Asymptotic Results for Decentralized Detection in Power Constrained Wireless Sensor Networks

Jean-François Chamberland, *Student Member, IEEE,* and Venugopal V. Veeravalli, *Senior Member, IEEE,*

***Abstract*—In this paper, we study a binary decentralized detection problem in which a set of sensor nodes provides partial information about the state of nature to a fusion center. Sensor nodes have access to conditionally independent and identically distributed observations, given the state of nature, and transmit their data over a wireless channel. Upon reception of the information, the fusion center attempts to accurately reconstruct the state of nature. Specifically, we extend existing asymptotic results about large sensor networks to the case where the network is subject to a joint power constraint, and where the communication channel from each sensor node to the fusion center is corrupted by additive noise. Large deviation theory is used to show that having identical sensor nodes, i.e., each node using the same transmission scheme, is asymptotically optimal. Furthermore, a performance metric by which sensor node candidates can be compared is established. We supplement the theory with examples to illustrate how the results derived in this paper apply to the design of practical sensing systems.**

***Index Terms*—Communication systems, decision-making, multi-sensor systems, radio communication, signal detection.**

I. Introduction

A TYPICAL binary decentralized detection sensing system is one where geographically dispersed sensor nodes receive information about the state of nature H. Each sensor node is required to transmit a summary of its own observation to a fusion center. Based on the received data, the fusion center produces an estimate of the state of nature. In the Bayesian problem formulation, the probability of error at the fusion center is minimized. While in the Neyman–Pearson formulation, the probability of type II error is minimized, subject to a constraint on the type I error probability.

The framework of detection theory was first extended to a decentralized setting by Tenney and Sandell [1]. Evidently, the performance of a decentralized system is suboptimal in comparison with its centralized counterpart, as information may be lost in local processing and transmission. Nonetheless, factors such as cost, communication bandwidth, and reliability may motivate the use of a decentralized detection system. Besides, in systems with a large number of sensor nodes, uncompressed information could flood and overwhelm the fusion center. The design goal of the decentralized detection problem is to jointly optimize the decision rules at the sensor nodes and at the fusion center, as to minimize the probability of detection error. The crux of the problem is to pick a quantization function for every sensor node. Once this is achieved the fusion center faces a classical hypothesis testing problem, and the optimal decision rule at the fusion center can be established based on the standard likelihood ratio test.

The problem of decentralized detection has received much attention in the literature. For instance, previous work on this topic includes decentralized detection by a large number of sensor nodes [2], [3], sequential decentralized detection [4], and decentralized detection for dependent observations [5]. The reader is referred to Tsitsiklis [6], Viswanathan and Varshney [7], Blum *et al.* [8], and to the references contained therein for a summary of the early work in this field. Most results on decentralized detection assume that each sensor node produces a finite-valued function of its observation, which is conveyed reliably to the fusion center. In a wireless system, this assumption of reliable transmission may fail as information is transmitted over a noisy channel [9]. We, therefore, consider the alternative problem formulation where the fusion center only has access to a noisy version of the sent messages, and where the wireless sensing system is limited by a joint power constraint on the sensor nodes. The joint power constraint on the sensor nodes will prove suitable for the design of energy efficient sensing systems.

The remainder of this paper is as follows. In Section II, we introduce a mathematical framework for the study of the decentralized detection problem in the context of power constrained wireless sensing systems. In Section III, we show how having identical sensor nodes is asymptotically optimal, as the power constraint increases to infinity. Moreover, we present a meaningful metric to compare the performance of sensor nodes in large systems. The subsequent section contains numerical examples that illustrate the pertinence of our results. In Section V, we discuss the Neyman–Pearson variant of the decentralized detection problem. Finally, we give our conclusions in the last section.

II. Statement of the Problem

Let $\{Y_\ell\}$ be a sequence of random observations, each taking values in a measurable space $(\mathcal{Y}, \mathcal{F})$. Sensor node ℓ has access to observation Y_ℓ and is required to send a summary $\gamma_\ell(Y_\ell)$ of its own observation to the fusion center. Information is transmitted

Manuscript received July 15, 2003; revised February 1, 2004. The work of J.-F. Chamberland was supported by a Motorola Center for Communication Fellowship. This paper was supported in part by the National Science Foundation (NSF) under the CAREER/PECASE Award CCR-0049089. This paper was presented in part at the IEEE International Symposium on Information Theory, Yokohama, Japan, June 2003.

The authors are with the Electrical and Computer Engineering Department and the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801 USA (e-mail: chmbrlnd@uiuc.edu; vvv@uiuc.edu).

Digital Object Identifier 10.1109/JSAC.2004.830894

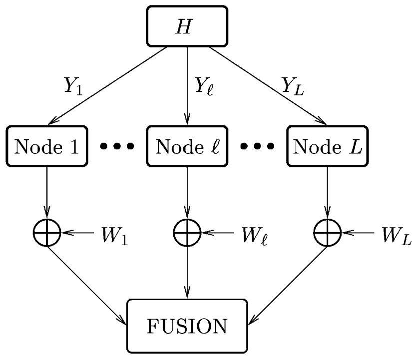

Fig. 1. Block diagram of a wireless sensing system where each sensor node transmits information to the fusion center over a wireless channel.

over a communication channel. The fusion center receives degraded information U_ℓ from node ℓ, which is of the form

$$U_\ell = \gamma_\ell(Y_\ell) + W_\ell \tag{1}$$

where W_ℓ is additive noise. We consider the simple situation, where the observations $\{Y_\ell\}$ are independent and identically distributed (i.i.d.), conditioned on the state of nature H, and where the additive noise is also i.i.d. across sensor nodes. The hypothesis testing problem consists of deciding, based on the sequence $\{U_\ell\}$, whether the law generating $\{Y_\ell\}$ is $\mathcal{P}_0$ corresponding to hypothesis H_0, or $\mathcal{P}_1$ corresponding to hypothesis H_1. This setting is illustrated in Fig. 1. We note that the number of active sensor nodes L is not fixed *a priori* in this framework.

Let Γ be a nonempty subset of the set of all measurable functions from observation space $\mathcal{Y}$ to transmission space $\mathcal{T}$. A transmission mapping γ is an element of Γ. A transmission strategy $\mathcal{G}$ is a vector function of the form $(\gamma_1, \dots, \gamma_L) : \mathcal{Y}^L \to \mathcal{T}^L$ with the interpretation that, upon receiving information $Y_\ell = y$, sensor node ℓ transmits summary $\gamma_\ell(y)$ to the fusion center.

