

# Decision Fusion Over Noncoherent Fading Multiaccess Channels

Feng Li, *Member, IEEE*, Jamie S. Evans, *Member, IEEE*, and Subhrakanti Dey, *Senior Member, IEEE*

**Abstract**—In this paper, we consider a distributed detection scenario where a number of remote sensors is linked to a decision fusion center by a fading multiaccess channel. The communication is assumed to be noncoherent meaning that channel gains are unknown at both sensors and the fusion center. Each sensor makes a binary local decision and communicates it to the fusion center simultaneously. We investigate the detection performance of the system in terms of error probability and error exponent under both Rayleigh and Rician fading scenarios. We reveal that ON-OFF keying is the most energy efficient modulation scheme when the channel is subject to Rayleigh fading and that optimizing the modulation scheme can lead to a gain in error exponent under Rician fading scenario. Under both fading scenarios, optimal decision fusion rules can be reduced to simple threshold tests.

**Index Terms**—Distributed detection, fading channels, multiaccess communication, multisensor systems, signal processing for communications.

## I. INTRODUCTION

IN this paper, we focus on distributed detection using geographically dispersed wireless sensors. The state of nature to be detected, for example, could be the presence of an enemy target or the leakage of poisonous gas. Intelligent sensors will make observations, which are dependent on the true state of nature, locally process their observations and communicate the results to a fusion center where a final decision will be made with regard to which hypothesis is true.

Classical distributed detection has been studied extensively [1]–[3]. However, to implement distributed detection in wireless sensor networks (WSNs), we meet some new challenges. One challenge is the stringent energy constraint. Normally, wireless sensors are powered by small batteries and it is often hard or not economic to replace those batteries when they run out. As a result, WSNs call for energy efficient computing and transmission schemes to extend the service time of the battery. To save transmission power, WSNs may sometimes be forced to operate with low SNR. Techniques to improve the detection performance in the low SNR region will then be a problem of interest.

Manuscript received December 23, 2010; revised March 24, 2011; accepted May 04, 2011. Date of publication May 23, 2011; date of current version August 10, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Biao Chen. This work was supported by the Australian Research Council. The results of this paper were presented in part at IEEE Global Communications Conference (GLOBECOM), New Orleans, LA, November 2008.

The authors are with the ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN, an affiliate of National ICT Australia), Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Victoria 3010, Australia (e-mail: fenli@unimelb.edu.au; jse@unimelb.edu.au; sdey@unimelb.edu.au).

Digital Object Identifier 10.1109/TSP.2011.2157501

## A. Related Works

When local sensors communicate with the fusion center, distributed data compression could be performed to reduce the communication bandwidth and to conserve energy (e.g., [4] and [5]). When designing distributed detection schemes for WSNs, we should also take into account the effect of unreliable wireless channels to yield a better performance. Disregarding the unreliability of the transmission channel will result in degradation of performance [6].

For a binary hypothesis testing problem, distributed detection schemes utilizing intelligent sensors with each of them making a local binary decision and communicating the uncoded binary decision to the fusion center have attracted considerable attention. Local decisions are normally transmitted in parallel to the fusion center without interfering with each other. Such simple schemes can dramatically reduce the communication bandwidth and the transmission power [7]–[11]. Reference [7] investigates the case where the channel state information is known, while [8] considers the case where only the knowledge of channel statistics is available. If necessary, local decisions may be transmitted to the fusion center via multiple hops [9]. In [10], however, a serial distributed detection structure is considered. The fusion of censored decisions is proposed in [11] which is motivated by [12].

Recently, the idea of distributed detection over multiaccess channels has emerged. For a set of discrete or quantized observations, Mergen, Naware, and Tong [13] as well as Liu and Sayeed [14] proposed type-based multiple access (TBMA) scheme. The TBMA scheme requires a number of orthogonal multiaccess channels. The actual number is equal to the cardinality of the set of all possible observations (or quantized observations). Each multiaccess channel is associated with a particular observation and each sensor will transmit a pulse over the multiaccess channel that corresponds to its observation type. Obviously, at least two orthogonal multiaccess channels are needed for a detection problem. The TBMA scheme is further developed by Anandkumar and Tong in [15]. For continuous sensor observations, several schemes that fuse analog signals (or soft decisions) over multiaccess channels are proposed and shown to be asymptotically optimal [14], [16] (i.e., could achieve the same error exponent as a centralized system can). Unlike the TBMA scheme, all the sensors transmit simultaneously over the same multiaccess channel in those schemes, no matter what their observations are. To achieve the asymptotic optimality, we need a nonfading channel or at least channel gains to be known at local sensors. However, wireless channels are always subject to fading and a lot of overhead

will be introduced for local sensors to obtain the channel state information. Another related reference is [17] where amplified local observations are communicated to the fusion center via code division multiple access.

Given nonfading channel, we have considered the design of local mapping rules, which map sensor observations to signals to be transmitted to the fusion center, for multiaccess schemes in our previous work [18], [19]. The multiaccess communication scheme also finds its application in distributed inference problems. The authors of [20] investigate a linear Gaussian sensor network for distributed estimation purpose, where sensors communicate with base stations over multiaccess channels, and derive a lower bound on the best achievable end-to-end distortion.

### B. Our Contributions

In this paper, we consider the fusion of local binary decisions over a noncoherent fading multiaccess channel. When the communication channel is not subject to fading or channel gains are known at sensors, having all sensors transmit simultaneously over the same multiaccess channel will not only reduce the communication bandwidth but also effectively suppress the channel noise and hence lead to good performance [14], [16]. However, the assumption of nonfading channel is often too strong for WSNs and it may take a lot of efforts for sensors to estimate the channel gain. As a result, we relax the assumption of nonfading channel or known channel gains, but instead assume that local decisions are simultaneously transmitted to the fusion center over a noncoherent fading multiaccess channel. Channel gains are not known at local sensors or the fusion center. The use of noncoherent detection can lift the burden caused by channel estimation, greatly reduce the computational load at sensor nodes and help to conserve energy.

Two types of fading multiaccess channels are considered, namely Rayleigh and Rician fading channels. For each fading scenario, we do the following:

- analyze the detection performance in terms of error probability;
- investigate the error exponent of the system with the help of the Continuity Theorem and Large Deviation Principle (LDP);
- and design energy efficient modulation scheme for local sensors.

Although the error probability is a special case of the Bayes risk, our results can be easily extended to general Bayesian and Neyman–Pearson tests. The performance of TBMA scheme over fading channel is investigated in [13], where channel gains are assumed to be real random variables. In this work, channel gains are complex and this makes the analysis more complicated.

The main results are summarized in the following.

*As the number of sensors increases, the system demonstrates an error floor under Rayleigh fading scenario, but can achieve an exponentially decreasing error probability under Rician fading scenario.* When the communication channel is subject to Rayleigh fading, the error probability can not be made arbitrarily small by increasing the number of sensors. We derive the closed-form expression of the error floor. This error floor

is due to the noncoherent addition of signals transmitted by different sensors. We also find that the multiaccess scheme can achieve an exponentially decreasing error probability when the communication channel is subject to Rician fading. We calculate the error exponent using LDP.

*Even under Rayleigh fading scenario, the multiaccess scheme can assist decision fusion when the system operates in the low SNR region.* Due to Rayleigh fading, signals transmitted by different sensors may cancel out each other. Surprisingly enough, we observe that the multiaccess scheme still outperforms the traditional parallel access scheme when there is a small number of sensors with low transmission power. We explain the intuition behind that in Section V. A short discussion on approaches to further improve the performance of the multiaccess scheme under Rayleigh fading scenario can be found in Section VI.

*Under Rayleigh fading scenario, ON–OFF keying (OOK) is the most energy efficient modulation scheme for local sensors to communicate their binary decisions; under Rician fading scenario, optimizing local sensors' modulation scheme will lead to a larger error exponent.* Given a certain energy budget per sensor, we find that OOK can provide the smallest error probability under Rayleigh fading scenario. However, OOK is in general not the optimal choice when Rician fading is considered. Under Rician fading scenario, a gain in error exponent can be achieved by optimizing the modulation scheme. In contrast, we can *not* obtain a larger error exponent by optimizing the modulation scheme when the communication channel is not subject to fading [19].

*Under both Rayleigh and Rician fading scenarios, both optimal and asymptotically optimal decision fusion rules can be reduced to threshold tests, which appear in much simpler forms than the likelihood ratio test.* Use  $y_I$  and  $y_Q$  to denote the in-phase and quadrature components of the received signal. At the fusion center, the likelihood ratio test is an optimal decision fusion rule. However, the expression of the likelihood ratio  $\frac{f(y_I, y_Q|H_1)}{f(y_I, y_Q|H_0)}$  becomes more and more complicated when the number of sensors increases. Under both Rayleigh and Rician fading scenarios, we show that the optimal decision fusion rule can be reduced to the following threshold test:

$$(y_I + A)^2 + y_Q^2 - A^2 \underset{H_0}{\overset{H_1}{\geq}} T \quad (1)$$

which means that the fusion center will decide hypothesis  $H_1$  to be true if the left-hand side of (1) is greater than  $T$  and decide hypothesis  $H_0$  otherwise. If the type of fading channel, the number of sensors and the modulation scheme are known, then  $A$  and  $T$  are all constants ( $A = 0$  under Rayleigh fading scenario). Asymptotically optimal decision fusion rule under Rician fading scenario has a similar form to (1) and is given by

$$(\tilde{y}_I + \tilde{A})^2 + \tilde{y}_Q^2 - \tilde{A}^2 \underset{H_0}{\overset{H_1}{\geq}} \tau^* \quad (2)$$

where  $\tilde{y}_I$  and  $\tilde{y}_Q$  are scaled versions of  $y_I$  and  $y_Q$ , and  $\tau^*$  does not depend on the number of sensors deployed. These discoveries allow us to simplify the implementation at the fusion center, and obtain the system performance through very simple

numerical calculation. Utilizing (2), we can even obtain the expression for the error exponent under Rician fading scenario, which is given by

$$\Lambda_1^* \left( \sqrt{\tau^* + \tilde{A}^2} - \tilde{A}, 0 \right). \quad (3)$$

$\Lambda_1^*(\cdot, \cdot)$  in (3) is some convex function (see Section IV for more information).

The rest of the paper is organized as follows. The problem formulation will be introduced in the next section. In Section III, we analyze the error probability of the system under both Rayleigh and Rician fading scenarios. The asymptotic behavior of the system under two fading scenarios is investigated in Section IV. Section V provides some numerical results, and we conclude in Section VI.

## II. PROBLEM FORMULATION

Consider the testing of two hypotheses  $H_0$  and  $H_1$  with the help of  $N$  distributed sensors. Let  $v_k$  be the observation obtained by the  $k$ th sensor. We do not assume any specific distribution for observations but do assume the observations are conditionally independent and identically-distributed (i.i.d.) given  $H_0$  or  $H_1$ . The prior probabilities of  $H_0$  and  $H_1$  are denoted by  $P_0$  and  $P_1$ . When observations are obtained, sensors will make local decisions according to a common local decision rule. As a result, all the sensors will have the same detection probability (denoted by  $P_d$ ) and the same false alarm probability (denoted by  $P_f$ ). We assume that  $P_d > P_f$ . A sensor will transmit a baseband symbol  $a \in \mathbb{R}^+$  (i.e.,  $a > 0$ ) to inform the fusion center that it decides  $H_1$  to be true and will transmit a baseband symbol  $-b \in \mathbb{R}$  (we assume  $|a| > |b|$ ) to indicate that it decides  $H_0$  to be true. Unlike the parallel access communication scheme where each sensor communicates with the fusion center through a dedicated channel, the multiaccess communication scheme allows all the sensors to transmit simultaneously over the same channel. The signal received by the fusion center will be the superposition of all the transmitted signals.

