Cooperative Sensing for Primary Detection in Cognitive Radio

Jayakrishnan Unnikrishnan, *Student Member, IEEE*, and Venugopal V. Veeravalli, *Fellow, IEEE*

***Abstract*—One of the main requirements of cognitive radio systems is the ability to reliably detect the presence of licensed primary transmissions. Previous works on the problem of detection for cognitive radio have suggested the necessity of user cooperation in order to be able to detect at the low signal-to-noise ratios experienced in practical situations. We consider a system of cognitive radio users who cooperate with each other in trying to detect licensed transmissions. Assuming that the cooperating nodes use identical energy detectors, we model the received signals as correlated log-normal random variables and study the problem of fusing the decisions made by the individual nodes. We design a linear-quadratic (LQ) fusion strategy based on a deflection criterion for this problem, which takes into account the correlation between the nodes. Using simulations we show that when the observations at the sensors are correlated, the LQ detector significantly outperforms the Counting Rule, which is the fusion rule that is obtained by ignoring the correlation.**

***Index Terms*—Cooperative sensing, correlated observations, decentralized detection, fusion, linear-quadratic detector.**

I. Introduction

TRADITIONALLY, the use of radio frequency bands has been regulated in most countries through the process of spectrum allocation in which the use of a particular frequency band is restricted to the license holders of the band. Within this framework, spectrum has often been viewed as a scarce resource in high demand. However, measurements conducted have suggested that most licensed spectrum is often under-utilized with large *spectral holes* at different places at different times [1]. Cognitive radio systems have been proposed as a possible solution to the spectrum crisis. The idea is to detect times when a specific licensed band is not used at a particular place and use the band for transmission without causing any significant interference to the transmissions of the license holder.

While detecting the presence of a particular transmission is in itself a well-studied communication problem, the specific case of cognitive radio introduces many more constraints on the detection system that make it more involved. First, the SNR of the signal from the licensed users (referred to as *primary users*) received by the cognitive users (also called *secondary users*) can be extremely small. This is because the secondary users have to ensure that they do not interfere even with the primary transmissions at the edge of the primary's coverage area. As illustrated in Fig. 1, the secondary users located within the primary's range or in the *guard band* around it (jointly referred to as the *protected region*) could potentially interfere with the primary users' communication. Hence, secondary users even at the edge of the guard band should be able to detect the primary signal even if decoding the signal may be impossible [2].

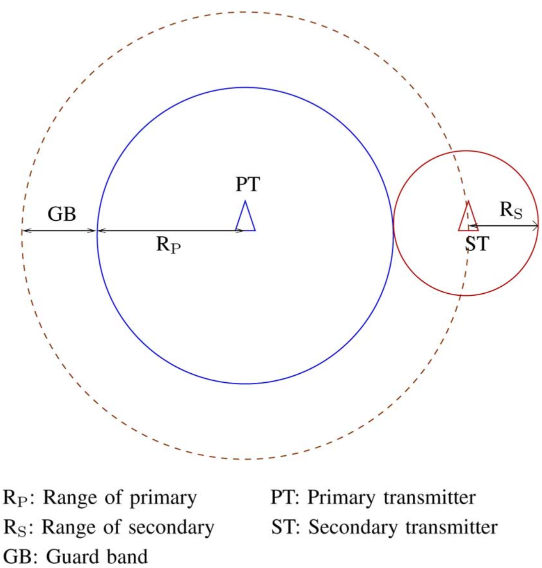


Fig. 1. Guard band. The interior of the primary's range and the guard band together form the protected region.

Secondly, the secondary users are in general not aware of the exact transmission scheme used by the primary users. Furthermore, the secondary users may not have access to training and synchronization signals for the primary transmission. This means that the secondary users are constrained to use noncoherent energy-based detectors (or feature detectors) that have much poorer performances than coherent receivers under low SNR.

Added to these issues of low SNR is the hidden-terminal problem that arises because of shadowing (see Fig. 2). Secondary users may be shadowed away from the primary user's transmitter but there may be primary receivers close to the secondary users that are not shadowed from the primary transmitter. Hence, if a secondary user transmits, it may interfere with the primary receiver's reception. This issue also needs to

Manuscript received April 13, 2007; revised November 9, 2007. This work was supported by the NSF under Award CCF 0049089, through the University of Illinois at Urbana-Champaign, and by a Vodafone Foundation Graduate Fellowship. This paper was presented in part at the Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, Nov. 2007 and at IEEE Globecom 2007, Washington, DC. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ananthram Swami.

The authors are with the Coordinated Science Laboratory and Dept. of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: junnikr2@uiuc.edu; vvv@uiuc.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTSP.2007.914880

1932-4553/$25.00 © 2008 IEEE

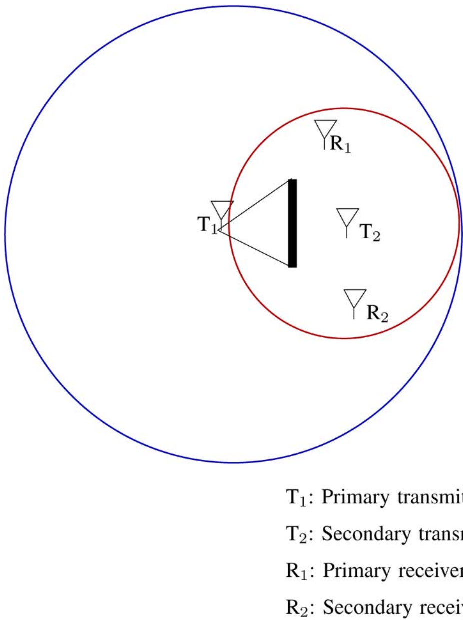


Fig. 2. Hidden terminal problem. The secondary transmitter is shadowed away from the primary transmitter but the primary receiver is not.

be addressed in order to design practical solutions to the detection problem.

Different approaches can be taken to handle the issue of low SNR. One idea is to average over longer durations of time while performing the detection. This scheme results in an increased effective SNR and hence in improved performance but at the expense of increased delay. Another possibility as suggested in [2] is to have the primary transmitter send a known pilot signal whenever it is ON. But this may not be feasible since it would require the license holders who own the band to redesign their transmit scheme throughout their network.

An alternative approach is to have users cooperating with each other to detect the primary signal. Cooperation between users follows almost as a necessary consequence of the above constraints. Having multiple cooperating users increases diversity by providing multiple measurements of the signal and thus guarantees better performance at low SNR. Additionally, having users cooperating over a wide area also provides us with a possible solution to the hidden-terminal problem, since this problem would arise only if all the secondary users are shadowed away from the primary. If the secondary users span a distance that is larger than the correlation distance of the shadow fading, it is unlikely that all of them are under a deep shadow simultaneously.