We assume that the probability measures $\mathcal{P}_0$ and $\mathcal{P}_1$ are known beforehand, that they are mutually absolutely continuous (equivalent), and that they are distinguishable. We also assume that the noise probability distribution is known. Then, under hypothesis H_i, the mapping γ induces a probability law $\mathcal{Q}_{i,\gamma}$ on the reception space $\mathcal{U}$.

A decision test $\mathcal{S}$ is a measurable map $\mathcal{S} : \mathcal{U}^L \mapsto \{0,1\}$ such that when $U_1 = u_1, \dots, U_L = u_L$ is observed, hypothesis H_0 is accepted if $\mathcal{S}(u_1, \dots, u_L) = 0$, while hypothesis H_1 is accepted if $\mathcal{S}(u_1, \dots, u_L) = 1$. The performance of decision test $\mathcal{S}$ is characterized by the error probabilities

$$\alpha = \mathcal{Q}_{0,\mathcal{G}}\{\mathcal{S}(U_1, \dots, U_L) = 1\}$$
$$\beta = \mathcal{Q}_{1,\mathcal{G}}\{\mathcal{S}(U_1, \dots, U_L) = 0\}$$

where $\mathcal{Q}_{i,\mathcal{G}}$ denotes the product measure on $\mathcal{U}^L$ induced by transmission strategy $\mathcal{G}$ under H_i. Since $\mathcal{P}_0$ and $\mathcal{P}_1$ are mutually absolutely continuous, so are the measures $\mathcal{Q}_{0,\gamma}$ and $\mathcal{Q}_{1,\gamma}$. The likelihood ratio between $\mathcal{Q}_{1,\gamma}$ and $\mathcal{Q}_{0,\gamma}$ (i.e., the Radon–Nikodym derivative of $\mathcal{Q}_{1,\gamma}$ with respect to $\mathcal{Q}_{0,\gamma}$)

$$\mathcal{L}_{\mathcal{Q},\gamma}(u) = \frac{d\mathcal{Q}_{1,\gamma}}{d\mathcal{Q}_{0,\gamma}}(u)$$

is, therefore, well-defined for every transmission map γ.

We consider the specific detection problem where a network of wireless sensor nodes is subject to a total power constraint. That is, the expected consumed power summed across all the sensor nodes may not exceed a given constraint A.

Definition 1: Fix *a priori* probabilities $\mathrm{P}(H_0)$ and $\mathrm{P}(H_1)$. An admissible transmission strategy $\mathcal{G}$ is a vector function $(\gamma_1, \dots, \gamma_L)$ such that

$$F(\mathcal{G}) \stackrel{\triangle}{=} \sum_{\ell=1}^{L} f(\gamma_\ell) \leq A$$

where $f(\gamma_\ell) > 0$ represents the expected power consumed by sensor node ℓ.

Note that $f(\gamma_\ell)$ implicitly depends on the *a priori* probabilities $\mathrm{P}(H_0)$ and $\mathrm{P}(H_1)$ since, for instance, the expected power radiated at the antenna of sensor node ℓ is given by

$$\mathrm{P}(H_0)\mathrm{E}_{\mathcal{P}_0}[|\gamma_\ell(Y)|^2] + \mathrm{P}(H_1)\mathrm{E}_{\mathcal{P}_1}[|\gamma_\ell(Y)|^2].$$

For transmission strategy $\mathcal{G}$, define

$$P_e(\mathcal{G}) \stackrel{\triangle}{=} \inf_{\mathcal{S}}[\alpha \mathrm{P}(H_0) + \beta \mathrm{P}(H_1)]$$

where the infimum is over the set of all decision tests. The design problem is to select an admissible transmission strategy $\mathcal{G}$ such that the Bayes probability of error at the fusion center $P_e(\mathcal{G})$ is minimized.

Being primarily interested in large sensor networks, we consider the asymptotic regime where the power constraint A increases to infinity. As we will see, this corresponds to the asymptotic regime where the number of sensor nodes and, possibly, the area covered by those nodes increase to infinity. For any reasonable collection of transmission strategies, the Bayes probability of error at the fusion center goes to zero exponentially fast as A tends to infinity. It is then natural to compare collections of strategies based on their exponential rate of convergence to zero

$$\liminf_{A\to\infty} \frac{\log P_e(\mathcal{G}_A)}{A}.$$

Throughout, we use $\mathcal{G}_A$ as a convenient notation for a transmission strategy subject to $F(\mathcal{G}_A) \leq A$.

III. Basic Concepts and Results

The following theorem demonstrates how the class of transmission strategies with identical sensor nodes is optimal in terms of exponential rate of convergence in error probability. Let $\mathbf{G}$ be the set of finite subsets of Γ. For $G \in \mathbf{G}$, let $G^{\mathbb{N}}$ be the set of transmission strategies of the form $\mathcal{G} = (\gamma_1, \dots, \gamma_L)$, where $L \in \mathbb{N}$ and $\gamma_\ell \in G$ for all ℓ. In other words, $G^{\mathbb{N}}$ is the set of all strategies with a finite number of sensor nodes, where each node employs a transmission mapping γ_ℓ contained in G. For example, $\{\gamma\}^{\mathbb{N}}$ denotes the set of strategies for which all sensor nodes use identical transmission mapping γ.

Theorem 1: Using identical transmission mappings for all the sensor nodes is asymptotically optimal

$$\inf_{G\in\mathbf{G}} \liminf_{A\to\infty} \min_{\mathcal{G}_A \in G^{\mathbb{N}}} \frac{\log P_e(\mathcal{G}_A)}{A} = \inf_{\gamma\in\Gamma} \liminf_{A\to\infty} \min_{\mathcal{G}_A \in \{\gamma\}^{\mathbb{N}}} \frac{\log P_e(\mathcal{G}_A)}{A}$$

where again $\mathcal{G}_A$ denotes admissible strategies for total power constraint A.

This theorem provides an extension to the work by Tsitsiklis [2], in which identical sensor nodes are shown optimal for decentralized detection by a large number of nodes. In the current formulation, the total power rather than the number of sensor nodes forms the fundamental constraint on the sensing system. Moreover, we relax the assumption that γ is a finite-valued function and that the communication channels between sensor nodes and the fusion center are noiseless.

This alternative framework where the system is limited by a total power constraint is suitable for wireless sensor networks, especially in view of the importance of power conservation in such systems. It establishes that using identical sensor nodes is asymptotically optimal, and also that optimal performance is achieved by selecting the sensor node candidates with the largest normalized Chernoff information. The second result is new and captures the relationship between power consumption and system performance in large sensor networks. It is made precise in the proof of Theorem 1, which we present below. We begin the proof with the introduction of some useful preliminary results.