The assumption that channel gains are known at local sensors will sometimes be too strong. Additional training overhead will be introduced to estimate the channel gain, which could be costly for energy-limited sensors. Here, we consider the non-coherent case. We assume that signals sent by different sensors are subject to independent but identical fading and channel gains are not known at the fusion center or local sensors. The received baseband symbol can be expressed as

$$y = \sum_{k=1}^N h_k x_k + n \quad (4)$$

where

$$x_k = u_k a - (1 - u_k) b,$$

$u_k \in \{0, 1\}$  is the local decision made by the  $k$ th sensor ( $u_k = i$  indicates that the  $k$ th sensor decides  $H_i$ ),  $h_k$  are i.i.d. complex channel gains and  $n$  is the complex Gaussian noise with zero mean and  $E[|n|^2] = 2\sigma_n^2$ . Such a model is illustrated in Fig. 1. Two fading scenarios will be considered. Under Rayleigh

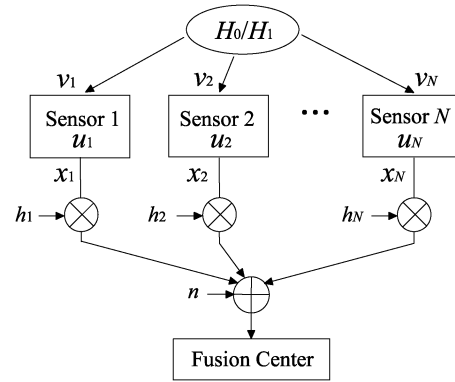


Fig. 1. System model.

fading scenario, the real and imaginary parts of  $h_k$  follow a joint Gaussian distribution with mean zero and covariance matrix  $\sigma_h^2 \mathbf{I}$ , where  $\mathbf{I}$  is a  $2 \times 2$  identity matrix. Under Rician fading scenario, the real and imaginary parts of  $h_k$  follow a joint Gaussian distribution with mean  $[\mu, 0]^T$  (note that we don't lose any generality by making this assumption and a detailed explanation can be found in Appendix A) and covariance matrix  $\sigma_h^2 \mathbf{I}$ . Channel statistics  $\sigma_h^2$  and  $\mu$  are assumed to be known at the fusion center. These channel statistics can be obtained via training or field measurements. For example, a sensor could transmit a series of pilot signals with pre-determined power to the fusion center. The fusion center could then calculate the channel statistics based on the received signals.

Based on the complex symbol  $y$ , the fusion center will make a final decision  $u \in \{0, 1\}$  (where  $u = i$  indicates that the fusion center decides  $H_i$ ). We are interested in the error probability  $P_e = \Pr\{u = 1 | H_0\} \times P_0 + \Pr\{u = 0 | H_1\} \times P_1$  and the error exponent  $-\lim_{N \rightarrow \infty} \frac{1}{N} \ln P_e$  of the system.

## III. PERFORMANCE ANALYSIS

In this section, we investigate system's error probability under the two fading scenarios.

### A. Rayleigh Fading

Under Rayleigh fading scenario, we can prove that OOK is the most energy efficient modulation scheme. A simplified proof can be found in Appendix B. For simplicity, we assume  $a = 1$  and  $b = 0$  throughout this sub-section (under this assumption, we have  $x_k = u_k$ ). So a sensor will transmit a pulse to inform the fusion center that it decides  $H_1$  and will remain silent when it decides  $H_0$ .

1) *Sufficient Statistic*: We first establish the fact that the received signal power is a sufficient statistic for final decision making. Decompose the complex symbol  $y$  into in-phase and quadrature components (denoted by  $y_I$  and  $y_Q$ , respectively)

$$\begin{bmatrix} y_I \\ y_Q \end{bmatrix} = \sum_{k=1}^N \begin{bmatrix} h_{Ik} \\ h_{Qk} \end{bmatrix} u_k + \begin{bmatrix} n_I \\ n_Q \end{bmatrix} \quad (5)$$

where  $h_{Ik}$  and  $h_{Qk}$  are real and imaginary parts of  $h_k$ ; and  $n_I$  and  $n_Q$  are i.i.d. Gaussian random variables with mean zero and variance  $\sigma_n^2$ . Define a random variable  $M = \sum_{k=1}^N u_k$  (i.e., the

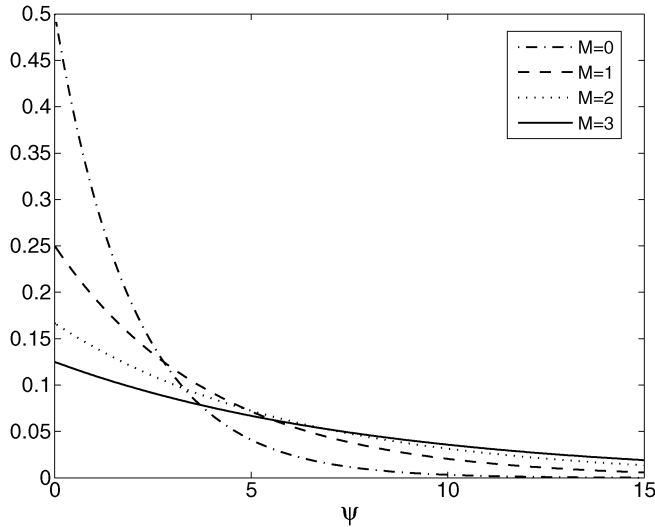


Fig. 2. Conditional probability density function of  $\psi$  given the number of simultaneously transmitting sensors.

total number of sensors that decide to transmit). Under  $H_i$ , we have

$$\begin{aligned} f(y_I, y_Q|H_i) &= \sum_{m=0}^N f(y_I, y_Q|m) \times \Pr\{M = m|H_i\} \\ &= \sum_{m=0}^N \frac{\Pr\{M = m|H_i\}}{2\pi(m\sigma_h^2 + \sigma_n^2)} \exp\left(-\frac{y_I^2 + y_Q^2}{2(m\sigma_h^2 + \sigma_n^2)}\right). \end{aligned} \quad (6)$$

Define  $\psi = |y|^2 = y_I^2 + y_Q^2$  which is the received signal power. Note that  $f(y_I, y_Q|H_i)$  depends on  $y_I$  and  $y_Q$  only through  $\psi$  [see (6)]. Hence, the received signal power  $\psi$  is a sufficient statistic. In quite a few detection problems, we find that the received signal power turns out to be a sufficient statistic [21].

Since  $\psi$  is a sufficient statistic, we are able to turn the detection problem in  $\mathbb{R}^2$  to a detection problem in  $\mathbb{R}$ . Given  $M = m$ ,  $\frac{\psi}{(m\sigma_h^2 + \sigma_n^2)}$  will be a random variable that follows a Chi-square distribution with two degrees of freedom. By performing a linear transformation to the Chi-square distributed random variable, we can obtain [22]

$$f(\psi|m) = \frac{1}{2(m\sigma_h^2 + \sigma_n^2)} \exp\left(-\frac{\psi}{2(m\sigma_h^2 + \sigma_n^2)}\right), (\psi > 0). \quad (7)$$

The probability density function of  $\psi$  given different values of  $M$  is plotted in Fig. 2.

Due to Rayleigh fading and the adoption of multiaccess communication scheme, the fusion center receives a mix of zero-mean faded signals that are transmitted by local sensors. At first glance, it seems to be difficult for the fusion center to retrieve useful information from the received signal. However, the ultimate goal of the fusion center is to distinguish between the two hypotheses based on the signal that it receives. For homogeneous sensors, it is more important for the fusion center to know how many sensors have decided  $H_1$  rather than to know

which sensors have decided  $H_1$ . Since a sensor will transmit only if it decides  $H_1$ , the more sensors that decide  $H_1$ , the more likely it is for the received signal power to be large as shown in Fig. 2. Although not precise, the superposed signal does provide the fusion center with some information about how many sensors have decided  $H_1$ .

2) *Optimal Decision Fusion Rule*: It is well known that the decision fusion rule that minimizes the error probability will be the following likelihood ratio test

$$\frac{f(\psi|H_1)}{f(\psi|H_0)} \underset{H_0}{\overset{H_1}{\geq}} \frac{P_0}{P_1} \quad (8)$$

which means that the fusion center will make a decision in favor of  $H_1$  when the likelihood ratio on the left-hand side of (8) is larger than  $\frac{P_0}{P_1}$  and make a decision in favor of  $H_0$  otherwise. As the number of sensors increases, the expression of the likelihood ratio becomes more and more complicated [see (10)].

Starting from (8), we shall now show how the optimal decision fusion rule can be reduced to a threshold test on the received signal power step by step. The distribution of  $\psi$  under  $H_i$  can be expressed as

$$f(\psi|H_i) = \sum_{m=0}^N f(\psi|m) \times \Pr\{M = m|H_i\}. \quad (9)$$

We have  $\Pr\{M = m|H_0\} = \binom{N}{m}(P_f)^m(1 - P_f)^{N-m}$ ; and if we replace  $P_f$  with  $P_d$  in the right-hand side of the equation we get  $\Pr\{M = m|H_1\}$ . Substituting (9) into (8), we get

$$\frac{\sum_{m=0}^N \binom{N}{m} \frac{(P_d)^m(1-P_d)^{N-m}}{2(m\sigma_h^2 + \sigma_n^2)} \exp\left(-\frac{\psi}{2(m\sigma_h^2 + \sigma_n^2)}\right)}{\sum_{m=0}^N \binom{N}{m} \frac{(P_f)^m(1-P_f)^{N-m}}{2(m\sigma_h^2 + \sigma_n^2)} \exp\left(-\frac{\psi}{2(m\sigma_h^2 + \sigma_n^2)}\right)} \underset{H_0}{\overset{H_1}{\geq}} \frac{P_0}{P_1}. \quad (10)$$

Define  $\Omega_1$  to be the set in the domain of  $\psi$  which contains all  $\psi$  that will make the left-hand side of (10) greater than its right-hand side and  $\Omega_0$  to be the complement set of  $\Omega_1$ .  $\Omega_0$  and  $\Omega_1$  are also known as the decision regions for  $H_0$  and  $H_1$  respectively. The boundaries of  $\Omega_0$  and  $\Omega_1$  can be found by solving

$$g(\psi) \triangleq P_0 f(\psi|H_0) - P_1 f(\psi|H_1) = 0 \quad (11)$$

because  $g(\psi)$  will not change its sign when  $\psi$  lies in between the two adjacent roots of  $g(\psi)$  due to the continuity of  $g(\psi)$  (or in other words, the set of all  $\psi$  that lie in between two adjacent roots of  $g(\psi)$  is a subset of either  $\Omega_0$  or  $\Omega_1$ ). Substitute (9) into (11) and we get

$$\sum_{m=0}^N \left\{ [(P_f)^m(1 - P_f)^{N-m}P_0 - (P_d)^m(1 - P_d)^{N-m}P_1] \times \binom{N}{m} \frac{1}{2(m\sigma_h^2 + \sigma_n^2)} x^{\frac{1}{2(m\sigma_h^2 + \sigma_n^2)}} \right\} = 0 \quad (12)$$

where  $x = \exp(-\psi)$ . Since  $\psi > 0$ , we require  $x < 1$ .

Before exploring the property of the optimal decision fusion rule, we provide one lemma and one theorem.

**Lemma 1:** Let  $A_0$  and  $A_1$  be two positive real numbers,  $K$  be a positive integer,  $0 < p < 1$  and  $0 < q < 1$ . The sequence  $S_k = A_0 p^k (1-p)^{K-k} - A_1 q^k (1-q)^{K-k}$  with  $k = 0, 1, \dots, K$  will change its sign at most once.

*Proof:* The proof can be found in Appendix C. ■

**Theorem 1:** Let  $p(x) = b_N x^{a_N} + b_{N-1} x^{a_{N-1}} + \dots + b_1 x^{a_1} + b_0 x^{a_0}$  and  $a_N > a_{N-1} > \dots > a_1 > a_0 > 0$ . If the sequence  $b_0, \dots, b_N$  changes its sign only once, then  $p(x)$  has exactly one positive root.

