-send censoring transmission scenario and expected cost are defined and randomized strategies are introduced. In Section III, the no-send-zone quantizer is introduced. The three sections that follow describe results for the censoring transmission scenario: Section IV describes the optimality of likelihood-ratio detectors and randomization, Section V describes the dual problems, and Section VI presents and demonstrates our approach for the special case of on-off sensor nodes. Finally, Section VII describes results for the no-send-zone quantizer transmission scenario.

## II. CENSORING PROBLEM FORMULATION

We begin by describing the decision problem faced by the sensor nodes. We are concerned with the binary hypothesis-testing problem of determining the true state of nature  $H$  as  $H_0$  (null) or  $H_1$  (target) based on a set of observations gathered at each sensor node. For  $i = 1, 2, \dots, N$ , let  $\mathbf{X}_i$  denote the set of observations available at sensor node  $i$ , and let  $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_N)$  denote the complete set of observations. Each sensor node computes a local output based on some subset  $X_{\lambda_i} \in \mathcal{X}_{\lambda_i}$  of its observations which are measured, where  $\lambda_i$  denotes the indices of the measured observations. Partial information about the local output  $g_{\lambda_i}(X_{\lambda_i})$  (defined more explicitly later) is then transmitted to a fusion node for global decision making.

We focus on the parallel sensor network topology shown in Fig. 1, where the sensor nodes communicate in parallel with a fusion node and neither communicate with each other nor receive feedback from the fusion node. The cost of communica-

tion is then due to transmission from each sensor node to the fusion node and reception at the fusion node.

We will consider two transmission scenarios in which partial observations are available at a fusion node for global decision making. In these scenarios, the cost due to transmission is determined by the send/no-send rate, also referred to as the transmit/censoring rate. We are motivated by previous results [4], [5] which indicate that sensor nodes do not need to transmit very frequently, particularly when one of the hypotheses is more likely. When a sensor node does transmit, the energy dissipation in the radio due to startup may be more significant than the energy dissipation due to transmission of a few packets [8]. Therefore, for the first transmission scenario, we consider transmission of a real-valued function when the observations are “informative” to the fusion node.

**Scenario C. Censoring a real-valued output:** Each sensor node computes a real-valued function of its observation

$$g_{\lambda_i}(X_{\lambda_i}) : \mathcal{X}_{\lambda_i} \rightarrow \mathbb{R} \quad (1)$$

and sends in the “informative” region  $R_{\lambda_i}$  and censors in the no-send region  $R_{\lambda_i}^c$  (i.e.,  $R_{\lambda_i} \cup R_{\lambda_i}^c = \mathcal{X}_{\lambda_i}$ ). We define the transmit rate of the censoring function as

$$P_{T,i}(R_{\lambda_i}) = P_0(X_{\lambda_i} \in R_{\lambda_i}) \quad (2)$$

where we consider transmission only under the null hypothesis, and  $P_0$  is the conditional distribution function under  $H_0$ . Let  $\pi_j = P(H_j)$  be the prior probability of hypothesis  $H_j$ ,  $j = 0, 1$ . Assuming that  $\pi_0 \gg \pi_1$ , transmission cost under the target hypothesis is relatively insignificant. (Note the uncanny use of the term rate to refer to the probability of transmission.)

We now describe the sensor decision rules and the fusion rule for the topology in Fig. 1. Let  $\mathcal{T}$  be the set of sensor nodes which transmit their real-valued outputs, and define  $\mathcal{T}^c$  as the set of sensor nodes which censor. The fusion node makes a global decision based on the received  $\{g_{\lambda_i}(X_{\lambda_i}), i \in \mathcal{T}\}$ , and the known no-send regions  $\{R_{\lambda_i}^c, i \in \mathcal{T}^c\}$ . Then, we can denote the fusion rule as  $\phi_0(\{g_{\lambda_i}(X_{\lambda_i})\}_{i \in \mathcal{T}}, \{R_{\lambda_i}^c\}_{i \in \mathcal{T}^c})$  and the complete decision strategy as

$$\phi = (\phi_0, \{\lambda_i, g_{\lambda_i}, R_{\lambda_i}\}_{i=1}^N). \quad (3)$$

We consider the cost of transmission and the cost of measurement to be additive at each sensor node, where, at each node, a subset of the observations may be measured. We define the total cost due to measurement and transmission across all sensor nodes as

$$C \triangleq \sum_{i=1}^N c_{\lambda_i} + c_{T,i} P_{T,i}(R_{\lambda_i}) \quad (4)$$

where  $c_{\lambda_i} \in [0, \infty)$  is the cost due to measurement of  $X_{\lambda_i}$ , and  $c_{T,i} \in [0, \infty)$  is the cost per transmission. By defining the relative costs at the individual sensor nodes, it is possible to account for a variety of node-dependent scenarios (e.g., a transmit cost that depends on the number of quantization levels (to be defined), or on the distance to the fusion node).

We now give the notation for the distribution of the observations and list our assumptions. The observations at the sensor nodes are described statistically by their conditional distributions, which are assumed to be known. Let  $P_j^i$  be the distribution function of  $X_i$  under  $H_j$ ,  $j = 0, 1$ , and  $p_j^i(x_i)$  be the corresponding probability density function (pdf) at  $X_i = x_i$ . We drop the superscript  $i$  when it can be determined from the context. We describe the joint distribution function over the observations as  $P_j = P_j^1 \times P_j^2 \times \dots \times P_j^N$ .

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

The conditional independence assumption is common in decentralized detection problem formulations and has been critical to the development of sensor decision rules based on likelihood ratios. Independence under the null hypothesis is satisfied in a variety of scenarios including measurement-dominated noise or noise due to local phenomena. Independence can also arise in practice under the target hypothesis. Some staple examples are signals arriving at each node with uniformly random phase offset or independent fading. When the observations are dependent, decentralized detection problems are less tractable and complexity of design is an issue [9]. While the observations are assumed to be conditionally independent across sensor nodes, the observations at a particular sensor node may be dependent (i.e., given  $\mathbf{X}_1 = [X_{1,1}, X_{1,2}]$ ,  $X_{1,1}$  may depend on  $X_{1,2}$ ). At a given sensor node, it is very likely that observations will be dependent across modalities. In particular, given dependent observations, measuring observations with lower measurement costs is preferable.

We impose the no-point-mass assumption as a smoothness condition on the distribution of the observations.