A. Large Deviations for I.I.D. Observations

For a sequence of independent observations, the exponential rate of convergence in the Bayes error probability can be computed based on Cramér's theorem [10]. Let $V_1, \ldots, V_n$ be a sequence of i.i.d. random variables and consider the hypothesis testing problem that consists in deciding whether the law generating $\{V_j\}$ is ν_0 or ν_1. Assume that ν_0 and ν_1 are mutually absolutely continuous, distinguishable measures and let $Z_j = \log \mathcal{L}_\nu(V_j)$. The logarithmic moment generating function of Z_j under ν_0 is defined as

$$\Lambda_0(\lambda) = \log \mathrm{E}_{\nu_0}[e^{\lambda Z_j}].$$

The large deviations associated with the empirical mean of i.i.d. random variables is characterized by the Fenchel–Legendre transform of $\Lambda_0(\lambda)$, which is defined as

$$\Lambda_0^*(s) \triangleq \sup_{\lambda \in \Re}\{\lambda s - \Lambda_0(\lambda)\}.$$

Write the conditional means of Z_j as $\bar{z}_0 = \mathrm{E}_{\nu_0}[Z_j], \bar{z}_1 = \mathrm{E}_{\nu_1}[Z_j]$. We note that because ν_0, ν_1 are distinguishable measures, Z_j is nonzero with positive measure and, as a consequence, $\bar{z}_1 > \bar{z}_0$. The following theorem, which we quote from Dembo and Zeitouni [10], characterizes the large deviations of the probabilities of error under likelihood ratio tests.

Theorem 2: Let α_n be the probability of declaring hypothesis H_1 under ν_0 given observations $V_1, \ldots, V_n$; similarly, let β_n be the probability of declaring H_0 under ν_1 given n observations. Then, the likelihood ratio test with constant threshold $\tau \in (\bar{z}_0, \bar{z}_1)$ satisfies

$$\lim_{n\to\infty} \frac{1}{n} \log \alpha_n = -\Lambda_0^*(\tau) < 0$$
$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n = \tau - \Lambda_0^*(\tau) < 0$$

where $\Lambda_0^*(\cdot)$ is the Fenchel–Legendre transform of $\Lambda_0(\cdot)$.

A useful corollary to this theorem states that the Chernoff bound on the exponential rate of convergence in the Bayes probability of error is tight.

Corollary 1 (Chernoff): If $0 < \mathrm{P}(H_0) < 1$, then

$$\liminf_{n\to\infty} \frac{1}{n} \log P_e^{(n)} = -\Lambda_0^*(0)$$

where $P_e^{(n)}$ is the probability of error at the fusion center given n observations and the infimum is over all decision tests.

We note that for hypothesis testing problems $\Lambda_0(0) = \Lambda_0(1) = 0$. Furthermore, the function $\Lambda_0(\lambda)$ is convex. These properties lead to a simpler expression for the Fenchel–Legendre transform of $\Lambda_0^*(s)$ evaluated at zero, namely

$$\Lambda_0^*(0) = -\inf_{\lambda\in\Re} \Lambda_0(\lambda) = -\min_{\lambda\in[0,1]} \Lambda_0(\lambda).$$

In the context of hypothesis testing, $\Lambda_0^*(0)$ is often called the Chernoff information.

A second result that can be derived from Theorem 2 is Stein's lemma [10].

Lemma 1 (Stein's Lemma): Let β_n^ϵ be the infimum of β_n among all tests with $\alpha_n < \epsilon$. Then, for any $\epsilon \in (0,1)$

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n^\epsilon = \mathrm{E}_{\nu_0}\left[\log \frac{d\nu_1}{d\nu_0}\right] = -D(\nu_0 \,\|\, \nu_1),$$

where $D(\nu_0 \,\|\, \nu_1)$ is the familiar Kullback–Leibler divergence or relative entropy between two probability measures.

Stein's lemma will prove useful in Section V, where we consider the Neyman–Pearson variant of the detection problem introduced in Section II.

B. Large Deviations in Decentralized Detection

The initial step in establishing Theorem 1 is to construct a lower bound on the exponential rate of convergence in the Bayes error probability.

Lemma 2: Let $G \in \mathbf{G}$ be fixed. There exist nonnegative real numbers $\{x_\gamma : \gamma \in G\}$ such that $\sum_{\gamma\in G} x_\gamma \le 1$ and

$$\liminf_{A\to\infty} \min_{\mathcal{G}_A \in G^{\mathbb{N}}} \frac{\log P_e(\mathcal{G}_A)}{A} \ge \sum_{\gamma\in G} \frac{x_\gamma}{f(\gamma)} \min_{\lambda\in[0,1]} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}]. \quad (2)$$

Proof: For transmission strategy $\mathcal{G}$, let $\mathrm{Num}(\gamma, \mathcal{G})$ represent the number of sensor nodes with transmission mapping γ in strategy $\mathcal{G}$ and define

$$x_{\gamma,A} \triangleq \frac{\mathrm{Num}(\gamma,\mathcal{G}) f(\gamma)}{A}.$$

In other words, $x_{\gamma,A}$ is the proportion of the power constraint A allocated to sensor nodes of type γ. Clearly, the inequality $\sum_{\gamma\in G} x_{\gamma,A} \le 1$ holds for any strategy $\mathcal{G}$ such that $F(\mathcal{G}) \le A$. For constraint A, define

$$\mathcal{G}_A^* = \arg \min_{\mathcal{G}\in G^{\mathbb{N}}, F(\mathcal{G})\le A} \frac{\log P_e(\mathcal{G})}{A}.$$

Note that the minimum is always achieved since there is only a finite number of strategies in $G^{\mathbb{N}}$ such that the condition $F(\mathcal{G}) \le A$ holds.

Consider an increasing sequence $A_1, A_2, A_3, \ldots$, with the following properties

$$\lim_{n\to\infty} A_n = \infty$$
$$\lim_{n\to\infty} \frac{\log P_e\left(\mathcal{G}^*_{A_n}\right)}{A_n} = \liminf_{A\to\infty} \min_{\mathcal{G}_A \in G^{\mathbb{N}}} \frac{\log P_e(\mathcal{G}_A)}{A}$$

and $\lim_{n\to\infty} x_{\gamma,A_n}$ exists for all $\gamma \in G$. Such a sequence exists since $x_{\gamma,A_n} \in [0,1]$ for all n and all $\gamma \in G$, and G only has a finite number of elements.