Previous works on user-cooperation for cognitive radio systems [3], [4] have considered two kinds of schemes: one where some kind of joint detection is employed among all the cooperating users and another where the final decision is made based on hard decisions made by each of the cooperating users. Gathering the entire received data at one place may be very difficult under practical communication constraints. Moreover, in practice, cooperation between the cognitive radio users cannot be guaranteed always, since a user can cooperate with others only when there are other users in its vicinity monitoring the same frequency band as itself.

In this paper, we focus on the more feasible system in which the individual secondary users make independent decisions about the presence of the primary signal in the frequency band that they are monitoring. They communicate their decisions to a fusion center that makes the final decision about the occupancy of the band by fusing the decisions made by all cooperating radios in that area that are monitoring the same frequency band. In practice, the fusion center could be some centralized controller that manages the channel assignment and scheduling for the secondary users. The system could also be one where the secondary users exchange their decisions and each secondary user performs its own fusion of all the decisions.

We assume that the fusion center knows the geographic locations of all cooperating secondary users and hence can learn the correlation between their observations. However, it is unaware of the primary's location. Since the decisions made by the secondary users contain just one bit of information each, and since we do not expect to have to keep track of the channel usage frequently, the data rates required for reliably communicating these decisions to the fusion center are expected to be within practical limits. Furthermore, the duration of data transmission is also not expected to affect the delay constraints of the spectrum sensing system.

In this paper, we address the problem of fusing decisions made at the cooperating sensors. Fusion of data observed at distributed sensors is an integral part of any decentralized detection procedure, and decentralized detection has been an active research area over the past 20 years. However, most of the significant research has been limited to the case where the sensor observations are conditionally independent under each hypothesis (see, e.g., [5]–[8] for an overview of these results). The case of correlated observations has also been studied [9], [10], [11], but the results are often not easy to implement in practice. We note that for the cognitive radio application we would have to deal with the fact that the sensors are going to observe conditionally dependent data due to correlated shadowing. The main contribution of this paper is a suboptimal fusion rule that handles correlation issues and at the same time is not heavily dependent on the model or on exact knowledge of the statistics of the signal. We show that a rule that uses only the knowledge of lower order moments of the quantized data yields good performance for different correlation structures. In the subsequent sections we first introduce the problem formulation, followed by our proposed solution, and then present our results and conclusions.

II. Problem Formulation

The basic task of the fusion center is to decide whether or not the secondary users are located inside the protected region shown in Fig. 1. As mentioned earlier, we assume that the secondary users employ energy detectors. Moreover, since the cooperating secondary users are expected to be located close to each other and are monitoring the same frequency band, the distributions of the received powers they see can be modeled as being identical, albeit not independent. So the problem now becomes a binary hypothesis testing problem to decide whether or not the mean received power at their location is higher than the power expected at the edge of the protected region. When the primary is ON and the secondary users are within the protected region, the power they receive is going to be the sum of the primary signal power and the noise[1] power. In this case, we model the received powers as being log-normally distributed. The log-normal distribution is a popular choice for modeling shadowing in wireless systems [12] in which it is assumed that the received power in decibels is distributed as Gaussian. Here, we are approximating the sum of the shadowed primary signal's power and the noise power to be log-normally distributed. We also adopt the popular correlation model [13] in which the correlation between the powers in decibels received at two different sensors decays exponentially with the distance between them.

When the secondary users are outside the protected region, the power they receive from the primary would be insignificant compared to the noise. This is particularly true if the primary is far away or is switched OFF. Under this scenario, the power at the output of the energy detectors will be simply the sum of noise power and the power of any interfering signals that may be present. We assume that both these powers can be measured with some accuracy at each of the secondary users. However, perfect knowledge of the noise powers is not feasible in practice due to the variations in the noise level at the receivers. We therefore model this uncertainty by modeling the power at the output of the energy detector as being log-normally distributed with some known variance as in [14]. This is equivalent to assuming that the uncertainty of the noise power is Gaussian in the decibels scale. Furthermore, we assume that the uncertainties are independent and identically distributed (i.i.d.) across the secondary users, with the understanding that the uncertainty in the received power under H_0 is dominated by the thermal noise at the individual receivers.

The two hypotheses of interest are H_1, the hypothesis that the primary is present and is located close to the secondary users, and H_0, the hypothesis that the primary is absent or is far away. Here, H_0 can also be viewed as the hypothesis that a spectral hole exists and hence the spectrum is free for secondary access. The cooperating secondary users subtract the estimated value of the sum of noise and interference powers (in dBm) from their received powers, to obtain their observations $\{Y_i\}_1^n$. Hence, we have the following statistical model for the vector $\underline{Y}$ of observations at the n cooperating secondary users (sensors) under the two hypotheses

$$\begin{aligned} H_0 &: \underline{Y} \sim \mathcal{N}\left(\underline{0}, \sigma_0^2 I\right) \\ H_1 &: \underline{Y} \sim \mathcal{N}(\theta \underline{1}, \Sigma) \text{ with } \theta \geq \mu_1 \end{aligned} \tag{1}$$

where $\mathcal{N}(\underline{v}, M)$ denotes a Gaussian vector distribution with mean $\underline{v}$ and covariance matrix M. Here, θ is a variable parameter representing the mean of the distributions observed under hypothesis H_1, while μ_1 is the mean total power in dBm received at the edge of the guard band minus the mean noise power in dBm, given by

$$\mu_1 = E[10 \log_{10}(1 + \mathsf{SNR})] \mathrm{dB}$$

[1]We are implicitly assuming that the noise includes interference from other distant users of the spectrum.

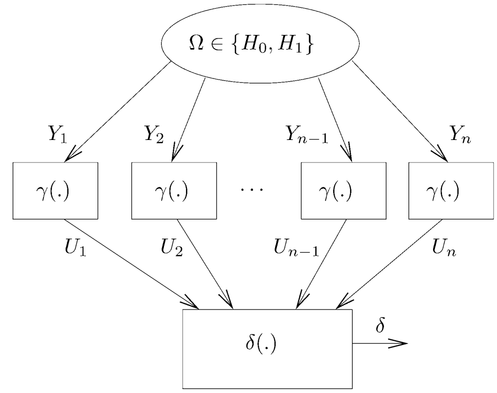


Fig. 3. Decentralized detection setup.

where SNR is the random signal-to-noise ratio at the edge of the guard band when the primary is ON, σ_0 quantifies the uncertainty in noise power, $\underline{1}$ is the vector of all ones, Σ is the matrix with elements $\Sigma_{ij} = \sigma_1^2 \rho^{d_{ij}}$ where d_{ij} is the distance between nodes indexed by i, and j, ρ is a measure of the correlation-coefficient between nodes separated by unit distance, and σ_1^2 is the net variance under H_1. The parameter ρ is related to the correlation distance $\mathsf{D_C}$ [13] by the relation $\rho = \exp(-1/\mathsf{D_C})$.

The constraint that our system should meet is to guarantee that the probability of interfering with the primary transmission is less than some pre-specified limit p_{I}. We assume that the secondary users use the spectrum for transmission whenever they detect a spectral hole. Hence, in our problem, the probability of interfering with the primary transmission would be equal to the probability of making an erroneous decision about the hypothesis under H_1. Hence, our system should guarantee that the probability of making an erroneous decision under hypothesis H_1 should be lower than the constraint on the probability of interference. Moreover, this constraint should be met for all values of θ greater than or equal to μ_1 in (1). This is a composite binary Neyman-Pearson hypothesis testing problem [15]. Now since we do not have any prior information about the distribution of the mean powers θ, we have to design our system under a *robust detection* or *universally most powerful* detection framework [15]. This means that we have to design our system such that it meets the interference probability constraint with equality for the least favorable value of θ, which in our case occurs when θ is equal to μ_1.

To summarize, we have reduced our detection problem (1) to a simple Neyman-Pearson hypothesis testing problem between the two modified hypotheses

$$\begin{aligned} H_0 &: \underline{Y} \sim \mathcal{N}(\underline{0}, \sigma_0^2 I) \\ H_1 &: \underline{Y} \sim \mathcal{N}(\mu_1 \underline{1}, \Sigma). \end{aligned} \tag{2}$$

In our system model, the final decision about the hypothesis is made at the fusion center which has access only to the binary-valued decisions made by the sensors based on their individual observations $\{Y_i\}_1^n$. We use $\{U_i\}_1^n$ to represent the decisions made at the individual sensors and $\underline{U}$ to represent the vector of decisions made by all sensors. Hence, we have the decentralized Neyman-Pearson hypothesis testing problem illustrated in Fig. 3.

We now introduce some additional notation. We use p_j to denote the distributions of the Y_i's, $i = 1, 2, \ldots, n$ and q_j for the distributions of the U_i's, $i = 1, 2, \ldots, n$, under hypothesis $H_j, j = 0, 1$. We also use P_j to denote the probability measure under hypothesis H_j and E_j to denote the expectation operator under hypothesis $H_j, j = 0, 1$.

III. Detection Rule at the Nodes

In the system we consider, each cognitive radio is designed without expecting cooperation from other users in the detection process, with the understanding that we do not always expect a particular radio to be close enough to other radios that are also monitoring the same primary signal. Hence, we assume that the detector employed at each radio meets the probability of interference constraint on its own. Since we further assume that the signals received at the cooperating users are identically distributed, the detectors used by the cooperating users will also be identical. Moreover, the optimal test used by sensor i to determine its decision U_i will be a likelihood ratio test on its observation Y_i [15]. It will be of the form

$$U_i = \mathcal{I}_{\{\log(L(Y_i)) > \tau\}}$$

where $\mathcal{I}_{\{\}}$ is the indicator function, which takes on value 1 when its argument is true and 0 otherwise, $L(Y_i) = (p_1(Y_i))/(p_0(Y_i))$ is the likelihood ratio of the observation at node i and τ is the threshold employed at every node. The threshold τ is chosen so that the probability of making an incorrect decision at the node under hypothesis H_1 is equal to the constraint on the interference probability, p_{I}. The identical likelihood ratio tests used at the sensors are represented by $\gamma()$ in Fig. 3.

For the Gaussian hypotheses described in (2), the log-likelihood ratio of the observations will be a quadratic function of Y_i in general [15]. Hence, node i would have to compare a quadratic function of its observation to a threshold and obtain its decision in the form of bit U_i. These bits are communicated to the fusion center where the final decision about the hypothesis is made. The symbol $\delta(.)$ in Fig. 3 represents the final decision made at the fusion center.

IV. Fusion of Decisions

The problem of optimal fusion of decentralized observations has been studied in many works [10], [11]. It is known that the optimal fusion rule is to compute the joint likelihood ratio of the bits and compare it with a threshold chosen so as to meet the interference probability requirement [10]. This solution, in general, requires the knowledge of the joint statistics of the bits under both hypotheses. The results of [11] show how the optimal fusion rule can be expressed in terms of the conditional correlation coefficients of all orders, which again requires the knowledge of the joint statistics of the bits.

However, in our problem, the U_i's are binary quantized versions of correlated Gaussian variables under H_1, and hence their joint statistics are not easy to compute especially for large values of n. In this paper, we present some simple suboptimal fusion strategies that circumvent this problem. The structure of these detectors can be obtained using only partial statistical information about the quantized observations. However, the threshold to be used at the fusion center for a target interference probability will need to be estimated using simulations. This would be the case with all possible fusion rules since analytical expressions for the error probabilities of the fusion rules can be obtained only when the joint statistics of the sensor decisions are available.

A. Counting Rule

One of the simplest suboptimal solutions to the data fusion problem is the *Counting Rule* [16] (also referred to as the *Voting Rule*), which just counts the number of sensor nodes that vote in favor of H_1 and compares it with a threshold. Equivalently, the decision is based solely on the *type* [17] of the received vector of bits. The threshold value has to be set using simulations since the joint statistics under H_1 are not available. It is easy to see that under the special scenario where the observations are i.i.d across the sensors under both hypotheses, this is the optimal rule since the joint likelihood ratio of the bits is a function of only the *type* of the received bit vector. Thus this would be a reasonable fusion strategy even when nothing is known about the correlation structure. How well a rule designed for a decentralized hypothesis test with correlated observations makes use of the correlation information could thus be quantified by comparing its performance with that of the counting rule for the same observations. Moreover, the fact that the fusion center threshold for even a simple fusion rule like the Counting Rule needs to be set using simulations suggests that the same is to be expected for more sophisticated fusion rules.

B. Linear Quadratic Detector

In this section, we present the main contribution of this paper—a general suboptimal solution to the fusion problem that uses partial statistical knowledge and gives better performance than the one obtained by ignoring the correlation information completely.

This solution makes use of the second-order statistics of the local decisions $\{U_i\}_1^n$ under H_1 and the fourth-order statistics under H_0 in the form of moments. Since the observations are independent under H_0, the moments under H_0 are easily calculated or estimated. The second-order moments under H_1 can be obtained by calculating or estimating just the pairwise statistics under H_1. We note that obtaining information about these moments is in general a lot easier than obtaining the entire joint statistics of the signals especially when there are a large number of cooperating nodes.

We consider fusion rules in the class of linear-quadratic (LQ) detectors, i.e., detectors that compare a linear-quadratic function of $\underline{U}$ with a threshold. Since we are including quadratic terms as well while computing our detection metric, we expect to see improved performance over the Counting Rule that is purely linear. Moreover, since we are using only moment information about $\underline{U}$, this detector is quite general and can be used for all classes of distributions of the signals. We seek to optimize over the class of LQ detectors using the *generalized signal-to-noise ratio* or *deflection* criterion [18]. If $\underline{X}$ represents the observations in some detection problem, the deflection of a detector that makes a decision by comparing a function $T(\underline{X})$ to a threshold is defined as

$$D_T = \frac{[E_1(T(\underline{X})) - E_0(T(\underline{X}))]^2}{\mathrm{Var}_0(T(\underline{X}))}. \quad (3)$$

Although deflection cannot be related directly to the error-probability for non-Gaussian observations, a detector with a higher value of deflection is expected to have better error-probability performance than one with a lower value of deflection. We show using simulations that the optimal deflection-based LQ detector that we derive in this section gives improved error performance over the Counting Rule in correlated environments.

Following [19], we solve for the optimal LQ detector. Since we now have quadratic terms as well, the values that we assign to the bits become significant. Values of 1 or 0 assigned to the decision random variables U_i do not make much sense simply because they are not representative of the actual values taken by the signal Y_i. The question then becomes: what values should be assigned to the decision variables?

For a binary hypothesis test involving two Gaussian vector distributions with equal variances and arbitrary covariances, it can be shown that [15] the optimal centralized detection rule is, in general, to compare a linear quadratic function of the observations with a threshold. In this case, the detection metric can be viewed as a linear quadratic function of the log-likelihood ratios of the individual random variables. This observation suggests that an intelligent choice of values to be assigned to the quantized observations in our problem would be the log-likelihood ratios of the bits themselves.

Hence, we express our decision metric as

$$T(\underline{X}) = \underline{h}^\top \underline{X} + \underline{X}^\top M \underline{X} \quad (4)$$

where $\underline{X}$ is the vector of log-likelihood ratios of the received bits with means under H_0 subtracted. The components of $\underline{X}$ are given by

$$X_i = \log\left(\frac{q_1(U_i)}{q_0(U_i)}\right) - E_0\left[\log\left(\frac{q_1(U_i)}{q_0(U_i)}\right)\right]$$

while $\underline{h}$ is a vector of length n and M is an $n \times n$ square matrix. It is important to note that X_i can be computed directly from U_i without any added information since the log-likelihood function and its expected value can be obtained directly from the first-order distributions of the bits which are assumed to be known at the fusion center.

We need to find the optimal LQ metric of the form (4) that maximizes the deflection given by (3). Clearly, this optimization will require the knowledge of up to the second-order statistics of the bits under H_1 and up to the fourth-order statistics of the bits under H_0 since these terms explicitly appear in the expression for the deflection (3).

Define matrix $C = E_0[\underline{X}\underline{X}^\top]$. Since adding a constant to the decision metric leaves the deflection unchanged, (4) can be replaced by a new decision metric given by

$$S(\underline{Z}) = \underline{x}^\top \underline{Z} \quad (5)$$

where $\underline{x}$ is now an $(n^2+n)\times 1$ vector and $\underline{Z}$ is an $(n^2+n)\times 1$ vector given by

$$\begin{aligned}\underline{Z} = \big[&X_1 \ldots X_n \quad X_1^2 - C_{11} \ldots X_1X_n - C_{1n} \\ &X_2X_1 - C_{21} \ldots X_2X_n - C_{2n} \\ &\ldots \\ &X_nX_1 - C_{n1} \ldots X_n^2 - C_{nn}\big]^\top.\end{aligned}$$

In other words, we form $\underline{Z}$ by appending $\underline{X}$ with the raster-scanned form of $\underline{X}\underline{X}^\top - C$. So the first n elements of $\underline{Z}$ are the elements of $\underline{X}$, the next n are the elements of the first row of $\underline{X}\underline{X}^\top - C$, followed by the elements of the second row of $\underline{X}\underline{X}^\top - C$, and so on. Similarly, $\underline{x}$ can be viewed as a vector formed by appending vector $\underline{h}$ with matrix M in raster-scanned form. So the problem of finding optimal $\underline{h}$ and M reduces to solving for the optimal $\underline{x}$ that maximizes the deflection for this decision metric.

From the construction of $\underline{Z}$ it is easy to see that $\underline{Z}$ has zero mean under H_0. Hence, applying (3) to (5), we have deflection for $S(\underline{Z})$ given by

$$\begin{aligned}D_S &= \frac{[E_1(\underline{x}^\top \underline{Z}) - E_0(\underline{x}^\top \underline{Z})]^2}{\mathrm{Var}_0(\underline{x}^\top \underline{Z})} \\ &= \frac{(\underline{x}^\top \underline{\mu})^2}{\underline{x}^\top K \underline{x}}\end{aligned} \quad (6)$$

where $\underline{\mu} = E_1(\underline{Z})$ and $K = E_0(\underline{Z}\underline{Z}^\top)$. Matrix K is a function of the second-, third-, and fourth-order moments of X_i's under H_0 and vector $\underline{\mu}$ is a function of the first- and second-order moments under H_1. Since these moments are assumed known, vector $\underline{\mu}$ and matrix K are known at the fusion center *a priori*. In general, K is positive semi-definite but not strictly positive definite. But it can be shown that the vector $\underline{x}$ that drives the denominator of (6) to zero drives the numerator also to zero. This is because if $\underline{x}^\top K \underline{x} = 0$, we have

$$E_0(\underline{x}^\top \underline{Z}\underline{Z}^\top \underline{x}) = E_0((\underline{x}^\top \underline{Z})^2) = 0$$

whence $\underline{x}^\top \underline{Z} = 0$ w.p. 1. Therefore,

$$E_1(\underline{x}^\top \underline{Z}) = \underline{x}^\top \underline{\mu} = 0$$

which would mean that the two distributions of $S(\underline{Z})$ have the same mean. Since we do not desire this, it is sufficient to perform the optimization of (6) over those $\underline{x}$ vectors that do not lie in the singular space of matrix K.

Since K is positive semi-definite, it can be diagonalized as $K = V\Lambda V^\top$ where V is a unitary matrix and Λ is a diagonal matrix with nonnegative entries. Therefore, (6) can be written as

$$\frac{(\underline{x}^\top \underline{\mu})^2}{\underline{x}^\top K \underline{x}} = \frac{(\underline{\tilde{x}}^\top \underline{\tilde{\mu}})^2}{\underline{\tilde{x}}^\top \Lambda \underline{\tilde{x}}} \quad (7)$$

which can be reduced to

$$D_S = \frac{(\underline{\widetilde{x_a}}^\top \underline{\widetilde{\mu_a}})^2}{\underline{\widetilde{x_a}}^\top \Lambda_a \underline{\widetilde{x_a}}} \quad (8)$$

where $\underline{\tilde{\mu}} = V^\top \underline{\mu}$ and $\underline{\tilde{x}} = V^\top \underline{x}$. Equation (8) follows by defining Λ_a as the diagonal matrix containing only the nonzero diagonal elements of Λ and defining $\underline{\widetilde{x_a}}$ and $\underline{\widetilde{\mu_a}}$ as the vectors composed of the corresponding elements of $\underline{\tilde{x}}$ and $\underline{\tilde{\mu}}$, respectively. Hence, we just have to optimize over $\underline{\widetilde{x_a}}$.

Now (8) can be written as

$$D_S = \frac{\left(\underline{\widetilde{x_a}}^\top \Lambda_a^{1/2} \Lambda_a^{-1/2} \underline{\widetilde{\mu_a}}\right)^2}{\left(\underline{\widetilde{x_a}}^\top \Lambda_a^{1/2} \Lambda_a^{1/2} \underline{\widetilde{x_a}}\right)} = \frac{\left(\left(\Lambda_a^{1/2}\underline{\widetilde{x_a}}\right)^\top \left(\Lambda_a^{-1/2}\underline{\widetilde{\mu_a}}\right)\right)^2}{\left(\Lambda_a^{1/2}\underline{\widetilde{x_a}}\right)^\top \left(\Lambda_a^{1/2}\underline{\widetilde{x_a}}\right)} = \left(\frac{\left(\Lambda_a^{1/2}\underline{\widetilde{x_a}}\right)^\top \left(\Lambda_a^{-1/2}\underline{\widetilde{\mu_a}}\right)}{\left|\left(\Lambda_a^{1/2}\underline{\widetilde{x_a}}\right)\right|}\right)^2. \quad (9)$$

The right hand side of (9) is of the form $((\underline{z}^\top \underline{w})/(|\underline{z}|))^2$ which is maximized when the vectors $\underline{z}$ and $\underline{w}$ are collinear. It follows that the optimal $\underline{\widetilde{x_a}}$ satisfies

$$\left(\Lambda_a^{1/2}\underline{\widetilde{x_a}}\right) = k\Lambda_a^{-1/2}\underline{\widetilde{\mu_a}} \quad (10)$$

where k in (10) is any constant, which can be chosen to be unity without loss of optimality, to give

$$(\underline{\widetilde{x_a}})_{\text{opt}} = \Lambda_a^{-1}\underline{\widetilde{\mu_a}}. \quad (11)$$

Since Λ_a contains only the positive eigenvalues of Λ as diagonal elements it is nonsingular and (11) is well-defined. Hence, the optimal decision metric from (5) has the form

$$S_{\text{opt}}(\underline{Z}) = \underline{x_{\text{opt}}}^\top \underline{Z} = \underline{\widetilde{x_{\text{opt}}}}^\top \underline{\tilde{Z}} = (\underline{\widetilde{x_a}})_{\text{opt}}^\top \underline{\widetilde{Z_a}} = \underline{\widetilde{\mu_a}}^\top \Lambda_a^{-1} \underline{\widetilde{Z_a}} \quad (12)$$

where $\underline{\tilde{Z}} = V^\top \underline{Z}$ and $\underline{\widetilde{Z_a}}$ is obtained by keeping only the terms of $\underline{\tilde{Z}}$ corresponding to those of $\underline{\tilde{\mu}}$ that appear in $\underline{\widetilde{\mu_a}}$, and the optimal deflection obtained is thus

$$(D_S)_{\text{opt}} = \underline{\widetilde{\mu_a}}^\top \Lambda_a^{-1} \underline{\widetilde{\mu_a}}. \quad (13)$$

Hence, the deflection-optimal LQ detector compares the metric given by (12) to a threshold chosen so as to obtain equality in the probability of interference constraint. This threshold would have to be set using simulations since the statistics of the detection metric are not available. Since the detection metric is discrete-valued, *randomization* may also be required to achieve equality in the interference probability constraint. Clearly, the computation of $\underline{\tilde{\mu}}$ and K requires only the knowledge of up to the fourth-order statistics of the decision variables under H_0 and up to the second-order statistics under H_1. Hence, the detector based on (12) can be used for all distributions of the original observations as long as these lower order statistics can be calculated.

C. Quantifying the Simulation Complexity

For both the LQ detector and the Counting Rule, the fusion center threshold has to be set using simulations since the joint statistics of the bits are unknown. In this section, we study the dependence of simulation complexity on the problem parameters.

Let W denote the test statistic at the fusion center in either the LQ detector or the Counting Rule. We are interested in determining τ so that $P_1(W < \tau)$ meets some target probability of interference requirement p_{I} with equality. In general, we need randomization between two different thresholds for gaining equality in the interference probability requirement.

For a known τ, we can estimate $p = P_1(W < \tau)$ by generating an i.i.d. sequence of variables $W_1, W_2, \dots, W_N$ under hypothesis H_1 and then obtain our estimate $\hat{p}$ of p as

$$\hat{p} = \frac{1}{N}\sum_{i=1}^{N} \mathcal{I}_{\{W_i < \tau\}}$$

where $\mathcal{I}$ is the indicator function. The ratio of the standard deviation σ of this estimate to the value of p can be calculated to be

$$\frac{\sigma}{p} = \sqrt{\frac{1-p}{pN}}.$$

Hence, in order to obtain a standard deviation of not more than αp we need to generate and average over $N = (1-p)/(\alpha^2 p)$ i.i.d. random variables with the same distribution as W. Since we are interested in threshold values that yield values of p close to p_{I}, we need average only over $N = (1-p_{\text{I}})/(\alpha^2 p_{\text{I}})$ random variables. In practice, the value of τ can be fixed by determining the point where the empirical cumulative distribution (c.d.f.) of the sequence $\{W_i\}_1^N$ equals p_{I}. Since the detection statistic for both the scenarios described earlier are discrete random variables, there may be no point where the c.d.f. becomes exactly equal to p_{I}. In such a case, we have to *randomize* [15] between two thresholds which correspond to the points where the c.d.f. gets closest to p_{I} from above and below, respectively. The randomization parameter can be set after the two thresholds are determined up to the desired level of accuracy.

Thus, the number of iterations required depends only on the target interference probability and the percentage accuracy that we desire and is independent of other parameters in the problem like the number of cooperating users. The number of cooperating users, however, affects the size of the alphabet to which W belongs. This alphabet size grows exponentially with the number of sensors n. A trivial upper bound on the size of this alphabet is 2^n since the decision metric is a function of n binary variables.

V. Simulation Results

Since analytical expressions for the error probabilities of these detectors cannot be obtained, we need to resort to simulations for estimating their performances. For our detection problem, the performance metrics of interest are the probability of successfully detecting the presence of spectral holes given by $P_0(\delta(\underline{U}) = H_0)$ and the probability of interference under H_1 given by $P_1(\delta(\underline{U}) = H_0)$, where δ represents the final decision about the hypothesis made at the fusion center. The performance of a detector can be illustrated by plotting

TABLE I
TABLE OF DETECTION PROBABILITIES

p_I (probability of interference)	Probability of detecting spectral holes				
	Single sensor	Counting rule	LQ detector with predetermined sensor threshold	LQ detector with sensor threshold varied	Centralized detector
0.001	0.01240	0.0218	0.0552	0.0663	0.2155
0.004	0.0458	0.0851	0.1522	0.1657	0.4920
0.007	0.0790	0.1486	0.2109	0.2181	0.6266
0.01	0.1127	0.2153	0.2698	0.2865	0.7130

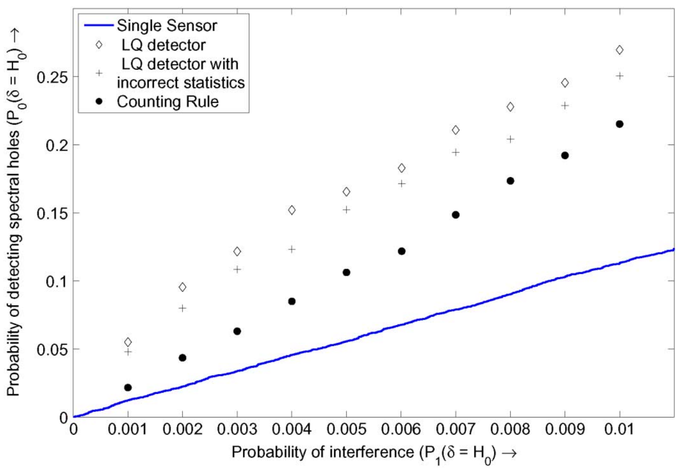


Fig. 4. Comparison of performance obtained with a single sensor and detection probabilities obtained with LQ detector and Counting Rule for nine sensors uniformly placed inside a unit square with parameters set to $\rho = 0.6, \mu_1 = 3.4$ dB, $\sigma_1 = 2.1$ dB and $\sigma_0 = 1$ dB. Also shown is the performance of the LQ detector when statistical information is inaccurate.

the detection probability under H_0 against the probability of interference under H_1 obtained with the detector.

A. Comparison of Performances

The two rules obtained in the previous section were simulated for a network of nine cooperating nodes uniformly placed inside a unit square with the distance between nearest neighbors kept at 0.5. The correlation parameter ρ appearing in the Σ matrix in (2) was kept at $\rho = 0.6$. This effectively amounts to assuming that the length of the side of the square is around half the correlation distance, since $\exp(-0.5) \simeq 0.6$. Assuming a mean received SNR of 0 dB at the edge of the guard band, a shadowing standard deviation of 4 dB and a noise uncertainty $\sigma_0 = 1$ dB, we get the value of the mean total power at the edge of the guard band, μ_1, to be 3.4 dB and the effective standard deviation of the received power under H_1, σ_1, to be 2.1 dB. We carried out simulations for the different cases outlined below.

1) Sensor Thresholds Fixed a Priori Based on Interference Probability Required: We simulated the two rules for probability of interference values in the range 0.001 to 0.01. The sensors use identical likelihood ratio tests for obtaining their decision variables U_i. The threshold used at the sensors is chosen such that the probability of making a wrong sensor decision under H_1 equals the constraint on the probability of interference. We allow for *randomization* [15] at the fusion center while performing the detection. The performance curve obtained with the single sensor detector is illustrated in Fig. 4. The points obtained by applying the LQ detector and Counting Rule for fusing the decisions of all sensors are also shown on the same graph. As expected, the performances of the detectors that make use of the information from all the sensors are better than the one that uses decisions made at a single sensor. In particular, the LQ detector is seen to give around two to three times the detection probability as that of the single sensor detector for the interference probability values considered even though the observations are highly correlated (the distance spanned by the nodes is equal to half the correlation distance of the shadow fading). It can also be inferred that the LQ detector yields a substantial gain over the Counting Rule, especially at low values of interference probability, which would be the region of interest for the cognitive radio application. We have listed the exact values obtained in Table I.

2) Incorrect Correlation Statistics: This simulation studies the effects of incorrect correlation information on the performance of the LQ detector. We keep the node configuration the same as before, but assume that the locations of the nodes are known only up to an accuracy within ± 0.3 of the correct value. In other words, the locations are known only up to an accuracy of 30% of the width of the square within which the cooperating secondaries are located. We further include a mismatch in the estimate of parameter ρ by assuming that the fusion center uses a incorrect value of 0.1 while computing the required statistical information. As we see in Fig. 4, the performance drop in the LQ detector is very little and it still gives a significant advantage over the Counting Rule detector. It was observed in other simulations that there is no perceptible drop in performance of the LQ detector for lower distance offsets.

3) Sensor Thresholds Varied: So far in this paper, we have assumed that the individual sensors have to set their thresholds assuming that they are not going to obtain any additional help from other users. We now relax this assumption and allow the users to choose their thresholds from a set of values close to the original predetermined threshold. We still restrict them to use the same threshold. The algorithm performs a limited search in a finite set of sensor thresholds and chooses the one that gives the best error performance at the fusion center. This choice would depend on both the configuration of the cooperating users as well as the target interference probability.

The detection probabilities for the LQ detector obtained by allowing the users to vary their thresholds is also listed in Table I. Clearly, the values in the table indicate that additional gains in

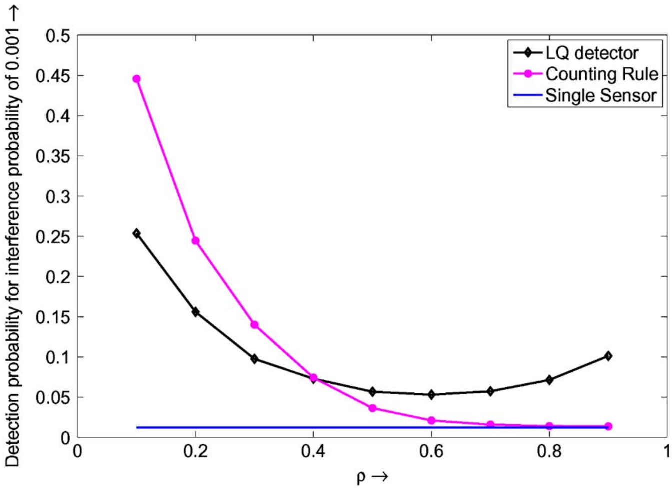


Fig. 5. Comparison of performances of LQ and Counting Rule detectors as a function of correlation parameter ρ.

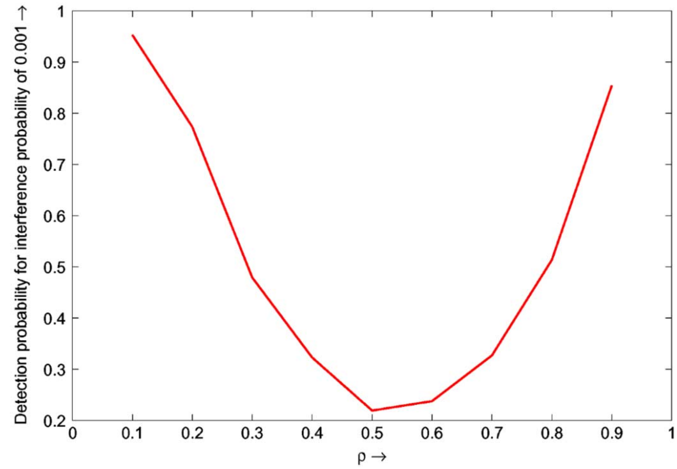


Fig. 6. Performance of optimal centralized detector as a function of correlation parameter ρ.

detection probability can be obtained, as expected. This, however, would require additional communication overheads between the fusion center and the individual users. Moreover, there is no simple criterion to optimally choose the threshold to be used at the individual users for a given interference probability and known correlation information. Searching over the entire real number line for possible threshold values is clearly not a feasible solution.

4) Centralized Detector: In Table I, we have also included the detection probabilities obtained by employing the centralized detection rule. This detector performs the joint likelihood ratio test on the entire vector of unquantized observations $\underline{Y}$. Clearly, it performs much better than the other detectors that use only the decision variables $\underline{U}$, thus illustrating the loss in performance incurred due to making binary decisions at the cooperating nodes.

B. Comparison of Performances as a Function of Correlation

In this simulation, we compare the performances of the LQ detector and the Counting Rule detector for different values of the correlation parameter ρ. Other signal parameters and the configuration of the nodes is kept the same as in Section V-A. The interference probability is kept fixed at 0.001 and the detection probability obtained with the detectors is plotted as a function of ρ. As in Section V-A1, the sensors use identical quantizers with threshold chosen such that the probability of making a wrong sensor decision under H_1 equals the constraint on the probability of interference. As seen in Fig. 5, the LQ detector outperforms the Counting Rule detector for all values of ρ greater than 0.4. For low values of correlation, the observations at the sensors are nearly independent under both hypotheses. Hence the Counting Rule, being the optimal detector for i.i.d. observations, performs better than the suboptimal LQ detector. For values of the correlation greater than 0.4, the performance of the Counting Rule detector steadily decreases and converges to that obtained with a single sensor, while the performance of the LQ detector starts increasing for higher correlation values. The nonmonotonic behavior of the performance of the LQ detector as a function of correlation is due to the fact that the value of the correlation parameter affects the detection problem in two different ways. There are primarily two different features in the signals that help us distinguish between the two hypotheses.

The first feature is the difference in the distributions of the signal at any single sensor. The ability to distinguish between the two hypotheses by exploiting only this feature decreases as the value of ρ increases since the amount of *diversity* or the amount of independent information that we get about the hypotheses decreases as ρ increases. This behavior is seen in the case of the Counting Rule detector, which relies primarily on the difference between the means of the quantized signals to distinguish between the two hypotheses.

The second feature that we can exploit is the difference in the amount of correlation existing between the observations under the two hypotheses. The ability to distinguish between the two hypotheses based on this feature clearly increases as ρ increases.

Hence, in detectors that try to exploit both these features, we expect to see a nonmonotonicity in performance as a function of ρ. This kind of behavior is seen in both the LQ detector (Fig. 5) which exploits both these features from the quantized observations as well as in the optimal centralized detector (Fig. 6), which makes optimal use of all the unquantized observations.

C. Averaging Over Topologies

Here, we average the performances obtained with the detection rules over different topologies of the network. We consider different random topologies obtained by randomly scattering 9 cooperating users within a unit square and compare the average performances obtained by employing the different detectors. We use the same signal parameters as in Section V-A. As seen in Fig. 7, the LQ detector still outperforms the Counting Rule significantly.

VI. Conclusion

Our results clearly suggest that even when the observations at the sensors are moderately correlated, it is important not to ignore the correlation between the nodes for fusing the local decisions made at the secondary users. In such scenarios, the LQ detector provides a simple fusion rule that yields significant

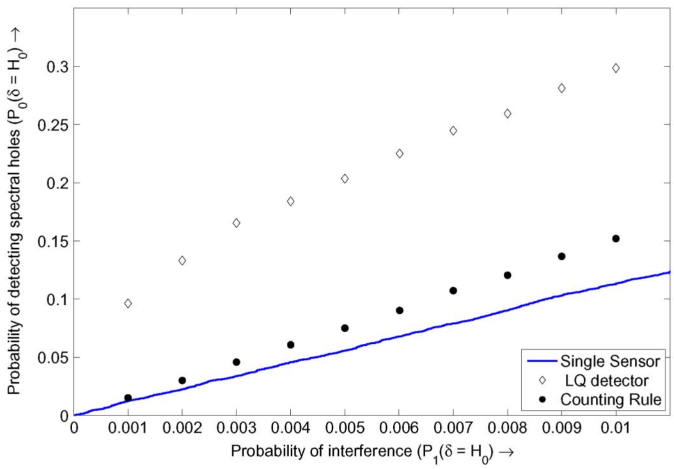


Fig. 7. Comparison of performances of various detectors averaged over random configurations of nine nodes within a unit square.

performance gains over the Counting Rule while still using only partial statistical knowledge about the correlated decision variables. The Counting Rule is useful in a system where the correlation between the observations at the users is small.

The LQ detector is also applicable for more general statistical models for the signals, since all it requires is information about the lower order moments about the correlated decision variables, which can be easily calculated for most standard signal models. It is expected that for a system comprising a larger number of cooperating nodes, better detection probabilities could be obtained with the LQ detector, even without varying the thresholds at the cooperating users. We could also generalize the LQ detector to the scenario where the cooperating users employ higher level quantizers. However, for such an application, the task of the fusion center would become more complex since computing the moments required for the LQ detector would become more involved. Moreover, designing the quantizers to be used at the sensors would also be nontrivial since we would now have to set multiple thresholds at each sensor.

The performance of the LQ detector can be further improved through some extensions. One such method as we illustrated in Section V-A3, is to allow the cooperating users to select their thresholds based on knowledge about the number of cooperating sensors and the correlation structure. However, as mentioned earlier, this scheme requires added communication overheads.

A different approach is possible in a system where the users act as fusion centers after obtaining the decisions made by the other users. In particular, each sensor could fuse its unquantized observation with the decisions that are made at the other sensors. Since we are using one unquantized observation, we expect to see an improvement in performance. The moment information required for this fusion method is different from what was required for the earlier method, but still requires only the pairwise statistics under H_1 and up to fourth-order statistics under H_0.

Another possible extension of this work is to combine such a cooperative spectrum sensing system with the system that controls the sensing and access policies of the cooperating users [20]. The joint design of sensing and access policies from a single user's perspective has been studied in previous works [21]. In our problem, we have a group of sensors cooperatively sensing the spectrum. We could extend the level of cooperation to include the access policies as well. In such a system the sensing and access policies to be used by all the cooperating users could be jointly designed so as to maximize the net throughput of the cooperating users.

The results on decentralized detection obtained in this paper are also useful in a more general context as we illustrate in [22]. They could, for instance, be used in sensor networks when the observations are conditionally dependent under either hypothesis.

Acknowledgment

The authors would like to thank J. Chang for assistance in performing the simulations.

References

[1] R. W. Broderson, A. Wolisz, D. Cabric, S. M. Mishra, and D. Willkomm, *CORVUS: A Cognitive Radio Approach for Usage of Virtual Unlicensed Spectrum*. Berkeley, CA: Univ. California Berkeley Whitepaper, Jul. 2004.

[2] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Proc. 42nd Allerton Conf. Communication, Control, and Computing*, Oct. 2004.

[3] S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative sensing among cognitive radios," in *Proc. IEEE Int. Conf. Communications*, vol. 4, pp. 1658–1663.

[4] E. Visotsky, S. Kuffner, and R. Peterson, "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," in *Proc. 1st IEEE Int. Symp. New Frontiers in Dynamic Spectrum Access Networks DySPAN*, 2005, pp. 338–345.

[5] J. N. Tsitsiklis, "Decentralized detection," *Advances in Statistical Signal Processing*, vol. 2, pp. 297–344, 1993.

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

[7] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Processing*, vol. 51, no. 2, pp. 407–416, Feb. 2003.

[8] V. V. Veeravalli and J.-F. Chamberland, , A. Swami, Ed. *et al.*, "Detection in sensor networks," in *Wireless Sensor Networks. Signal Processing and Communications Perspectives*. New York: Wiley, 2007.

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

[10] E. Drakopoulos and C.-C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 5–14, Jul. 1991.

[11] M. Kam, Q. Zhu, and W. S. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 916–920, 1992.

[12] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2001.

[13] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.

[14] K. V. Cai, V. Phan, and R. J. O'Connor, "Energy detector performance in a noise fluctuating channel," in *Proc. Military Commun. Conf.*, 1989, vol. 1, pp. 85–89.

[15] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer-Verlag, 1994.

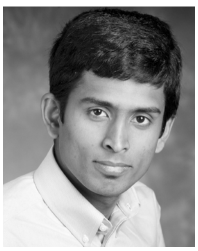
[16] V. Aalo and R. Viswanathan, "Asymptotic performance of a distributed detection system in correlated Gaussian noise," *IEEE Trans. Signal Processing*, vol. 40, pp. 211–213, 1992.

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

[18] B. Picinbono, "On deflection as a performance criterion in detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 3, pp. 1072–1081, 1995.

[19] B. Picinbono and P. Duvaut, "Optimal linear-quadratic systems for detection and estimation," *IEEE Trans. Inform. Theory*, vol. 34, no. 2, pp. 304–311, 1988.

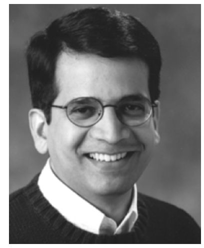
[20] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE J. Select. Areas Commun.*, vol. 25, no. 3, pp. 589–600, Apr. 2007.
[21] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Nov. 2006, pp. 696–700.
[22] J. Unnikrishnan and V. V. Veeravalli, "Decentralized detection with correlated observations," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Nov. 2007.



Jayakrishnan Unnikrishnan (S'06) received the B. Tech. degree in electrical engineering from the Indian Institute of Technology Madras, Chennai, India, in 2005 and the M.S. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign in 2007, where he is currently pursuing the Ph.D. degree.

His research interests include communication systems and detection theory.

Mr. Unnikrishnan is a recipient of the Vodafone Graduate Fellowship Award from the University of Illinois at Urbana-Champaign for 2007–2008.



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

He joined the University of Illinois at Urbana-Champaign in 2000, where he is currently a Professor in the Department of Electrical and Computer Engineering, and a Research Professor in the Coordinated Science Laboratory. He served as a Program Director for communications research at the U.S. National Science Foundation in Arlington, VA, from 2003 to 2005. He has held academic positions at Harvard University, Cambridge, MA, Rice University, Houston, TX, and Cornell University, Ithaca, NY. His research interests include distributed sensor systems and networks, wireless communications, detection and estimation theory, and information theory.

Dr. Veeravalli is on the Board of Governors of the IEEE Information Theory Society. He was an Associate Editor for Detection and Estimation for the IEEE TRANSACTIONS ON INFORMATION THEORY from 2000 to 2003, and an associate editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 1999 to 2000. Among the awards he has received for research and teaching are the IEEE Browder J. Thompson Best Paper Award, the National Science Foundation CAREER Award, and the Presidential Early Career Award for Scientists and Engineers (PECASE).