*Assumption 2.2 (No-Point Mass):* For each sensor node  $i$  and every choice of  $\lambda_i$ , the likelihood ratio defined as  $l(X_{\lambda_i}) = p_1(x_{\lambda_i})/p_0(x_{\lambda_i})$  has no point mass under either hypothesis

$$P_j(l(X_{\lambda_i}) = t) = 0, \quad t \in [0, \infty), \quad j = 0, 1. \quad (5)$$

### A. Randomization

For centralized detection problems, the N-P Lemma [3] states that randomization of the threshold in the likelihood-ratio test may be required to meet a desired false-alarm rate, while minimizing the miss probability. In decentralized detection problems, randomization of the decision strategy may also be required due to the constraint on the expected cost. In decentralized problems, we must also distinguish between dependent and independent randomization due to communication costs. In a dependent or coordinated strategy, the fusion node may generate a random variable to determine the sensor rules for all of the sensor nodes and, subsequently, broadcast the sensor rules. (Note that sensor nodes may avoid reception by following a preassigned randomization schedule.) In an independent strategy, each sensor node generates a random variable independent of other sensor nodes to determine its sensor rule and, subsequently, informs the fusion node. In both types of randomization, the fusion node requires knowledge of which sensor decision rules are in use.

We define the set of independently and dependently randomized decision strategies [7]. Let  $\Phi_i$  be the set of all deterministic sensor rules for sensor node  $i$ , where a particular deterministic sensor rule is the fixed choice of observation  $\lambda_i$ , censoring function  $g_{\lambda_i}$  and censoring region  $R_{\lambda_i}^c$  for all time. Let  $\Phi_0$  be the set of all deterministic fusion rules. Then, the set of all decision strategies can be described as  $\Phi = \Phi_0 \times \Phi_1 \times \Phi_2 \times \dots \times \Phi_N$ . Let  $\bar{\Phi}$  be the set of independently randomized decision strategies in the sense that each sensor node has a finite set of candidate sensor rules  $\{\phi_i^{(k)}\}_{k=1}^{K_i}$ , where  $K_i$  is the number of available rules, and chooses to use rule  $\phi_i^{(k)}$  with probability  $\mu_i^{(k)}$ , where  $\mu_i^{(k)}$  are determined independently at each sensor node. Let  $\hat{\Phi}$  be the set of dependently randomized decision strategies in the same sense as for  $\bar{\Phi}$ , except that  $\mu_i^{(k)}$  may be dependent across sensor nodes. Based on the definitions, it is easy to see that deterministic strategies are a subset of independently randomized strategies which are a subset of dependently randomized strategies (i.e.,  $\Phi \subset \bar{\Phi} \subset \hat{\Phi}$ ).

Design over the set of randomized strategies is generally more difficult since the complete decision strategy includes the choice of sensor decision rules, no-send regions, fusion rule, and randomization probability for each strategy (i.e.,  $(\phi^{(1)}, \dots, \phi^{(K)}, \mu^{(1)}, \dots, \mu^{(K)})$ ). The expected communication rate  $P_{T,i}$  and expected cost  $C$  are obtained by averaging over the candidate rules.

### B. System-Level Optimization Problems

Consider the system-level optimization problems given in (6)–(8), where  $P_M$  is the miss probability,  $P_F$  is the false-alarm probability, the error probability is defined as  $P_E = \pi_0 P_F + \pi_1 P_M$ , and  $D_f$  is the A-S distance measure to be defined. We denote the resource level and maximum false-alarm rate to be satisfied as  $c_0$  and  $\alpha$ , respectively

$$\text{PB-C : } \min_{\phi} P_E(\phi) \text{ s.t. } C(\phi) \leq c_0 \quad (6)$$

$$\text{PN-C : } \min_{\phi} P_M(\phi) \quad (7)$$

$$\text{s.t. } P_F(\phi) \leq \alpha, \text{ and } C(\phi) \leq c_0$$

$$\text{PD-C : } \max_{\phi} \sum_{i=1}^N D_f(\phi_i) \text{ s.t. } C(\phi) \leq c_0. \quad (8)$$

Suppose that a sensor network is to be designed for detection that lasts for a given period of time, while minimizing the detection error probability. If the transmission costs at each node are best modeled by how frequently a node transmits, then we are interested in the Bayesian problem PB-C (P denotes problem and B denotes Bayesian). If no prior probability of target occurrence is known, then we may consider keeping the false-alarm rate below some maximum level as in the N-P problem PN-C. When ease of design is an issue such as in nonstationary noise/target environments, a constraint on the A-S distance could be considered as in the distance problem PD-C.

We now define the A-S distance measure. Ali and Silvey described a general class of coefficients for measuring the divergence or closeness between two distributions which satisfies various desirable properties [10]. The A-S distance for the

sensor decision rule  $\phi_i$  is  $D_f(\phi_i) = E_0[f(\phi_i)]$ , where  $f$  is a continuous, convex function, and  $E_0$  denotes the expected value under  $H_0$ .

### III. NO-SEND-ZONE QUANTIZER PROBLEM FORMULATION

As an alternative to Scenario C, we introduce Scenario N which is useful for digital transmission. In Scenario N, the output of sensor node  $i$  is quantized to  $D_i$  discrete levels, as in the canonical decentralized detection problem [11]. We extend the canonical problem by considering a no-send value at each sensor node. Each sensor node and the fusion node can agree on a particular value that is not transmitted and can thereby cut transmission costs.

**Scenario N. No-send-zone quantizer:** Each sensor node maps its observation  $X_{\lambda_i}$  to one of  $D_i$  values

$$\gamma_{\lambda_i}(X_{\lambda_i}) : \mathcal{X}_{\lambda_i} \rightarrow \{1, \dots, D_i\}. \quad (9)$$

We assume that only one of the quantizer levels  $\tilde{d}_{\lambda_i}$  is not transmitted, so that the quantizer maps the no-send region to  $\tilde{d}_{\lambda_i}$

$$P_{T,i}(R_{\lambda_i}) = P_0(\gamma_{\lambda_i}(X_{\lambda_i}) \neq \tilde{d}_{\lambda_i}). \quad (10)$$

In Scenario N,<sup>1</sup> for a fixed choice of  $\gamma_{\lambda_i}$ , the quantizer value which has largest probability would be chosen as the no-send value. We denote the quantizer outputs as  $\mathbf{U} = (U_1, \dots, U_N)$ , where  $U_i = \gamma_{\lambda_i}(X_{\lambda_i})$  is the output of the  $i$ th quantizer. We denote the decision strategy as

$$\gamma = \left( \gamma_0, \{\lambda_i, \gamma_{\lambda_i}, \tilde{d}_{\lambda_i}\}_{i=1}^N \right). \quad (11)$$

For the quantizer transmission Scenario N, there are several relevant system-level optimization problems

$$\text{PB-N} : \min_{\gamma} P_E(\gamma) \text{ s.t. } C(\gamma) \leq c_0 \quad (12)$$

$$\text{PN-N} : \min_{\gamma} P_M(\gamma) \quad (13)$$

s.t.  $P_F(\gamma) \leq \alpha$ , and  $C(\gamma) \leq c_0$

$$\text{PD-N} : \max_{\gamma} \sum_{i=1}^N D_f(\gamma_i) \text{ s.t. } C(\gamma) \leq c_0. \quad (14)$$

### IV. OPTIMAL DECISION STRATEGY FOR CENSORING

In the decentralized detection literature, under Assumption 2.1, results for various criteria have shown the optimality of tests based on the likelihood ratio of the observations [4], [6], [13]. In problems PB-C, PN-C, and PD-C, we can obtain similar results for the optimal decision strategy (3) for the deterministic case. We develop the results for transmission Scenario C completely before considering Scenario N in Section VII. For flow of presentation, proofs of the theorems to be presented are given in the Appendix.