Let $\delta > 0$ be given. For $x_\gamma = \lim_{n\to\infty} x_{\gamma,A_n}$, select a real number r_γ such that $(r_\gamma/f(\gamma))$ is a rational number and $0 \le r_\gamma - x_\gamma < \delta$. Then, choose an integer N such that $(r_\gamma N/f(\gamma)) \in \mathbb{N}$ for all $\gamma \in G$. Define the sequence $B_1, B_2, \ldots$, componentwise by

$$B_n = N\left\lceil \frac{A_n}{N} \right\rceil$$

and let $\mathcal{G}_{B_n}$ be a transmission strategy with $(r_\gamma B_n/f(\gamma))$ sensor nodes of type $\gamma \in G$. By construction, we have $P_e(\mathcal{G}^*_{A_n}) \ge P_e(\mathcal{G}_{B_n})$. It follows that

$$\begin{aligned}
&\lim_{n\to\infty} \frac{\log P_e\left(\mathcal{G}^*_{A_n}\right)}{A_n} \\
&\quad\ge \lim_{n\to\infty} \frac{\log P_e(\mathcal{G}_{B_n})}{A_n} = \lim_{n\to\infty} \frac{\log P_e\left(\mathcal{G}_{B_n}\right)}{B_n} \\
&\quad\overset{(a)}{=} \min_{\lambda\in[0,1]} \left\{ \frac{1}{N} \sum_{\gamma\in G} \frac{r_\gamma N}{f(\gamma)} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}] \right\} \\
&\quad\ge \sum_{\gamma} \frac{r_\gamma}{f(\gamma)} \min_{\lambda\in[0,1]} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}] \\
&\quad\ge \sum_{\gamma} \frac{x_\gamma + \delta}{f(\gamma)} \min_{\lambda\in[0,1]} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}]
\end{aligned}$$

where (a) follows from Corollary 1 applied to vector observations $V_j = (U_{jN+1}, \ldots, U_{(j+1)N})$ and where the number of sensor nodes of type γ in $\{U_{jN+1}, \ldots, U_{(j+1)N}\}$ is $(r_\gamma N/f(\gamma))$. Since

$$\min_{\lambda\in[0,1]} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}]$$

is bounded for all $\gamma \in G$ and $\delta > 0$ is arbitrary, we obtain (2) as desired. □

The form of an optimal solution is considered next. Lemma 3 establishes that the exponential rate of convergence in error probability is lower bounded by the performance of the best transmission mappings in G.

Lemma 3: Let $G \in \mathbf{G}$ be fixed. The constrained optimization problem

$$\min_{x_\gamma} \sum_{\gamma\in G} \frac{x_\gamma}{f(\gamma)} \min_{\lambda\in[0,1]} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}] \tag{3}$$

subject to $\sum_{\gamma\in G} x_\gamma \le 1$ has an optimal solution, where x_γ is equal to one for some $\gamma \in G$, and the remaining x_γ are equal to zero.

Proof: Let

$$\gamma^* = \arg\min_{\gamma\in G} \left\{ \frac{1}{f(\gamma)} \min_{\lambda\in[0,1]} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}] \right\}.$$

For any choice of $\{x_\gamma\}$ such that $\sum_{\gamma\in G} x_\gamma \le 1$, we obtain

$$\begin{aligned}
&\sum_{\gamma\in G} \frac{x_\gamma}{f(\gamma)} \min_{\lambda\in[0,1]} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}] \\
&\quad\ge \sum_{\gamma\in G} \frac{x_\gamma}{f(\gamma^*)} \min_{\lambda\in[0,1]} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma^*}}[e^{\lambda X_{\gamma^*}}] \\
&\quad\ge \frac{1}{f(\gamma^*)} \min_{\lambda\in[0,1]} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma^*}}[e^{\lambda X_{\gamma^*}}].
\end{aligned}$$

Hence, having $x_{\gamma^*} = 1$ and $x_\gamma = 0$ for all $\gamma \in G\backslash\{\gamma^*\}$ is optimal. □

By comparing (2) and (3), we see that the solution to the minimization problem in Lemma 3 gives a lower bound on the right-hand side of (2). Hence, to complete the Proof of Theorem 1, it suffices to show that the lower bound derived in Lemma 3 is achieved by a sequence of admissible transmission strategies.

Lemma 4: Let $G \in \mathbf{G}$ be fixed. For any $\gamma \in G$, there exists a set of transmission strategies $\{\mathcal{G}_A\}$ such that

$$\lim_{A\to\infty} \frac{\log P_e(\mathcal{G}_A)}{A} = \frac{1}{f(\gamma)} \min_{\lambda\in[0,1]} \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}]. \tag{4}$$

Proof: Consider the strategy $\mathcal{G}_A$ composed of $\lfloor (A/f(\gamma)) \rfloor$ sensor nodes with transmission mapping γ. Then, $F(\mathcal{G}_A) \le A$ and by Corollary 1

$$\begin{aligned}
&\liminf_{A\to\infty} \frac{\log P_e(\mathcal{G}_A)}{A} \\
&\quad= \lim_{A\to\infty} \frac{\left\lfloor \frac{A}{f(\gamma)} \right\rfloor}{A} \liminf_{A\to\infty} \frac{\log P_e(\mathcal{G}_A)}{\left\lfloor \frac{A}{f(\gamma)} \right\rfloor} \\
&\quad= \frac{1}{f(\gamma)} \min_{\lambda\in[0,1]} \left\{ \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}] \right\}.
\end{aligned}$$

That is, there exists a subsequence of transmission strategies converging to the right-hand side of (4), as desired. □

Collecting the results of Lemma 2 through Lemma 4, we immediately obtain Theorem 1. That is, using identical transmission mappings for all the sensor nodes is asymptotically optimal. Furthermore, for wireless sensor networks with a large power constraint, prospective sensor types should be compared according to the normalized Chernoff information

$$-\frac{1}{f(\gamma)} \min_{\lambda\in[0,1]} \left\{ \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}] \right\}.$$

The normalized Chernoff information captures the natural tradeoff between power consumption and information rendering in large sensing systems. Intuitively, allocating more power per node implies receiving more reliable information from each node at the fusion center. On the other hand, for a fixed power constraint A, a reduction in power consumption per sensor node allows the network to contain more active nodes. This tradeoff is expressed in mathematical terms by normalizing the Chernoff information by the consumed power. For example, doubling the Chernoff information provided by each sensor node present in the network results in the same gain in overall performance as reducing the power consumption per node by half and doubling the number of nodes.