*Proof:* The proof can be found in Appendix D. ■

Based on Lemma 1 and Theorem 1, we have

**Theorem 2:** Under Rayleigh fading scenario, the optimal decision fusion rule can be reduced to a threshold test on  $\psi$

$$\psi \underset{H_0}{\overset{H_1}{\geq}} T_1. \quad (13)$$

*Proof:* The proof can be found in Appendix E. ■

**Remark 1:** Let  $x_{Ik}$  and  $x_{Qk}$  be the real and imaginary parts of  $h_k u_k$ . We can show that

$$\frac{f(x_{Ik}^{(1)}, x_{Qk}^{(1)} | H_1)}{f(x_{Ik}^{(1)}, x_{Qk}^{(1)} | H_0)} \geq \frac{f(x_{Ik}^{(2)}, x_{Qk}^{(2)} | H_1)}{f(x_{Ik}^{(2)}, x_{Qk}^{(2)} | H_0)}$$

for any  $x_{Ik}^{(1)}, x_{Qk}^{(1)}, x_{Ik}^{(2)}, x_{Qk}^{(2)}$  such that  $(x_{Ik}^{(1)})^2 + (x_{Qk}^{(1)})^2 \geq (x_{Ik}^{(2)})^2 + (x_{Qk}^{(2)})^2$ . However, this monotone likelihood ratio property is not a sufficient condition for a threshold test on  $\psi$  as shown in (13) (i.e., we can not derive Theorem 2 purely based on the monotone likelihood ratio property). This should convince the reader that Theorem 2 is a nontrivial result.

Suppose (12) has one positive root  $x_r < 1$ . Then the threshold  $T_1$  for the test on  $\psi$  is given by  $T_1 = -\ln x_r$ . Due to the uniqueness of the positive root,  $x_r$  can be numerically calculated using bisection method.

3) **Error Probability:** With the help of Theorem 2, we can derive the error probability

$$\begin{aligned} P_e &= P_1 \int_0^{T_1} f(\psi | H_1) d\psi + P_0 \int_{T_1}^{\infty} f(\psi | H_0) d\psi \quad (14) \\ &= P_1 - P_1 \sum_{m=0}^N \binom{N}{m} P_d^m (1 - P_d)^{N-m} \\ &\quad \times \exp\left(-\frac{T_1}{2(m\sigma_h^2 + \sigma_n^2)}\right) \\ &\quad + P_0 \sum_{m=0}^N \binom{N}{m} P_f^m (1 - P_f)^{N-m} \\ &\quad \times \exp\left(-\frac{T_1}{2(m\sigma_h^2 + \sigma_n^2)}\right). \quad (15) \end{aligned}$$

We can easily obtain  $P_e$  by numerically calculating  $T_1$  and substituting it into (15). Fig. 3 shows the numerically calculated error probabilities for two different cases. The two cases have the same  $\sigma_n$  but different  $\sigma_h$ . A larger  $\sigma_h$  implies a better chance to see a large channel gain and hence a smaller error probability.

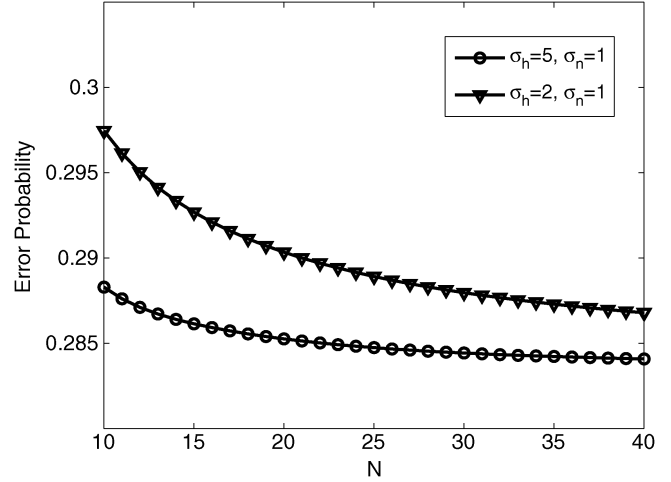


Fig. 3. Detection performance.

### B. Rician Fading

Under Rician fading scenario, we assume that the real and imaginary parts of  $h_k$  follow a joint Gaussian distribution with mean  $[\mu, 0]^T$  and covariance matrix  $\sigma_h^2 \mathbf{I}$ .

1) **Sufficient Statistic and Optimal Decision Fusion Rule:** Under Rician fading scenario, the conditional joint distribution  $f(y_I, y_Q | H_0)$  is given by

$$\sum_{m=0}^N \binom{N}{m} \frac{(P_f)^m (1 - P_f)^{N-m}}{2\pi [Nb^2\sigma_h^2 + (a^2 - b^2)\sigma_h^2 m + \sigma_n^2]} \times \exp\left(-\frac{(y_I - [(a+b)\mu m - b\mu N])^2 + y_Q^2}{2[Nb^2\sigma_h^2 + (a^2 - b^2)\sigma_h^2 m + \sigma_n^2]}\right). \quad (16)$$

By replacing  $P_f$  with  $P_d$  in (16), we get  $f(y_I, y_Q | H_1)$ . At the fusion center, the optimal fusion rule will be the following likelihood ratio test

$$\frac{f(y_I, y_Q | H_1)}{f(y_I, y_Q | H_0)} \underset{H_0}{\overset{H_1}{\geq}} \frac{P_0}{P_1}. \quad (17)$$

Again, we use  $\Omega_0$  and  $\Omega_1$  to denote the decision regions for  $H_0$  and  $H_1$  respectively. Due to the continuity of  $f(y_I, y_Q | H_0)$  and  $f(y_I, y_Q | H_1)$ , the boundaries of  $\Omega_0$  and  $\Omega_1$  can be found by solving

$$g(y_I, y_Q) \triangleq f(y_I, y_Q | H_0) P_0 - f(y_I, y_Q | H_1) P_1 = 0. \quad (18)$$

Substituting (16) and the expression of  $f(y_I, y_Q | H_1)$  into (18), we get

$$\begin{aligned} &\sum_{m=0}^N \left\{ [(P_f)^m (1 - P_f)^{N-m} P_0 - (P_d)^m (1 - P_d)^{N-m} P_1] \right. \\ &\quad \times \binom{N}{m} \frac{1}{2\pi [Nb^2\sigma_h^2 + (a^2 - b^2)\sigma_h^2 m + \sigma_n^2]} \\ &\quad \times \exp\left(-\frac{[(a+b)\mu m - b\mu N]^2}{2[Nb^2\sigma_h^2 + (a^2 - b^2)\sigma_h^2 m + \sigma_n^2]}\right) \\ &\quad \left. \times x^{\frac{1}{2[Nb^2\sigma_h^2 + (a^2 - b^2)\sigma_h^2 m + \sigma_n^2]}} \right\} = 0 \quad (19) \end{aligned}$$

where  $x = \exp(-[(y_I + A)^2 + y_Q^2 - A^2])$  and  $A = \frac{(abN\sigma_h^2\mu + \sigma_n^2\mu)}{(a-b)\sigma_h^2}$ . Based on the above result, we can establish the following theorem.

*Theorem 3:* Under Rician fading scenario, the optimal decision fusion rule can be reduced to the following threshold test:

$$(y_I + A)^2 + y_Q^2 - A^2 \underset{H_0}{\overset{H_1}{\geq}} T_2. \quad (20)$$

*Proof:* The approach used to prove Theorem 2 can be applied to prove this theorem. ■

2) *Error Probability:* Once the threshold  $T_2$  is obtained, we can calculate the error probability as

$$P_e = P_0 \iint_{(y_I+A)^2 + y_Q^2 - A^2 > T_2} f(y_I, y_Q|H_0) dy_I dy_Q + P_1 \iint_{(y_I+A)^2 + y_Q^2 - A^2 < T_2} f(y_I, y_Q|H_1) dy_I dy_Q. \quad (21)$$

When the channel is subject to Rician fading, OOK is in general not the most energy efficient modulation scheme. Notice that the transmission power per sensor is given by  $P_w = (P_d P_1 + P_f P_0) a^2 + (1 - P_d P_1 - P_f P_0) b^2$ . We may wish to optimize  $a$  and  $b$  subject to the power constraint  $P_w \leq c$ . The fact that the error probability associated with a particular pair of  $a$  and  $b$  can be numerically calculated via (21) makes it a lot easier for us to design an energy efficient modulation scheme under Rician fading scenario.

#### IV. ASYMPTOTIC BEHAVIOR

In this section, we will evaluate large system performance under both Rayleigh and Rician fading scenarios.

##### A. Rayleigh Fading

As can be observed from Fig. 3, the curves of error probability flatten out as the number of sensors grows. Will similar phenomenon be observed when parameters such as  $\sigma_h$  and  $\sigma_n$  change? We will answer the question in this subsection. Again, we assume that  $a = 1$  and  $b = 0$ .

To evaluate the system performance as the number of sensors goes large, we construct

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \frac{1}{\sqrt{P_f N \sigma_h^2}} \begin{bmatrix} y_I \\ y_Q \end{bmatrix}$$

which is merely a scaled version of  $[y_I, y_Q]^T$ . When the number of sensors goes to infinity, we have

*Theorem 4:* If both  $\sigma_n^2$  and  $\sigma_h^2$  are finite, then  $f(\mathbf{z}|H_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I})$  and  $f(\mathbf{z}|H_1) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \frac{P_d}{P_f} \mathbf{I})$  as  $N \rightarrow \infty$ , where  $\xrightarrow{d}$  denotes convergence in distribution.

*Proof:* The detailed proof can be found in Appendix F. ■

Since  $\mathcal{N}(\mathbf{0}, \mathbf{I})$  and  $\mathcal{N}(\mathbf{0}, \frac{P_d}{P_f} \mathbf{I})$  can not be separated without error, an error floor occurs (similar results are reported in the study of type-based schemes [13], [15] when the channel gain

has zero mean). The error that results from distinguishing between the two joint Gaussian distributions (or the error floor) is given by

$$P_e = P_1 \iint_{\Omega_0} \frac{1}{2\pi \frac{P_d}{P_f}} \exp\left(-\frac{t_1^2 + t_2^2}{2 \frac{P_d}{P_f}}\right) dt_1 dt_2 + P_0 \iint_{\Omega_1} \frac{1}{2\pi} \exp\left(-\frac{t_1^2 + t_2^2}{2}\right) dt_1 dt_2 \quad (22)$$

where

$$\hat{\Omega}_1 = \left\{ t_1, t_2 \left| \frac{P_1}{2\pi \frac{P_d}{P_f}} \exp\left(-\frac{t_1^2 + t_2^2}{2 \frac{P_d}{P_f}}\right) > \frac{P_0}{2\pi} \exp\left(-\frac{t_1^2 + t_2^2}{2}\right) \right. \right\}$$

or equivalently

$$\hat{\Omega}_1 = \left\{ t_1, t_2 \left| t_1^2 + t_2^2 > \frac{2 \ln \frac{P_0 P_d}{P_1 P_f}}{\left(1 - \frac{P_f}{P_d}\right)} \right. \right\}$$

and  $\hat{\Omega}_0$  is the complement set of  $\hat{\Omega}_1$ . Under the assumption of  $P_d > P_f$ , if  $P_0 P_d < P_1 P_f$ ,  $\hat{\Omega}_1$  will be the entire  $\mathbb{R}^2$  plane and the error probability will be equal to  $P_0$ . Otherwise,  $P_e$  is given by

$$P_e = \int_0^{2\pi} d\theta \int_0^{\sqrt{T_3}} \frac{P_1}{2\pi \frac{P_d}{P_f}} \exp\left(-\frac{r^2}{2 \frac{P_d}{P_f}}\right) r dr + \int_0^{2\pi} d\theta \int_{\sqrt{T_3}}^{\infty} \frac{P_0}{2\pi} \exp\left(-\frac{r^2}{2}\right) r dr = P_1 \left[ 1 - \exp\left(-\frac{T_3}{2 \frac{P_d}{P_f}}\right) \right] + P_0 \exp\left(-\frac{T_3}{2}\right) \quad (23)$$

where

$$T_3 = \frac{2 \ln \frac{P_0 P_d}{P_1 P_f}}{1 - \frac{P_f}{P_d}}.$$

Note that the integrals in (22) are calculated in polar coordinates in (23). The error floor is plotted in Fig. 4 as a function of  $P_1$  for different values of  $P_d$  while  $P_f = 0.02$  is kept unchanged. Since the error probability can not be driven arbitrarily close to zero by increasing the number of sensors, the error exponent is zero.