<sup>1</sup>Scenario N is a special case of the problem of maximizing detection performance under an average bit-rate constraint where the no-send symbol is entropy coded to zero bits. Observe that in the canonical decentralized detection problem, one could easily use entropy coding on the output of the sensor quantizers to reduce the average bit rate. Since the fusion node is assumed to know the distribution of  $l(U_i)$ , the entropy-coded values could be mapped back to the original  $\gamma_{\lambda_i}$ . A practical approach would be to choose the quantizer according to a distance criterion [12] and Huffman code the quantizer output.

**Theorem 4.1:** Given a constraint on the expected cost (and a false-alarm rate in PN-C) to be satisfied, over the set  $\Phi$ , the optimal decision strategy for PB-C, PN-C, and PD-C has sensor decision rules  $\phi_{\lambda_i}$  which are censored likelihood ratios of the form

$$\phi_{\lambda_i}(X_{\lambda_i}) = \begin{cases} l(X_{\lambda_i}), & l(X_{\lambda_i}) \in R_{\lambda_i} \\ \rho_{\lambda_i}, & \text{otherwise} \end{cases} \quad (15)$$

where  $R_{\lambda_i}^c = \{t_{1,i} < l(X_{\lambda_i}) \leq t_{2,i}\}$  is a single interval of  $l(X_{\lambda_i})$  and

$$\rho_{\lambda_i} = \frac{P_1(l(X_{\lambda_i}) \in R_{\lambda_i}^c)}{P_0(l(X_{\lambda_i}) \in R_{\lambda_i}^c)}. \quad (16)$$

The optimal fusion rule is of the form

$$\phi_0(\cdot) = \begin{cases} 1, & \\ \xi, & \prod_{i=1}^N \phi_{\lambda_i}(X_{\lambda_i}) = \tau \\ 0, & \end{cases} \quad (17)$$

where  $\xi$  is the randomization parameter, independent of  $X_{\lambda_i}$ , and arbitrary for problems PB-C and PN-C, which is chosen to meet the false-alarm rate constraint. (Note that randomization at the fusion node is due to point mass where all sensor nodes censor their likelihoods).

Theorem 4.1 implies that each sensor node chooses a single interval in which not to send its likelihood ratio. The no-send interval is deemed uninformative to the fusion node since it likely corresponds to a null event. At the fusion node, the product of the received likelihoods and the average likelihoods for the censoring nodes (these are known to the fusion node) is compared against a threshold to make a global decision.

Over the set of deterministic strategies, an expected cost constraint can be met with equality by choosing the rate of transmission, since  $P_{T,i} \in [0, 1]$  is continuous, while measurement cost  $c_{\lambda_i}$  typically has discrete levels. For each subset of the measurements, the sensor node could transmit anywhere from very infrequently to always. With the introduction of randomized strategies, an expected cost constraint can be met by also randomizing over the choice of measurement.

In Sections IV-A–C, we show that the optimal sensor decision rules  $g_{\lambda_i}$  over the set of randomized strategies are censored likelihood ratios. The results on randomization given here can be derived from the results for the centralized sequential problem with generalized cost given in [2]. The particular cost function we have considered allows us to address decentralized problems and explore the tradeoff between cost due to measurement and cost due to transmission. While we have chosen to use linear programming (LP) arguments, supporting hyperplane theorem arguments could also be used to prove the results [2], [7].

#### A. Dependent Randomization

Over the set of randomized strategies, we are interested in the complexity of the decision strategy in terms of the number of deterministic decision strategies  $K$  over which to randomize. By considering person-by-person-optimal (pbpo) sensor decision rules and using simple linear-programming arguments, we can show that for dependently randomized strategies,  $K \leq 2$  for problems PB-C and PD-C since they have a single constraint, and  $K \leq 3$  for problem PN-C since it has two constraints.

*Theorem 4.2:* Given a constraint on the expected cost to be satisfied, consider the problems PB-C, and PD-C. Over the set  $\bar{\Phi}$ , the optimal decision strategy randomizes between at most two deterministic decision strategies  $\phi^{(1)}$  and  $\phi^{(2)}$  with corresponding probabilities  $\mu^{(1)}$  and  $1 - \mu^{(1)}$ , where  $\phi^{(1)}$  and  $\phi^{(2)}$  are as given in (15) and (17).

In PN-C, the introduction of a false-alarm constraint increases the number of constraints in the LP by one, so we have the following result.

*Theorem 4.3:* Given expected cost and false-alarm rate constraints to be satisfied, over the set  $\bar{\Phi}$ , the optimal decision strategy for problem PN-C randomizes between at most three deterministic decision strategies.

### B. Independent Randomization

The performance of the optimal independently randomized strategy is at best as good as the performance of the optimal dependently randomized strategy, since the former is a subset of the latter. However, because of the ease of implementation, independently randomized strategies may be more relevant from a system design perspective. We find the number of sensor decision rules  $\mathcal{K}_i$ , over which to randomize at each node by considering pbpo sensor decision rules.

*Theorem 4.4:* Given a constraint on the expected cost (and a false-alarm rate constraint in PN-C) to be satisfied.

- 1) Over the set  $\bar{\Phi}$ , the optimal decision strategy for PB-C and PD-C randomizes between at most two deterministic sensor decision rules,  $\phi_{\lambda_i}^{(1)}$  and  $\phi_{\lambda_i}^{(2)}$ , with corresponding probabilities  $\mu_i$  and  $1 - \mu_i$  at each sensor node, where  $\phi_{\lambda_i}^{(1)}$  and  $\phi_{\lambda_i}^{(2)}$  are as given in (15).
- 2) Over the set  $\bar{\Phi}$ , the optimal decision strategy for PN-C randomizes between at most three deterministic sensor decision rules at each sensor node.

The optimal independently randomized strategy may involve randomization over two send rates. Consider fixing the choice of measurement at all sensor nodes and the choice of send rate at all sensor nodes except for node  $k$ . The error probability  $P_E$  is monotone decreasing in the send rate for node  $k$ , but not necessarily convex. It is possible then that randomization over two send rates lowers error probability compared with selecting a fixed send rate.

### C. System Design Perspective

Theorems 4.2–4.4 provide important guidelines on the design of sensor networks. For the proposed transmission scenario, each sensor node need only have two modes of operation corresponding to two sensor decision rules. In general, the choice of randomization probability and send rate at each sensor node requires joint optimization over all sensor nodes. From the system design perspective, we are particularly interested in settings in which the joint optimization separates into independent optimization at each sensor node. We point out two settings in which the joint optimization separates:

- 1) PD-C, where the distance measure and expected cost at each sensor node are considered;

- 2) problems of maximizing the deflection at the fusion node subject to a constraint on the expected cost at each sensor node.