IV. Applications and Numerical Examples

In this section, we illustrate how the normalized Chernoff information can be employed to compute optimal transmit power levels for wireless sensor nodes. We study three distinct scenarios. In the first scenario, we compute an upper bound for the asymptotic decay in error exponent based on the power radiated at the antenna of each sensor node. We then provide a more realistic analysis which takes into account the radiated power together with the power consumed at the sensor nodes themselves. Finally, we discuss how the results presented in this paper can be applied to the classical decentralized detection framework, where each sensor node reliably transmits a finite-valued function of its observation to the fusion center.

A. Upper Bound on Error Exponent

To obtain an upper bound on the normalized Chernoff information

$$-\frac{1}{f(\gamma)}\min_{\lambda\in[0,1]}\left\{\log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}]\right\}$$

we let $f(\gamma)$ be the expected radiated power at the antenna of a sensor node. This is only an approximation since the analysis disregards the power consumed at the sensor nodes themselves. In fact, this produces an upper bound on the performance of the system since, in general, the total power drained from the battery includes the power radiated by the communication unit, as well as the power consumed by the processing unit of the sensor. An upper bound based on radiated power quantifies the ultimate performance of a sensing system in a specific environment. As technology improves, power consumption at the sensor nodes decreases and, as a consequence, the aforementioned upper bound on performance becomes increasingly tight.

We consider the scenario where the wireless sensor nodes have access to Gaussian observations with shifted means. Mathematically, the random variable available at each sensor node has conditional distribution

$$p_{Y\,|\,H}(y\,|\,H_0) \sim \mathcal{N}\left(-m_y, \sigma_y^2\right)$$
$$p_{Y\,|\,H}(y\,|\,H_1) \sim \mathcal{N}\left(m_y, \sigma_y^2\right)$$

where $\mathcal{N}(m_y, \sigma_y^2)$ denotes a Gaussian distribution with mean m_y and variance σ_y^2. The fusion center receives a noisy version of the data transmitted by the sensor nodes as described in (1). We also assume that the communication noise W_ℓ corresponding to sensor node ℓ has a Gaussian distribution $\mathcal{N}(0, \sigma_w^2)$. Based on the received data, the fusion center makes a final decision regarding the state of nature.

Finding a transmission mapping γ that maximizes the normalized Chernoff information over all measurable functions from $\mathcal{Y}$ to $\mathcal{T}$ is, in general, hard. This is partially due to the fact that the search space for an optimal γ tends to be vast. Also, the Chernoff information with its minimization over $\lambda \in [0,1]$ is a complicated performance metric; it is not suitable for most standard optimization techniques. Yet the normalized Chernoff information can readily be employed to assess the performance of practical sensing systems, where an optimal transmission mapping γ is to be selected from a reasonable pool of candidates Γ. This is illustrated below, where we study two classes of sensor nodes.

1) Binary Sensor Nodes: In this first class of sensor nodes, each node computes and sends a 1-bit summary of its own observation. We assume that the transmission mapping γ_m employed by the nodes is a binary threshold function of the form

$$\gamma_m(y) = \begin{cases} m: & y \geq 0 \\ -m: & y < 0 \end{cases} \tag{5}$$

where $m > 0$. The probability measures on the reception space $\mathcal{U}$ are incidentally absolutely continuous with respect to the Lebesgue measure and have probability density functions given by

$$\mathcal{Q}_{0,\gamma_m}(u) = \frac{1}{\sqrt{2\pi\sigma_w^2}} Q\left(\frac{m_y}{\sigma_y}\right)\exp\left(-\frac{(u-m)^2}{2\sigma_w^2}\right) + \frac{1}{\sqrt{2\pi\sigma_w^2}} Q\left(-\frac{m_y}{\sigma_y}\right)\exp\left(-\frac{(u+m)^2}{2\sigma_w^2}\right)$$
$$\mathcal{Q}_{1,\gamma_m}(u) = \mathcal{Q}_{0,\gamma_m}(-u)$$

where $Q(\,\cdot\,)$ is the complementary Gaussian cumulative distribution function $Q(x) = \int_x^\infty (1/\sqrt{2\pi})e^{-(\xi^2/2)}d\xi$. We note that the radiated power per sensor node in this example is independent of *a priori* probabilities $\mathrm{P}(H_0)$ and $\mathrm{P}(H_1)$, and that it is given by $f(\gamma_m) = m^2$. The normalized Chernoff information can be computed as

$$-\frac{1}{m^2}\log\left(\int_{-\infty}^{\infty}\sqrt{\mathcal{Q}_{0,\gamma_m}(u)\mathcal{Q}_{1,\gamma_m}(u)}\,du\right). \tag{6}$$

Although (6) does not admit a closed form expression, it can easily be computed numerically. It is also possible to derive an upper bound for the Chernoff information of (6). First, we note from Fig. 2 that the normalized Chernoff information is monotone decreasing in m. It follows that the normalized Chernoff information corresponding to transmission mapping γ_m is upper bounded by its limiting value as m approaches zero

$$\lim_{m\downarrow 0} -\frac{1}{m^2}\log\left(\int_{-\infty}^{\infty}\sqrt{\mathcal{Q}_{0,\gamma_m}(u)\mathcal{Q}_{1,\gamma_m}(u)}\,du\right) = \frac{1}{2\sigma_w^2}\left(Q\left(-\frac{m_y}{\sigma_y}\right) - Q\left(\frac{m_y}{\sigma_y}\right)\right)^2.$$

Fig. 2 shows the normalized Chernoff information along with the corresponding upper bound for the transmission mapping of (5) and the following channel parameters: $m_y^2/\sigma_y^2 \in \{(1/16), 1, 16\}$, $f(\gamma) \in [10^{-2}, 10^2]$, and $\sigma_w^2 = 1$.

