Type-Based Random Access for Distributed Detection Over Multiaccess Fading Channels

Animashree Anandkumar, *Student Member, IEEE*, and Lang Tong, *Fellow, IEEE*

***Abstract*—The problem of distributed detection in a sensor network over multiaccess fading channels is considered. A random-access transmission scheme referred to as the type-based random access (TBRA) is proposed and analyzed. Error exponents of TBRA under noncoherent detection are characterized with respect to the mean transmission rate and the channel-coherence index. For the zero-mean multiaccess fading channels, it is shown that there exists an optimal mean-transmission rate that maximizes the detection-error exponents. The optimal mean-transmission rate can be calculated numerically or estimated using the Gaussian approximation, and it gives a sensor-activation strategy that achieves an optimal allocation of transmission energy to spatial and temporal domains. Numerical examples and simulations are used to compare TBRA with the conventional centralized time-division multiple access (TDMA) scheme. It is shown that for the zero-mean multiaccess fading channels, TBRA gives substantial improvement in the low signal-to-noise ratio (SNR) regime whereas for the nonzero mean fading channels, TBRA performs better over a wide range of SNR.**

***Index Terms*—Distributed detection, multisensor systems, performance analysis, signal processing for communications.**

I. Introduction

In the classical setting of distributed detection, sensors in the field sense a certain physical phenomenon and transmit their observations to a fusion center, and the fusion center makes decisions on the underlying phenomenon. The channel between a sensor and the fusion center may be rate-constrained, and local compressions of information are necessary. Typically, sensors transmit their quantized measurements to the fusion center through orthogonal channels, and the transmissions are assumed perfect. The focus of the classical distributed detection is on the design of the quantization rule at sensors and the global decision rule at the fusion center. Such formulations are suitable for the cases when sensors have wired connections to the fusion center, and the number of sensors is not large. See [2]–[4].

The design of large wireless-sensor networks (WSN), however, must deal with challenges beyond the optimization of the local and the global decision rules. Bandwidth has to be allocated to accommodate a large number of sensor nodes; transmissions must be made energy efficient to prolong network lifetime. Wireless transmissions make the medium-access control a crucial component. To this end, well-known deterministic scheduling schemes such as the time-division multiple access (TDMA) may not be appropriate; nodes may be sleeping, faulty, or placed in locations with poor transmission conditions. It is thus desirable to consider MAC schemes in the context of detection that facilitate effective delivery of information from a random number of nodes to the fusion center.

Manuscript received December 24, 2005; revised December 28, 2006. This work was supported in part through the collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011 and by the Army Research Office under Grant ARO-W911NF-06-1-0346. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jaume Riba.

The authors are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: aa332@cornell.edu; ltong@ece.cornell.edu).

Digital Object Identifier 10.1109/TSP.2007.896302

A. Scope of Work

We consider in this paper the problem of distributed detection over a wireless-fading channel via random access. We will not deal with the design of local quantization rules, which is a challenging problem even for the classical distributed detection. Our focus is on the communication (or the random access) aspect of the distributed detection, which to our knowledge has not been treated in the past.

We model the number of sensors involved in each transmission as random with a certain average transmission rate λ. There are several reasons to consider random access. The sensors may use a simple probabilistic wake-up strategy in which a sensor decides to participate in transmission based on a simple coin-flip. The sensor may also decide if a transmission is warranted according to its measurement, transmitting only when the data is "significant" [5]. Yet another possibility is that the fusion center is a mobile-access point, and it travels to different regions of the field to collect data, in which case the number of sensors involved in each collection is random.

A schematic of our problem is illustrated in Fig. 1 with detailed model definition and assumptions given in Section II. The fusion center collects data in multiple slots, each involving a random number of transmitting sensors. We couple the so-called type-based multiple access (TBMA) [6], [7] with a simple random access protocol analogous to the ALOHA. Referred to as the type-based random access (TBRA), sensors transmit probabilistically using a set of orthogonal waveforms keyed to their measurements. Specifically, sensors with the same data value will transmit (if they decide to do so) using the same waveform on a multiaccess fading channel. The bandwidth requirement of TBMA in the absence of fading, is proportional to the number of local quantization levels, not to the number of sensors. The use of orthogonal waveforms eliminates interference among users with different data values and makes it possible to have coherent combining of transmissions in the absence of fading. We will see, however, that simultaneous transmissions in fading is much more complicated, and it may not always be desirable.

1053-587X/$25.00 © 2007 IEEE

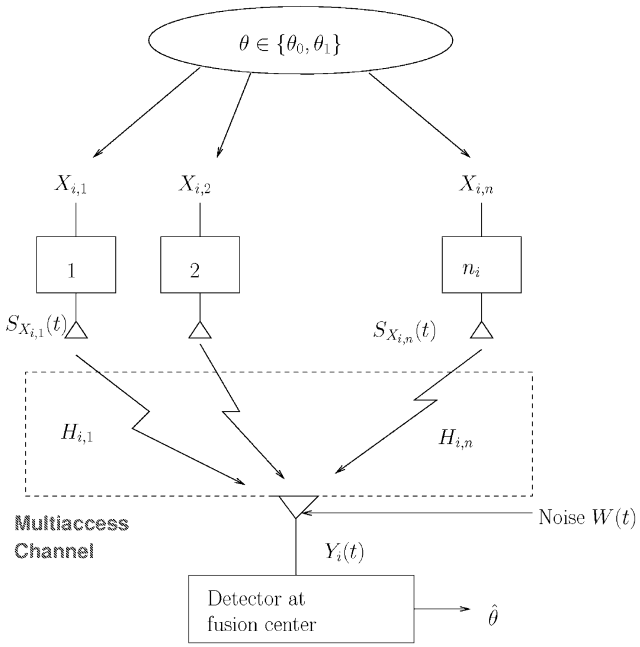

Fig. 1. Distributed detection over multiaccess fading channel.

B. Summary of Main Results

Given the fixed local quantization rule and the available set of orthogonal waveforms for transmission, the design of TBRA reduces to the optimal choice of the mean transmission rate λ. Intuitively, if λ is too small, not enough sensors transmit, and performance suffers. On the other hand, if too many sensors transmit, since they transmit on a multiaccess channel, it is not obvious that the transmissions will not interfere with each other, resulting in poor detection performance.

In searching for the optimal transmission rate λ_*, we use the detection error exponent E_λ, a function of λ, to characterize performance. We first establish that given the expected number of transmissions ρ in l collections, $\rho \triangleq \lambda l$, the detection error probability P_e decays exponentially in the form

$$P_e = e^{-\rho E_\lambda + o(\rho)} \quad (\rho \to \infty) \tag{1}$$

where $(o(\rho)/\rho) \to 0$ as $\rho \to \infty$.

The form of E_λ varies depending on the type of detectors (Bayesian or Neyman-Pearson) and the fading characteristics of the multiaccess channel.

Next, we characterize the behavior of the error exponent E_λ for different cases. It turns out that E_λ crucially depends on the *coherence index* γ defined by

$$\gamma = \frac{|\mathbb{E}(H)|^2}{\mathrm{Cov}(H)} \tag{2}$$

where H is the effective fading coefficient between a sensor and the fusion center.[1] Intuitively, higher γ leads to better SNR gain from simultaneous transmissions at the fusion center. Illustrated in Fig. 2 are sketches of error exponents as functions

[1]The dependencies of inference performance on the coherence index has been observed in the past [6], [8].

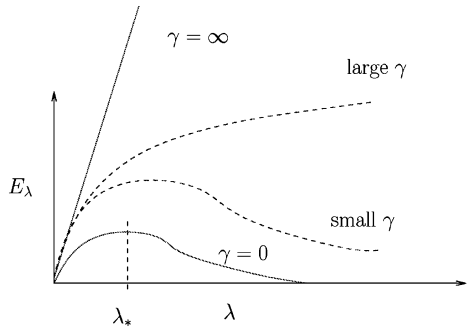

Fig. 2. Error exponent E_λ as a function of transmission rate λ for different channel-coherence indices γ [see (2)].

of λ and γ. The shapes of these curves will be justified by analytical and numerical results in Sections III and IV. We see that for low coherence indices, there exists an optimal λ_* for which the error exponent is maximized. This implies that there is an optimal sensor-activation probability so that the average number of transmitting sensors is optimal. The intuition is that for fading channels with zero-mean $(\gamma = 0)$, sensors transmitting simultaneously using the same waveform tend to cancel each other (in the mean), which is the reason that TBMA schemes involving a single data collection fail [6], [9], [7]. A sharp contrast is the extreme case when the channel is deterministic without fading $(\gamma = \infty)$. We show that there does not exist an optimal λ_*, which means that the optimal strategy is to have simultaneous transmissions, in order to take advantage of the channel coherency. This paper aims to provide insights into the optimal tradeoff, for the case when the expected number of transmissions ρ goes to infinity.

We show the existence of an optimal average transmission rate λ_* when the channel-coherence index γ is small. We also provide the characterization of the error exponent when λ is large. It is in fact the behavior of E_λ as $\lambda \to \infty$ that helps us to describe the shape of error exponent curve in Fig. 2. By letting $\lambda \to \infty$, we employ a version of the central limit theorem (CLT) involving a random number of summands. The limiting distribution allows us to characterize E_∞ analytically. For large transmission rates λ, Gaussian approximation can be used to obtain estimates of the error exponent. Perhaps more importantly in practice, the Gaussian approximation provides $\tilde{\lambda}_*$, an approximation to the optimal rate λ_*.

Our numerical evaluation and simulations are also informative. We numerically evaluate E_λ under different conditions to confirm our theory. We present a performance comparison between TBRA and TDMA, under a fixed energy constraint. The simulation confirms the analysis and our intuition: the two schemes have different operation regimes for the zero-mean $(\gamma = 0)$ multiaccess channels (if complexity is not part of the consideration). At low SNR, TBRA performs considerably better than TDMA because of its optimal allocation of transmissions over time and across sensors, to obtain a significant SNR gain. At high SNR, on the other hand, SNR gain is not needed and the deterministic scheduling of TDMA shows an advantage as it avoids the possibility of interfering transmissions due to random access. On the other hand, under large-γ regime, TBRA performs better than TDMA for a wide range of SNR values, by exploiting channel coherency.

C. Related Work

The problem of classical distributed detection has been investigated in considerable detail. See [2]–[4]. These earlier results, however, assume perfect channels between sensors and the fusion center. In the context of power and bandwidth-constrained WSN, Chamberland and Veeravalli used large-deviation techniques for the optimal design of local quantization rules [10], [11]. See also Aldosari and Moura [12]. We too use large-deviation techniques, but for the design of multiaccess communications. Distributed detection in the presence of channel fading is considered in [13] and [14], where each user has a dedicated channel to the fusion center.

The problem of distributed detection on multiaccess channels are more recent [6], [7], [9], [15]. The transmission scheme used is the so-called TBMA proposed independently by Mergen and Tong [6], [9] and by Liu and Sayeed [7]. The positive result of TBMA is that when there is no fading, the asymptotic performance of TBMA (as the number of sensors approaches infinity) is same as that when the fusion center has direct access to sensor observations. The negative result, however, is that when the channel has zero-mean fading, TBMA fails to be consistent for a single data collection. Furthermore, these results apply only for a fixed number of sensor. In [16], we proposed TBRA as a multiaccess scheme for nonzero mean fading channels, incorporating random number of sensors. We used large-deviation approaches and compared the detection error exponents of TBRA and TBMA for nonzero mean fading channels.

While TBRA is used in this paper among transmitting sensors, it differs from the existing approaches in several significant aspects. This paper allows the fading channels to have zero mean and detectors to be noncoherent. This scenario is relevant since it may be difficult to estimate a large number of fading coefficients at the receiver. Also, it may be difficult to synchronize transmissions among geographically distributed nodes to achieve phase coherency at the receiver. By having the expected number of transmissions ρ go to infinity, the exponential decay of error probabilities is achieved. Under the formulation of this paper, the large-deviation approaches considered in [6], [7], [9], [16] are not applicable.

D. Organization

In Section II, we explain the system model in detail and give the problem statement. In Section III, we explain the receiver structure and present the expressions of the error exponent with respect to the expected number of transmissions. Both Neyman-Pearson and Bayesian detectors are considered. We present a characterization of asymptotic behavior of error exponents when the average number of transmissions goes to infinity. We also discuss the use of Gaussian approximation to investigate behaviors of error exponents. In Section IV, we provide numerical evaluations of error exponents and simulation results on the detection error probability where we compare TBRA with deterministic TDMA scheduling. Conclusions and comments are made in Section V.

II. Model and Problem Formulation

A. TBRA: Transceiver and Sufficient Statistics

We assume the global statistical model as simple binary hypotheses

$$\mathcal{H}_0 : \theta = \theta_0 \quad \text{versus} \quad \mathcal{H}_1 : \theta = \theta_1.$$

As illustrated in Fig. 1, the fusion center collects data in multiple time slots indexed by i. In each collection, there are N_i sensors involved in the transmission, where N_i is a random variable with mean λ and probability-mass function (pmf) $g(n, \lambda) \triangleq \Pr(N_i = n)$. We assume that the sequence N_i is i.i.d.

In the ith data collection, a sensor involved in the transmission,[2] say sensor k, has measurement $X_{i,k} \in \{1, \ldots, M\}$, i.e., quantized to M levels. We assume that the sensor data $\{X_{i,k}\}$ are conditionally i.i.d. across time and sensors, given θ, with pmf $p_\theta(\cdot)$. In vector notation, we have

$$X_{i,k} \overset{\text{i.i.d.}}{\sim} \mathbf{p}_\theta = (p_\theta(1), \cdots, p_\theta(M)), \quad \theta \in \{\theta_0, \theta_1\}.$$

In the ith collection, sensor k encodes its measurement $X_{i,k}$ to a certain waveform and transmits it over a multiaccess fading channel. As in TBMA, a set of M orthonormal waveforms $\{\phi_m(t), m = 1, \ldots, M\}$ are used, each corresponding to a specific data value. Specifically, the baseband signal transmitted by sensor k in collection i is given by

$$S_{i,k}(t) = \sqrt{\mathcal{E}}\phi_{X_{i,k}}(t)$$

where $\mathcal{E}$ is the energy of the transmission.

The channel coefficients ($\tilde{H}_{i,k} \in \mathbb{C}$) are time-varying, i.i.d. across sensors and time. We assume coarse synchronization in the sense that at the fusion center, there is no intercollection interference. Such synchronization can be derived by letting fusion center transmit a synchronization beacon. It can also be accomplished by adding sufficient guard time between consecutive data collections. For low rate applications, this assumption is reasonable.

The received complex-baseband signal after collecting l samples is given by

$$Y_i(t) = \sum_{k=1}^{N_i} \tilde{H}_{i,k} S_{i,k}(t - \tau_{i,k}) + W_i(t), \quad i = 1, \ldots, l \quad (3)$$

where we assume that the channel-state information $\{\tilde{H}_{i,k}\}$ is not known at the receiver. $\tau_{i,k}$ are the random delays for different sensor transmissions and the noise $W_i(t)$ is assumed to be complex white zero-mean Gaussian, with power density σ^2. We define the sensor signal-to-noise ratio (SNR) by $\text{SNR} \triangleq (\mathcal{E}/\sigma^2)$.

Under the narrow-band signal assumption, the flat-fading approximation which neglects the time dispersion in the signal is valid. Therefore, the delay is only through the carrier phase i.e., $S_{i,k}(t - \tau_{i,k}) \approx S_{i,k}(t) \exp(-j2\pi f_c \tau_{i,k})$, where f_c is the carrier frequency. Denoting the effective fading statistic by $H_{i,k} \triangleq \tilde{H}_{i,k} \exp(-j2\pi f_c \tau_{i,k})$ with mean $\mu_H \triangleq \mathbb{E}(H_{i,k})$ and

[2]Without loss of generality, we will only consider those sensors involved in the transmission.

covariance $\sigma_H^2 \triangleq \mathrm{Cov}(H_{i,k})$, the received signal is thus given by

$$Y_i(t) = \sum_{k=1}^{N_i} H_{i,k} S_{i,k}(t) + W_i(t), \quad i = 1, \ldots, l \tag{4}$$

where we assume that $\{H_{i,k}\}$ are proper complex Gaussian, and are unknown at the fusion center.

Sufficient statistics $\{\mathbf{Y}_i\}$ are generated from the bank of filters matched to the orthogonal basis $\{\phi_m(t)\}$. For the ith collection

$$\begin{aligned}\mathbf{Y}_i &\triangleq \frac{1}{\sqrt{\mathcal{E}}}[\langle Y_i(\cdot), \phi_1(\cdot)\rangle, \ldots, \langle Y_i(\cdot), \phi_M(\cdot)\rangle] \\ &= \sum_{k=1}^{N_i} H_{i,k} \mathbf{e}_{X_{i,k}} + \mathbf{W}_i \end{aligned} \tag{5}$$

where $\langle Y_i(\cdot), \phi_m(\cdot)\rangle$ is the output of the matched filter corresponding to $\phi_m(t)$, $\mathbf{e}_m$ the unit vector with nonzero entry at the mth position, and $\mathbf{W}_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}_c(0, (1/\mathrm{SNR})\mathbf{I})$.

To see the intuition behind the coherence index γ defined in (2), we explicitly write the mth entry of $\mathbf{Y}_i = [Y_{i,1}, \ldots, Y_{i,M}]^{\mathrm{T}}$

$$Y_{i,m} = \sum_{k=1}^{N_i} H_{i,k} 1_{\{X_{i,k}=m\}} + W_{i,m} \tag{6}$$

where $1_{\mathcal{A}}$ is the event-indicator function. The extreme case is when the channel is deterministic with $H_{i,k} \equiv 1 (\gamma \to \infty)$. Transmissions from those sensors observing data value m add up coherently, and $Y_{i,m}$ is the number of sensors that observe data level m (plus noise), which gives rise to notion of type-based transmission.[3] On the other hand, when $\gamma = 0, (\mu_H = 0)$, the transmissions add up noncoherently, and the mean of $\mathbf{Y}_i$ contains no information of the model.

B. Spatio-Temporal Tradeoff and Problem Formulation

The design of TBRA reduces to finding the optimal-activation strategy that minimizes detection-error probability. For sensors that are activated probabilistically either by themselves or by the beacon from the fusion center, a TBRA scheme reduces to finding the mean number of transmissions $\lambda = \mathbb{E}(N_i)$. To this end, we need to connect λ with the detection error probability.

If the fusion center collects data using TBRA for l time slots, the expected number of transmissions is $\rho \triangleq \lambda l$, which is also proportional to the total energy consumption. Fixing ρ, there is a spatio-temporal tradeoff between the average number of transmissions per slot and the total number of time slots. Should energy be allocated mostly to simultaneous transmissions by making λ large? Or should we rely on taking more data collections by choosing a large l. The optimal design of TBRA is to achieve optimal tradeoff between λ and l.

We will consider two types of detector: the Bayesian detector and the Neyman-Pearson detector. The explicit characterizations of error probabilities for these two cases are not tractable. We thus examine the case when the expected number of transmissions ρ is large. Let $P_e(\rho, \lambda)$ be the detection-error probability (either the misdetection probability of the Neyman-Pearson detector or the average of the misdetection and the false-alarm probabilities in the Bayesian setup). We will optimize TBRA through the error exponent

$$E_\lambda \triangleq -\lim_{\rho \to \infty} \frac{1}{\rho} \log P_e(\rho, \lambda) \tag{7}$$

which is equivalent to saying that $P_e(\rho, \lambda)$ decays exponentially with respect to ρ with rate E_λ, a function of λ, as in (1). The justification of exponential decay of P_e will be given in Section III using standard arguments involving the Cramér's theorem and the Stein's lemma. Next, we optimize TBRA by seeking

$$\lambda_* = \arg\sup_{\lambda > 0} E_\lambda. \tag{8}$$

Although E_λ can be evaluated numerically for a given statistical model of hypotheses and fading, it is of theoretical and practical significance to establish that λ_* is finite and bounded. To this end, we need to characterize E_λ as $\lambda \to 0$ and $\lambda \to \infty$.

III. Optimal TBRA

The key step towards optimal TBRA is the characterization of detection error exponent defined in (7). The form of error exponent is well known in the theory of large-deviation analysis [17]: the Chernoff information for the Bayesian detector, and the Kullback-Leibler (KL) distance (relative entropy) for the Neyman-Pearson detector. We first present the optimal detector and then give the general characterizations of the error exponents with respect to the expected number of transmissions ρ. Next, we state a result on the existence of optimal λ_* that maximizes the error exponent. We then consider the limiting case when $\lambda \to \infty$. The asymptotic analysis not only gives the key argument for the existence of a finite optimal λ_*, but also provides a qualitative assessment of the error exponents and a computationally tractable way of estimating λ_*.

A. Optimal Noncoherent Detector

The optimal detector given the matched filter output $\{\mathbf{Y}_i\}$ is the likelihood-ratio detector under both the Bayesian and Neyman-Pearson settings. With the i.i.d. assumption, the detector is given by

$$T_l = \frac{1}{l} \sum_{i=1}^{l} \log \frac{f_{1,\lambda}(\mathbf{Y}_i)}{f_{0,\lambda}(\mathbf{Y}_i)} \gtrless \tau \tag{9}$$

where $f_{k,\lambda}(\mathbf{y})$ is the probability density function (pdf)[4] of $\mathbf{Y}_i$ under hypothesis $\mathcal{H}_k$, l is the number of data collections and the threshold τ is chosen according to the prior for the Bayesian detector or the false-alarm rate for the Neyman-Pearson detector.

The receiver only needs to compute the likelihood ratio in (9). In practice, the likelihood ratio may not have a closed-form expression,[5] and numerical evaluation is necessary. However, since the receiver is noncoherent, it does not have a RAKE structure. Therefore, the complexity of the receiver is not limited by the number of simultaneous transmissions.

[3] Given $X_{i,k} = x_{i,k}$, $N_i = n_i$ and the observation $\mathbf{Y}_i = \mathbf{y}_i$, in the absence of noise, the type of $x_{i,k}$ is $(1/n_i)\mathbf{y}_i$. [17], [18].

[4] We will assume that the likelihood function is well defined.

[5] For the special case of Poisson number of sensors with Gaussian fading, an infinite-sum expression is available.

B. Detection-Error Exponents

In the following theorem, we give the expressions for the detection-error exponents. These are direct applications of the Cramér's theorem and the Stein's lemma with only a trivial modification that changes time index in the standard setting to the expected number of transmissions ρ.

Theorem 1 (Error Exponents): Given expected number of transmissions ρ and mean transmission rate λ, let $P_e^{\mathrm{B}}(\rho,\lambda)$ be the average error probability of the Bayesian detector under any prior, and $P_e^{\mathrm{NP}}(\rho,\lambda)$ be the misdetection-error probability of the Neyman-Pearson detector under any fixed size α. The error exponents for the two detectors are given by

$$E^{\mathrm{NP}}(\lambda) \triangleq -\lim_{\rho\to\infty}\frac{1}{\rho}\log P_e^{\mathrm{NP}}(\rho,\lambda) = \frac{1}{\lambda}D_\lambda(f_0\|f_1) \quad (10)$$

$$E^{\mathrm{B}}(\lambda) \triangleq -\lim_{\rho\to\infty}\frac{1}{\rho}\log P_e^{\mathrm{B}}(\rho,\lambda) = \frac{1}{\lambda}C_\lambda(f_0,f_1) \quad (11)$$

where $D_\lambda(f_0\,\|\,f_1)$ is the Kullback-Leibler distance and $C_\lambda(f_0,f_1)$ the Chernoff information.

Proof: See [17] and [18, pp. 92–94]. □

While this theorem provides the basis for investigating error exponents, it says little about the behavior of error exponents as functions of λ, especially about whether there exists an optimal λ_*. The following theorem gives the results for the two extreme cases: $\gamma=0$ and $\gamma=\infty$.

Theorem 2 (Existence of Optimal λ): Let λ be the mean transmission rate and let $f_{i,\lambda}(\mathbf{y})$ be the pdf of the matched-filter output $\mathbf{Y}$, under the hypothesis $\mathcal{H}_i$. Assume the following:

1) for the Neyman-Pearson detection, pdf $f_{0,\lambda}$ and $f_{1,\lambda}$ are differentiable functions of λ almost everywhere;
2) for the Bayesian detection, the above assumption and in addition, the optimizing parameter ν_* is differentiable in λ almost everywhere, given by

$$\nu_*(\lambda) = \arg\min_{\nu\in[0,1]}\log\int_{\mathbf{y}} f_{0,\lambda}^{\nu}(\mathbf{y})f_{1,\lambda}^{1-\nu}(\mathbf{y})\mathbf{dy}. \quad (12)$$

In addition, assume that the pmf of $N, g(n,\lambda)$ is differentiable in λ and satisfies the following properties:

$$\lim_{\lambda\to 0} g(n,\lambda) = 1_{\{n=0\}} \quad (13)$$

$$\lim_{\lambda\to 0}\frac{d}{d\lambda}g(n,\lambda) = -a1_{\{n=0\}} + a1_{\{n=1\}}, \quad a>0 \quad (14)$$

where $1_{\mathcal{A}}$ is the event-indicator function. The following results hold:

1) if the channel has zero-mean fading, i.e., $\gamma=0$, then

$$\lim_{\lambda\to 0} E^{\mathrm{NP}}(\lambda) = \lim_{\lambda\to\infty} E^{\mathrm{NP}}(\lambda) = 0 \quad (15)$$

$$\lim_{\lambda\to 0} E^{\mathrm{B}}(\lambda) = \lim_{\lambda\to\infty} E^{\mathrm{B}}(\lambda) = 0 \quad (16)$$

which imply that there exist $\lambda_*^{\mathrm{NP}}, \lambda_*^{\mathrm{B}}$ such that $0<\lambda_*^{\mathrm{NP}}, \lambda_*^{\mathrm{B}}<\infty$ and

$$\sup_{\lambda>0} E^{\mathrm{NP}}(\lambda) = \frac{1}{\lambda_*^{\mathrm{NP}}}D(f_{0,\lambda_*^{\mathrm{NP}}}\,\|\,f_{1,\lambda_*^{\mathrm{NP}}}) \quad (17)$$

$$\sup_{\lambda>0} E^{\mathrm{B}}(\lambda) = \frac{1}{\lambda_*^{\mathrm{B}}}C(f_{0,\lambda_*^{\mathrm{B}}}\,\|\,f_{1,\lambda_*^{\mathrm{B}}}); \quad (18)$$

2) if the channel is deterministic, i.e., $\mathrm{Cov}(H)=0$ or $\gamma=\infty$, then there does not exist a bounded optimal λ_* that maximizes the error exponents. In particular

$$E^{\mathrm{NP}}(\lambda)=\Theta(\lambda), \quad E^{\mathrm{B}}(\lambda)=\Theta(\lambda), \quad \text{as} \quad \lambda\to\infty \quad (19)$$

where the notation Θ means that λ is an exponentially tight bound.[6]

Proof: The proof for the above theorem relies on the analysis of the extreme cases under the regularity assumptions of the KL distance and the Chernoff information. See Appendix, Section A. □

In the above theorem, the assumptions on the pmf of N (13) and (14) imply that at low λ, there is utmost one transmission. Examples include Poisson distribution and the binomial distribution $B(n,p)$ for a fixed n with $p\to 0$.

The continuity assumption for the Neyman-Pearson detection is easily satisfied for many well-behaved distributions. On the other hand, the assumption for Bayesian detection in (12) is harder to satisfy in practice.

Theorem 2 establishes the general shape of E_λ as shown in Fig. 2, for the extreme values of coherence index, $\gamma=0$ and $\gamma=\infty$. The role of γ in E_λ is embedded in the KL distance or the Chernoff information through the pdfs $(f_{i,\lambda}(\mathbf{y}), \quad i=0,1)$, which are typically continuous functions of γ. Therefore, one can infer the behavior of E_λ for small and large γ.

C. Asymptotic Distribution and Gaussian Approximation

A key step in proving Theorem 2 is the investigation of E_λ as $\lambda\to\infty$. The idea is to use the continuity argument coupled with a version of the central limit theorem (CLT) to calculate the KL distance and the Chernoff information. We elaborate this approach here for two reasons. First, the calculation of E_∞ is needed in proving Theorem 2; this is accomplished by the use of CLT. Second, from a practical stand-point, the Gaussian approximation via CLT gives a computationally tractable way to approximate the error exponent. This is especially useful when we try to find the optimal sensor activation rate λ_*. The accuracy of such an approximation of course, depends on the specific distributions of the sensor measurements and the channel, and we will demonstrate its performance in Section IV.

We shall focus in this section on the single collection model, and evaluate the error exponents using the limiting distribution as $\lambda\to\infty$. For ease of notation, we drop the time index i in (5), and consider the model

$$\mathbf{Y} = \sum_{k=1}^{N} H_k\mathbf{e}_{X_k} + \mathbf{W} \quad (20)$$

where we have a random summand N, with pmf $g(n,\lambda)$ and mean $\mathbb{E}(N)=\lambda$.

Theorem 3 (Limiting Distribution of $\mathbf{Y}$): Assume that the effective channel gains $\{H_k\}$ are independent and identically distributed (i.i.d.) with mean μ_H and covariance σ_H^2, and that the number of sensors N has pmf $g(n,\lambda)$, with mean λ. Also,

[6] $\Theta(a(\lambda)) = \{b(\lambda): 0\le c_1a(\lambda)\le b(\lambda)\le c_2a(\lambda), \forall\lambda>\lambda_o\}$ for some $c_1, c_2, \lambda_o>0$.

assume that N/λ converges to a constant $\eta > 0$ in distribution, i.e.,

$$\frac{N}{\lambda} \xrightarrow{d} \eta, \quad \text{as } \lambda \to \infty. \tag{21}$$

When the pmf of sensor measurements $p_\theta(m) > 0, \forall\, m = \{1, \ldots, M\}$ and $\theta = \{\theta_0, \theta_1\}$, the shifted and scaled matched-filter output $\mathbf{Y}$ satisfies the central limit theorem and has the limiting complex-normal distribution according to

$$\frac{\mathbf{Y} - \eta\lambda\mu_H \mathbf{p}_\theta}{\sqrt{\eta\lambda}} \xrightarrow{d} \mathcal{N}_c(0, \mathrm{Cov}(H_1 \mathbf{e}_{X_1})) \quad \text{as} \quad \lambda \to \infty. \tag{22}$$

In addition, if N is Poisson, then each entry of vector $\mathbf{Y}$ is independently asymptotically Gaussian, given by

$$\frac{Y(m) - \lambda\mu_H p_\theta(m)}{\sigma_H \sqrt{\lambda p_\theta(m)}} \xrightarrow{d} \mathcal{N}_c(0,1), \quad \text{as} \quad \lambda \to \infty \tag{23}$$

$\forall\, m = 1, \ldots, M$.

Proof: The proof of the above is an application of CLT with a random summand. We use the fact that $\eta = 1$ for Poisson. See Appendix, Section B for details. □

Evaluating the covariance matrix in (22), we have

$$\mathrm{Cov}(H_1 \mathbf{e}_{X_1}) = \sigma_H^2 \mathrm{Diag}(\mathbf{p}_\theta) + |\mu_H|^2 \left(\mathrm{Diag}(\mathbf{p}_\theta) - \mathbf{p}_\theta \mathbf{p}_\theta^T\right). \tag{24}$$

However, the result for the Poisson distribution in (23) is stronger than for a general pmf $g(n, \lambda)$ in (22), since the asymptotic distribution is independent across the quantization levels. This is due to the property of marking, which implies that the number of sensors transmitting a particular data-level is independently Poisson. Moreover, by (24), under zero-mean fading ($\mu_H = 0$), the asymptotic distribution is independent across the data-levels for any general pmf $g(n, \lambda)$. In this section, we assume that N is Poisson, for the ease of evaluation of the exponents under the limiting distribution.

Since $\mathbf{Y}$ is asymptotically Gaussian, in the large-λ regime, the hypothesis-testing problem can be cast as the testing of binary hypotheses with

$$\tilde{\mathcal{H}}_i : \mathbf{Y} \sim \mathcal{N}_c(\mu_i, \Sigma_i), i = 0, 1 \tag{25}$$

$$\mu_i = \lambda\mu_H \mathbf{p}_{\theta_i}, \quad \Sigma_i = \lambda\sigma_H^2 \mathrm{diag}(\mathbf{p}_{\theta_i}) + \frac{\sigma^2}{\mathcal{E}}\mathbf{I}. \tag{26}$$

The effect of the channel coherency is more evident in the asymptotic distribution. For any positive γ ($\mu_H > 0$), the mean of $\mathbf{Y}$ in (26) contains information about the underlying hypothesis $\mathcal{H}_i$. As λ increases, the two hypotheses are more separated and the error probability decays exponentially. When $\mu_H = 0$ ($\gamma = 0$), on the other hand, the information is in the covariance, and the error probability for a single-data collection converges to a constant value as $\lambda \to \infty$ [6]. However, as mentioned in Section III-B, in this paper, we let the expected number of transmissions ρ go to infinity. In this case, there is exponential decay of error probability, and we, therefore, need to characterize the exponents.

We now characterize the error exponents, defined by (10) and (11), for the Gaussian distribution. It turns out that the Gaussian Chernoff-information $\tilde{C}$ and the Kullback-Leibler distance $\tilde{D}$ have closed-form expressions, enabling us to evaluate the asymptotic limits of the true exponents.

Lemma 1 (Gaussian Error Exponents): Let σ_H^2 be the channel variance, γ is the channel-coherence index, SNR be the sensor SNR and $\Delta \triangleq p_{\theta_0}(k) - p_{\theta_1}(k)$. Denote

$$\alpha_k \triangleq \frac{\lambda\sigma_H^2\, \mathrm{SNR}\, p_{\theta_0}(k) + 1}{\lambda\sigma_H^2\, \mathrm{SNR}\, p_{\theta_1}(k) + 1}, \quad \omega_k \triangleq \lambda\gamma\Delta \tag{27}$$

and let β_k be the positive root of the quadratic equation,[7]

$$\omega_k \alpha_k \beta_k^2 + (\alpha_k - 1)\beta_k - \log\alpha_k - \omega_k = 0. \tag{28}$$

The error exponents $\tilde{E}^{\mathrm{NPD}}(\lambda)$ and $\tilde{E}^{\mathrm{BD}}(\lambda)$ under the Neyman-Pearson and the Bayesian settings, for the Gaussian distribution in (25), are given by

$$\tilde{E}^{\mathrm{NPD}}(\lambda) = \frac{\tilde{D}_{\lambda,\gamma}}{\lambda} = \frac{1}{\lambda}\sum_{k=1}^{M}(-\log\alpha_k + (1+\omega_k)(\alpha_k - 1)), \tag{29}$$

$$\tilde{E}^{\mathrm{BD}}(\lambda) = \frac{\tilde{C}_{\lambda,\gamma}}{\lambda} = \frac{1}{\lambda}\sum_{k=1}^{M}\left(-\log\beta_k + \beta_k + \frac{(\beta_k - 1)^2 \alpha_k \omega_k}{(\alpha_k - 1)} - 1\right). \tag{30}$$

Proof: The proof is derived using Kullback-Leibler distances for Gaussian distributions. See Appendix, Section C. □

Given the closed-form expressions for the Gaussian error exponents $\tilde{E}^{\mathrm{NPD}}(\lambda)$ and $\tilde{E}^{\mathrm{BD}}(\lambda)$, we can evaluate the various limits.

Theorem 4 (Limiting Properties of Error Exponents): The Chernoff information $\tilde{C}_{\lambda,\gamma}$ and the KL distance $\tilde{D}_{\lambda,\gamma}$ are monotonically increasing functions of the coherence index γ, transmission rate λ and sensor SNR. For finite γ, the error exponents $\tilde{E}^{\mathrm{NPD}}(\lambda)$ and $\tilde{E}^{\mathrm{BD}}(\lambda)$ converge to a finite limit, proportional to the coherence index γ, as $\lambda \to \infty$ given by

$$\lim_{\lambda\to\infty} \tilde{E}^{\mathrm{NPD}}(\lambda) = \gamma \sum_{k=1}^{M} \frac{\Delta}{p_{\theta_1}(k)} \tag{31}$$

$$\lim_{\lambda\to\infty} \tilde{E}^{\mathrm{BD}}(\lambda) = \gamma \sum_{k=1}^{M} (\sqrt{p_{\theta_0}(k)} - \sqrt{p_{\theta_1}(k)})^2. \tag{32}$$

Proof: We establish the monotonicity and evaluate the limits using expressions in (29) and (30). Also note, (32) is a scaled version of Hellinger distance [19]. See Appendix, Section D. □

We also investigate the case when the channel is perfectly coherent: $\mu_H = 1$ and $\sigma_H \to 0$, or $\gamma \to \infty$.

[7]Exactly one positive root exists for the case when $p_{\theta_0}(k) \neq p_{\theta_1}(k)$.

Theorem 5 (Error Exponent for Perfectly Coherent Channels): In the absence of fading, the error exponents of the NP and Bayesian detectors for the limiting distribution are given by

$$\lim_{\gamma\to\infty} \tilde{E}^{\mathrm{NPD}}(\lambda) = \lambda \mathrm{SNR} \sum_{k=1}^{M} \Delta^2. \tag{33}$$

$$\lim_{\gamma\to\infty} \tilde{E}^{\mathrm{BD}}(\lambda) = \frac{\lambda \mathrm{SNR}}{4} \sum_{k=1}^{M} \Delta^2. \tag{34}$$

Proof: Substituting $\sigma_H = 0$, we derive the expressions by finding the KL distance and Chernoff information between the distributions, $\mathcal{N}_c(\lambda \mathbf{p}_{\theta_0}, \sigma^2)$ and $\mathcal{N}_c(\lambda \mathbf{p}_{\theta_1}, \sigma^2)$. □

To contrast the perfectly coherent case, we examine the case when the channel is noncoherent, i.e., $\mu_H = 0$ ($\gamma = 0$). Interestingly, the dependencies of the Chernoff information and the KL distance on the transmission rate λ, the sensor SNR, and the channel variance σ_H^2 can be summarized using a single parameter—the average SNR at the receiver

$$\chi \triangleq \lambda \sigma_H^2 \mathrm{SNR}. \tag{35}$$

Theorem 6 (Error Exponents for Noncoherent Channels): For the noncoherent channels ($\mu_H = 0$), the Chernoff information $\tilde{C}_\chi$ and the KL distance $\tilde{D}_\chi$ are functions of $\chi \triangleq \lambda \sigma_H^2 \mathrm{SNR}$, and have the following properties.

1) $\tilde{C}_\chi$ and $\tilde{D}_\chi$ are monotonically increasing concave functions of χ.
2) As $\chi \to \infty, \tilde{C}_\chi$ and $\tilde{D}_\chi$ converge to finite limits when $p_{\theta_i}(k) > 0 \; \forall i, k$.
3) Normalized functions $(\tilde{C}_\chi/\chi)$ and $(\tilde{D}_\chi/\chi)$ have unique maxima, which are only functions of $\mathbf{p}_{\theta_0}$ and $\mathbf{p}_{\theta_1}$.

Proof: See Appendix, Section E.

The compact parameter χ provides additional insights and design options. The optimal error exponents can be achieved with a combination of choices of sensor activation and sensor SNR. At small SNR, for example, more sensors are needed to obtain SNR gain. On the other hand, at high SNR, fewer sensors transmit to avoid the noncoherent cancelation of the signals (on an average). The optimal λ_* is chosen so as to balance these opposing effects of the multiaccess fading channel.

The error exponents $\tilde{E}^{\mathrm{NPD}}(\lambda)$ and $\tilde{E}^{\mathrm{BD}}(\lambda)$ in (29) and (30) of the limiting Gaussian distributions are good approximations for the true exponents at large λ, due to the continuity property. We shall demonstrate this with a numerical example below, which is a qualitative representation for the general case.

IV. Numerical Results and Simulations

In this section, we resort to numerical and simulation techniques to validate the theories developed in this paper. We consider binary quantized measurements with pmf

$$\mathbf{p}_{\theta_0} = [0.7, 0.3], \quad \mathbf{p}_{\theta_1} = [0.3, 0.7].$$

For the Bayesian setting, we assume equally likely priors.

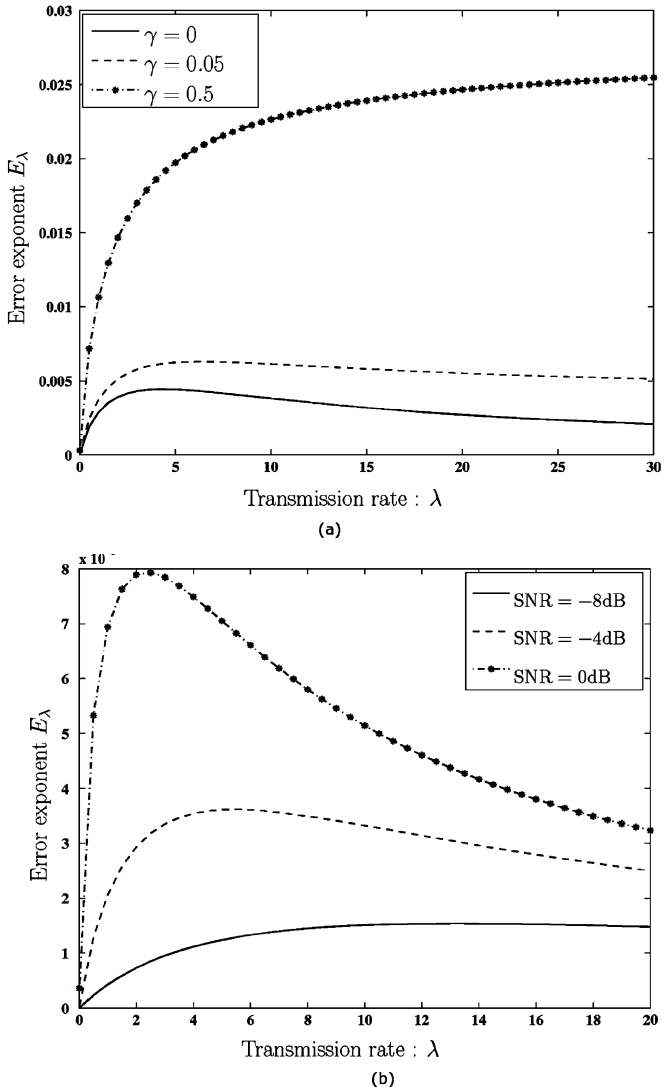

Fig. 3. Error exponent versus transmission rate. (a) $\mathrm{SNR} = 7$ dB, $\sigma_H^2 = 0.1$. (b) $\gamma = 0, \sigma_H^2 = 1$.

We assume that the channel fading is Gaussian[8] $H_{i,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_H, \sigma_H^2)$ with the mean and the variance varying according to different simulation conditions. The number of sensors involved in each transmission N_i is i.i.d. Poisson.

The error exponents are evaluated numerically (without using the Gaussian approximation), and the detection-error probabilities are estimated using Monte Carlo simulations.

A. Evaluation of Error Exponents

Since the central limit theorem is applicable only in the large-λ regime, in order to draw conclusions for finite λ we numerically evaluate the Chernoff information and the Kullback-Leibler distance. We found that the Chernoff information and Kullback-Leibler distance have similar shapes. Therefore, only the behavior of Chernoff information is presented here.

Fig. 3(a) and (b) show the behavior of the actual error exponent E_λ versus λ, one Fig. 3(a) with varying channel-coherence

[8]For proper Gaussian variables, the real and imaginary parts are independent. Therefore, it suffices to limit to real variables

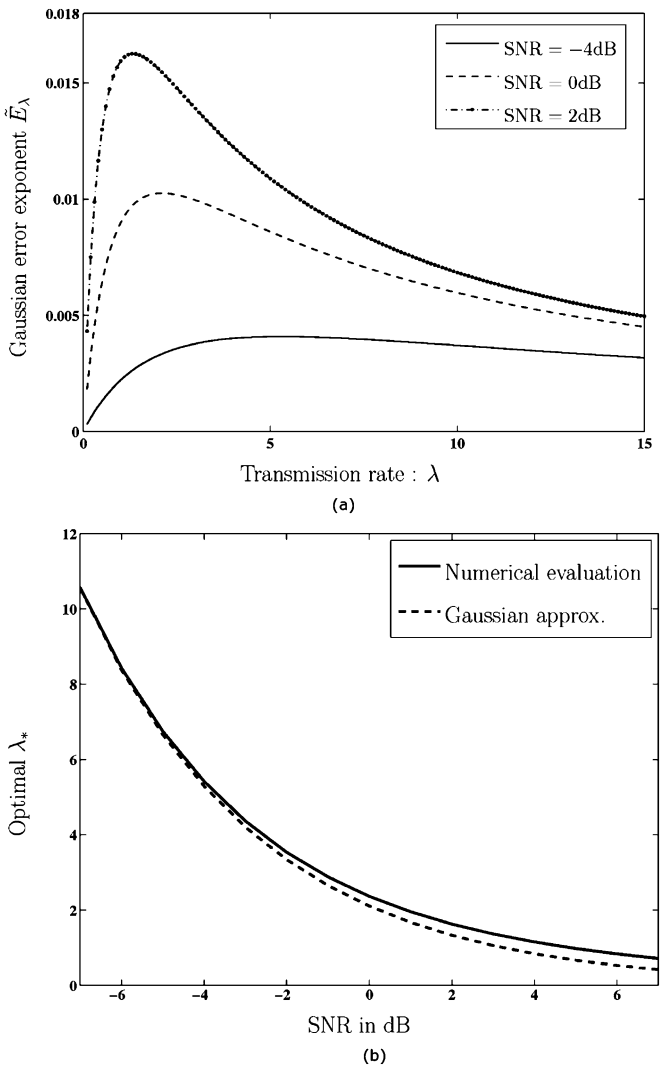

Fig. 4. Gaussian approximation. (a) Gaussian error exponent versus transmission rate. ($\gamma = 0, \sigma_H^2 = 1.$). (b) Optimal transmission rate versus SNR in dB. ($\gamma = 0, \sigma_H^2 = 1.$).

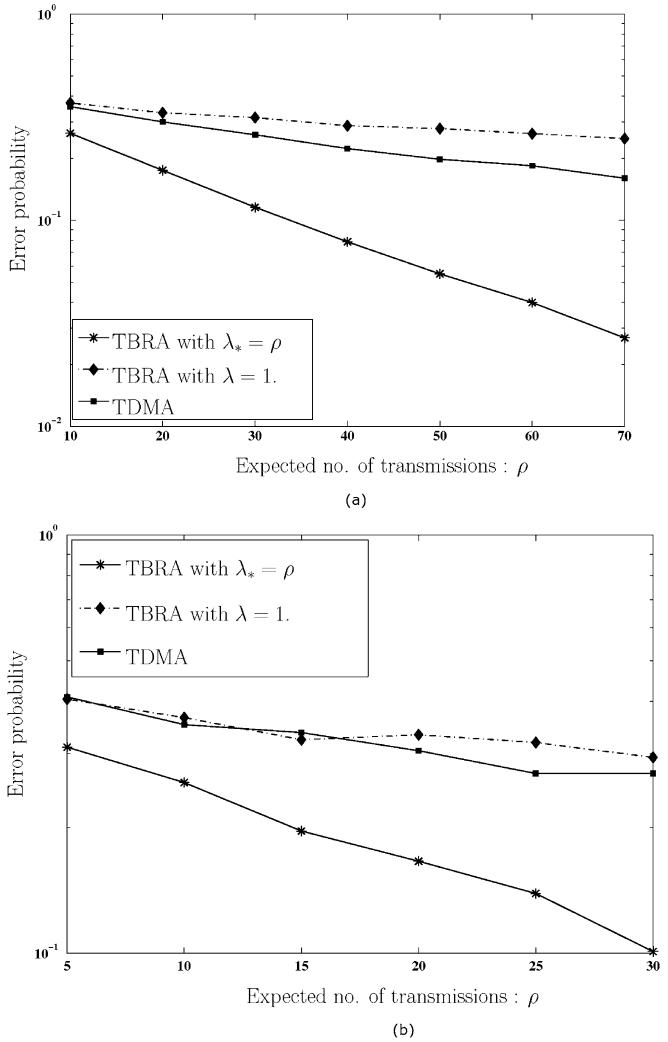

Fig. 5. Error probability versus expected number of transmissions. Fading with nonzero mean. (a) $\gamma = 0.5, \mathrm{SNR} = -4$ dB, $\sigma_H^2 = 1$. (b) $\gamma = 0.5, \mathrm{SNR} = 7$ dB, $\sigma_H^2 = 0.1$.

index γ for a fixed SNR, and the other with varying SNR for $\gamma = 0$. The existence of optimal λ_* is evident for small γ. To see the similarity and difference between the actual E_λ and the Gaussian approximated $\tilde{E}_\lambda$, we plot $\tilde{E}_\lambda$ in Fig. 4(a). The curves in Figs. 3(b) and 4(a) have similar shapes and share the same trend with respect to both λ and SNR. The actual values of the error exponents are indeed different, with the Gaussian approximation giving a more conservative estimate of the true error exponent.

Fig. 4(b) shows the actual optimal transmission rate λ_* and the suboptimal rate $\tilde{\lambda}_*$ [obtained via optimizing the Gaussian error exponents in (29) and (30)]. The optimal $\tilde{\lambda}_*$ obtained from the Gaussian approximation appears to provide a reasonable approximation for the true λ_*. We observe that as SNR decreases, the suboptimal $\tilde{\lambda}_*$ approaches the true λ_*. This behavior is the result of CLT, since the value of λ_* increases as the SNR decreases. Moreover, even at high SNR, the absolute error of approximation is fortunately quite small, since the value of λ_* decreases as the SNR increases.

B. Performance and Discussion

We compare the TBRA scheme with the conventional TDMA scheme, in which one sensor is scheduled to transmit in a data collection, with energy $\mathcal{E}$. We fix the expected number of transmissions $\rho = l\lambda$ in the comparison. Our comparison is fair since both TBRA and TDMA have the same total energy budget.

We run simulations with values specified in the beginning of the Section IV. We consider two regimes of γ and SNR, with four possible scenarios. For the TBRA scheme, from Section IV-A, the optimal performance is at λ_*. We also consider the performance of TBRA under $\tilde{\lambda}_*$, obtained by Gaussian approximation. We also include the TBRA scheme with $\lambda = 1$, which enables us to study the random-access aspect of TBRA under different conditions.

Fig. 5(a) and (b) are simulations for nonzero mean fading channels. For the case shown in Fig. 5(a), the optimal strategy is single-shot transmission (all sensors transmitting simultaneously), since $\lambda_* > \rho$, for the values of ρ used in the simulation. We see that the optimal TBRA scheme performs better than TDMA. However, for the TBRA scheme with $\lambda = 1$ the

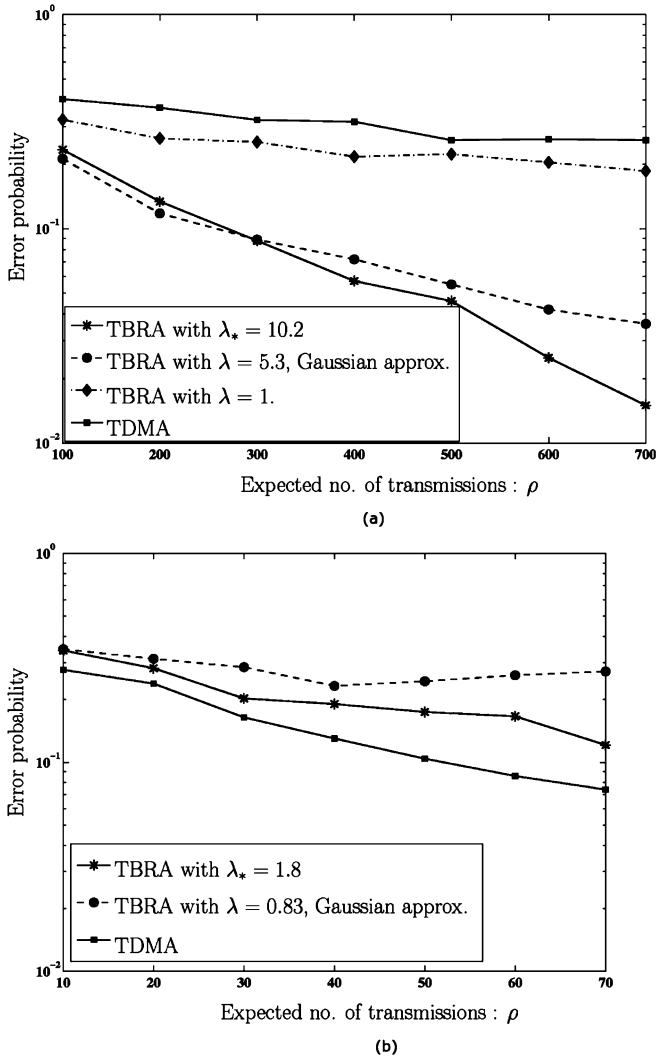

Fig. 6. Error probability versus expected number of transmissions. Fading with zero mean. (a) $\gamma = 0, \sigma_H^2 = 1, \mathrm{SNR} = -4$ dB. (b) $\gamma = 0, \sigma_H^2 = 1, \mathrm{SNR} = 4$ dB.

performance is similar to TDMA. This suggests that the gain for TBRA comes mainly from coherence. In Fig. 5(b), at high SNR, the optimal strategy is still single-shot transmission, due to the high value of γ. This suggests that at large γ, the optimal strategy is to have as many simultaneous transmissions as allowed by the network, in order to exploit the channel coherency.

Fig. 6(a) and (b) shows simulations under zero-mean fading. Fig. 6(a) shows that TBRA performs better than TDMA at low SNR. In Fig. 6(b), we see that TDMA does slightly better at high SNR, under zero-mean fading.

We observe that there is not much performance gap between the optimal rate λ_* and suboptimal $\tilde{\lambda}_*$, from Gaussian approximation. Also, we have ignored the communication overheads involved in each data collection. This is significantly higher for TDMA, since it has more data collections than TBRA, for a fixed ρ.

Some intuitions on the comparison of TBRA and TDMA are in order. At large γ, the optimal rate λ_* is also large. The gain from coherence suggests that at large γ, TBRA will do better than TDMA for a wide range of SNR values. Moreover, the total number of dimensions used by TBRA is just the number of data levels M (for single-shot transmission), which is far less than the number of dimensions used by TDMA, given by $M\rho$. Thus, TBRA is also bandwidth efficient at large γ.

Under the zero-mean fading, λ_* is inversely proportional to the sensor SNR. Therefore, at high SNR, λ_* is small, in order to avoid canceling effects of zero-mean fading Fig. 4(b). However, there are still some collisions between the transmissions, due to random access. In contrast, TDMA schedules exactly one sensor to transmit. We, therefore, expect TDMA to do better at high SNR, under zero-mean fading.

However, at low SNR, simultaneous transmissions counter noise more effectively. Therefore, we expect TBRA to perform better than TDMA. Moreover, at low SNR, TBRA has significantly lower number of data collections than TDMA leading to a quicker detection at fusion center. Also, the total number of dimensions used by TBRA $(M\rho/\lambda_*)$ is far less than the number of dimensions used by TDMA $(M\rho)$. Thus, at low SNR, TBRA is also bandwidth efficient under zero-mean fading.

V. Conclusion

In this paper, we focus on the communication aspect—random access in particular—of distributed detection for large sensor networks. We employ TBRA which inherits most of the attractive features of TBMA, e.g., the efficient bandwidth scaling and the asymptotic optimality under ideal channel conditions. The main advance of this paper is the removal of the requirement of channel coherency and the ability to handle random number of sensors, transmitting simultaneously in a slot. By examining a number of extreme cases, we are able to obtain a general characterization of the error exponents as illustrated in Fig. 2. From a practical stand-point, the approaches using the Gaussian approximation, presented in Section III-C, seem to give the correct insight into an optimal design. Such a characterization is a valuable guide, as a network designer pursues practical solutions.

In this paper, we assume homogeneity in the signal field and hence the spatio-temporal samples are i.i.d. This may not be true for many applications, e.g., those requiring sensor ID or location information; or when the sensors are placed in different noise levels. Although restrictive, the i.i.d. assumption gives us some important insights (e.g., spatio-temporal tradeoffs). The large-deviation analysis of error exponents applies to a large class of models [17], although closed-form expressions may be intractable. In [20] and [21], we obtain the expression for the error exponent, when the sensor measurements are correlated according to a Gauss-Markov random field model.

We have left several important problems open. We have considered a spatio-temporal allocation scheme under the total energy constraint. For applications under other constraints (e.g., time), our formulation does not hold. We have not dealt with the design of local quantization rule. Given that the large-deviation analysis is used in this paper, as well as in several related work [7], [11], [12] dealing with different aspects of the problem a "cross-layer" optimization of local quantization, communications and global inference should be of interest.

APPENDIX

A. Proof of Theorem 2

Let $\mathfrak{o}(\lambda)$ represent a function such that $(\mathfrak{o}(\lambda)/\lambda) \to 0$ as $\lambda \to 0$.

For the pmf of $N, g(n,\lambda)$, applying Taylor's expansion for λ near zero, we have $\mathbb{P}(N_i = 1) = a\lambda + \mathfrak{o}(\lambda)$ and $\mathbb{P}(N_i > 1) = \mathfrak{o}(\lambda)$.

Define the conditional pdf of matched filter output $\mathbf{Y}_i$ under hypothesis $\mathcal{H}_k$ as $f_k(\cdot|N=0) \triangleq w(\cdot), f_k(\cdot|N=1), \triangleq h_k(\cdot)$ and $f_k(\cdot|N>1) \triangleq c_k(\cdot)$, where $w(\cdot)$ is the pdf of White Gaussian noise. Marginalizing over N, we have the pdf of $\mathbf{Y}_i$ under hypothesis $\mathcal{H}_k$,

$$f_k(\cdot) = (1 - \lambda a - \mathfrak{o}(\lambda))w(\cdot) + (\lambda a + \mathfrak{o}(\lambda))h_k(\cdot) + \mathfrak{o}(\lambda)c_k(\cdot). \quad (36)$$

Now we have the KL distance [18]

$$E^{\mathrm{NPD}}(\lambda) = \frac{D(f_0 \,||\, f_1)}{\lambda} = \frac{1}{\lambda}\int_{\mathbf{y}} f_0(\mathbf{y}) \log \frac{f_0(\mathbf{y})}{f_1(\mathbf{y})} \mathbf{dy}. \quad (37)$$

Using (36) we have

$$\lim_{\lambda\to 0} E^{\mathrm{NPD}}(\lambda) = a\int_{\mathbf{y}} (h_0(\mathbf{y}) - h_1(\mathbf{y}))\mathbf{dy} = 0.$$

For Chernoff information, we have [18]

$$C = D(f_{\nu_*} \,||\, f_0) = D(f_{\nu_*} \,||\, f_1)$$
$$f_\nu(\cdot) = \frac{f_0^\nu(\cdot) f_1^{1-\nu}(\cdot)}{\int_{\mathbf{y}} f_0^\nu(\mathbf{y}) f_1^{1-\nu}(\mathbf{y})\mathbf{dy}}. \quad (38)$$

Thus, we have, for $k = 0, 1$

$$\lim_{\lambda\to 0} \frac{D(f_\nu||f_k)}{\lambda} = \int_{\mathbf{y}} \nu a(h_0(\mathbf{y}) + h_1(\mathbf{y}) - 2w(\mathbf{y}))\mathbf{dy} = 0.$$

Therefore, from (38), we have $\lim_{\lambda\to 0} E^{\mathrm{BD}}(\lambda) = 0$. For the case when $\lambda \to \infty$, we first show that the limit coincides with exponents of limiting distribution. From assumption (1) stated in Theorem 2, the integrand in (37) is differentiable with respect to λ implying that the integral is differentiable [22]. This implies that the limit in (37) exists as $\lambda \to \infty$ and coincide with the Gaussian exponent. Similar argument holds for (38). For expressions of Gaussian exponents for $\gamma = 0$ and $\gamma = \infty$, refer Theorems 6 and 5. Although these theorems have results for Poisson N, the exponents for any general pmf $g(n,\lambda)$ coincide with Poisson, for $\gamma = 0$. For $\gamma = \infty$, the claim can be easily shown for a general pmf $g(n,\lambda)$. □

B. Proof of Theorem 3

Recall the CLT with random number of summands [23, p. 369]. Let $X_1, X_2, \ldots,$ be i.i.d. random variables with mean 0 and variance σ^2, and $S_n = \sum_{i=1}^n X_i$. For each positive t, let ν_t be a random variable assuming positive integers as values; not necessarily independent of X_n. Suppose, there exist positive constants a_t and η such that $a_t \to \infty, (\nu_t/a_t) \xrightarrow{d} \eta$ as $t \to \infty$. Then

$$\frac{S_{\nu_t}}{\sigma\sqrt{\nu_t}} \xrightarrow{d} \mathcal{N}(0,1) \quad \frac{S_{\nu_t}}{\sigma\sqrt{\eta a_t}} \xrightarrow{d} \mathcal{N}(0,1). \quad (39)$$

In our case, parameter a_t corresponds to λ, ν_t to N. We have $(N/\lambda) \xrightarrow{d} \eta > 0$ and $(1/\sqrt{\eta\lambda})\mathbf{W} \xrightarrow{p} \mathbf{0}$ as $\lambda \to \infty$. By Slutsky's theorem [23], $(\mathbf{W}/\sqrt{\eta\lambda}) \xrightarrow{d} \mathbf{0}$ as $\lambda \to \infty$. Extending the above to complex domain and to random vectors using *multivariate central limit theorem* [23, p. 385], we obtain

$$\frac{\sum_{k=1}^N (H_k \mathbf{e}_{X_k} - \mu_H \mathbf{p}_\theta)}{\sqrt{\eta\lambda}} \xrightarrow{d} \mathcal{N}_c\left(0, \mathrm{Cov}\left(H_1 e_{X_1}\right)\right).$$

When N is Poisson, let $N^{(m)}$ be the number of sensors transmitting data level m. Since $N^{(m)}$ is a thinning Poisson process [24, p. 317], $N^{(m)}$ is independent for different data levels and

$$N^{(m)} \sim \mathrm{Poiss}(\lambda p_\theta(m)).$$

Therefore, the vector $\mathbf{Y}$ has independent entries. Applying the aforementioned central limit theorem for random summands, to each entry of the vector we obtained the needed result.

C. Proof of Lemma 1

Let $f_i \sim \mathcal{N}_c(\mu_i, \Sigma_i)$. The KL distance $D(f_0||f_1)$ is given by

$$D(f_0||f_1) = \sum_k \left(\log \frac{\Sigma_1(k,k)}{\Sigma_0(k,k)} + \left(\frac{\Sigma_0(k,k) + |\mu_0(k) - \mu_1(k)|^2}{\Sigma_1(k,k)} - 1\right)\right).$$

Define density function f_γ by

$$f_\gamma(\mathbf{y}) = \frac{f_0(\mathbf{y})^\gamma f_1(\mathbf{y})^{1-\gamma}}{\int f_0(\mathbf{y})^\gamma f_1(\mathbf{y})^{1-\gamma} d\mathbf{y}}.$$

The Chernoff information is given by [18],

$$C(f_0, f_1) = D(f_{\gamma^*}||f_0) = D(f_{\gamma^*}||f_1).$$

Solving the above equation yields the expression for Chernoff information. For our setup, we find expressions for μ_i and Σ_i. □

D. Proof of Lemma 4

a) To prove the monotonicity of $\tilde{D}_{\lambda,\gamma}$ and $\tilde{C}_{\lambda,\gamma}$ with respect to λ and γ, we see that

$$\frac{\partial \tilde{D}}{\partial \lambda} = \sum_{k=1}^M \left(\frac{\partial \tilde{D}}{\partial \alpha_k}\frac{\partial \alpha_k}{\partial \lambda} + \frac{\partial \tilde{D}}{\partial \omega_k}\frac{\partial \omega_k}{\partial \lambda}\right).$$

From the expressions of α_k and ω_k we find that

$$\frac{\partial \tilde{D}}{\partial \alpha_k} = \left(-\frac{1}{\alpha_k} + 1 + \omega_k\right), \frac{\partial \tilde{D}}{\partial \omega_k} = \alpha_k - 1.$$

We see that $\Delta < 0$ implies $\alpha_k < 1, \omega_k < 0, (\partial\alpha_k/\partial\lambda) < 0$ and $(\partial\omega_k/\partial\lambda) < 0$. Similar results can be obtained for $\alpha_k > 1$. Combining the above, $(\partial\tilde{D}/\partial\lambda) > 0$. Similarly we obtain

$$\frac{\partial\tilde{D}}{\partial\gamma} = \sum_k (\alpha_k - 1)\lambda\Delta > 0$$
$$\frac{\partial\tilde{D}}{\partial_{\mathrm{SNR}}} = \sum_k \left(\frac{\partial\tilde{D}}{\partial\alpha_k}\frac{\partial\alpha_k}{\partial_{\mathrm{SNR}}}\right) > 0.$$

Similar results can be obtained for $\tilde{C}_{\lambda,\gamma}$.

b) As $\lambda \to \infty, \alpha_k \to (p_{\theta_0}(k)/p_{\theta_1}(k))$ and $\lim_{\lambda\to\infty}(\omega_k(\alpha_k - 1)/\lambda) = (\gamma\Delta^2/p_{\theta_1}(k))$. This gives the result for $\tilde{E}^{\mathrm{NPD}}(\lambda)$. As $\lambda \to \infty, \beta_k \to (1/\sqrt{\alpha_k})$ and, hence, the result for $\tilde{E}^{\mathrm{BD}}(\lambda)$.

E. Proof of Theorem 6

Define

$$f(x) \triangleq (-\log x + x - 1)$$
$$g(x) \triangleq \frac{\log x}{x-1}, \quad h(x,a,b) \triangleq \frac{xa+1}{xb+1}.$$

For the zero-mean fading, we have

$$D = \sum_{k=1}^{M} f(\alpha_k) \quad \alpha_k = h(\chi, p_{\theta_0}(k), p_{\theta_1}(k))$$
$$C = \sum_{k=1}^{M} f(\beta_k) \quad \beta_k = g(\alpha_k).$$

a) *Monotonicity*: $g'(x) < 0$ and

$$f'(x)\begin{cases} < 0 & x < 1, \\ \geq 0 & o.w. \end{cases} \quad \frac{\partial h}{\partial x}(x,a,b)\begin{cases} > 0 & a > b \\ \leq 0 & o.w. \end{cases}$$

Combining the above results, we obtain $C_\chi^{'} > 0$ Similarly, we obtain $C_\chi^{''} < 0$.

b) *Limits*: We find C_∞ by substituting, $\lim_{\chi\to\infty}\alpha_k = (p_{\theta_0}(k)/p_{\theta_1}(k))$ and see that $C_\infty < \infty$ for $p_{\theta_i}(k) > 0\ \forall i,k$.

c) *Extremal Points*: Let $E_\chi = (C_\chi/\chi)$. Now, $E_\chi^{'} = 0$ implies

$$\chi C_\chi^{'} = C_\chi. \tag{40}$$

We have $E_0 = 0$ and $E_\infty = 0$ and E is differentiable. This implies that solution exists for (40). Let χ_* be the solution. Then the double derivative at χ_* is given by

$$E_{\chi_*}^{''} = \frac{C_{\chi_*}^{''}}{\chi_*} < 0.$$

Therefore, χ_* is the unique maximum. Similar results follow for D_χ. □

ACKNOWLEDGMENT

The authors would like to thank Dr. J.-F. Chamberland, Dr. C. Tepedelenlioglu, and P. Gao for their valuable comments.

REFERENCES

[1] A. Anandkumar and L. Tong, "Distributed statistical inference using type based random access over multiaccess fading channels," presented at the CISS'06, Princeton, NJ, Mar. 2006.
[2] P. K. Varshney, *Distributed Detection and Data Fusion*. New York: Springer, 1997.
[3] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I—Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.
[4] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors: Part II—Advanced topics," *Proc. IEEE*, vol. 85, no. 1, pp. 64–79, Jan. 1997.
[5] P. Willett and L. Tong, "One aspect to cross-layer design in sensor networks," presented at the MILCOM, Monterey, CA, Oct. 2004.
[6] 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.
[7] K. Liu and A. M. Sayed, "Optimal distributed detection strategies for wireless sensor networks," in *42nd Ann. Allerton Conf. Commun., Contr., Comp.*, Oct. 2004.
[8] G. Mergen and L. Tong, "Type based estimation over multiaccess channels," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 613–626, Feb. 2006.
[9] G. Mergen and L. Tong, "Estimation over deterministic multiaccess channels," presented at the 42nd Allerton Conf. Commun., Contr., Comp., Monticello, IL, Sep. 28–Oct. 1 2003.
[10] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE J. Sel. Areas Commun. (Special Issue on Wireless Sensor Networks)*, vol. 22, no. 6, pp. 1007–1015, Aug. 2004.
[11] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, Feb. 2003.
[12] S. Aldosari and J. Moura, "Detection in decentralized sensor networks," presented at the ICASSP '04 Conf., Montreal, QC, Canada, May 2004.
[13] B. Chen, R. Jiang, T. Kasetkasem, and P. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.
[14] B. Chen and P. Willett, "On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 693–700, Feb. 2005.
[15] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access in sensor networks," in *Proc. SPAWC'05*, New York, Jun. 2005, pp. 1018–1022.
[16] A. Anandkumar and L. Tong, "A large deviation analysis of detection over multiaccess channels with random number of sensors," presented at the ICASSP'06, Toulouse, France, May 2006.
[17] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer, 1998.
[18] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[19] F. Topsøe, "Some inequalities for information divergence and related measures of discrimination," *IEEE Tran. Inf. Theory*, vol. 46, no. 4, pp. 1602–1609, Jul. 2000.
[20] A. Anandkumar, L. Tong, and A. Swami, "Detection of Gauss-Markov random fields with nearest-neighbor dependency," in *Proc. ICASSP'07*, HI, Apr. 2007.
[21] A. Anandkumar, L. Tong, and A. Swami, "Detection of Gauss-Markov random fields with nearest-neighbor dependency," Tech. Rep. ACSP TR 01-07-01, Jan. 2007 [Online]. Available: http://arxiv.org/abs/0706.1588v1
[22] R. S. Strichartz, *The Way of Analysis*. Sudbury, MA: Jones and Bartlett, 2000, vol. 2.
[23] P. Billingsley, *Probability and Measure*. New York: Wiley Inter-Sci., 1995, vol. 3.
[24] S. Resnick, *Adventures in Stochastic Processes*. New York: Springer-Verlag, 2002.

Animashree Anandkumar (S'02) received the B.Tech. degree in electrical engineering from Indian Institute of Technology Madras, Chennai, in 2004.

She is currently pursuing the Ph.D. degree in electrical engineering at Cornell University, Ithaca, NY. She has been a member of the Adaptive Communications and Signal Processing Group (ACSP), Cornell University, since August 2004. Her research interests are in the area of statistical-signal processing, information theory, and networking. Specifically, she has been working on detection and estimation, asymptotic analysis, and networking, in the context of wireless-sensor networks.

Ms. Anandkumar received the Student Paper Award at the 2006 International Conference on Acoustic, Speech and Signal Processing (ICASSP). She was named as a finalist of the Google Anita-borg Scholarship 2007–2008.

Lang Tong (S'87–M'91–SM'01–F'05) received the B.E. degree from Tsinghua University, Beijing, China, in 1985, and the M.S. and Ph.D. degrees in electrical engineering in 1987 and 1991, respectively, from the University of Notre Dame, Notre Dame, IN.

He is the Irwin and Joan Jacobs Professor in Engineering at Cornell University Ithaca, NY. Prior to joining Cornell University, he was on the faculty of West Virginia University and the University of Connecticut. He was also the 2001 Cor Wit Visiting Professor at the Delft University of Technology, The Netherlands. He was a Postdoctoral Research Affiliate with the Information Systems Laboratory, Stanford University, Stanford, CA, in 1991. His research is in the general area of statistical signal processing, wireless communications and networking, and information theory.

Dr. Tong received the 1993 Outstanding Young Author Award from the IEEE Circuits and Systems Society, the 2004 Best Paper Award (with M. Dong) from the IEEE Signal Processing Society, and the 2004 Leonard G. Abraham Prize Paper Award from the IEEE Communications Society (with P. Venkitasubramaniam and S. Adireddy). He is a coauthor of five Student Paper Awards. He also received the Young Investigator Award from the Office of Naval Research. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON INFORMATION THEORY, and IEEE SIGNAL PROCESSING LETTERS.