In setting 1), the detection performance is additive over the sensor nodes, so each node maximizes its distance measure subject to its own cost. Setting 2) is important in many applications and was shown to be separable in [5], since the deflection measure can be written as a particular A-S distance at each sensor node.

## V. DUAL PROBLEMS

In many applications, we are interested in maximizing the system lifetime subject to constraints on the detection performance. For example, we may be interested in finding the smallest battery or lowest cost solar cell which will achieve a given detection performance. If transmission cost is dominated by how frequently nodes transmit, then we are interested in Scenario C, while if data rate is also relevant, then Scenario N is of interest. We may also be interested in constraining the false-alarm rate and miss probability individually instead of constraining the error probability.

Such problems of minimizing expected cost subject to performance constraints are dual to the problems considered in (6)–(8) and include the following, where DB-C represents the dual Bayesian problem in Scenario C, and so on:

$$\text{DB-C : } \min_{\phi} C(\phi) \text{ s.t. } P_E(\phi) \leq e_0 \quad (18)$$

$$\text{DN-C : } \min_{\phi} C(\phi) \quad (19)$$

$$\text{s.t. } P_F(\phi) \leq \alpha, \text{ and } P_M(\phi) \leq \beta$$

$$\text{DD-C : } \min_{\phi} C(\phi) \text{ s.t. } \sum_{i=1}^N D_f(\phi_i) \geq d_0 \quad (20)$$

where  $e_0$ ,  $\alpha$ , and  $\beta$  are the largest tolerable error-probability, false-alarm rate, and miss probability, respectively, and  $d_0$  is the minimum acceptable A-S distance. For the dual problems, extensions of the results on optimality of likelihood-ratio tests and randomization follow from LP duality arguments.

## VI. ON-OFF SENSOR NODES

The simplest application of decentralized detection problems with a constraint on the expected cost is for sensor nodes which are either awake or asleep (on-off sensor nodes). The sensor node may have only one modality, or the cost of measurement across all modalities may be small compared with the cost of computing likelihoods. In such problems, it is reasonable to consider the total cost of measurement *and* computation as the cost of the node being awake. In practice, the cost of the detector implementation (i.e., complexity) is related to properties of the likelihood ratio. Although the cost of implementation is a relevant issue, the largest cost savings will most likely be achieved from turning the entire sensor node off.

We assume that when a sensor node is awake, it measures all sensing modalities, so the measurement is the entire space of observations (i.e.,  $X_{\lambda_i} = \mathbf{X}_i$ ). We include the cost of computing local decisions with the cost of measuring  $\mathbf{X}_i$  as  $c_i$ . Since there is no choice of measurement, only randomization over the

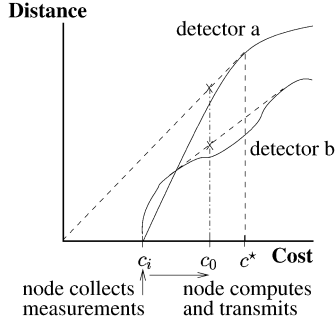


Fig. 2. Hypothetical curves for distance as a function of expected cost.

choice of send rate and asleep or awake is necessary. When  $c_i \gg c_{T,i}$ , we would expect randomization between awake and asleep to play a large role in achieving the optimal detection performance and meeting a constraint on the expected cost. When  $c_{T,i} \gg c_i$ , randomization over the choice of send rates is likely to play the important role.

As mentioned before, in many problems, likelihood-ratio-based approaches are frequently substituted for tests which maximize the A-S distance or deflection measure. We elaborate on distance-measure approaches for on-off sensor nodes. In the on-off scenario, the choice of measurement is fixed, so the sensor node computes the likelihood ratio when it is awake. Based on Theorem 4.4, it is clear that randomization over two sensor decision rules is optimal. Randomization between sleeping and a fixed send rate is possible, as well as randomization between two send rates and corresponding no-send regions  $R_i^{(1)}$  ( $\tilde{d}^{(1)}$ ), and  $R_i^{(2)}$  ( $\tilde{d}^{(2)}$ ).

For Scenario C, the problem of maximizing distance for on-off sensor nodes can be described as

$$\begin{aligned} \max_{\phi_i} \quad & D_f(\phi_i) \\ \text{s.t.} \quad & C(\phi_i) \leq c_0 \end{aligned}$$

where the decision rule  $\phi_i = (g_i, R_i^{(1)}, \mu_i^{(1)}, R_i^{(2)})$  is to be obtained,  $\mu_i^{(1)}$  is the fraction of time that the sensor node uses the send region  $R_i^{(1)}$ ,  $1 - \mu_i^{(1)}$  is the fraction of time that the sensor node uses the send region  $R_i^{(2)}$ , and the desired resource level is denoted as  $c_0$ .

The choice of transmit rate when awake, and how often a sensor node should sleep, can be determined graphically and depends on the curvature of  $D_f$  and the ratio  $c_i/c_{T,i}$ . Fig. 2 shows the distance measure as a function of the expected cost for two possible detectors. For detector B, it is optimal to randomize between two choices of send rate. For detector A, if  $c_0 < c^*$ , then it is optimal for the sensor node to sleep for the fraction of time  $1 - (c_0/c^*)$  and transmit at the rate  $(c^* - c_i)/c_{T,i}$  when awake. For  $c_0 > c^*$ , the sensor node stays awake with the fixed transmit rate  $(c_0 - c_i)/c_{T,i}$ . Since the sensor node stays awake for  $c_0 > c^*$ , we define  $c^*$  as the always-awake cost, and the corresponding always-awake transmit rate as  $\kappa^*$ . We note that convexity of  $D_f$  is a sufficient but not necessary condition for the existence of an always-awake region.

### A. Example: Kullback–Leibler Divergence

Consider the Kullback–Leibler (K-L) divergence, which is the particular case of A-S distance with  $f(Y) = -\log(Y)$ . K-L divergence is relevant for N-P testing when the number of observations tends to infinity [14, Stein’s Lemma, p. 309]. Let us examine the decentralized hypothesis-testing problem of detecting a mean shift in Gaussian noise with  $N$  censoring sensor nodes

$$H_0 : X_i \sim \mathcal{N}(0, \sigma_i^2) \text{ vs. } H_1 : X_i \sim \mathcal{N}(\theta_i, \sigma_i^2) \quad (21)$$

where  $\theta_i > 0$  is the mean and  $\sigma_i^2$  is the noise variance at sensor node  $i$ . For the mean-shift problem, the likelihood ratio of the observation is monotone in  $X_i$ , so the likelihood does not have to be computed locally and there is no computational cost at the sensor node. It is, hence, sufficient to censor the observation  $X_i$  directly.