2) Analog Sensor Nodes: The second class of sensor nodes we study is the collection of nodes where each unit retransmits an amplified version of its own observation. In this setup, a sensor node acts as an analog relay amplifier with a transmission mapping of the form

$$\gamma_a(y) = ay \tag{7}$$

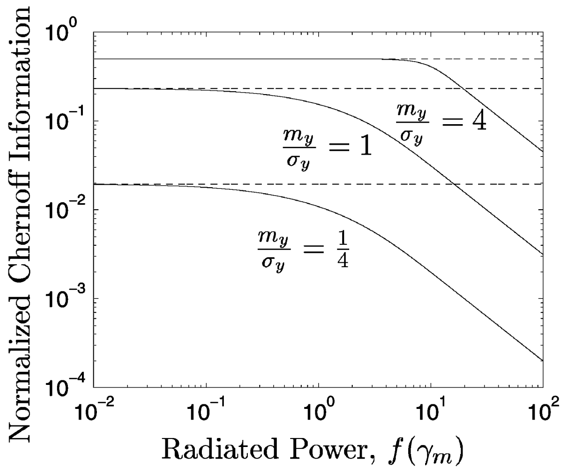

Fig. 2. Normalized Chernoff information corresponding to wireless sensor nodes with binary transmission mapping γ_m and radiated power $f(\gamma_m) = m^2$.

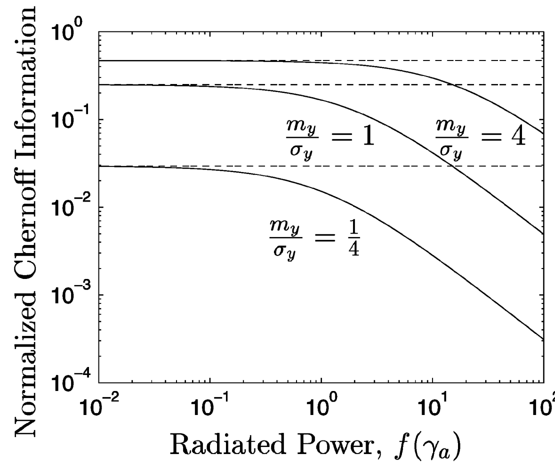

Fig. 3. Normalized Chernoff information corresponding to wireless sensor nodes with analog transmission mapping γ_a and radiated power $f(\gamma_a) = a^2(m_y^2 + \sigma_y^2)$.

where $a > 0$. Transmission map γ_a induces the following probability laws on reception space $\mathcal{U}$

$$\mathcal{Q}_{0,\gamma_a} \sim \mathcal{N}\left(-am_y, a^2\sigma_y^2 + \sigma_w^2\right)$$
$$\mathcal{Q}_{1,\gamma_a} \sim \mathcal{N}\left(am_y, a^2\sigma_y^2 + \sigma_w^2\right).$$

The associated radiated power per sensor node, which is again independent of *a priori* probabilities $\mathrm{P}(H_0)$ and $\mathrm{P}(H_1)$, is given by

$$f(\gamma_a) = \mathrm{E}[a^2y^2] = a^2m_y^2 + a^2\sigma_y^2$$

where the expectation is taken over the random variables Y and H. We can write the corresponding normalized Chernoff information as

$$\frac{m_y^2}{2\left(m_y^2 + \sigma_y^2\right)\left(a^2\sigma_y^2 + \sigma_w^2\right)}.$$

Again, we see that the normalized Chernoff information is a monotone decreasing function of radiated power. Moreover, for any transmission mapping γ_a, where γ_a has the form of (7), the asymptotic rate of decay in error exponent is upper bounded by

$$-\liminf_{A\to\infty} \frac{\log P_e(\mathcal{G}_A)}{A} \leq \frac{m_y^2}{2\sigma_w^2\left(m_y^2 + \sigma_y^2\right)}.$$

Fig. 3 plots the normalized Chernoff information along with the corresponding upper bound for the transmission mapping of (7).

3) Performance Comparison: It is instructive to compare the transmission schemes introduced in the previous two sections. Fig. 4 plots the asymptotic performance of sensor nodes with binary transmission mapping γ_m against the asymptotic performance of sensor nodes with analog transmission mapping γ_a. As seen in Fig. 4, there is a crossover between the two functions, with analog sensor nodes performing better below a threshold SNR. This precludes an early dismissal of analog sensor nodes in favor of the more studied digital nodes. Indeed, for some detection applications, wireless sensor nodes with continuous transmission mappings may outperform sensor nodes with finite-valued transmission mappings.

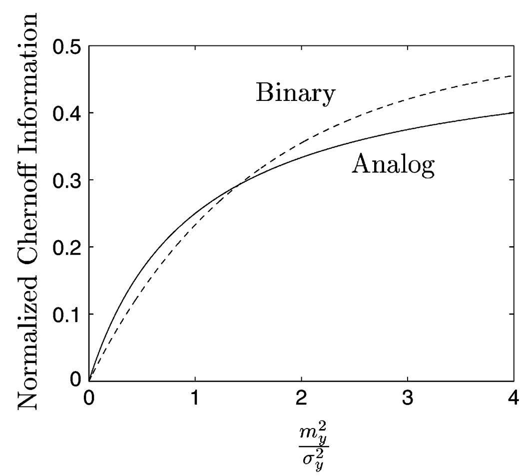

Fig. 4. Comparison of the normalized Chernoff information for wireless sensing systems with binary sensor nodes and analog sensor nodes.

B. Optimal Radiated Power

In Section IV-A, we derived upper bounds on the asymptotic decay in error exponent for two classes of sensor nodes. These bounds were obtained under the optimistic assumption that the power consumed at the sensor node is negligible compared with the power radiated by the communication unit. However, this may not be true, especially if sensor nodes transmit their data at very modest power levels.

In this section, we show how the analysis of the previous section can be modified to account for the power consumed at the sensor nodes. In particular, we reconsider the analysis of binary sensor nodes with transmission mapping γ_m, where γ_m is as defined in (5), under the assumption that the power consumed at the nodes is nonnegligible. When the power consumed at a node itself P_{node} is nonnegligible, the normalized Chernoff information becomes

$$-\frac{1}{P_{\text{node}} + m^2} \min_{\lambda\in[0,1]} \left\{\log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}]\right\}.$$

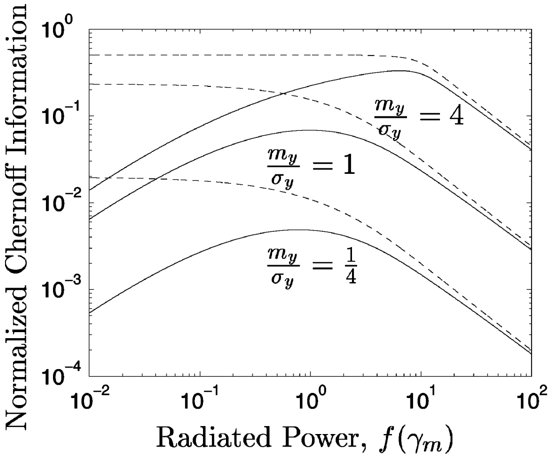

Fig. 5. Plot of the normalized Chernoff information for binary sensor nodes, taking into account the power consumed at the nodes themselves. Optimal radiated power are finite and correspond to global maxima in the normalized Chernoff information curves.