##### B. Rician Fading

Unlike the Rayleigh fading channel, a Rician fading channel can lead to a positive error exponent. This can be understood by closely examining (4). Under Rayleigh fading scenario, the law of large numbers suggests that

$$\begin{bmatrix} \text{Re} \left[ \frac{1}{N} \sum_{k=1}^N h_k x_k \right] \\ \text{Im} \left[ \frac{1}{N} \sum_{k=1}^N h_k x_k \right] \end{bmatrix} \quad (24)$$

will converge to  $\mathbf{0}$  under both hypotheses as the number of sensors increases (because of the fact that the mean of

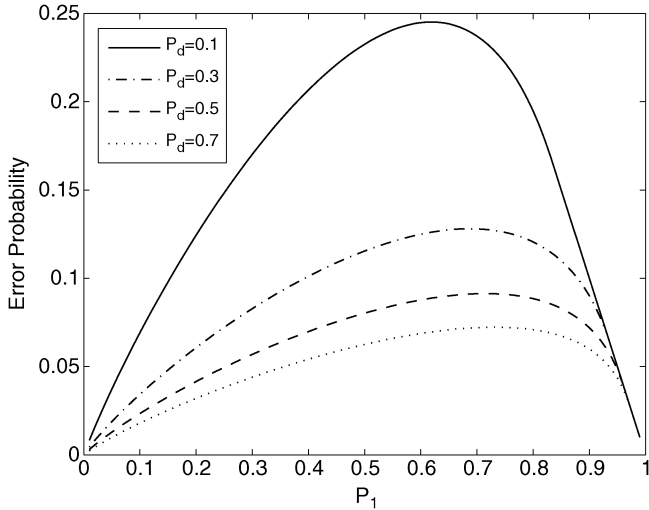


Fig. 4. Limitations in performance.

$[\text{Re}[h_k x_k], \text{Im}[h_k x_k]]^T$  equals to  $\mathbf{0}$  under both  $H_0$  and  $H_1$ ). However, under Rician fading scenario, the vector in (24) will converge to different vectors under  $H_0$  and  $H_1$ . In the following, we shall evaluate the asymptotic behavior of the system under Rician fading scenario.

We start by examining the asymptotic behavior of a simple scheme. Define the scaled received symbol  $\tilde{y}$  as  $\tilde{y} = \frac{1}{N} [\sum_{k=1}^N h_k x_k + n]$ . Use  $\tilde{y}_I$  and  $\tilde{y}_Q$  to denote the real and imaginary parts of  $\tilde{y}$ . Consider the following threshold test:

$$(\tilde{y}_I + \tilde{A})^2 + \tilde{y}_Q^2 - \tilde{A}^2 \underset{H_0}{\overset{H_1}{\geq}} \tau \quad (25)$$

where the threshold  $\tau$  is a constant (i.e., not a function of  $N$ ) and  $\tilde{A} = \frac{ab\mu}{(a-b)}$ . The error exponent  $E_\tau$  associated with the decision fusion rule (25) can be written as

$$E_\tau = \min(E_{F,\tau}, E_{M,\tau}) \quad (26)$$

where  $E_{F,\tau}$  and  $E_{M,\tau}$  are exponents of false alarm and miss probabilities. When there are  $N$  sensors deployed, the logarithmic moment generating function associated with  $[\tilde{y}_I, \tilde{y}_Q]^T$  under  $H_i$  is defined as

$$\Lambda_i^{(N)}(\lambda_1, \lambda_2) \triangleq \ln E[\exp(\lambda_1 \tilde{y}_I + \lambda_2 \tilde{y}_Q) | H_i].$$

We further define

$$\Lambda_i(\lambda_1, \lambda_2) \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_i^{(N)}(N\lambda_1, N\lambda_2).$$

We can show that

$$\begin{aligned} \Lambda_0(\lambda_1, \lambda_2) &= \ln \left( P_f \exp \left( a\mu\lambda_1 + \frac{1}{2}a^2\sigma_h^2\lambda_1^2 \right) \right. \\ &\quad \times \exp \left( \frac{1}{2}a^2\sigma_h^2\lambda_2^2 \right) + (1 - P_f) \\ &\quad \times \exp \left( -b\mu\lambda_1 + \frac{1}{2}b^2\sigma_h^2\lambda_1^2 \right) \\ &\quad \left. \times \exp \left( \frac{1}{2}b^2\sigma_h^2\lambda_2^2 \right) \right). \end{aligned} \quad (27)$$

The detailed derivation can be found in Appendix G. By replacing  $P_f$  with  $P_d$  in (27), we get  $\Lambda_1(\lambda_1, \lambda_2)$ . Let  $\Lambda_0^*(v_1, v_2)$  and  $\Lambda_1^*(v_1, v_2)$  be the Fenchel–Legendra transform of  $\Lambda_0(\lambda_1, \lambda_2)$  and  $\Lambda_1(\lambda_1, \lambda_2)$ , i.e.,

$$\Lambda_i^*(v_1, v_2) \triangleq \sup_{\lambda_1, \lambda_2 \in \mathbb{R}} [\lambda_1 v_1 + \lambda_2 v_2 - \Lambda_i(\lambda_1, \lambda_2)]. \quad (28)$$

Both  $\Lambda_i(v_1, v_2)$  and  $\Lambda_i^*(v_1, v_2)$  are convex functions [23]. Since  $\Lambda_0(\lambda_1, \lambda_2)$  and  $\Lambda_1(\lambda_1, \lambda_2)$  are essentially smooth and lower semicontinuous functions, we can apply Gartner–Ellis theorem [23] to show that

$$\begin{aligned} E_{F,\tau} &= - \lim_{N \rightarrow \infty} \frac{1}{N} \ln \Pr \{ [\tilde{y}_I, \tilde{y}_Q]^T \in \tilde{\mathcal{S}}_\tau | H_0 \} \\ &= \inf_{[v_1, v_2]^T \in \tilde{\mathcal{S}}_\tau} \Lambda_0^*(v_1, v_2) \end{aligned} \quad (29)$$

$$\begin{aligned} E_{M,\tau} &= - \lim_{N \rightarrow \infty} \frac{1}{N} \ln \Pr \{ [\tilde{y}_I, \tilde{y}_Q]^T \in \mathcal{S}_\tau | H_1 \} \\ &= \inf_{[v_1, v_2]^T \in \mathcal{S}_\tau} \Lambda_1^*(v_1, v_2) \end{aligned} \quad (30)$$

where

$$\begin{aligned} \mathcal{S}_\tau &= \{ [v_1, v_2]^T \in \mathbb{R}^2 | (v_1 + \tilde{A})^2 + v_2^2 - \tilde{A}^2 \leq \tau \} \\ \tilde{\mathcal{S}}_\tau &= \{ [v_1, v_2]^T \in \mathbb{R}^2 | (v_1 + \tilde{A})^2 + v_2^2 - \tilde{A}^2 \geq \tau \}. \end{aligned}$$

The mean of the real and imaginary parts of  $h_k x_k$  under  $H_0$  is  $\tilde{\mu}_0 = -b\mu + (a+b)P_f\mu$  and 0, respectively. The mean of the real and imaginary parts of  $h_k x_k$  under  $H_1$  is  $\tilde{\mu}_1 = -b\mu + (a+b)P_d\mu$  and 0, respectively. Hence, as  $N$  goes to infinity,  $[\tilde{y}_I, \tilde{y}_Q]^T$  will converge to  $\boldsymbol{\mu}_0 = [\tilde{\mu}_0, 0]^T$  under  $H_0$  and  $\boldsymbol{\mu}_1 = [\tilde{\mu}_1, 0]^T$  under  $H_1$ . To make sure that neither  $E_{F,\tau}$  nor  $E_{M,\tau}$  is zero (or equivalently  $E_\tau \neq 0$ ), we need  $\boldsymbol{\mu}_0 \notin \tilde{\mathcal{S}}_\tau$  and  $\boldsymbol{\mu}_1 \notin \mathcal{S}_\tau$ , or equivalently  $\tilde{\mu}_0^2 + 2\tilde{\mu}_0\tilde{A} < \tau < \tilde{\mu}_1^2 + 2\tilde{\mu}_1\tilde{A}$ . Let  $\mathcal{S} = \{ \tau | \tilde{\mu}_0^2 + 2\tilde{\mu}_0\tilde{A} < \tau < \tilde{\mu}_1^2 + 2\tilde{\mu}_1\tilde{A} \}$ .

As  $\tau \in \mathcal{S}$  increases, the area of  $\mathcal{S}_\tau$  becomes larger and the area of  $\tilde{\mathcal{S}}_\tau$  becomes smaller. Hence,  $\inf_{[v_1, v_2]^T \in \tilde{\mathcal{S}}_\tau} \Lambda_0^*(v_1, v_2)$  is a nondecreasing function of  $\tau$  and  $\inf_{[v_1, v_2]^T \in \mathcal{S}_\tau} \Lambda_1^*(v_1, v_2)$  is a nonincreasing function of  $\tau$  for  $\tau \in \mathcal{S}$ . We gradually increase  $\tau$  until  $E_{F,\tau} = E_{M,\tau}$ . Let  $\tau^* \in \mathcal{S}$  be the value of  $\tau$  such that  $E_{F,\tau^*} = E_{M,\tau^*}$ . We have the following theorem.

**Theorem 5:** The following decision fusion rule with a constant threshold  $\tau^*$

$$(\tilde{y}_I + \tilde{A})^2 + \tilde{y}_Q^2 - \tilde{A}^2 \underset{H_0}{\overset{H_1}{\geq}} \tau^* \quad (31)$$

can achieve the same error exponent as the optimal decision fusion rule (20) can.

**Proof:** The detailed proof can be found in Appendix H. ■

Theorem 5 suggests that the decision fusion rule (31) is asymptotically optimal. It is also worth noting that  $T_2$  in the optimal decision fusion rule (20) is a function of  $N$  while  $\tau^*$  in (31) is not.

Based on the convexity of  $\Lambda_0(\lambda_1, \lambda_2)$  and  $\Lambda_1(\lambda_1, \lambda_2)$ , we have the following theorem.

**Theorem 6:** For any  $\tau \in \mathcal{S}$ , we have

$$E_{F,\tau} = \inf_{[v_1, v_2]^T \in \tilde{\mathcal{S}}} \Lambda_0^*(v_1, v_2) \quad (32)$$

$$E_{M,\tau} = \Lambda_1^*(\sqrt{\tau + \tilde{A}^2} - \tilde{A}, 0) \quad (33)$$

where

$$\hat{S} = \left\{ [v_1, v_2]^T \in \mathbb{R}^2 \mid (v_1 + \tilde{A})^2 + v_2^2 - \tilde{A}^2 = \tau \right\}. \quad (34)$$

*Proof:* The detailed proof can be found in Appendix I. ■

For any given  $v_1$  and  $v_2$ ,  $\Lambda_i^*(v_1, v_2)$  can be obtained by solving a convex optimization problem. This is because of the concavity of  $\lambda_1 v_1 + \lambda_2 v_2 - \Lambda_i(\lambda_1, \lambda_2)$ . So for any given  $\tau$ ,  $E_{M,\tau}$  can be obtained by solving a convex optimization problem as suggested by (33). Given  $\tau$ , we can calculate  $E_{F,\tau}$  by performing a one-dimensional search (because feasible  $v_1$  and  $v_2$  are dependent as shown in (34)). To find  $\tau^*$ , we only need to gradually increase  $\tau$  until  $E_{F,\tau} = E_{M,\tau}$ . Once we find  $\tau^*$ , the error exponent is given by

$$E_{\tau^*} = E_{F,\tau^*} = E_{M,\tau^*} = \Lambda_1^*(\sqrt{\tau^* + \tilde{A}^2} - \tilde{A}, 0).$$

According to Theorem 5, this error exponent is also the error exponent associated with the optimal decision fusion rule (20).

Given a certain energy constraint, we can optimize  $a$  and  $b$  such that the largest error exponent is achieved. An example is given in Section V.

## V. NUMERICAL RESULTS

In this section we show some simulation results. We first focus on the Rayleigh fading scenario. As before, we assume that  $a = 1$  and  $b = 0$  when Rayleigh fading is considered. It is pointed out in Section IV that the error probability will not converge to zero even when we deploy a large number of sensors. So we are interested in the performance when a small number of sensors is deployed. In the following, we shall compare the multiaccess scheme with the traditional parallel access scheme, where we assume different sensors communicate with the fusion center at different time slots, in order to understand the pros and cons of the multiaccess scheme. Define  $\rho = \frac{\sigma_h^2}{\sigma_n^2}$ . Increasing transmission power at local sensors is equivalent to increasing  $\sigma_h^2$ . So  $\rho$  can be viewed as a measure of effective transmission power level. Fig. 5 shows the simulated performance of the two access schemes for  $N = 5$  and  $N = 10$ . As  $\rho$  increases, both schemes will achieve a better performance. We can also observe that the multiaccess scheme outperforms the parallel access scheme when  $\rho$  is small and loses superiority when  $\rho$  is large.

There are two major reasons why the multiaccess scheme performs better when sensors' transmission power is low. First, the multiaccess scheme will provide a higher SNR at the receiver of the fusion center than the parallel access scheme can. Suppose there are  $M = m$  sensors transmitting. The SNR at the fusion center will be

$$\frac{\int_0^\infty \psi f(\psi|m) d\psi}{2\sigma_n^2} = \frac{m\sigma_h^2}{\sigma_n^2}. \quad (35)$$

In the parallel access scheme, one time slot will be assigned to only one sensor. If it transmits, then the SNR will be  $\frac{\sigma_h^2}{\sigma_n^2}$  which is only  $\frac{1}{m}$  of (35). Such an increment in the SNR is extremely beneficial when the noise power at the fusion center is large. The

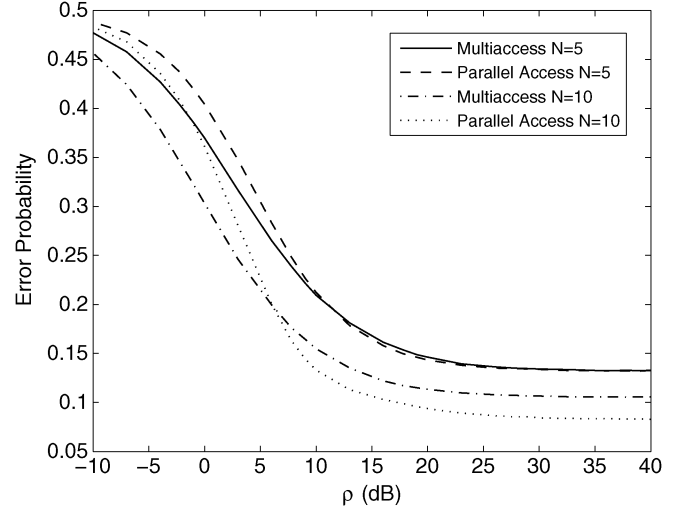


Fig. 5. Comparisons between two schemes with  $P_1 = 0.5$ ,  $P_d = 0.3$  and  $P_f = 0.02$ .

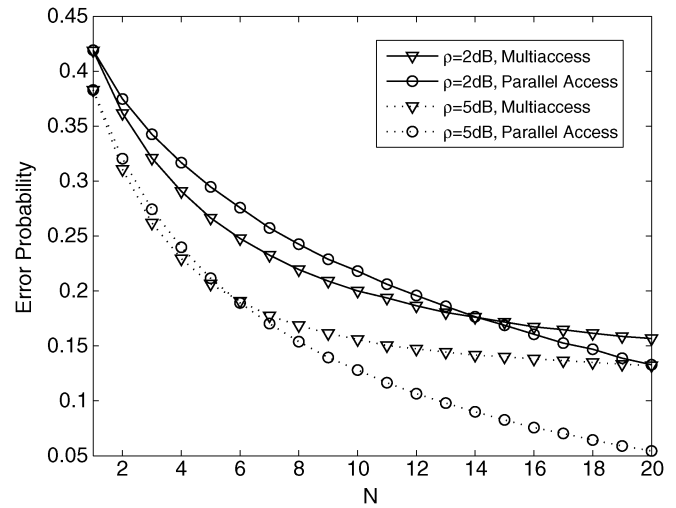


Fig. 6. Comparisons between two schemes with  $P_1 = 0.5$ ,  $P_d = 0.5$  and  $P_f = 0.02$ .

second reason, which is already stated in Section III, is that the fusion center will still be able to roughly infer the total number of sensors that decide  $H_1$  based on the received signal power. When the transmission power is extremely large, we could tell almost surely how many sensors have decided  $H_1$  if the parallel access scheme is used. Due to the fading channel, transmitted pulses may cancel out each other in the multiaccess scheme. Although the SNR is high, it is still difficult to tell exactly how many sensors have decided  $H_1$ . This suggests that the multiaccess scheme will be a good choice when the sensor's transmission power is low but not a good choice when the transmission power is high.

In Fig. 6, we find that the parallel access scheme has a better performance when the number of sensors is large. Also notice that the number of sensors required for the parallel access scheme to beat the multiaccess scheme increases when  $\rho$  decreases.



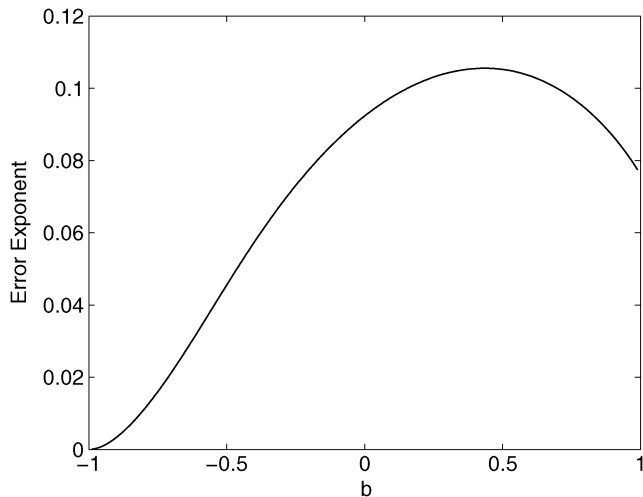


Fig. 7. Error exponent as  $b$  varies ( $\sigma_n^2 = \sigma_n^2 = 1$ ,  $\mu = 1$ ,  $P_d = 0.5$ ,  $P_f = 0.02$ ,  $P_1 = 0.5$ ).

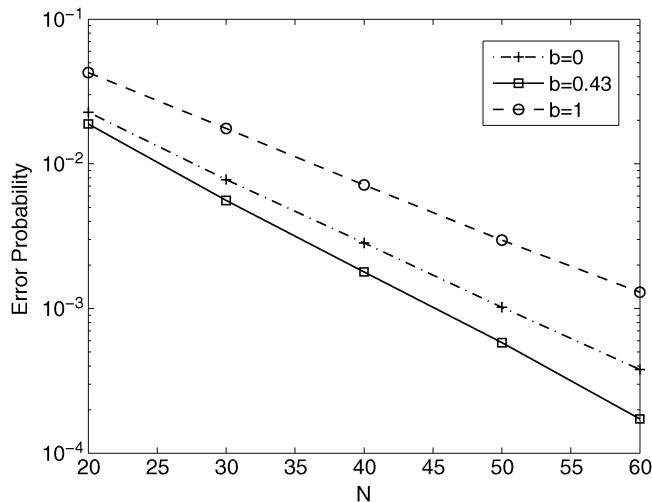


Fig. 8. Error probabilities of different modulation schemes.

Next, we focus on the Rician fading scenario and show how the design of modulation scheme could affect the performance. The transmission power per sensor is given by  $P_w = (P_d P_1 + P_f P_0)a^2 + (1 - P_d P_1 - P_f P_0)b^2$ . We are free to choose  $a$  and  $b$  as long as the constraint

$$(P_d P_1 + P_f P_0)a^2 + (1 - P_d P_1 - P_f P_0)b^2 = 1 \quad (36)$$

is satisfied (i.e., each sensor has unit transmission power). For different values of  $b$  [notice that for any given  $b$ ,  $a$  can be determined from (36)], Fig. 7 shows the corresponding error exponent. It can be observed from Fig. 7 that by choosing  $b = 0.43$  we can achieve the maximum error exponent. In Fig. 8, Monte Carlo simulations are used to examine the performance of three modulation schemes. In the first scheme,  $b = 0$  (which corresponds to OOK); in the second scheme,  $a = b = 1$  (which corresponds to BPSK); in the third scheme,  $b = 0.43$ . As can be observed, the third scheme offers the best performance.

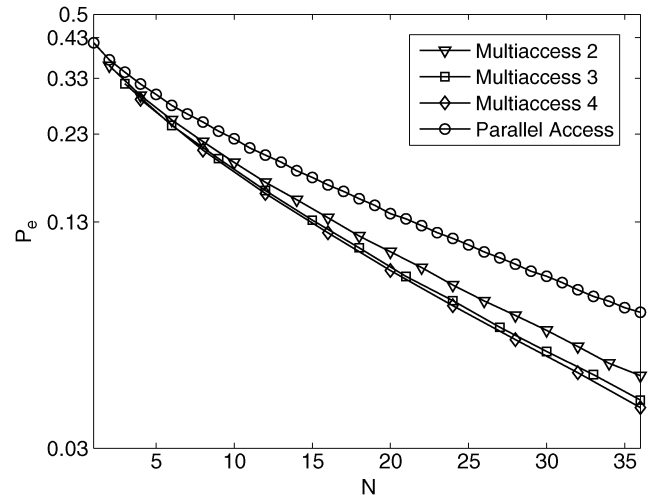


Fig. 9. Multiaccess scheme with multiple time slots with  $P_1 = 0.5$ ,  $P_d = 0.7$ ,  $P_f = 0.02$ , and  $\rho = 0$  dB.

## VI. CONCLUSION AND DISCUSSION

In this work, we consider the fusion of local binary decisions over multiaccess channels. Under Rayleigh fading scenario, although the system exhibits an error floor, the multiaccess scheme still outperforms the parallel access scheme when a small number of sensors is deployed and each sensor's transmission power is low. Under Rician fading scenario, an exponentially decreasing error probability can be achieved. Under both fading scenarios, we optimize local sensors' modulation schemes and simplify decision fusion rules at the fusion center.

Finally, we wish to say a few words about how to improve the performance of the multiaccess scheme when we have a large number of sensors and the channel is subject to Rayleigh fading. One solution is to divide sensors into several groups and allocate each group a dedicated time slot to communicate with the fusion center. A group of sensors can transmit simultaneously in the time slot allocated to that group. The number of groups grows as the total number of sensors increases. Such a multiaccess scheme can achieve an exponentially decreasing error probability. Fig. 9 compares the performance of multiaccess schemes with group size of two, three and four ("group size of  $L$ " means that each group will consist of  $L$  sensors) with that of the parallel access scheme. According to our simulation results, the multiaccess scheme described above can perform better than the parallel access scheme when sensors' transmission power is low. Moreover, the multiaccess scheme requires less bandwidth.

## APPENDIX A

### AN ASSUMPTION WITHOUT LOSS OF GENERALITY

Use  $h_{Ik}$  and  $h_{Qk}$  to denote the real and imaginary parts of  $h_k$ . Suppose  $h_{Ik}$  and  $h_{Qk}$  follow a joint Gaussian distribution with mean  $[\mu_I, \mu_Q]^T$  and covariance matrix  $\sigma_h^2 \mathbf{I}$  and  $\sqrt{\mu_I^2 + \mu_Q^2} = \mu$ . The real and imaginary parts of the received complex symbol  $y$  (denoted by  $y_I$  and  $y_Q$ , respectively) are given by

$$\begin{bmatrix} y_I \\ y_Q \end{bmatrix} = \sum_{k=1}^N \begin{bmatrix} h_{Ik} \\ h_{Qk} \end{bmatrix} x_k + \begin{bmatrix} n_I \\ n_Q \end{bmatrix} \quad (37)$$

where  $n_I$  and  $n_Q$  are the real and imaginary parts of the noise  $n$  and follow a joint Gaussian distribution with mean  $\mathbf{0}$  and covariance matrix  $\sigma_n^2 \mathbf{I}$ . Multiplying both sides of (37) by a rotation matrix, we get

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} y_I \\ y_Q \end{bmatrix} = \sum_{k=1}^N \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} h_{Ik} \\ h_{Qk} \end{bmatrix} x_k + \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} n_I \\ n_Q \end{bmatrix} \quad (38)$$

where  $\cos \theta = \frac{\mu_I}{\mu}$  and  $\sin \theta = -\frac{\mu_Q}{\mu}$ . Let the left-hand side of (38) be  $[\hat{y}_I, \hat{y}_Q]^T$ . Equation (38) now becomes

$$\begin{bmatrix} \hat{y}_I \\ \hat{y}_Q \end{bmatrix} = \sum_{k=1}^N \begin{bmatrix} \hat{h}_{Ik} \\ \hat{h}_{Qk} \end{bmatrix} x_k + \begin{bmatrix} \hat{n}_I \\ \hat{n}_Q \end{bmatrix} \quad (39)$$

where  $\hat{h}_{Ik}$  and  $\hat{h}_{Qk}$  follow a joint Gaussian distribution with mean  $[\mu, 0]^T$  and covariance matrix  $\sigma_h^2 \mathbf{I}$ , and  $\hat{n}_I$  and  $\hat{n}_Q$  follow a joint Gaussian distribution with mean  $\mathbf{0}$  and covariance matrix  $\sigma_n^2 \mathbf{I}$ . We can write (39) in a more compact form as follows:

$$\hat{y} = \sum_{k=1}^N \hat{h}_k x_k + \hat{n}. \quad (40)$$

It is obvious that (40) and (4) are two identical models, because the distributions of  $h_k$  and  $n$  in (4) are exactly the same as the distributions of  $\hat{h}_k$  and  $\hat{n}$  in (40). As a result, we do not lose generality by assuming that the real and imaginary parts of  $h_k$  in (4) follow a joint Gaussian distribution with mean  $[\mu, 0]^T$ .

#### APPENDIX B OPTIMALITY OF OOK

Suppose sensors have the capability to transmit quadrature amplitude modulation (QAM) signals. We now consider a more generalized scheme. A sensor will send a complex baseband symbol  $s_0 = s_{0I} + js_{0Q}$  instead of staying silent when it decides  $H_0$  to be true. When  $H_1$  is decided,  $s_1 = s_{1I} + js_{1Q}$  (we assume that  $|s_1|^2 > |s_0|^2$ ) is sent. The average transmission power per sensor is  $P_w = (P_d P_1 + P_f P_0) |s_1|^2 + (1 - P_d P_1 - P_f P_0) |s_0|^2$ . However, for any  $s_0 \neq 0$ , we can always construct an OOK modulation scheme (e.g., each sensor sends  $\sqrt{|s_1|^2 - |s_0|^2} + 0j$  (or  $0 + 0j$ ) when  $H_1$  (or  $H_0$ ) is decided), which consumes less power but offers better or equivalent performance.

#### APPENDIX C PROOF OF LEMMA 1

Rewrite  $S_k$  as  $S_k = A_0 \left(\frac{p}{1-p}\right)^k (1-p)^K - A_1 \left(\frac{q}{1-q}\right)^k (1-q)^K$ . Without loss of generality, we assume  $0 < p < q < 1$ . We have  $\frac{q}{1-q} > \frac{p}{1-p}$ . Suppose for a given  $\kappa$  we have  $S_\kappa < 0$ , then for all  $k > \kappa$ , we have  $S_k < 0$ . This can be proved by induction. First, we examine  $S_{\kappa+1}$ .  $S_\kappa < 0$  implies that  $A_1 \left(\frac{q}{1-q}\right)^\kappa (1-q)^K > A_0 \left(\frac{p}{1-p}\right)^\kappa (1-p)^K$ . Due to the fact that  $\frac{q}{1-q} > \frac{p}{1-p}$ , we have

$$\begin{aligned} & A_1 \left(\frac{q}{1-q}\right)^{(\kappa+1)} (1-q)^K \\ &= A_1 \left(\frac{q}{1-q}\right)^\kappa (1-q)^K \left(\frac{q}{1-q}\right) \end{aligned}$$

$$\begin{aligned} & > A_0 \left(\frac{p}{1-p}\right)^\kappa (1-p)^K \left(\frac{q}{1-q}\right) \\ & > A_0 \left(\frac{p}{1-p}\right)^\kappa (1-p)^K \left(\frac{p}{1-p}\right) \\ &= A_0 \left(\frac{p}{1-p}\right)^{(\kappa+1)} (1-p)^K. \end{aligned}$$

As a result,  $S_{\kappa+1}$  is also smaller than zero. Similarly, we can prove that  $S_{\kappa+2}$  must be negative and so on. From the above results, we know that  $S_k$  will be always nonnegative, or always negative, or first nonnegative and then negative. So the sign of  $S_k$  will change at most once.

#### APPENDIX D PROOF OF THEOREM 1

Suppose  $p(x)$  has more than one positive root. Arbitrarily pick two positive roots and denote them by  $r_1$  and  $r_2$ . Without loss of generality, we assume  $r_1 < r_2$ . Since the sequence  $b_0, \dots, b_N$  changes its sign only once, we can assume that  $b_0 > 0$  and  $b_N < 0$  (the reason is that by multiplying  $p(x)$  by  $-1$ , the case where  $b_0 < 0$  and  $b_N > 0$  can be treated in a similar way). Then there exists an integer  $k$  such that  $b_i < 0 \forall i > k$  and  $b_i \geq 0 \forall i \leq k$ . Since  $r_2 > r_1$ , we have  $r_2^{a_i - a_k} > r_1^{a_i - a_k} \forall i > k$  and  $r_2^{a_i - a_k} \leq r_1^{a_i - a_k} \forall i \leq k$ . Combining these results, we get

$$\sum_{i=0}^k b_i r_2^{a_i - a_k} + \sum_{i=k+1}^N b_i r_2^{a_i - a_k} < \sum_{i=0}^k b_i r_1^{a_i - a_k} + \sum_{i=k+1}^N b_i r_1^{a_i - a_k}. \quad (41)$$

Since  $r_1$  is the root of  $p(x)$ , we have  $\sum_{i=0}^k b_i r_1^{a_i} + \sum_{i=k+1}^N b_i r_1^{a_i} = 0$  or equivalently  $\sum_{i=0}^k b_i r_1^{a_i - a_k} + \sum_{i=k+1}^N b_i r_1^{a_i - a_k} = 0$ . Similarly, we have  $\sum_{i=0}^k b_i r_2^{a_i - a_k} + \sum_{i=k+1}^N b_i r_2^{a_i - a_k} = 0$ . But these contradict (41). Now we only have to show that  $p(x)$  has at least one positive root. When  $x > 0$  is sufficiently small, the sign of  $p(x)$  is determined by  $b_0$  which is positive. When  $x > 0$  is sufficiently large, the sign of  $p(x)$  is determined by  $b_N$  which is negative. Due to the continuity of  $p(x)$  on  $(0, \infty)$ ,  $p(x)$  will have at least one positive root. The theorem follows.

#### APPENDIX E PROOF OF THEOREM 2

Since  $\binom{N}{m} \frac{1}{2(m\sigma_h^2 + \sigma_n^2)} > 0$ , the coefficients of  $x^{\frac{1}{2(m\sigma_h^2 + \sigma_n^2)}}$  on the left-hand side of (12) will change signs at most once according to Lemma 1. If the signs don't change, the fusion center will either always decide  $H_0$  or always decide  $H_1$ . So the threshold of the test will be either  $\infty$  or  $0$ . If the signs of the coefficients change once, then (12) will have one positive root according to Theorem 1. If the positive root is smaller than 1,  $g(\psi)$  will only have one positive root, which is due to the fact that  $x$  is a strictly decreasing function of  $\psi$ . The sign of  $g(\psi)$  changes at this unique root and this results in a threshold test on  $\psi$ . If the positive root is no smaller than 1, then the fusion center will always decide one hypothesis to be true (because  $\exp(-\psi) < 1$  for  $\psi > 0$ ).

Suppose (12) has one positive root  $x_r < 1$ . Due to the assumption of one positive root  $x_r$ , the sign of the sequence

$(P_f)^m(1-P_f)^{N-m}P_0 - (P_d)^m(1-P_d)^{N-m}P_1$  changes once. Let  $A_0 = P_0, A_1 = P_1, p = P_f, q = P_d$  in Lemma 1. Since  $P_d > P_f$ , the signs of  $(P_f)^m(1-P_f)^{N-m}P_0 - (P_d)^m(1-P_d)^{N-m}P_1$  will first be positive and then be negative as  $m$  increases, which can be easily understood from the proof of Lemma 1. For  $x > 0$  sufficiently small, the sign of the left-hand side of (12) will be the same as the sign of the coefficient of  $x^{\frac{2(N\sigma_h^2 + \sigma_n^2)}{1}}$ . The coefficient of  $x^{\frac{2(N\sigma_h^2 + \sigma_n^2)}{1}}$  is negative, because  $(P_f)^N P_0 - (P_d)^N P_1$  is negative. Therefore,  $g(\psi) < 0$  for all sufficiently large  $\psi$ . The theorem follows.

#### APPENDIX F PROOF OF THEOREM 4

Here, we only prove the case when  $H_0$  is true. The result for the case when  $H_1$  is true can be derived by a similar approach. Under  $H_0$ , the joint characteristic function of  $z_1$  and  $z_2$  is given by

$$\begin{aligned}
 \varphi_{\mathbf{z};N|H_0}(\mathbf{t}) &= E[\exp(jt_1 z_1 + jt_2 z_2)|H_0] \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(jt_1 z_1 + jt_2 z_2) \\
 &\quad \times \left[ \sum_{m=0}^N f(z_1, z_2|m) \times \Pr\{M = m|H_0\} \right] dz_1 dz_2 \\
 &= \sum_{m=0}^N [\Pr\{M = m|H_0\} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(jt_1 z_1 + jt_2 z_2) \\
 &\quad \times f(z_1, z_2|m) dz_1 dz_2] \\
 &= \sum_{m=0}^N \left[ \Pr\{M = m|H_0\} \int_{-\infty}^{\infty} \exp(jt_1 z_1) f(z_1|m) dz_1 \right. \\
 &\quad \left. \times \int_{-\infty}^{\infty} \exp(jt_2 z_2) f(z_2|m) dz_2 \right]. \quad (42)
 \end{aligned}$$

The last equality is due to the fact that  $z_1$  and  $z_2$  are conditionally independent given  $m$ . For any particular  $m$ ,  $z_1$  is the sum of  $m + 1$  independent Gaussian random variables with  $m$  of them contributed by sensors which are transmitting and 1 of them contributed by the noise. So is  $z_2$ . Thus, we have

$$\begin{aligned}
 \int_{-\infty}^{\infty} \exp(jt_1 z_1) f(z_1|m) dz_1 \\
 = \varphi_x^m \left( \frac{t_1}{\sqrt{P_f N \sigma_h^2}} \right) \varphi_n \left( \frac{t_1}{\sqrt{P_f N \sigma_h^2}} \right) \quad (43)
 \end{aligned}$$

and

$$\begin{aligned}
 \int_{-\infty}^{\infty} \exp(jt_2 z_2) f(z_2|m) dz_2 \\
 = \varphi_x^m \left( \frac{t_2}{\sqrt{P_f N \sigma_h^2}} \right) \varphi_n \left( \frac{t_2}{\sqrt{P_f N \sigma_h^2}} \right) \quad (44)
 \end{aligned}$$

where  $\varphi_x(t) = \exp\left(-\frac{\sigma_x^2 t^2}{2}\right)$  is the characteristic function of the Gaussian distribution with mean zero and variance  $\sigma_x^2$  and  $\varphi_n(t) = \exp\left(-\frac{\sigma_n^2 t^2}{2}\right)$  is the characteristic function of the Gaussian distribution with mean zero and variance  $\sigma_n^2$ . Substituting (43) and (44) into (42), we get

$$\begin{aligned}
 \varphi_{\mathbf{z};N|H_0}(\mathbf{t}) &= \left\{ 1 + P_f \left[ \varphi_x \left( \frac{t_1}{\sqrt{P_f N \sigma_h^2}} \right) \varphi_x \left( \frac{t_2}{\sqrt{P_f N \sigma_h^2}} \right) - 1 \right] \right\}^N \\
 &\quad \times \varphi_n \left( \frac{t_1}{\sqrt{P_f N \sigma_h^2}} \right) \varphi_n \left( \frac{t_2}{\sqrt{P_f N \sigma_h^2}} \right) \\
 &= \left[ 1 + P_f \left( \exp \left( -\frac{\sigma_h^2 (t_1^2 + t_2^2)}{2P_f N \sigma_h^2} \right) - 1 \right) \right]^N \\
 &\quad \times \exp \left( -\frac{\sigma_n^2 (t_1^2 + t_2^2)}{2P_f N \sigma_h^2} \right).
 \end{aligned}$$

For any  $\mathbf{t} \in \mathbb{R}^2$ , we have

$$\lim_{N \rightarrow \infty} \exp \left( -\frac{\sigma_n^2 (t_1^2 + t_2^2)}{2P_f N \sigma_h^2} \right) = 1, \quad (45)$$

$$\begin{aligned}
 \lim_{N \rightarrow \infty} \left[ 1 + P_f \left( \exp \left( -\frac{\sigma_h^2 (t_1^2 + t_2^2)}{2P_f N \sigma_h^2} \right) - 1 \right) \right]^N \\
 = \exp \left( -\frac{t_1^2 + t_2^2}{2} \right). \quad (46)
 \end{aligned}$$

Due to (45) and (46), we eventually have  $\lim_{N \rightarrow \infty} \varphi_{\mathbf{z};N|H_0}(\mathbf{t}) = \exp\left(-\frac{t_1^2 + t_2^2}{2}\right)$ . Applying the Continuity Theorem [24], we have  $f(\mathbf{z}|H_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I})$ .

#### APPENDIX G DERIVATION OF EQUATION (27)

Let  $x_{Ik}$  and  $x_{Qk}$  be the real and imaginary parts of  $h_k x_k$ . We have

$$\begin{aligned}
 &\frac{1}{N} \Lambda_0^{(N)}(N\lambda_1, N\lambda_2) \\
 &= \frac{1}{N} \ln E[\exp(N\lambda_1 \tilde{y}_I + N\lambda_2 \tilde{y}_Q)|H_0] \\
 &= \frac{1}{N} \ln E \left[ \exp \left( \sum_{k=1}^N (x_{Ik} \lambda_1 + x_{Qk} \lambda_2) \right. \right. \\
 &\quad \left. \left. + n_I \lambda_1 + n_Q \lambda_2 \right) \middle| H_0 \right] \\
 &= \frac{1}{N} \ln \left( \prod_{k=1}^N E[\exp(x_{Ik} \lambda_1 + x_{Qk} \lambda_2)|H_0] \right. \\
 &\quad \left. \times E[\exp(n_I \lambda_1 + n_Q \lambda_2)|H_0] \right) \\
 &= \ln E[\exp(x_{Ik} \lambda_1 + x_{Qk} \lambda_2)|H_0] \\
 &\quad + \frac{1}{N} \ln E[\exp(n_I \lambda_1 + n_Q \lambda_2)|H_0]. \quad (47)
 \end{aligned}$$

The joint distribution of  $n_I$  and  $n_Q$  is a Gaussian distribution with mean zero and covariance matrix  $\sigma_n^2 \mathbf{I}$ . If we take the limit as  $N \rightarrow \infty$ ,  $\frac{1}{N} \ln E[\exp(n_I \lambda_1 + n_Q \lambda_2) | H_0]$  disappears. Since the joint distribution of  $x_{Ik}$  and  $x_{Qk}$  under  $H_0$  is

$$f(x_{Ik}, x_{Qk} | H_0) = \frac{P_f}{2\pi a^2 \sigma_h^2} \exp\left(-\frac{(x_{Ik} - a\mu)^2 + x_{Qk}^2}{2a^2 \sigma_h^2}\right) + \frac{1 - P_f}{2\pi b^2 \sigma_h^2} \exp\left(-\frac{(x_{Ik} + b\mu)^2 + x_{Qk}^2}{2b^2 \sigma_h^2}\right),$$

we can easily show that  $\ln E[\exp(x_{Ik} \lambda_1 + x_{Qk} \lambda_2) | H_0]$  equals to (27).

#### APPENDIX H PROOF OF THEOREM 5

We first consider a noise-free case (i.e.,  $\sigma_n^2 = 0$ ). If we set  $\sigma_n^2 = 0$ , then  $A$  in (20) becomes  $\frac{ab\mu N}{(a-b)}$ . Hence, the optimal decision fusion rule for the noise-free case becomes

$$\left(y_I + \frac{ab\mu N}{a-b}\right)^2 + y_Q^2 - \left(\frac{ab\mu N}{a-b}\right)^2 \underset{H_0}{\geq} T_2. \quad (48)$$

Let  $\tilde{y}_I = \frac{1}{N} y_I$  and  $\tilde{y}_Q = \frac{1}{N} y_Q$ , which are scaled versions of  $y_I$  and  $y_Q$ . Decision rule (48) then becomes

$$\left(\tilde{y}_I + \frac{ab\mu}{a-b}\right)^2 + \tilde{y}_Q^2 - \left(\frac{ab\mu}{a-b}\right)^2 \underset{H_0}{\geq} \frac{T_2}{N^2}. \quad (49)$$

Instead of using the optimal decision fusion rule (49), we consider the following decision fusion rule for the noise-free case:

$$\left(\tilde{y}_I + \frac{ab\mu}{a-b}\right)^2 + \tilde{y}_Q^2 - \left(\frac{ab\mu}{a-b}\right)^2 \underset{H_0}{\geq} \tau^*. \quad (50)$$

Notice that in the noise-free case,  $[\tilde{y}_I, \tilde{y}_Q]^T$  is the empirical mean of i.i.d. random vectors  $[x_{Ik}, x_{Qk}]^T$  ( $k = 1, \dots, N$ ), where  $x_{Ik}$  and  $x_{Qk}$  are the real and imaginary parts of  $h_k x_k$ . The logarithmic moment generating function associated with  $[x_{Ik}, x_{Qk}]^T$  under  $H_0$  is given by

$$\begin{aligned} \tilde{\Lambda}_0(\lambda_1, \lambda_2) &\triangleq \ln E[\exp(x_{Ik} \lambda_1 + x_{Qk} \lambda_2) | H_0] \\ &= \ln \left( P_f \exp\left(a\mu \lambda_1 + \frac{1}{2} a^2 \sigma_h^2 \lambda_1^2\right) \right. \\ &\quad \times \exp\left(\frac{1}{2} a^2 \sigma_h^2 \lambda_2^2\right) + (1 - P_f) \\ &\quad \times \exp\left(-b\mu \lambda_1 + \frac{1}{2} b^2 \sigma_h^2 \lambda_1^2\right) \\ &\quad \left. \times \exp\left(\frac{1}{2} b^2 \sigma_h^2 \lambda_2^2\right) \right). \end{aligned} \quad (51)$$

And  $\tilde{\Lambda}_1(\lambda_1, \lambda_2)$ , the logarithmic moment generating function associated with  $[x_{Ik}, x_{Qk}]^T$  under  $H_1$ , can be obtained by replacing  $P_f$  with  $P_d$  in (51).

Due to the fact that logarithmic moment generating functions  $\tilde{\Lambda}_0(\lambda_1, \lambda_2)$  and  $\tilde{\Lambda}_1(\lambda_1, \lambda_2)$  have identical expressions as  $\Lambda_0(\lambda_1, \lambda_2)$  and  $\Lambda_1(\lambda_1, \lambda_2)$  and that (50) is identical to (31) (because  $\tilde{A} = \frac{ab\mu}{(a-b)}$ ), the error exponent achieved by decision fusion rule (31) in the noise-free case is equal to the error exponent achieved by decision fusion rule (31) in the noisy case.

Based on the above result, we next prove that the error exponent achieved by the optimal decision fusion rule (49) in the noise-free case is equal to the error exponent achieved by the decision fusion rule (31) in the noisy case. We prove this by showing that in the noise-free case the error exponent achieved by decision fusion rule (50) is equal to the error exponent achieved by the optimal decision fusion rule (49). Notice that  $\tau^*$  will make  $\tilde{E}_{F, \tau^*}$  equal to  $\tilde{E}_{M, \tau^*}$ , where  $\tilde{E}_{F, \tau^*}$  is the exponent of false alarm probability associated with decision fusion rule (50) in the noise-free case and  $\tilde{E}_{M, \tau^*}$  is the exponent of miss probability associated with decision fusion rule (50) in the noise-free case. For any given  $N$ , the optimal decision fusion rule (49) is a threshold test with optimal threshold  $\frac{T_2}{N^2}$ , which is a function of  $N$ . When (49) is used, the false alarm probability and the miss probability can be expressed as

$$P_F^* = \Pr \left\{ (\tilde{y}_I + \tilde{A})^2 + \tilde{y}_Q^2 - \tilde{A}^2 > \frac{T_2}{N^2} \middle| H_0 \right\} \\ P_M^* = \Pr \left\{ (\tilde{y}_I + \tilde{A})^2 + \tilde{y}_Q^2 - \tilde{A}^2 \leq \frac{T_2}{N^2} \middle| H_1 \right\}.$$

Use  $P_{F, \tau^*}$  and  $P_{M, \tau^*}$  to denote the false alarm probability and the miss probability associated with the decision rule (50). If  $\tau^* < \frac{T_2}{N^2}$ , we will have  $P_M^* \geq P_{M, \tau^*}$ . If  $\tau^* > \frac{T_2}{N^2}$ , we will have  $P_F^* \geq P_{F, \tau^*}$ . Hence

$$P_1 P_M^* + P_0 P_F^* \geq \min(P_0, P_1) \times \min(P_{M, \tau^*}, P_{F, \tau^*}).$$

Use  $E_o$  to denote the error exponent of the optimal scheme. We have

$$\begin{aligned} E_o &= - \lim_{N \rightarrow \infty} \frac{1}{N} \ln (P_1 P_M^* + P_0 P_F^*) \\ &\leq - \lim_{N \rightarrow \infty} \frac{1}{N} \ln \min(P_0, P_1) \\ &\quad - \lim_{N \rightarrow \infty} \frac{1}{N} \ln \min(P_{M, \tau^*}, P_{F, \tau^*}) \\ &= \tilde{E}_{M, \tau^*} \end{aligned} \quad (52)$$

where the last equality is due to the fact that  $\tilde{E}_{F, \tau^*} = \tilde{E}_{M, \tau^*}$ . Equation (52) suggests that  $E_o$  is less than or equal to the error exponent associated with decision fusion rule (50). However, the optimal decision fusion rule should always have the largest error exponent. As a result, we can conclude that  $E_o$  is equal to the error exponent associated with the decision fusion rule (50).

Finally, notice that the error exponent achieved by the optimal decision fusion rule (49) in the noise-free case should be greater than or equal to the error exponent achieved by the optimal decision fusion rule (20) in the noisy case. So in the noisy case, the error exponent achieved by the decision fusion rule (31) is greater than or equal to the error exponent achieved by the optimal decision fusion rule (20). Obviously, we cannot see the error exponent achieved by the decision fusion rule (31) to be greater than the error exponent achieved by the optimal decision fusion rule (20). This finishes the proof.

#### APPENDIX I PROOF OF THEOREM 6

From definition (28), it is easy to see that

$$\Lambda_i^*(v_1, v_2) \geq 0 \times v_1 + 0 \times v_2 - \Lambda_i(0, 0) = 0. \quad (53)$$

Let  $x_{Ik}$  and  $x_{Qk}$  be the real and imaginary parts of  $h_k x_k$ . Equations (27) and (51) suggest that

$$\begin{aligned} \Lambda_0(\lambda_1, \lambda_2) &= \ln E[\exp(x_{Ik}\lambda_1 + x_{Qk}\lambda_2)|H_0] \\ &\geq E[x_{Ik}\lambda_1 + x_{Qk}\lambda_2|H_0] \end{aligned} \quad (54)$$

$$\begin{aligned} &= E[x_{Ik}\lambda_1|H_0] + E[x_{Qk}\lambda_2|H_0] \\ &= \tilde{\mu}_0\lambda_1 \end{aligned} \quad (55)$$

where (54) is due to Jensen's inequality. Due to (55), we have for any  $\lambda_1$  and  $\lambda_2$

$$\tilde{\mu}_0\lambda_1 + 0 \times \lambda_2 - \Lambda_0(\lambda_1, \lambda_2) \leq 0. \quad (56)$$

Inequality (56) together with the definition (28) suggests that  $\Lambda_0^*(\tilde{\mu}_0, 0) \leq 0$ . However, (53) suggests that  $\Lambda_0^*(\tilde{\mu}_0, 0) \geq 0$ . Hence,  $\Lambda_0^*(\tilde{\mu}_0, 0) = 0$ . Similarly, we can prove that  $\Lambda_1^*(\tilde{\mu}_1, 0) = 0$ .

We prove (32) by showing that for any point  $[v_a, v_b]^T \in \tilde{\mathcal{S}}_\tau \setminus \hat{\mathcal{S}}_\tau$  we can always find  $[v_c, v_d]^T \in \hat{\mathcal{S}}_\tau$  such that  $\Lambda_0^*(v_c, v_d) \leq \Lambda_0^*(v_a, v_b)$ . Use a straight line to connect  $[v_a, v_b]^T$  with  $[\tilde{\mu}_0, 0]^T$ . The resulting line segment will intersect with  $\hat{\mathcal{S}}$  at a certain point  $[v_c, v_d]^T$ . We can write  $[v_c, v_d]^T = \theta[\tilde{\mu}_0, 0]^T + (1-\theta)[v_a, v_b]^T$  for some  $\theta \in [0, 1]$ . Due to (53) and  $\Lambda_0^*(\tilde{\mu}_0, 0) = 0$ , we know  $\Lambda_0^*(v_a, v_b) \geq \Lambda_0^*(\tilde{\mu}_0, 0)$ . Due to the convexity of  $\Lambda_0^*(v_1, v_2)$ , we have

$$\begin{aligned} \Lambda_0^*(v_a, v_b) &= \theta\Lambda_0^*(v_a, v_b) + (1-\theta)\Lambda_0^*(v_a, v_b) \\ &\geq \theta\Lambda_0^*(\tilde{\mu}_0, 0) + (1-\theta)\Lambda_0^*(v_a, v_b) \\ &\geq \Lambda_0^*(v_c, v_d). \end{aligned}$$

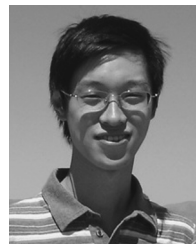
To prove (33), we show that for any point  $[v_a, v_b]^T \in \mathcal{S}_\tau \setminus \{\sqrt{\tau + \tilde{A}^2} - \tilde{A}, 0\}^T$  we have  $\Lambda_1^*(v_a, v_b) \geq \Lambda_1^*(\sqrt{\tau + \tilde{A}^2} - \tilde{A}, 0)$ . For any given  $\lambda_1$ ,  $\Lambda_1(\lambda_1, \lambda_2)$  is symmetric about  $\lambda_2 = 0$ . So it is not difficult to see that  $\Lambda_1^*(v_a, v_b) = \Lambda_1^*(v_a, -v_b)$ . Due to the convexity of  $\Lambda_1^*(v_1, v_2)$ , we have  $\Lambda_1^*(v_a, 0) \leq \Lambda_1^*(v_a, v_b)$ .  $[\sqrt{\tau + \tilde{A}^2} - \tilde{A}, 0]^T$  can be written as  $\theta[\tilde{\mu}_1, 0]^T + (1-\theta)[v_a, 0]^T$  for some  $\theta \in [0, 1]$ . Thus,

$$\begin{aligned} \Lambda_1^*(v_a, v_b) &\geq \Lambda_1^*(v_a, 0) \\ &= \theta\Lambda_1^*(v_a, 0) + (1-\theta)\Lambda_1^*(v_a, 0) \\ &\geq \theta\Lambda_1^*(\tilde{\mu}_1, 0) + (1-\theta)\Lambda_1^*(v_a, 0) \\ &\geq \Lambda_1^*(\sqrt{\tau + \tilde{A}^2} - \tilde{A}, 0). \end{aligned}$$

REFERENCES

[1] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I—Fundamentals," *Proc. IEEE*, vol. 85, pp. 54–63, Jan. 1997.  
 [2] 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.  
 [3] P. K. Varshney, *Distributed Detection and Data Fusion*. New York: Springer Verlag, 1996.  
 [4] M. Longo, T. D. Lookabaugh, and R. M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. Inf. Theory*, vol. 36, pp. 241–255, Mar. 1990.  
 [5] C. Yu and P. K. Varshney, "Bit allocation for discrete signal detection," *IEEE Trans. Commun.*, vol. 46, pp. 173–175, Feb. 1998.

[6] B. Liu and B. Chen, "Channel-optimized quantizers for decentralized detection in sensor networks," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3349–3358, Jul. 2006.  
 [7] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.  
 [8] R. Niu, B. Chen, and P. K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1018–1027, Mar. 2006.  
 [9] Y. Lin, B. Chen, and P. K. Varshney, "Decision fusion rules in multi-hop wireless sensor networks," *IEEE Trans. Aerosp. Elect. Syst.*, vol. 41, pp. 475–488, Apr. 2005.  
 [10] I. Bahceci, G. Al-Regib, and Y. Altunbasak, "Serial distributed detection for wireless sensor networks," in *Proc. ISIT*, Adelaide, Australia, Sep. 2005, pp. 830–834.  
 [11] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2668–2673, Nov. 2005.  
 [12] 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.  
 [13] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access over multiaccess fading channels," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 1081–1092, Mar. 2007.  
 [14] K. Liu and A. M. Sayeed, "Type-based decentralized detection in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1899–1910, May 2007.  
 [15] A. Anandkumar and L. Tong, "Type-based random access for distributed detection over multiaccess fading channels," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 5032–5043, Oct. 2007.  
 [16] W. Li and H. Dai, "Distributed detection in wireless sensor networks using a multiple access channel," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 822–833, Mar. 2007.  
 [17] S. K. Jayaweera, "Bayesian fusion performance and system optimization for distributed stochastic Gaussian signal detection under communication constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1238–1250, Apr. 2007.  
 [18] F. Li and J. S. Evans, "Optimal strategies for distributed detection over multiaccess channels," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Las Vegas, NV, Apr. 2008, pp. 2417–2420.  
 [19] F. Li, J. S. Evans, and S. Dey, "Design of distributed detection schemes for multiaccess channels," *IEEE Trans. Aerosp. Electron. Syst.*, to be published.  
 [20] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 745–754, Apr. 2005.  
 [21] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer-Verlag, 1994.  
 [22] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.  
 [23] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer Verlag, 1998.  
 [24] A. F. Karr, *Probability*. New York: Springer-Verlag, 1993.



**Feng Li** (M'09) was born in Shanghai, China, in 1981. He received the B.E. degree from Shanghai Jiao Tong University, China, in 2004 and the Ph.D. degree from the Department of Electrical and Electronic Engineering, the University of Melbourne, Australia, in 2010.

From 2006 to 2007, he worked at NEC Australia. From 2009 to 2010, he was a Lecturer in the Department of Electrical and Electronic Engineering, the University of Melbourne, where he is currently a Research Fellow. His research interests include statistical signal processing, signal processing for wireless sensor networks, communication systems, and information theory.

Dr. Li received the Best Student Paper Award in Australian Communications Theory Workshop twice in 2008 and 2009. He served as the Local Chair of Australian Communications Theory Workshop in 2011.



**Jamie S. Evans** (M'98) was born in Newcastle, Australia, in 1970. He received the B.S. degree in physics and the B.E. degree in computer engineering from the University of Newcastle, in 1992 and 1993, respectively, where he received the University Medal upon graduation, and the M.S. and the Ph.D. degrees from the University of Melbourne, Australia, in 1996 and 1998, respectively, both in electrical engineering.

From March 1998 to June 1999, he was a Visiting Researcher in the Department of Electrical Engineering and Computer Science, University of California, Berkeley. He returned to Australia to take up a position as Lecturer at the University of Sydney, Australia, where he stayed until July 2001. Since that time, he has been with the Department of Electrical and Electronic Engineering, University of Melbourne, where he is currently a Professor. His research interests are in communications theory, information theory, and statistical signal processing with a focus on wireless communications networks.

Dr. Evans was awarded the Chancellor's Prize for excellence for his Ph.D. thesis from the University of Melbourne.



**Subhrakanti Dey** (M'96–SM'06) was born in Calcutta, India, in 1968. He received the B.Tech. and M.Tech. degrees from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, in 1991 and 1993, respectively, and the Ph.D. degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia, in 1996.

He has been with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia, since February 2000, where he is currently a Full Professor. From September 1995 to September 1997 and September 1998 to February 2000, he was a postdoctoral Research Fellow with the Department of Systems Engineering, Australian National University. From September 1997 to September 1998, he was a Postdoctoral Research Associate with the Institute for Systems Research, University of Maryland, College Park. His current research interests include networked control systems, wireless communications and networks, signal processing for sensor networks, and stochastic and adaptive estimation and control.

Dr. Dey currently serves on the Editorial Board of the *Elsevier Systems and Control Letters*. He was also an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL from 2005 to 2007 and an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2007 to 2010.