Define the complementary Gaussian cumulative distribution function as

$$Q(t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-v^2/2} dv$$

where  $v$  is the variable of integration. The K-L distance for the censoring test  $\phi_i$  can then be written as

$$\begin{aligned} D_f(\phi_i) = & \frac{\theta_i}{\sigma_i \sqrt{2\pi}} \left( \exp\left(-\frac{\tau_2^2}{2}\right) - \exp\left(-\frac{\tau_1^2}{2}\right) \right) \\ & + \frac{\theta_i^2}{2\sigma_i^2} (1 - Q(\tau_1) + Q(\tau_2)) \\ & + \log\left(\frac{1 - \kappa_{1,i}}{1 - \kappa_{0,i}}\right) (1 - \kappa_{1,i}) \end{aligned}$$

where  $\tau_k = (t_{k,i} - \theta_i)/\sigma_i$ ,  $k = 1, 2$ , and  $R_i^c = \{X_i : X_i \in [t_{1,i}, t_{2,i}]\}$  is the censoring region, and  $\kappa_{j,i} = P_j^i(R_i)$  is the probability of the send region under  $H_j$ ,  $j = 0, 1$

$$\begin{aligned} \kappa_{1,i} &= 1 - Q(\tau_1) + Q(\tau_2) \\ \kappa_{0,i} &= 1 - Q(t_{1,i}) + Q(t_{2,i}). \end{aligned}$$

Fig. 3 shows the optimal K-L divergence (found by brute-force search) for each transmit rate, over various values of  $\sigma_i$  and  $\theta_i = 1$ . For small transmit rates, the censoring region maximizing the K-L distance was found to be  $R_i^c = \{X_i \leq t_i\}$  (i.e., only a single threshold is required), so no search is required, and  $t_i$  is chosen according to the transmit rate.

Let  $c_T$  be the cost of transmission, and  $c_S$  be the cost of the sensor node (measurement and local computation) in the on-off scenario. Consider  $\sigma_i = \theta_i = 1$ ,  $c_T = 1/3$ , and  $c_S = 2/3$ , corresponding to the cost of the sensor node twice as large as the cost of transmission, and the total cost  $c_S + c_T$  normalized to one. By drawing a tangent to the K-L divergence plot in Fig. 3, we find that the always-awake cost and always-awake transmit rate are  $c^* = 0.79$  and  $\kappa^* = 0.36$ , respectively. When the desired resource level is less than 79% of the total cost of staying awake and transmitting every observation, (i.e.,  $c_0 < 0.79$ ), the sensor node switches between awake with transmit rate of 0.36 and asleep. For example, if  $c_0 = 2/3$ , the sensor node is awake

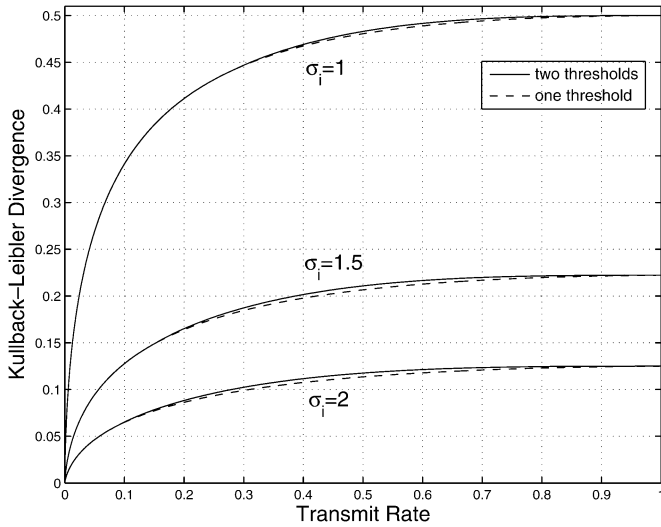


Fig. 3. K-L divergence as a function of transmit rate for detecting a mean shift in Gaussian noise using one or two thresholds for censoring.

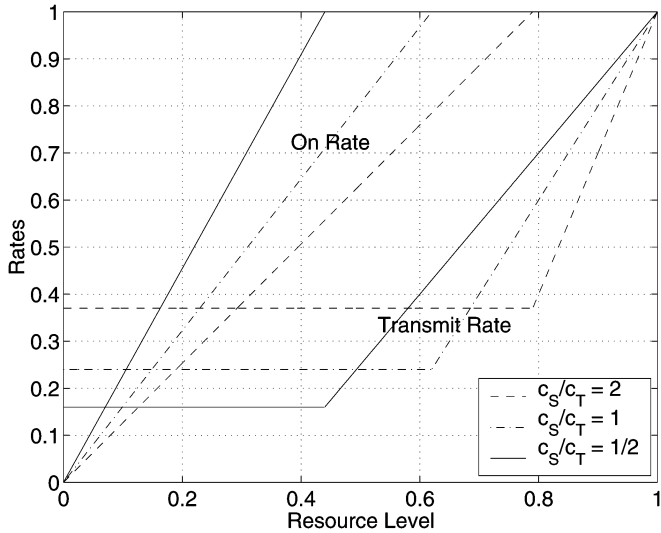


Fig. 4. K-L-distance-optimal on and transmit rates for detecting a mean shift in Gaussian noise.

for 85% of the time. For  $c_0 = 0.83$ , the sensor node stays awake and transmits 50% of the time.

Similarly, we find the optimal on-off and transmit rates for  $c_S/c_T = 1$ , and  $1/2$ , and summarize the results in Fig. 4. We observe that as the resource level increases (e.g., due to larger batteries): 1) below the always-awake region, the transmit rate is fixed and the on rate increases and 2) in the always-awake region, the transmit rate increases and the sensor node stays awake. As we consider different sensor node hardware and the relative cost of transmission to sensor node operation increases, we find that below the always-awake region, the transmit rate decreases and the fraction of time the sensor node is awake increases.

### B. Example: Error Probability

Consider design in terms of error probability in which one, two, three, or four sensor nodes using Scenario C for transmission and a fusion node, which has its own observations, are

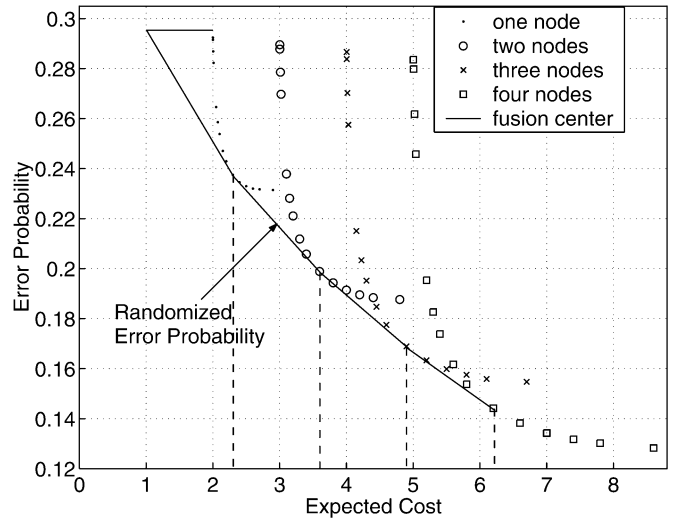


Fig. 5. Error probability for detecting a mean shift in Gaussian noise with one to four nodes, where the fusion node has its own observations,  $\pi_0 = 0.6$ ,  $c_S = c_T = 1$ , and  $c_F = 1$ .

faced with the detection problem (21) (i.e., problem PB-C). Let  $\sigma_i = \theta_i = 1$  and fix the cost of operating the fusion node to  $c_F = 1$ , the cost of transmission from the sensor node to the fusion node to  $c_T = 1$ , and the cost of operating each sensor node to  $c_S = 1$ . At the fusion node, there is no cost due to transmission, and complete observations are available for decision making.

Fig. 5 shows the minimum error probability as a function of expected cost for  $\pi_0 = 0.6$ . When no fusion node is available, the prior probability determines the best guess, so the error probability is fixed at  $\min\{\pi_0, 1 - \pi_0\}$ . By simulation, we obtain the error probability and associated cost when one sensor node or both sensor nodes are awake. It is simple to determine the censoring intervals by applying results in [5] in which it is shown that the lower thresholds of the censoring intervals is zero for sufficiently small communication rates and sufficiently large  $\pi_0$ . In general, assigning equal rates of transmission to the sensor nodes may not be optimal (see [15] for a counterexample in the canonical problem). For this particular example, we find experimentally that when two sensor nodes are on, it is not advantageous for the sensor nodes to choose unequal rates of transmission.

By setting the nodes to sleep and transmitting only a fraction of the time, it is possible to meet constraints on the expected cost, while minimizing the error probability. Let  $c_0$  denote the maximum resource level across all sensor nodes, as in (4). By drawing tangents to the error-probability curves, we determine the approximate always-awake cost for one, two, three, or four sensor nodes as  $c_1^* = 2.3$ ,  $c_2^* = 3.6$ ,  $c_3^* = 4.9$ , and  $c_4^* = 6.2$ , respectively, for  $\pi_0 = 0.6$ . The corresponding always-awake transmit rates are  $\kappa_i^* = 0.3$ ,  $i = 1, 2, 3, 4$ .

The on and censoring rates obtained in Section VI-A for K-L divergence are similar to the rates obtained in this section for error probability. For K-L divergence, the rates for each sensor node are identical since the sensor observations are identically distributed. Consider a resource level of  $c_0 = 0.8$  to be satisfied for  $c_S = c_T = 1$ , and  $c_F = 1$ . In terms of K-L divergence,

the optimal on rate and transmit rate are (from Fig. 4 at a normalized resource level of 0.4) 0.65 and 0.24, respectively. In terms of error probability, for  $\pi_0 = 0.6$ , the optimal on rate and transmit rate are (from Fig. 5) 0.62 and 0.3, respectively, when considering only a single node. For two nodes, one node is completely awake and the other is awake 20% of the time (averages to 60% awake across the two nodes), both with a transmit rate of 0.3.

## VII. OPTIMAL DECISION STRATEGY FOR NO-SEND-ZONE QUANTIZER

The no-send-zone quantizer belongs to a class of general quantization problems examined by Tsitsiklis [6]. Applying definitions and theorems from [6], results for decentralized detection problems with a constraint on the expected cost can be derived. See the Appendix for a proof outline.

*Theorem 7.1:* Given a constraint on the expected cost (and a false-alarm rate in PN-N) to be satisfied, over the set  $\Gamma$  or  $\bar{\Gamma}$ , the optimal decision strategy for PB-N, PN-N, and PD-N has sensor decision rules  $\gamma_{\lambda_i}$  which are likelihood-ratio quantizers (LRQs) of the form

$$\gamma_{\lambda_i}(X_{\lambda_i}) = \begin{cases} 1, & l(X_{\lambda_i}) < t_{1,i} \\ 2, & t_{1,i} < l(X_{\lambda_i}) < t_{2,i} \\ \vdots & \\ D, & t_{D_i-1,i} < l(X_{\lambda_i}) \end{cases} \quad (22)$$

where the set of thresholds,  $0 \leq t_{1,i} \leq \dots \leq t_{D_i-1,i} \leq \infty$ , depend on the choice of  $\lambda_i$ , one of the quantizer values  $\tilde{d}_{\lambda_i}$  is censored, and the assignment of  $\gamma_{\lambda_i}$  is arbitrary when  $l(X_{\lambda_i})$  is equal to one of the thresholds.

The optimal fusion rule has the form

$$\gamma_0(\mathbf{U}) = \begin{cases} 1, & \prod_{i=1}^N l(U_i) > \tau \\ \xi, & \prod_{i=1}^N l(U_i) = \tau \\ 0, & \prod_{i=1}^N l(U_i) < \tau \end{cases} \quad (23)$$

where  $l(U_i) = P(U_i|H_1)/P(U_i|H_0)$ ,  $\xi$  is a randomization parameter that does not depend on  $\mathbf{U}$ , and the  $\{\tilde{d}_{\lambda_i}\}_{i=1}^N$  are known at the fusion node.

Theorem 7.1 states that the optimal sensor rule at each node is a LRQ. At each sensor node, one of the quantization levels is a no-send level, which is known to the fusion node. At the fusion node, a likelihood ratio is computed based on the received quantized values (if a value is not received, it is assumed to be  $\tilde{d}_{\lambda_i}$ ), and the known distribution of  $l(X_i)$ .

Whereas in Scenario C, independent randomization may improve performance, in Scenario N, Theorem 7.1 describes the equivalence of deterministic and independently randomized strategies. For dependently randomized strategies  $\hat{\Gamma}$ , we have the following result.

*Theorem 7.2:* Given a constraint on the expected cost (and a false-alarm rate constraint in PN-N) to be satisfied.

- 1) Over the set  $\hat{\Gamma}$ , the optimal decision strategy for PB-N and PD-N randomizes between at most two deterministic decision strategies  $\gamma^{(1)}$  and  $\gamma^{(2)}$  with corresponding probabilities  $\mu^{(1)}$  and  $1 - \mu^{(1)}$ , where  $\gamma^{(1)}$  and  $\gamma^{(2)}$  are as given by (22) and (23).

- 2) Over the set  $\hat{\Gamma}$ , the optimal decision strategy for PN-N randomizes between at most three deterministic decision strategies.

For completeness, we also describe the dual problems for Scenario N

$$\text{DB-N} : \min_{\gamma} C(\gamma) \text{ s.t. } P_E(\gamma) \leq e_0 \quad (24)$$

$$\text{DN-N} : \min_{\gamma} C(\gamma) \text{ s.t. } P_F(\gamma) \leq \alpha, \text{ and } P_M(\gamma) \leq \beta \quad (25)$$

$$\text{DD-N} : \min_{\gamma} C(\phi) \text{ s.t. } \sum_{i=1}^N D_f(\gamma_i) \geq d_0. \quad (26)$$

## VIII. CONCLUSION

Detection problems with a constraint on the expected cost arising from measurement have been examined by various authors. We have considered a constraint on the expected cost arising from measurement and transmission. A send/no-send scenario for transmission led to a tractable and realistic problem formulation where likelihood-ratio-based detectors are optimal. We found that randomization of the measurement and choice of transmit rate could be used to meet a constraint on the expected cost and to optimize detection performance. In the examples, some reasonable settings for the detection problem and the sensor node led to interesting results on the design of sensor nodes. For on-off sensor nodes, we found an always-awake region above a particular resource level. When the cost of measurement is large, randomization over measurement was particularly useful, while censoring effectively reduces transmission cost independent of the particular cost per transmission. For the problem of detecting a mean shift in Gaussian noise, we found that design in terms of error probability gives similar results to design in terms of K-L distance. While we have considered conditionally independent observations, a future area of work is to develop optimal decision strategies for dependent observations. In particular, formulations based on distance metrics may be tractable and lead to approaches for determining optimal system parameters.

## APPENDIX

### A. Proof Outline of Theorem 4.1

Fix a particular subset  $\lambda_i$  of the observations to be measured at each sensor node and apply results from [4] and [5]. Consider a censoring region consisting of two nonoverlapping no-send intervals of  $l(X_i)$ . By the no-point mass assumption, there exists a consolidation from two no-send intervals to one no-send interval that preserves the no-send probability under both hypotheses  $P_j(R_i^c)$ ,  $j = 0, 1$ . It is possible to show that the error probability does not increase (and similarly for PN-C and PD-C) due to the consolidation. The complete proof is rather lengthy and can be found in the [4], [5]. ■

We clarify a mistake in the result for A-S distances provided in [4]. It is claimed that maximizing the A-S distance  $D_f\left(\prod_{i=1}^N \phi_{\lambda_i}(X_{\lambda_i})\right)$  is the same as maximizing



$\prod_{i=1}^N D_f(\phi_{\lambda_i}(X_{\lambda_i}))$  due to the conditional independence of the observations. However, such a claim is not true in general (e.g.,  $f(X) = X \log X$ ). Instead, for the detection problems under consideration, we maximize the sum of the A-S distances over the individual sensor decision rules. For the particular case of maximizing the deflection at the fusion node, the equivalence with maximizing the A-S distance  $f(Y) = Y^2$  at each sensor node is shown in [5].

#### B. Proof of Theorem 4.2

Consider the pbpo solution obtained by fixing all sensor decision rules  $\phi_{\lambda_j}(X_{\lambda_j})$ ,  $j \neq i$  and the fusion rule. We can write the average error probability as

$$P_E = \sum_{k=1}^{K_i} \mu_i^{(k)} P_E^{(k)} \quad (27)$$

where  $P_E^{(k)}$  is the error probability, which depends on the choice of all the sensor decision rules and, in particular, on the decision rule for sensor  $i$ ,  $\phi_i^{(k)}$ . Similarly, we can write  $C(\phi) = \sum_{k=1}^{K_i} \mu_i^{(k)} C(\phi^{(k)})$ , where  $\phi^{(k)}$  is the decision strategy employing the sensor decision rule  $\phi_i^{(k)}$  for sensor  $i$ . Problem PB-C is then a LP problem in  $\mu_i = (\mu_i^{(1)}, \dots, \mu_i^{(K_i)})$  with the additional constraints  $\mu_i^{(k)} \geq 0$ , and  $\sum_{k=1}^{K_i} \mu_i^{(k)} = 1$ . In standard form, the LP problem has two constraints, which implies that a basic feasible solution has at most two components of  $\mu_i$  that are strictly nonzero. Since the form of the pbpo solution is independent of the choice of the other sensor decision rules and the fusion rule, the class of pbpo solutions is also globally optimal. The same arguments can be applied to problem PD-C by replacing  $P_E$  with  $D_f$ . ■

#### C. Proof of Theorem 4.4

The proof is very similar to the proof of Theorem 4.2. Consider the pbpo solution obtained by fixing all sensor decision rules  $\phi_{\lambda_j}(X_{\lambda_j})$ ,  $j \neq i$ , and the fusion rule. We can write the average error probability as

$$P_E(\phi) = \sum_{k=1}^{K_i} \mu_i^{(k)} P_E^{(k)} \quad (28)$$

where  $P_E^{(k)}$  is the error probability, which depends on the choice of all the sensor decision rules and, in particular, the decision rule for sensor node  $i$ ,  $\phi_i^{(k)}$ . Similarly, we can write  $C(\phi) = \sum_{k=1}^{K_i} \mu_i^{(k)} C(\phi^{(k)})$ , where  $\phi^{(k)}$  is the decision strategy employing the sensor decision rule  $\phi_i^{(k)}$  for sensor node  $i$ . Problem PB-C is then a LP problem in  $\mu_i = (\mu_i^{(1)}, \dots, \mu_i^{(K_i)})$  with the additional constraints  $\mu_i^{(k)} \geq 0$ , and  $\sum_{k=1}^{K_i} \mu_i^{(k)} = 1$ . In standard form, the LP problem has two constraints, which implies that a basic feasible solution has at most two components of  $\mu_i$  that are strictly nonzero. For the N-P problem PN-C, the number of constraints increases by one, so at most three components of  $\mu_i$  are strictly nonzero. ■

#### D. Proof of Theorem 7.1

Fix a particular subset  $\lambda_i$  of the observations to be measured at each sensor. The main idea of the proof is to show that the sensor decision rule at sensor  $i$  can be selected from the class of LRQs when all other sensor decision rules  $\gamma_{\lambda_j}(X_{\lambda_j})$ ,  $j \neq i$ , and the fusion rule  $\gamma_0$  is fixed (i.e., the pbpo solution). Since the class of LRQs is optimal at each sensor independent of the choice of other sensor rules, LRQs are also globally optimal.

In the pbpo setting, the expressions for  $P_M$ ,  $P_F$ , and  $D_f$  simplify considerably and depend only on  $\gamma_{\lambda_i}$  and  $R_{\lambda_i}$  (e.g., see [7, p. 311]). We have

$$\begin{aligned} P_F &= E_{X_{\lambda_1}, \dots, X_{\lambda_N} | H_0} [\gamma_0(\gamma_{\lambda_1}(X_{\lambda_1}), \dots, \gamma_{\lambda_N}(X_{\lambda_N}))] \\ &= E_{X_{\lambda_i} | H_0} [Z_0(\gamma_{\lambda_i}(X_{\lambda_i}))] \\ &= \sum_{d=1}^{D_i} Z_0(d) P_0^i(\gamma_{\lambda_i}(X_{\lambda_i}) = d) \end{aligned} \quad (29)$$

where  $Z_0(d)$  depends on the output,  $U_i = d$ , of the  $i$ th quantizer, and the third equality is almost always equal (a.e.) to the second equality (i.e., except on a set of probability zero). Similarly, we can obtain

$$\begin{aligned} P_M &= \sum_{d=1}^{D_i} Z_1(d) P_1^i(\gamma_{\lambda_i}(X_{\lambda_i}) = d) \\ D_f(\gamma_i) &= E_{X_{\lambda_i} | H_0} [f(l(U_i))] \\ &= \sum_{d=1}^{D_i} P_0^i(\gamma_{\lambda_i}(X_{\lambda_i}) = d) \\ &\quad \times f\left(\frac{P_1^i(\gamma_{\lambda_i}(X_{\lambda_i}) = d)}{P_0^i(\gamma_{\lambda_i}(X_{\lambda_i}) = d)}\right) \end{aligned} \quad (30)$$

where  $Z_1(d)$  depends on the output  $U_i = d$  of the  $i$ th quantizer. The constraint on the expected cost (4) becomes

$$P_0^i(\gamma_{\lambda_i}(X_{\lambda_i}) \neq \tilde{d}_{\lambda_i}) \leq \tilde{\alpha} \quad (31)$$

where  $\tilde{\alpha}$  is independent of the choice of the  $i$ th sensor decision rule.

Applying [6, Proposition 3.5] directly, we can conclude the optimality of LRQs in PB-N, PN-N, and PD-N under the no-point mass assumption. An alternative approach for PB-N and PN-N is to consider the following problem P1 in which the optimality of LRQs has been established [16]:

$$\begin{aligned} \text{P1 : } \min_{\gamma_{\lambda_i}} & \sum_{d=1}^{D_i} Z_1(d) P_1^i(\gamma_{\lambda_i}(X_{\lambda_i}) = d) \\ \text{s.t.} & P_0^i(\gamma_{\lambda_i}(X_{\lambda_i}) = d) = \alpha_d \quad \forall d, \\ & \text{and } \sum_{d=1}^{D_i} \alpha_d = 1. \end{aligned} \quad (32)$$

Let  $\gamma^\dagger$  be the optimal solution to PN-N, and let  $\alpha_d^\dagger$  be the value of  $P_0^i(\gamma^\dagger(X_{\lambda_i}) = d)$ ,  $d = 1, \dots, D_i$ , achieved by  $\gamma^\dagger$ . Fix  $\alpha_d = \alpha_d^\dagger$ ,  $d = 1, \dots, D_i$ , in P1. Since LRQs are optimal in P1 for this choice of  $\{\alpha_d\}_{d=1}^{D_i}$ , LRQs are optimal in PN-N as well. A similar argument works for relating PB-N to P1.

Determining the fusion rule is easy since we have a centralized detection problem in the sensor outputs  $\mathbf{U}$ . When the sensor

does not send, the fusion center attributes no-send to the quantizer value  $\hat{d}_i$ .

Using results from [6], it is possible to argue that PB-N, PN-N, and PD-N have the same solution over  $\bar{\Gamma}$  as over  $\Gamma$ . In [6], a more general definition of LRQs is given where the assignment of the quantizer value  $\gamma_{\lambda_i}$  when  $l(X_{\lambda_i})$  is equal to one of the thresholds, is not arbitrary. It is shown that over the set of dependently randomized strategies, the general class of LRQs is optimal for PB-N, PN-N, and PD-N. Under the no-point-mass assumption, the probability of the event that  $l(X_{\lambda_i})$  is equal to one of the thresholds is zero, so decision strategies in  $\bar{\Gamma}$  have the same detection performance as decision strategies in  $\Gamma$  except on a set of probability zero. ■

## REFERENCES

- [1] K.-S. Fu, Y. T. Chien, and G. P. Cardillo, "A dynamic programming approach to sequential pattern recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, pp. 313–326, May 1986.
- [2] E. Ertin, "Polarimetric processing and sequential detection for automatic target recognition systems," Ph.D. dissertation, Ohio State Univ., Columbus, OH, 1999.
- [3] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
- [4] 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.
- [5] S. Appadwedula, "Energy-efficient sensor networks for detection applications," Ph.D. dissertation, Univ. Illinois at Urbana–Champaign, Urbana, IL, 2003.
- [6] J. N. Tsitsiklis, "Extremal properties of likelihood-ratio quantizers," *IEEE Trans. Commun.*, vol. 41, no. 4, pp. 550–558, Apr. 1993.
- [7] ———, "Decentralized detection," in *Advances in Statistical Signal Processing*, H. V. Poor and J. B. Thomas, Eds. Greenwich, CT: JAI Press, 1991, vol. 2, Signal Detection, pp. 297–344.
- [8] S. Cho and A. P. Chandrakasan, "Energy efficient protocols for low duty cycle wireless microsensor networks," in *Proc. ICASSP*, vol. 4, Salt Lake City, UT, May 2001, pp. 2041–2044.
- [9] J. N. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Trans. Autom. Control*, vol. 30, no. 5, pp. 440–446, May 1985.
- [10] S. M. Ali and S. D. Silvey, "A general class of coefficients of divergence of one distribution from another," *J. Royal Statist. Soc.*, ser. B, vol. 28, no. 1, pp. 131–142, 1966.
- [11] R. R. Tenney and N. R. Sandell Jr, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 17, pp. 501–510, Jul. 1981.
- [12] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Trans. Inf. Theory*, vol. 33, no. 6, pp. 773–787, Nov. 1987.
- [13] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I—Fundamentals," in *Proc. IEEE*, vol. 85, Jan. 1997, pp. 54–63.
- [14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [15] M. Cherikh and P. B. Kantor, "Counterexamples in distributed detection," *IEEE Trans. Inf. Theory*, vol. 38, no. 1, pp. 162–165, Jan. 1992.
- [16] R. Viswanathan, A. Ansari, and S. C. A. Thomopoulos, "Optimal partitioning of observations in distributed detection," Dept. Elec. Eng., Southern Illinois Univ., Tech. Rep. TR-SIU-EE-3-87, 1987.



**Swaroop Appadwedula** (M'04) 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 at Lincoln Laboratory, Massachusetts Institute of Technology (MIT), Lexington, where he currently works on signal processing problems in antenna array. 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) received the B. Tech. degree (Silver Medal Honors) from the Indian Institute of Technology, Bombay, in 1985, the M.S. degree from Carnegie-Mellon University, Pittsburgh, PA, in 1987, and the Ph.D. degree from the University of Illinois at Urbana–Champaign (UIUC), in 1992, all in electrical engineering.

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

Dr. Veeravalli has received the following awards for research and teaching: 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 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*. 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 at 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 in the Department of Electrical and Computer Engineering, the Coordinated Science Laboratory, and the Beckman Institute. He was on sabbatical leave at the University of Washington in Spring 1995 and at the University of California at Berkeley in Spring 2002. In the Spring semester of 1999, he served as the Texas Instruments Visiting Professor at Rice University. He is an author of two DSP laboratory textbooks. 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.

Dr. Jones was selected as the 2003 Connexions Author of the Year. He is currently serving on the Board of Governors of the IEEE Signal Processing Society.