It is interesting to note that when P_{node} is taken into account the optimal transmit power has a finite, nonzero value. Indeed for $P_{\text{node}} > 0$, we have the following limits:

$$\lim_{m\downarrow 0} -\frac{\log\left(\int_{-\infty}^{\infty}\sqrt{\mathcal{Q}_{0,\gamma_m}(u)\mathcal{Q}_{1,\gamma_m}(u)}\,du\right)}{P_{\text{node}}+m^2} = 0$$
$$\lim_{m\uparrow \infty} -\frac{\log\left(\int_{-\infty}^{\infty}\sqrt{\mathcal{Q}_{0,\gamma_m}(u)\mathcal{Q}_{1,\gamma_m}(u)}\,du\right)}{P_{\text{node}}+m^2} = 0$$

and, by continuity, this implies that the optimal transmit power belongs to the open interval $(0,\infty)$. This behavior is seen in Fig. 5, where the normalized Chernoff information is plotted for node power $P_{\text{node}} = 0.5$ W. For reference, the figure also includes the upper bound derived under the assumption that the power consumed at the node is negligible.

C. Finite Quantizers With Reliable Channels

The final scenario we consider is the classical decentralized detection framework, where each sensor node transmits a finite-valued function of its observation to the fusion center over a reliable communication channel. We assume that the nodes share a multiple-access channel that is capable to carry A bits of information per unit time. In this framework, $f(\gamma) = \lceil \log_2(D_\gamma)\rceil$, where D_γ is the number of candidate messages potentially sent by transmission mapping γ. From Theorem 1, we know that using identical transmission mappings for all the sensor nodes is asymptotically optimal. Therefore, an optimal transmission mapping is a discrete map γ^* that maximizes the normalized Chernoff information

$$-\frac{1}{\lceil \log_2(D_\gamma)\rceil}\min_{\lambda\in[0,1]}\left\{\log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}]\right\}.$$

From this observation, we obtain a result analogous to our previous result on capacity constrained sensor networks [11].

Theorem 3: Assume that the probability measures $\mathcal{P}_0$ and $\mathcal{P}_1$ on the observation space $\mathcal{Y}$ are mutually absolutely continuous. If there exists a binary transmission mapping $\tilde{\gamma}$ such that

$$-\min_{\lambda\in[0,1]}\left\{\log \mathrm{E}_{\mathcal{Q}_{0,\tilde{\gamma}}}[e^{\lambda X_{\tilde{\gamma}}}]\right\} \geq -\frac{1}{2}\min_{\lambda\in[0,1]}\left\{\log \mathrm{E}_{\mathcal{P}_0}\left[\left(\frac{d\mathcal{P}_1}{d\mathcal{P}_0}\right)^{\lambda}\right]\right\}$$

then having identical sensor nodes, each sending 1 bit of information, is asymptotically optimal.

Proof: Let γ be any admissible transmission mapping with $D_\gamma > 2$. Then, from the sequence of inequalities

$$\begin{aligned}
&-\frac{1}{\lceil \log_2(D_\gamma)\rceil}\min_{\lambda\in[0,1]}\left\{\log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}]\right\}\\
&\leq -\frac{1}{2}\min_{\lambda\in[0,1]}\left\{\log \mathrm{E}_{\mathcal{P}_0}\left[\left(\frac{d\mathcal{P}_1}{d\mathcal{P}_0}\right)^{\lambda}\right]\right\}\\
&\leq -\min_{\lambda\in[0,1]}\left\{\log \mathrm{E}_{\mathcal{Q}_{0,\tilde{\gamma}}}[e^{\lambda X_{\tilde{\gamma}}}]\right\}
\end{aligned}$$

we conclude that having binary sensor nodes is asymptotically optimal. □

Since the conditions of Theorem 3 hold for Gaussian observations and exponential observations (see [11]), having identical binary sensor nodes is asymptotically optimal in these two situations. Note that the results presented in this paper assume only one observation per sensor node and are valid as long as the wireless sensing system is large enough. This is in contrast with our initial results [11], where we assumed that each sensor node receives a long sequence of observations.

V. Neyman–Pearson Problem

Extending the results of the preceding sections to the Neyman–Pearson variant of the detection problem introduced in Section II requires little effort. In the latter problem formulation, $\mathrm{P}(H_0)$ and $\mathrm{P}(H_1)$ are unknown and $f(\gamma_\ell)$ denotes the expected power consumed by sensor node ℓ under hypothesis H_0. The power constraint A is then a constraint on the behavior of the system under hypothesis H_0.

Again, one can show that using identical transmission mappings is asymptotically optimal as the power constraint A tends to infinity. For $\epsilon \in (0,1)$, let $\beta^\epsilon(\mathcal{G})$ be the infimum of $\beta(\mathcal{G})$ among all decision tests $\mathcal{S}$ such that $\alpha(\mathcal{G}) < \epsilon$, i.e.,

$$\beta^\epsilon(\mathcal{G}) = \inf_{\mathcal{S}}\{\mathcal{Q}_{1,\mathcal{G}}\{\mathcal{S}(U_1,\ldots U_L)=0\}: \mathcal{Q}_{0,\mathcal{G}}\{\mathcal{S}(U_1,\ldots U_L)=1\} < \epsilon\}.$$

Theorem 4: Using identical transmission mappings for all the sensor nodes is asymptotically optimal

$$\inf_{G\in\mathbf{G}}\liminf_{A\to\infty}\min_{\mathcal{G}_A\in\mathbf{G}^{\mathbb{N}}}\frac{\log\beta^\epsilon(\mathcal{G}_A)}{A} = \inf_{\gamma\in\Gamma}\liminf_{A\to\infty}\min_{\mathcal{G}_A\in\{\gamma\}^{\mathbb{N}}}\frac{\log\beta^\epsilon(\mathcal{G}_A)}{A}$$

where $\mathcal{G}_A$ denotes admissible strategies for total power constraint A.

The steps required to prove Theorem 4 are essentially the same as the steps used to prove Theorem 1. For this reason, we omit the proof. It is interesting to note that the normalized relative entropy

$$\frac{1}{f(\gamma)} D(\mathcal{Q}_{0,\gamma} \,\|\, \mathcal{Q}_{1,\gamma})$$

plays the role of the normalized Chernoff information for the Neyman–Pearson variant of the detection problem introduced in Section II. That is, in a Neyman–Pearson framework, prospective sensor types for a sensor network with a large power constraint should be compared according to the normalized relative entropy.

VI. Conclusion and Discussion

We considered a decentralized detection problem in which a network of wireless sensor nodes provides relevant information about the state of nature to a fusion center. We addressed the specific case where observations are conditionally i.i.d., given H, where the sensor network is subject to a total power constraint, and where sensor nodes are relaying information to the fusion center over a noisy communication channel.

We focused primarily on the Bayesian problem in which the probability of error at the fusion center is minimized. Having identical sensor nodes was found to be optimal in the asymptotic regime where the total power constraint tends to infinity. Moreover, our analysis showed that the normalized Chernoff information is an appropriate metric in comparing prospective sensor nodes for large systems. The optimality of wireless sensor networks with identical sensor nodes is encouraging as it simplifies the design of such systems. In particular, for large systems, transmission mapping candidates should be compared according to the normalized Chernoff information criterion

$$-\frac{1}{f(\gamma)} \min_{\lambda \in [0,1]} \left\{ \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}}[e^{\lambda X_\gamma}] \right\}.$$

Alternatively, in the Neyman–Pearson problem formulation where the probability of type II error is minimized subject to a constraint on the type I error probability, the normalized relative entropy

$$\frac{1}{f(\gamma)} D(\mathcal{Q}_{0,\gamma} \,\|\, \mathcal{Q}_{1,\gamma})$$

should be used as a performance metric.

We also provided examples on how to apply these results to practical systems. In particular, we have derived upper bounds on the performance of binary sensor nodes and analog sensor nodes. We have shown how optimal transmission power levels can be derived for binary sensor nodes having access to Gaussian observations. Finally, we have used the normalized Chernoff information to derive conditions that insure the optimality of binary sensor nodes in the classical decentralized detection framework, where sensor nodes transmit discrete information over a reliable multiple access channel. In particular, binary sensor nodes are asymptotically optimal for Gaussian observations and exponential observations in this scenario.

We now discuss some future avenues of research. Wireless communication channels are often subject to fading. If sensor nodes are to be scattered around somewhat randomly, it is conceivable that their respective communication channels will be subject to fading, with certain nodes having much better channels than others. It would be interesting to develop techniques that quantify the performance loss due to fading in sensor networks, especially in the scenario where sensor nodes can adapt to different fade levels.

Another interesting topic is to derive meaningful performance metrics for sensing systems where observations are not conditionally independent. This would be very relevant for systems where sensor nodes are densely packed together, as they are more likely to observe dependent random variables. For instance, it may be possible to exploit the dependence among sensor observations to reduce the transmit power per sensor without sacrificing overall performance.

Finally, in our analysis, we have implicitly assumed that the bandwidth allocated to the sensing system is large or, equivalently, that the time period during which sensor nodes transmit their data is long. This assumption results in each sensor node transmitting data to the fusion center over a dedicated communication channel. If the system bandwidth and transmission period are severely constrained, then interference among sensor nodes needs to be considered. It would be interesting to see what types of communication strategies maximize the performance of sensing systems under such additional constraints.

References

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

[2] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Math. Control Signals Syst.*, vol. 1, no. 2, pp. 167–182, 1988.

[3] P.-N. Chen and A. Papamarcou, "New asymptotic results in parallel distributed detection," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1847–1863, Nov. 1993.

[4] V. V. Veeravalli, T. Başar, and H. V. Poor, "Decentralized sequential detection with a fusion center performing the sequential test," *IEEE Trans. Inform. Theory*, vol. 39, pp. 433–442, Mar. 1993.

[5] P. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad and ugly: Distributed detection of a known signal in dependent gaussian noise," *IEEE Trans. Signal Processing*, vol. 48, pp. 3266–3279, Dec. 2000.

[6] J. N. Tsitsiklis, "Decentralized detection," *Adv. Statist. Signal Processing*, vol. 2, pp. 297–344, 1993.

[7] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I—Fundamentals," *Proc. IEEE*, vol. 85, pp. 54–63, Jan. 1997.

[8] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors—Part II: Advanced topics," *Proc. IEEE*, vol. 85, pp. 64–79, Jan. 1997.

[9] T. M. Duman and M. Salehi, "Decentralized detection over multiple-access channels," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, pp. 469–476, Apr. 1998.

[10] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Boston, MA: Jones & Barlett, 1993.

[11] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Processing*, vol. 51, pp. 407–416, Feb. 2003.

[12] M. Longo, T. D. Lookabaugh, and R. M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. Information Theory*, vol. 36, pp. 241–255, Mar. 1990.

[13] J. N. Tsitsiklis, "Extremal properties of likelihood-ratio quantizers," *IEEE Trans. Commun.*, vol. 41, pp. 550–558, Apr. 1993.

[14] W. W. Irving and J. N. Tsitsiklis, "Some properties of optimal thresholds in dencentralized detection," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 835–838, Apr. 1994.

[15] P. Willett and P. F. Swaszek, "On the performance degradation from one-bit quantized detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1997–2003, Nov. 1995.

[16] 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, pp. 554–568, Apr. 1996.
[17] T. Kasetkasem and P. K. Varshney, “Communication structure planning for multisensor detection systems,” in *Proc. Inst. Elect. Eng.—Radar, Sonar, Navigation*, vol. 148, Feb. 2001, pp. 2–8.
[18] V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava, “Energy-aware wireless microsensor networks,” *IEEE Signal Processsing Mag.*, vol. 19, pp. 40–50, Mar. 2002.
[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[20] R. Durrett, *Probability: Theory and Examples*, 2nd ed. New York: Duxbury, 1996.

Jean-François Chamberland (S’98) received the B.Eng. degree in electrical engineering from McGill University, Montreal, QC, Canada, in 1998, and the M.S. degree from the School of Electrical Engineering, Cornell University, Ithaca, NY, in 2000. He is currently working toward the Ph.D. degree at the University of Illinois at Urbana–Champaign.

His research interests include communication systems, detection and estimation, and statistical signal 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, in 1992, all in electrical engineering.

He joined the University of Illinois 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 was an Assistant Professor at Cornell University, Ithaca, NY, from 1996 to 2000. He is on the Editorial Boards for *Communications in Information and Systems* (CIS), *Sensor Letters*, and the *Journal of Advances of Information Fusion*. His research interests include mobile and wireless communications, distributed sensor systems, detection and estimation theory, and information theory.

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, and the Xerox Award for facutly 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. He is currently an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY.