Optimal Linear Cooperation for Spectrum Sensing in Cognitive Radio Networks

Zhi Quan, *Student Member, IEEE*, Shuguang Cui, *Member, IEEE*, and Ali H. Sayed, *Fellow, IEEE*

***Abstract*—Cognitive radio technology has been proposed to improve spectrum efficiency by having the cognitive radios act as secondary users to opportunistically access under-utilized frequency bands. Spectrum sensing, as a key enabling functionality in cognitive radio networks, needs to reliably detect signals from licensed primary radios to avoid harmful interference. However, due to the effects of channel fading/shadowing, individual cognitive radios may not be able to reliably detect the existence of a primary radio. In this paper, we propose an optimal linear cooperation framework for spectrum sensing in order to accurately detect the weak primary signal. Within this framework, spectrum sensing is based on the linear combination of local statistics from individual cognitive radios. Our objective is to minimize the interference to the primary radio while meeting the requirement of opportunistic spectrum utilization. We formulate the sensing problem as a nonlinear optimization problem. By exploiting the inherent structures in the problem formulation, we develop efficient algorithms to solve for the optimal solutions. To further reduce the computational complexity and obtain solutions for more general cases, we finally propose a heuristic approach, where we instead optimize a modified *deflection* coefficient that characterizes the probability distribution function of the global test statistics at the fusion center. Simulation results illustrate significant cooperative gain achieved by the proposed strategies. The insights obtained in this paper are useful for the design of optimal spectrum sensing in cognitive radio networks.**

***Index Terms*—Cognitive radio, cooperative communications, energy detection, nonlinear optimization, spectrum sensing.**

I. Introduction

COGNITIVE radios [2] have emerged as a potential technology to revolutionize spectrum utilization. According to the Federal Communications Commission (FCC) [3], cognitive radios (CR) are defined as radio systems that continuously perform spectrum sensing, dynamically identify unused spectrum, and then operate in those spectrum holes where the licensed (primary) radio systems are idle. This new communication paradigm can dramatically enhance spectrum efficiency, and is also referred to as the neXt Generation (XG) or Dynamic Spectrum Access (DSA) network.

Manuscript received April 14, 2007; revised October 29, 2007. This work was supported in part by the NSF under Grants ECS-0601266, ECS-0725441, CNS-0627118, and CNS-0721935, and by DoD under Grant HDTRA-07-1-0037. This work was presented in part at the IEEE Globecom Conference, Washington, DC, November 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Akbar Sayeed.

Z. Quan and A. H. Sayed are with the Electrical Engineering Department at University of California, Los Angeles, CA 90095 USA (e-mail: quan@ee.ucla.edu; sayed@ee.ucla.edu).

S. Cui is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: cui@ece.tamu.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.2007.914882

Spectrum sensing, as a key enabling functionality in cognitive radio networks, needs to reliably detect weak primary radio (PR) signals of possibly-unknown types [4]. Spectrum sensing should also monitor the activation of primary users in order for the secondary users to vacate the occupied spectrum segments. However, it is difficult for a cognitive radio to capture such information instantaneously due to the absence of cooperation between the primary and secondary users. Thus, recent research efforts on spectrum sensing have focused on the detection of ongoing primary transmissions by cognitive radio devices. Generally, spectrum sensing techniques fall into three categories: energy detection [5], coherent detection [6], and cyclostationary feature detection [7]. If the secondary user has limited information on the primary signals (e.g., only the local noise power is known), then the energy detector is optimal [8]. When certain primary signal features are known to the CRs (such as pilots, preambles, or synchronization messages), the optimal detector usually applies the matched filter structure to maximize the probability of detection. On the other hand, cyclostationary feature detectors differentiate the primary signal energy from the local noise energy by exploiting certain periodicity exhibited by the mean and autocorrelation of a particular modulated signal. In this paper, we assume that the primary signaling is unknown and we adopt energy detection as the building block for the proposed cooperative spectrum sensing scheme.

The detection performance of spectrum sensing schemes is usually compromised by destructive channel conditions between the target-under-detection and the cognitive radios, since it is hard to distinguish between a white spectrum and a weak signal attenuated by deep fading. In order to improve the reliability of spectrum sensing, radio cooperation exploiting spatial diversity among secondary users has been proposed in [4] and [9]. In such scenarios, a network of cooperative cognitive radios, which experience different channel conditions from the target, would have a better chance of detecting the primary radio if they combine the sensing information jointly. In other words, cooperative spectrum sensing can alleviate the problem of corrupted detection by exploiting spatial diversity, and thus reduce the probability of interfering with primary users. Since cooperative sensing is generally coordinated over a separate control channel, efficient cooperation schemes should be designed to reduce bandwidth and power requirements while maximizing the sensing reliability.

A. Prior Work

Although distributed detection has a rich literature (see, e.g., [10], [11] and the references therein), the results might not be directly applicable to cognitive radios, and the study of cooperative spectrum sensing for cognitive radio networks is rather limited. The scheme based on voting rules [12] is one of the simplest suboptimal solutions, which counts the number of sensor nodes that vote for the presence of the signal and compares it against a given threshold. In [13], a fusion rule known as the OR logic operation was used to combine decisions from several secondary users. In [14], two decision-combining approaches were studied: hard decision with the AND logic operation and soft decision using the likelihood ratio test (LRT) [11]. It was shown that the soft decision combination of spectrum sensing results yields gains over hard decision combining. In [15], the authors exploited the fact that summing signals from two secondary users can increase the signal-to-noise ratio (SNR) and detection reliability if the signals are correlated. This cooperative method is different from those discussed in [12]–[14], since it requires a wide-band control channel.

B. Contribution

In this paper, we develop an efficient linear cooperation framework for spectrum sensing, where the global decision is based on simple energy detection over a linear combination of the local statistics from individual nodes. Although the LRT-based optimal fusion rule [5] involves a quadratic form as in (26), the use of a linear detector in this paper is motivated by several considerations.

- First, the proposed linear detector has less computational complexity than does a quadratic detector, and the difference becomes more significant as the number of nodes increases.
- Second, the probabilities of detection and false alarm based on the linear detector have closed-form solutions, which could lead to intuitive system design guidelines. On the other hand, the performance evaluation and threshold computation of the quadratic detector are mathematically more intractable, since the computation involves many integrals. Thus, one has to turn generally to Monte Carlo simulations in studying the LRT-based quadratic detectors. With the linear detector, the designer can use the closed-form expressions to make quick adaptations when some network parameters change during the operation.
- Unlike the LRT-based detector for which the simulation complexity becomes prohibitively high at a low probability of false alarm and a high probability of detection, the optimal linear detector has a fixed complexity for any chosen probabilities of false alarm and detection due to its closed-form solutions.
- The optimal linear detection provides performance comparable to that achievable by the optimal LRT-based fusion in many situations, at least for the set of parameters used in this paper.

Note that even the LRT-based fusion structure (over analog forwarded local statistics) is only an optimal fusion rule, but not an optimal system-wide distributed detection rule. It has been shown in [16] that for a distributed detection problem with nonideal communication channels between distributed nodes and the fusion center, the globally optimal structure is to perform LRT both at individual nodes and at the fusion center. However, how to efficiently find the optimal LRT thresholds for individual nodes and for the fusion center is still unknown. As such, in this paper we propose a simple but efficient distributed detection scheme, which does not need to find optimal thresholds for individual nodes. Instead, we transmit the local test statistics to the fusion center, in which we conduct linear combination and simple energy detection. By doing so, the optimal threshold at the fusion center can be jointly determined with the optimal linear combining weights.

Our objective is to maximize the probability of detection (P_d) while satisfying a requirement on the probability of false alarm (P_f). In cognitive radio networks, a larger P_d leads to less interference to primary radios and a smaller P_f results in higher spectrum efficiency. This interpretation is based on the assumption that if a primary signal is detected (possibly a false alarm), cognitive radios are restrained to use the channel (such that spectrum is wasted in case of false alarms); if no primary signals are detected, cognitive radios use the channel (such that interference is generated in case of miss-detection). To achieve the above goal, we first derive bounds on the probability of detection for a given probability of false alarm. In order to find the optimal weights that achieve the maximum probability of detection, we define three classes of spectrum sensing schemes for cognitive radio networks based on their keenness towards opportunistic spectrum usage: *conservative*, *aggressive*, and *hostile*, with respect to different target values of P_f and P_d. For the *conservative* system with a high P_f target, we develop an efficient algorithm to search for the optimal combining weights within the derived bounds, by solving the dual problems of a sequence of quadratic constraint quadratic programs (QCQP). For the *aggressive* cognitive radio system with a medium P_f target, finding the exact optimal weights is transformed into a convex optimization problem, which can be solved using known algorithms [17]. If the system is *hostile* (with extremely low P_f target), we will show that the bounds obtained in this paper are tight enough to approximate the optimal operating point, although there might not exist an efficient method to solve for the optimal solution.

Furthermore, we propose a heuristic approach to control the combining weights, which optimizes a modified *deflection* coefficient (MDC) that characterizes the probability distribution function (PDF) of the global test statistic at the fusion center. This approach slightly compromises the detection performance with less computational complexity and provides near-optimal solutions for general systems (i.e., the same model applies to conservative, aggressive, and hostile systems).

These optimized cooperation schemes improve the sensing reliability while relaxing the harsh requirements on the RF front-end sensitivity and signal processing gain at individual CR nodes. Simulation results illustrate that the proposed cooperation schemes achieve superior sensing performance.

The rest of the paper is organized as follows. In Section II, we introduce the system model and notation. Section III describes the cooperative spectrum sensing approach in details. In Section IV, we first derive bounds on the probability of detection for different cognitive scenarios. Based on these bounds, we develop efficient algorithms to find the optimal weights for the conservative and aggressive systems. In addition, a heuristic

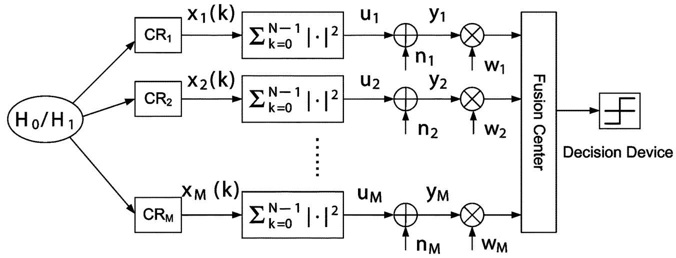

Fig. 1. Schematic representation of weighting cooperation for spectrum sensing in cognitive radio networks.

approach based on the MDC maximization is proposed to reduce the computational complexity and provide general solutions. Simulation results illustrating the effectiveness of the proposed approaches are given in Section V. Section VI concludes the paper.

II. Model and Notation

A. System Model

We consider a cognitive radio network with M secondary users. The binary hypothesis test for spectrum sensing at the kth time instant is formulated as follows:

$$\begin{aligned} \mathcal{H}_0 &: x_i(k) = v_i(k) \quad i = 1, 2, \ldots, M \\ \mathcal{H}_1 &: x_i(k) = h_i s(k) + v_i(k) \quad i = 1, 2, \ldots, M \end{aligned} \tag{1}$$

where $s(k)$ denotes the signal transmitted by the primary user and $x_i(k)$ is the received signal by the ith secondary user. The signal $s(k)$ is distorted by the channel gain h_i, which is assumed to be constant during the detection interval, and is further corrupted by the zero-mean additive white Gaussian noise (AWGN) $v_i(k)$, i.e., $v_i(k) \sim \mathcal{CN}\left(0, \sigma_i^2\right)$. We call $\{v_i(k)\}$ the sensing noises and collect their variances into a vector $\boldsymbol{\sigma} = \left[\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2\right]^T$. Without loss of generality, $s(k)$ and $\{v_i(k)\}$ are assumed to be independent of each other.

As illustrated in Fig. 1, each secondary user calculates a summary statistic u_i over a detection interval of N samples, i.e.,

$$u_i = \sum_{k=0}^{N-1} |x_i(k)|^2 \quad i = 1, 2, \ldots, M \tag{2}$$

where N is determined from the time-bandwidth product. The summary statistics $\{u_i\}$ are then transmitted to the fusion center through a control channel in an orthogonal manner, and represented as

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix}}_{\mathbf{u}} + \underbrace{\begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_M \end{bmatrix}}_{\mathbf{n}} \tag{3}$$

where the channel noises $\{n_i\}$ are zero-mean, spatially uncorrelated Gaussian variables with variances $\left\{\delta_i^2\right\}$; the variances are collected into the vector form $\boldsymbol{\delta} = \left[\delta_1^2, \delta_2^2, \ldots, \delta_M^2\right]^T$. The use of the AWGN channel model in (3) is justified by assumptions on analog-forwarding schemes and the slow-changing nature of the channels between the secondary users and the fusion center. We assume that the channel coherence time is much larger than the estimation period such that once the fusion center has estimated the channel gains from the secondary users, these channels could be treated as constant AWGN channels.

The fusion center computes a global test statistic, y_c as in (13), from the outputs $\{y_i\}$ of the individual secondary users in a linear manner; and y_c is then used by the spectrum sensor to make a global decision. The main purpose of this paper is to design the optimal linear fusion rules at the fusion center in order to maximize the detection sensitivity while meeting a given requirement on the probability of false alarm.

B. Notation

We define the following notations, which will be used in the paper:

$Q(\cdot)$	*complementary cumulative distribution* function, which calculates the tail probability of a zero mean unit variance Gaussian variable, i.e., $Q(x) = \int_x^{+\infty} \exp(-t^2/2) dt / \sqrt{2\pi}$;
$\|\cdot\|$	Euclidean norm of a vector;
$\mathrm{diag}(\cdot)$	square diagonal matrix with the elements of a given vector on the diagonal;
$\lambda(\cdot)$	eigenvalues of a matrix; specifically, we use $\lambda_{\max}$ and $\lambda_{\min}$ to represent the maximum and minimum eigenvalues of a given matrix, respectively;
$\succeq$	matrix inequality, i.e., $\mathbf{A} \succeq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semi-definite; for a vector, it represents the component-wise inequality, with '$\succ$' denoting the strict inequality;
$\dagger$	Moore-Penrose generalized inverse;
$\mathcal{R}(\cdot)$	range space of a given matrix.

III. Cooperative Spectrum Sensing

In this section, we develop a linear cooperation scheme for spectrum sensing. In particular, we adopt energy detection (i.e., radiometry) as the local sensing rule, which will be explained as follows.

A. Local Sensing

We first consider local spectrum sensing at individual secondary users. The test statistic of the ith secondary user using energy detection is given by (2). Since u_i is the sum of the squares of N Gaussian random variables, it can be shown that u_i/σ_i^2 follows a central chi-square χ^2 distribution with N degrees of freedom if $\mathcal{H}_0$ is true; otherwise, it would follow a noncentral χ^2 distribution with N degrees of freedom and parameter η_i. That is,

$$\frac{u_i}{\sigma_i^2} \sim \begin{cases} \chi_N^2 & \mathcal{H}_0 \\ \chi_N^2(\eta_i) & \mathcal{H}_1 \end{cases} \tag{4}$$

where

$$\eta_i = \frac{E_s|h_i|^2}{\sigma_i^2} \tag{5}$$

is the local SNR at the ith secondary user and the quantity

$$E_s = \sum_{k=0}^{N-1} |s(k)|^2 \tag{6}$$

represents the transmitted signal energy over a sequence of N samples during each detection interval. Please note that the so-defined local SNR is N times the average SNR at the output of the local energy detector, which should be equal to $E_s|h_i|^2/N\sigma_i^2$.

According to the central limit theorem [18], if the number of samples N is large enough (e.g., ≥ 10 in practice), the test statistics $\{u_i\}$ are asymptotically normally distributed with mean

$$\mathrm{E}u_i = \begin{cases} N\sigma_i^2 & \mathcal{H}_0 \\ (N+\eta_i)\,\sigma_i^2 & \mathcal{H}_1 \end{cases} \tag{7}$$

and variance

$$\mathrm{Var}\,(u_i) = \begin{cases} 2N\sigma_i^4 & \mathcal{H}_0 \\ 2(N+2\eta_i)\sigma_i^4 & \mathcal{H}_1 \end{cases} \tag{8}$$

which can be compactly represented as $u_i \sim \mathcal{N}\left[\mathrm{E}u_i, \mathrm{Var}\,(u_i)\right]$ for N large enough.

For a single-CR spectrum sensing scheme, the decision rule at each secondary user is given by

$$u_i \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma_i \quad i = 1, 2, \ldots, M \tag{9}$$

where γ_i is the corresponding decision threshold. Therefore, secondary user i will have the following probabilities of false alarm and detection:

$$P_f^{(i)} = \Pr\left(u_i > \gamma_i|\mathcal{H}_0\right) = Q\left[\frac{\gamma_i - \mathrm{E}\left(u_i|\mathcal{H}_0\right)}{\sqrt{\mathrm{Var}\left(u_i|\mathcal{H}_0\right)}}\right] \tag{10}$$

and

$$P_d^{(i)} = \Pr\left(u_i > \gamma_i|\mathcal{H}_1\right) = Q\left[\frac{\gamma_i - \mathrm{E}\left(u_i|\mathcal{H}_1\right)}{\sqrt{\mathrm{Var}\left(u_i|\mathcal{H}_1\right)}}\right]. \tag{11}$$

As we see from the above, spectrum sensing with a single CR is quite simple, but may suffer from destructive channel effects such as fading or shadowing. Therefore, there is a necessity that several CRs cooperate with each other to jointly detect the existence of the primary transmission in order to improve the inference accuracy.

B. Global Decision

To allow multiple secondary users to collaborate, we transmit the test statistics $\{u_i\}$ directly to the fusion center via a dedicated control channel. According to (3), the received statistics $\{y_i\}$ are normally distributed with means $\mathrm{E}y_i = \mathrm{E}u_i$ and variances

$$\mathrm{Var}\,(y_i) = \begin{cases} 2N\sigma_i^4 + \delta_i^2 & \mathcal{H}_0 \\ 2(N+2\eta_i)\sigma_i^4 + \delta_i^2 & \mathcal{H}_1. \end{cases} \tag{12}$$

Once the fusion center receives $\{y_i\}$, a global test statistic is calculated linearly as follows:

$$y_c = \sum_{i=1}^{M} w_i y_i = \mathbf{w}^T\mathbf{y} \tag{13}$$

where

$$\mathbf{w} \triangleq \left[w_1, w_2, \ldots, w_M\right]^T, \quad w_i \geq 0 \tag{14}$$

is the weight vector used to control the global spectrum detector. The combining weight for the signal from a particular user represents its contribution to the global decision. For example, if a CR generates a high-SNR signal that may lead to correct detection on its own, it should be assigned a larger weighting coefficient. For those secondary users experiencing deep fading or shadowing, their weights are decreased in order to reduce their negative contribution to the decision fusion.

Since the $\{y_i\}$ are normal random variables, their linear combination is also normal. Consequently, y_c has mean

$$\bar{y}_c = \mathrm{E}y_c = \begin{cases} N\boldsymbol{\sigma}^T\mathbf{w} & \mathcal{H}_0 \\ \left(N\boldsymbol{\sigma} + E_s\boldsymbol{g}\right)^T\mathbf{w} & \mathcal{H}_1 \end{cases} \tag{15}$$

where

$$\boldsymbol{g} \triangleq \left[|h_1|^2, |h_2|^2, \ldots, |h_M|^2\right]^T \tag{16}$$

represent the squared amplitudes of the channel gains, and variance

$$\begin{aligned}\mathrm{Var}\,(y_c) &= \mathrm{E}\left(y_c - \bar{y}_c\right)^2 \\ &= \mathbf{w}^T\mathrm{E}\left[\left(\mathbf{y}-\bar{\mathbf{y}}\right)\left(\mathbf{y}-\bar{\mathbf{y}}\right)^T\right]\mathbf{w}.\end{aligned} \tag{17}$$

The variances under different hypotheses are respectively given by

$$\begin{aligned}\mathrm{Var}\,(y_c|\mathcal{H}_0) &= \mathbf{w}^T\mathrm{E}\left[\left(\mathbf{y}-\bar{\mathbf{y}}_{\mathcal{H}_0}\right)\left(\mathbf{y}-\bar{\mathbf{y}}_{\mathcal{H}_0}\right)^T|\mathcal{H}_0\right]\mathbf{w} \\ &= \sum_{i=1}^{M}\left(2N\sigma_i^4 + \delta_i^2\right)w_i^2 \\ &= \mathbf{w}^T\boldsymbol{\Sigma}_{\mathcal{H}_0}\mathbf{w}\end{aligned} \tag{18}$$

with

$$\boldsymbol{\Sigma}_{\mathcal{H}_0} \triangleq 2N\mathrm{diag}^2(\boldsymbol{\sigma}) + \mathrm{diag}(\boldsymbol{\delta}) \tag{19}$$

and

$$\begin{aligned}\mathrm{Var}\,(y_c|\mathcal{H}_1) &= \mathbf{w}^T\mathrm{E}\left[\left(\mathbf{y}-\bar{\mathbf{y}}_{\mathcal{H}_1}\right)\left(\mathbf{y}-\bar{\mathbf{y}}_{\mathcal{H}_1}\right)^T|\mathcal{H}_1\right]\mathbf{w} \\ &= \sum_{i=1}^{M}\left(2N\sigma_i^4 + 4\eta_i\sigma_i^4 + \delta_i^2\right)w_i^2 \\ &= \mathbf{w}^T\boldsymbol{\Sigma}_{\mathcal{H}_1}\mathbf{w}\end{aligned} \tag{20}$$

with

$$\boldsymbol{\Sigma}_{\mathcal{H}_1} \triangleq 2N\mathrm{diag}^2(\boldsymbol{\sigma}) + \mathrm{diag}(\boldsymbol{\delta}) + 4E_s\mathrm{diag}\left(\boldsymbol{g}\right)\mathrm{diag}(\boldsymbol{\sigma}). \tag{21}$$

Since $\mathbf{\Sigma}_{\mathcal{H}_1}$ is positive semi-definite and diagonal, its square root can be represented as

$$\mathbf{\Sigma}_{\mathcal{H}_1}^{1/2} = \mathrm{diag}\begin{bmatrix} \sqrt{2N\sigma_1^4 + 4E_s|h_1|^2\sigma_1^2 + \delta_1^2} \\ \sqrt{2N\sigma_2^4 + 4E_s|h_2|^2\sigma_2^2 + \delta_2^2} \vdots \\ \sqrt{2N\sigma_M^4 + 4E_s|h_M|^2\sigma_M^2 + \delta_M^2} \end{bmatrix} \quad (22)$$

which will be used throughout the rest of this paper. We would like to point out that the received variables $\{y_i\}$ do not have to be conditionally independent though here we use the independent case for illustration purpose, i.e., with $\mathbf{\Sigma}_{\mathcal{H}_0}$ and $\mathbf{\Sigma}_{\mathcal{H}_1}$ diagonal. If the received variables $\{y_i\}$ are correlated with each other, then the covariance matrices $\mathbf{\Sigma}_{\mathcal{H}_0}$ and $\mathbf{\Sigma}_{\mathcal{H}_1}$ are generally non-diagonal and $\mathbf{\Sigma}_{\mathcal{H}_1}^{1/2}$ can be chosen as the square root obtained from the Cholesky decomposition [19] and the subsequent analysis will continue to hold.

Considering the linear rule at the fusion center with a test threshold γ_c, we have

$$y_c \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma_c. \quad (23)$$

As such, the performance of the proposed cooperative spectrum detection scheme can be evaluated as

$$P_f = Q\left[\frac{\gamma_c - N\boldsymbol{\sigma}^T\mathbf{w}}{\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{w}}}\right] \quad (24)$$

and

$$P_d = Q\left[\frac{\gamma_c - (N\boldsymbol{\sigma} + E_s\boldsymbol{g})^T\mathbf{w}}{\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{w}}}\right]. \quad (25)$$

We see that the sensing performance of the linear detector depends largely on the weighting coefficients and the decision threshold. We next show how to design the optimal weight vector $\mathbf{w}$ and the optimal decision threshold γ_c in order to maximize the sensing sensitivity under certain requirement on the probability of false alarm.

On the other hand, the LRT-based fusion rule [5] has a quadratic form given by

$$y_{\mathrm{LRT}} = \mathbf{y}^T\left(\mathbf{\Sigma}_{\mathcal{H}_0}^{-1} - \mathbf{\Sigma}_{\mathcal{H}_1}^{-1}\right)\mathbf{y} + 2\left[(N\boldsymbol{\sigma} + E_s\boldsymbol{g})^T\mathbf{\Sigma}_{\mathcal{H}_1}^{-1} - N\boldsymbol{\sigma}^T\mathbf{\Sigma}_{\mathcal{H}_0}^{-1}\right]\mathbf{y} \quad (26)$$

which is difficult to numerically evaluate since finding the probability distribution of y_{LRT} involves many integrals. Thus, the optimal threshold for y_{LRT} is not mathematically tractable.

Compared with other detection approaches such as feature and coherent detection, the energy-detection based model (23) is more suitable for cognitive radio networks because it requires minimum *a priori* knowledge about primary users. Due to the lack of collaboration between the primary and cognitive users, it is difficult for cognitive radios to obtain the exact channel gains $\{h_i\}_{i=1}^M$. In our system model, to evaluate the two system parameter matrices $\mathbf{\Sigma}_{\mathcal{H}_0}$ and $\mathbf{\Sigma}_{\mathcal{H}_1}$, we only need to know *a priori* the local noise variance at each cognitive radio, i.e., $\{\sigma_i^2\}_{i=1}^M$. The proposed cooperation scheme uses the local SNR at each cognitive radio, i.e., $\{\eta_i\}_{i=1}^M$, instead of the exact channel gains h_i's (i.e., amplitudes and phases) between the primary and cognitive users. The quantities $\{\eta_i\}_{i=1}^N$ can be estimated in practice with the assumption that the channel coherence time is large enough. For instance, if the local noise levels $\{\sigma_i^2\}$ are known *a priori*, possibly from experimental measurements when the primary system is turned off or from some previous experience, then the received primary signal power $E_s|h_i|^2/N$ can be calculated as the total power at the RF front-end minus the noise power. Furthermore, the exact channel power gains $\{|h_i|^2\}_{i=1}^M$ can be estimated if the secondary users know the primary transmit power. Such information is indeed obtainable in some circumstances. For example, the main target spectrum of current IEEE standardization activities (i.e., IEEE 802.22) for cognitive radio technologies is in TV bands. In such a case, it is possible for secondary users to have *a priori* information about the primary signal power, since most current TV stations transmit at fixed power levels. In addition, this mechanism is also applicable to the downlinks in certain cellular networks, where base stations periodically transmit pilot signals at known power levels.

IV. Performance Optimization

For cognitive radio networks, the probabilities of false alarm and detection have unique implications. Specifically, $1 - P_d$ measures the probability of interference from secondary users on the primary users. On the other hand, P_f determines an upper bound on the spectrum efficiency, where a large P_f usually results in low spectrum utilization. This is based on a typical assumption that if primary signals are detected, the secondary users do not use the corresponding channel, and if no primary signals are detected, the secondary users use the corresponding channel. In this section, we maximize the P_d by controlling the weight vector while meeting a certain requirement on the P_f. Before we proceed, we define three classes of cognitive radio systems in terms of their keenness to use the targeted frequency bands.

Definition 1 (Conservative System): A *conservative* CR system has an opportunistic spectrum utilization rate less than or equal to 50% and a probability of interference less than 1/2. That is, the targeted probability of false alarm satisfies $P_f \geq 1/2$ and the probability of detection has $P_d > 1/2$.

Definition 2 (Aggressive System): An *aggressive* CR system expects to achieve more than 50% opportunistic spectrum utilization while maintaining less than 50% probability to interfere with the primary radio. This corresponds to a targeted probability of false alarm $P_f < 1/2$ and a probability of detection $P_d > 1/2$.

Definition 3 (Hostile System): A *hostile* CR system targets more than 50% opportunistic spectrum utilization and is likely to cause interference to the primary radio with a probability greater than or equal to 50%. This system has a targeted false alarm probability $P_f < 1/2$ and a probability of detection $P_d \leq 1/2$.

A. Universal Bounds on the Probability of Detection

Given a targeted probability of false alarm P_f, we would like to maximize the probability of detection (25). From (24), the test threshold in terms of the target P_f is given by

$$\gamma_c = N\boldsymbol{\sigma}^T\mathbf{w} + Q^{-1}(P_f)\sqrt{\mathbf{w}^T\boldsymbol{\Sigma}_{\mathcal{H}_0}\mathbf{w}}. \tag{27}$$

Substituting (27) into (25), we get

$$P_d = Q\left[\frac{Q^{-1}(P_f)\sqrt{\mathbf{w}^T\boldsymbol{\Sigma}_{\mathcal{H}_0}\mathbf{w}} - E_s\boldsymbol{g}^T\mathbf{w}}{\sqrt{\mathbf{w}^T\boldsymbol{\Sigma}_{\mathcal{H}_1}\mathbf{w}}}\right]. \tag{28}$$

As such, to maximize P_d with a given P_f, the optimization of the detection threshold γ_c is integrated into the optimization of $\mathbf{w}$. However, finding the exact maximum probability of detection is generally difficult since the function in (28) is not concave. Nevertheless, we can first derive some bounds on P_d for a given P_f as follows.

Lemma 1 (Bounds for Aggressive and Hostile Systems): For an aggressive or hostile sensing system that has a targeted probability of false alarm $P_f < 1/2$, the probability of detection can be bounded as

$$Q[\phi_a(P_f)] \le P_d \le Q[\psi_a(P_f)] \tag{29}$$

where

$$\phi_a(P_f) = Q^{-1}(P_f)\lambda_{\max}^{1/2}\left(\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{\Sigma}_{\mathcal{H}_0}\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right) \tag{30}$$

and

$$\psi_a(P_f) = Q^{-1}(P_f)\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{\Sigma}_{\mathcal{H}_0}\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right) - E_s\left\|\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\right\|. \tag{31}$$

Proof: Refer to Appendix A. ∎

These bounds become quite tight when the probability of false alarm P_f is small, i.e.,

$$Q[\psi_a(P_f)] - Q[\phi_a(P_f)] \to 0 \text{ as } P_f \to 0. \tag{32}$$

Although the asymptotic result itself is not informative since P_d decreases to zero as P_f gets close to zero for any rationally designed detectors, it is interesting to investigate how fast the bound gap shrinks as P_f decreases. As such, we can at least predict how tight the bound is when P_f is reasonably small. We will show that the gap between the upper and lower bounds decreases exponentially with respect to $Q^{-1}(P_f)$ in the asymptotic sense. This can be shown by using the *large deviation* theory [20]. In particular, the Q-function has the following asymptotic log-similarity property [21]:

$$Q(x) \overset{\log}{\sim} \exp(-\xi x). \tag{33}$$

Here, we use $f(x) \overset{\log}{\sim} g(x)$ if $\log f(x) \sim \log g(x)$, where $f(x) \sim g(x)$ means $\lim_{x\to\infty} f(x)/g(x) = 1$. The positive constant ξ is typically called the *asymptotic decay rate*. This result leads to the well known approximation (e.g., see [22] and references therein)

$$Q(x) \approx e^{-\xi x} \tag{34}$$

for large values of x. Applying this approximation to the bounds in (29) for small values of P_f (large values of $Q^{-1}(P_f)$), we have

$$\begin{aligned} &Q[\psi_a(P_f)] - Q(\phi_a(P_f)) \\ &\quad\approx e^{-\xi\psi_a(P_f)} - e^{-\xi\phi_a(P_f)} \\ &\quad= e^{\xi E_s\|\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\|}e^{-\xi Q^{-1}(P_f)\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{\Sigma}_{\mathcal{H}_0}\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right)} \\ &\qquad - e^{-\xi Q^{-1}(P_f)\lambda_{\max}^{1/2}\left(\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{\Sigma}_{\mathcal{H}_0}\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right)} \\ &\quad\le e^{\xi E_s\|\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\|}e^{-\xi\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{\Sigma}_{\mathcal{H}_0}\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right)Q^{-1}(P_f)} \end{aligned} \tag{35}$$

where the last inequality follows from the fact that $\xi\lambda_{\max}^{1/2}\left(\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{\Sigma}_{\mathcal{H}_0}\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right) > 0$ and $Q^{-1}(P_f) > 0$. It can be seen that the gap between the upper and lower bounds decays at least exponentially with respect to $Q^{-1}(P_f)$ as P_f gets close to zero.

Remark 1: For hostile systems with the probability of false alarm (P_f) set small (close to zero) in order for the secondary users to occupy the targeted frequency bands more frequently, both bounds in (29) would provide good approximations to the probability of detection (P_d), and the choice of weight vectors has negligible influence on detection performance. As such, in this paper, we mainly focus on optimizing conservative and aggressive cognitive systems.

Lemma 2 (Bounds for Conservative Systems): For a conservative system with the targeted probability of false alarm $P_f \ge 1/2$, the lower and upper bounds on the probability of detection are given by

$$Q[\phi_c(P_f)] \le P_d \le \min\{Q[\omega_c(P_f)], Q[\psi_c(P_f)]\} \tag{36}$$

where

$$\psi_c(P_f) = Q^{-1}(P_f)\lambda_{\max}^{1/2}\left(\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{\Sigma}_{\mathcal{H}_0}\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right) - E_s\left\|\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\right\| \tag{37}$$

$$\omega_c(P_f) = -\sqrt{E_s^2 + Q^{-2}(P_f)} \times \lambda_{\max}^{1/2}\left[\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\left(\boldsymbol{g}\boldsymbol{g}^T + \boldsymbol{\Sigma}_{\mathcal{H}_0}\right)\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right] \tag{38}$$

and

$$\phi_c(P_f) = Q^{-1}(P_f)\lambda_{\min}^{1/2}\left(\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{\Sigma}_{\mathcal{H}_0}\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right). \tag{39}$$

Proof: See Appendix B. ∎

We would like to point out that the universal lower bounds, $Q[\phi_a(P_f)]$ and $Q[\phi_c(P_f)]$, are the worst-case probability of detection, because any rational combining weight vector will result in a larger probability of detection given the probability of false alarm. Nevertheless, they provide insight on what the worst performance would be if the weight vector is not appropriately chosen. In Section IV-C, we will present a tighter lower bound on the maximum probability of detection, which can be achieved by an easily found weight vector.

B. Maximum Probability of Detection

In the following, we will show how to find the optimal weight vector $\mathbf{w}_{\rm opt}$ that maximizes (28), where the optimal detection threshold will be defined according to (27). Define the function

$$f(\mathbf{w}) = \frac{Q^{-1}(P_f)\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{w}} - E_s\boldsymbol{g}^T\mathbf{w}}{\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{w}}}. \quad (40)$$

Since the Q-function in (28) is a monotonically decreasing function, maximizing P_d is equivalent to minimizing $f(\mathbf{w})$. Therefore, we formulate the following optimization problem:

$$\min_{\mathbf{w}} f\ (\mathbf{w}) \text{(P1)}$$

whose optimal solution is denoted by $\mathbf{w}_1^o$, which is also the optimal solution for maximizing the probability of detection given a fixed probability of false alarm. Note that any scaled version of $\mathbf{w}_1^o$ is also an optimal solution given the special structure of $f(\mathbf{w})$ with a fixed P_f. Later we will discuss how to solve for the optimal solutions and the one with unit norm will be taken as the final solution.

Remark 2: Since at the fusion center we collect local statistics based on energy detection and then do a joint energy detection, the combining weight coefficients should be nonnegative. Hence, problem (P1) has an implicit constraint that $\mathbf{w} \succeq \mathbf{0}$. It does not need to be explicitly included since the numerator in the objective function, $Q^{-1}(P_f)\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{w}} - E_s\boldsymbol{g}^T\mathbf{w}$, has $E_s > 0$ and $\boldsymbol{g} \succeq \mathbf{0}$ such that positive elements in $\mathbf{w}$ are always better than negative choices in term of minimizing the objective function in (P1).

Directly solving (P1) for general cases is challenging. As such, we apply a divide-and-conquer strategy to solve the problem for the conservative and aggressive systems individually, and then combine the solutions together for the original problem.

1) Conservative Systems: For a conservative system where $Q^{-1}(P_f) \le 0$, we have

$$E_s\boldsymbol{g}^T\mathbf{w} - Q^{-1}(P_f)\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{w}} \ge 0. \quad (41)$$

Thus, problem (P1) has the equivalent form

$$\begin{aligned} \max_{w_0,\mathbf{w}} \quad & \frac{E_s\boldsymbol{g}^T\mathbf{w} - Q^{-1}(P_f)\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{w}}}{w_0} \quad \text{(P2)} \\ \text{st.} \quad & \mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{w} \le w_0^2 \\ & w_0 > 0 \end{aligned}$$

which can be written as follows by changing the optimization variables to $\mathbf{z} = \mathbf{w}/w_0$

$$\begin{aligned} \min_{\mathbf{z}} \quad & Q^{-1}(P_f)\sqrt{\mathbf{z}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{z}} - E_s\boldsymbol{g}^T\mathbf{z} \quad \text{(P3)} \\ \text{st.} \quad & \mathbf{z}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{z} \le 1 \end{aligned}$$

where we denote the optimal objective value and the optimal solution by p_3^o and z_3^o, respectively. Note that solving over the new variable $\mathbf{z}$ is equivalent to solving a scaled version of the original design variable $\mathbf{w}$, where the norm of $\mathbf{z}$ is bounded by the constraint in (P3). As such, there will be a unique solution for z_3^o, which can map back to an infinite number of optimal solutions in terms of $\mathbf{w}$ (any scaled version of the optimal $\mathbf{w}$ is still optimal). For convenience, we can take the optimal $\mathbf{w}$ with unit norm as the optimal solution for the original problem.

Since it can be shown that $Q^{-1}(P_f)\sqrt{\mathbf{z}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{z}}$ is a concave function if $Q^{-1}(P_f) \le 0$, problem (P3) requires minimizing a concave function over an ellipsoid, which is unfortunately not a convex problem. To solve this problem, we can iteratively approximate the optimal value p_3^o by finding a tighter bound on p_3^o at each step.

Remark 3: For a conservative cognitive system where $Q^{-1}(P_f) \le 0$, we have $Q^{-1}(P_f)\sqrt{\mathbf{z}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{z}} - E_s\boldsymbol{g}^T\mathbf{z} < 0$. Thus, the optimal value p_3^o of (P3) should satisfy $p_3^o < 0$. Note that for any $\alpha < 0$

$$\begin{aligned} & Q^{-1}(P_f)\sqrt{\mathbf{z}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{z}} - E_s\boldsymbol{g}^T\mathbf{z} \le \alpha \Longleftrightarrow \\ & \qquad -Q^{-1}(P_f)\sqrt{\mathbf{z}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{z}} \ge -E_s\boldsymbol{g}^T\mathbf{z} - \alpha. \quad (42) \end{aligned}$$

If the following condition is satisfied:

$$-E_s\boldsymbol{g}^T\mathbf{z} - \alpha \ge 0 \quad (43)$$

then squaring both sides of (42) results in the quadratic inequality

$$\mathbf{z}^T\left[E_s^2\boldsymbol{g}\boldsymbol{g}^T - Q^{-2}(P_f)\mathbf{\Sigma}_{\mathcal{H}_0}\right]\mathbf{z} + 2\alpha E_s\boldsymbol{g}^T\mathbf{z} + \alpha^2 \le 0. \quad (44)$$

Therefore, if the problem

$$\begin{aligned} \text{find} \quad & \mathbf{z} \quad \text{(P4)} \\ \text{st.} \quad & \mathbf{z}^T\left[E_s^2\boldsymbol{g}\boldsymbol{g}^T - Q^{-2}(P_f)\mathbf{\Sigma}_{\mathcal{H}_0}\right]\mathbf{z} + 2\alpha E_s\boldsymbol{g}^T\mathbf{z} + \alpha^2 \le 0 \\ & \mathbf{z}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{z} \le 1 \end{aligned}$$

is feasible, then we have $p_3^o \le \alpha$. On the other hand, if (P4) is not feasible under the condition (43), then we can conclude $p_3^o > \alpha$.

Furthermore, the feasibility problem (P4) can be transformed into a quadratic constraint quadratic program as follows:

$$\begin{aligned} \min_{\mathbf{z}} \quad & \mathbf{z}^T\left[E_s^2\boldsymbol{g}\boldsymbol{g}^T - Q^{-2}(P_f)\mathbf{\Sigma}_{\mathcal{H}_0}\right]\mathbf{z} \\ & \quad + 2\alpha E_s\boldsymbol{g}^T\mathbf{z} + \alpha^2 \quad \text{(P5)} \\ \text{st.} \quad & \mathbf{z}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{z} \le 1 \end{aligned}$$

where both $E_s^2\boldsymbol{g}\boldsymbol{g}^T - Q^{-2}(P_f)\mathbf{\Sigma}_{\mathcal{H}_0}$ and $\mathbf{\Sigma}_{\mathcal{H}_1}$ are symmetric matrices. Let p_5^o and z_5^o denote the optimal value and the optimal solution of (P5), respectively. If $p_5^o \le 0$, then (P4) is feasible; otherwise, (P4) is not. Note that the matrix $E_s^2\boldsymbol{g}\boldsymbol{g}^T - Q^{-2}(P_f)\mathbf{\Sigma}_{\mathcal{H}_0}$ is indefinite, and hence, (P5) is not a convex optimization problem in general.

Remark 4: Since we have $\alpha < 0$, problem (P5) contains the implicit constraint $\mathbf{z} \succeq \mathbf{0}$ such that the objective function is minimized. Thus, the optimal solution $\mathbf{z}_5^o$ should be nonnegative.

To solve problem (P5), we can derive its dual problem. Specifically, the Lagrangian of (P5) is given by

$$L(\mathbf{z},\mu) = \mathbf{z}^T \left[E_s^2 \boldsymbol{g}\boldsymbol{g}^T - Q^{-2}\left(P_f\right)\boldsymbol{\Sigma}_{\mathcal{H}_0} + \mu\boldsymbol{\Sigma}_{\mathcal{H}_1}\right]\mathbf{z} + 2\alpha E_s \boldsymbol{g}^T\mathbf{z} + \alpha^2 - \mu \quad (45)$$

with $\mu \geq 0$, and its dual function is given by

$$q(\mu) = \inf_{\mathbf{z}} L(\mathbf{z},\mu). \quad (46)$$

In particular, if the following two conditions are satisfied:

$$E_s^2 \boldsymbol{g}\boldsymbol{g}^T - Q^{-2}\left(P_f\right)\boldsymbol{\Sigma}_{\mathcal{H}_0} + \mu\boldsymbol{\Sigma}_{\mathcal{H}_1} \succeq 0 \quad (47)$$

and

$$\boldsymbol{g} \in \mathcal{R}\left[E_s^2 \boldsymbol{g}\boldsymbol{g}^T - Q^{-2}\left(P_f\right)\boldsymbol{\Sigma}_{\mathcal{H}_0} + \mu\boldsymbol{\Sigma}_{\mathcal{H}_1}\right] \quad (48)$$

then

$$q(\mu) = -\alpha^2 E_s^2 \boldsymbol{g}^T \left[E_s^2 \boldsymbol{g}\boldsymbol{g}^T - Q^{-2}\left(P_f\right)\boldsymbol{\Sigma}_{\mathcal{H}_0} + \mu\boldsymbol{\Sigma}_{\mathcal{H}_1}\right]^{\dagger}\boldsymbol{g} + \alpha^2 - \mu. \quad (49)$$

Otherwise, $q(\mu) = -\infty$.

As a result, the dual problem can be expressed as

$$\begin{aligned} \max_{\mu} \quad & q(\mu) \quad (\text{P6}) \\ \text{st.} \quad & E_s^2 \boldsymbol{g}\boldsymbol{g}^T - Q^{-2}\left(P_f\right)\boldsymbol{\Sigma}_{\mathcal{H}_0} + \mu\boldsymbol{\Sigma}_{\mathcal{H}_1} \succeq 0 \\ & \boldsymbol{g} \in \mathcal{R}\left[E_s^2 \boldsymbol{g}\boldsymbol{g}^T - Q^{-2}\left(P_f\right)\boldsymbol{\Sigma}_{\mathcal{H}_0} + \mu\boldsymbol{\Sigma}_{\mathcal{H}_1}\right] \\ & \mu \geq 0. \end{aligned}$$

Using Schur's complement, we can further transform the dual problem into a semi-definite program (SDP)

$$\begin{aligned} \min_{\mu,\beta} \quad & \beta \quad (\text{P7}) \\ \text{st.} \quad & \begin{bmatrix} E_s^2 \boldsymbol{g}\boldsymbol{g}^T - Q^{-2}\left(P_f\right)\boldsymbol{\Sigma}_{\mathcal{H}_0} + \mu\boldsymbol{\Sigma}_{\mathcal{H}_1} & \boldsymbol{g} \\ \boldsymbol{g}^T & \frac{\beta-\mu}{\gamma} \end{bmatrix} \succeq 0 \end{aligned}$$

where $\gamma = \alpha^2 E_s^2$. This problem can be solved easily like a linear program for the optimal value β^o and the dual solution μ^o.

If the Slater's constraint qualification [17] is satisfied, i.e., if there exists a feasible vector $\mathbf{z}$ such that $\mathbf{z}^T\boldsymbol{\Sigma}_{\mathcal{H}_1}\mathbf{z} < 1$, then strong duality holds between problem (P5) and its Lagrange dual problem (P6) [23]. Moreover, if $p_5^o \leq 0$, then

$$\mathbf{z}_5^o = -\alpha E_s \left[E_s^2 \boldsymbol{g}\boldsymbol{g}^T - Q^{-2}\left(P_f\right)\boldsymbol{\Sigma}_{\mathcal{H}_0} + \mu^o\boldsymbol{\Sigma}_{\mathcal{H}_1}\right]^{\dagger}\boldsymbol{g} \quad (50)$$

is the optimal solution of (P5). Moreover, $\mathbf{z}_5^o$ is also the solution of (P4).

However, the condition (43) may not be satisfied by some α. If this is the case, then (42) can still hold if $-E_s\boldsymbol{g}^T\mathbf{z} - \alpha < 0$, i.e., the following problem is feasible:

$$\begin{aligned} \text{find} \quad & \mathbf{z} \quad (\text{P8}) \\ \text{st.} \quad & -E_s\boldsymbol{g}^T\mathbf{z} - \alpha < 0 \\ & \mathbf{z} \succeq \mathbf{0}. \end{aligned}$$

This problem can also be transformed into the following linear program

$$\begin{aligned} \min_{\mathbf{z}} \quad & -E_s\boldsymbol{g}^T\mathbf{z} \quad (\text{P9}) \\ \text{st.} \quad & \mathbf{z} \succeq \mathbf{0} \end{aligned}$$

where if the optimal value $p_9^o < \alpha$, (P8) is feasible and (42) holds.

Consequently, if any of the two problems (P4) and (P8) is feasible, we can conclude that $p_3^o \leq \alpha$. Otherwise, we should have $p_3^o > \alpha$. As such, we can find p_3^o by solving problem (P4) or (P8) for an evolving sequence of α, and stop until decreasing α a little further (say, $-\epsilon/2$) causes both problems to be infeasible. When the algorithm stops, it is guaranteed that the final α value is at most ϵ-away from the real optimal value p_3^o for problem (P3).

Specifically, to find the solution of (P3), we can use the *bisection* search method to update α and solve the feasibility problem (P4) or (P8) at each step. We start from an interval $[L, U]$ containing the optimal value, where the interval can be determined by (36). We first solve the feasibility problem at its midpoint $\alpha = (L+U)/2$, to determine whether the optimal value is in the lower or upper half of the interval. We then update the interval and the optimal solution $\mathbf{z}' = \mathbf{z}_4^o$ or $\mathbf{z}_8^o$ accordingly. We now have a new interval containing the optimal value but with half the width of the initial interval. This procedure is repeated until the width of the interval is small enough, and then $\mathbf{z}'$ is a good approximation to the optimal solution $\mathbf{z}_3^o$. As a result, the optimal weight vector is given by

$$\mathbf{w}_{\text{opt}} = \frac{\mathbf{w}'_{\text{opt}}}{\|\mathbf{w}'_{\text{opt}}\|} \quad (51)$$

where

$$\mathbf{w}'_{\text{opt}} = \mathbf{z}_3^o \sqrt{\mathbf{z}_3^{oT}\boldsymbol{\Sigma}_{\mathcal{H}_1}\mathbf{z}_3^o} \quad (52)$$

with the optimal test threshold calculated by (56). The algorithm for calculating the optimal combining weights in a conservative cognitive radio network is summarized with pseudo-code in Algorithm 1.

Optimal Cooperation for Conservative Systems

Given $L = \max\left\{\psi_c\left(P_f\right), \omega_c\left(P_f\right)\right\}$, $U = \phi_c\left(P_f\right)$, and tolerance $\varepsilon > 0$.

While $U - L > \varepsilon$

$\alpha = (L+U)/2$

if (P8) is feasible for α

$U = \alpha$ and $\mathbf{z}_{\text{opt}} = \mathbf{z}_o$

else if (P4) is feasible **then**

$U = \alpha$ and $\mathbf{z}_{\text{opt}} = \mathbf{z}_o$

else

$L = \alpha$

end if

end while

Compute $\mathbf{w}_{\text{opt}}$ using **(51)** and γ_c using **(27)**.

Return $\mathbf{w}_{\text{opt}}$ and γ_c.

2) Aggressive Systems: For an aggressive cognitive radio network, the probabilities of false alarm and detection satisfy $P_f < 1/2$ and $P_d > 1/2$, respectively. According to (28), we should have

$$Q^{-1}(P_f)\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{w}} - E_s\boldsymbol{g}^T\mathbf{w} < 0 \tag{53}$$

where $Q^{-1}(P_f) > 0$. Following the same procedure as in the previous subsection, we can obtain an optimization problem similar to (P3), which is written as follows:

$$\begin{aligned} \min_{\mathbf{z}} \quad & Q^{-1}(P_f)\sqrt{\mathbf{z}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{z}} - E_s\boldsymbol{g}^T\mathbf{z} \quad \text{(P10)} \\ \text{st.} \quad & \mathbf{z}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{z} \le 1. \end{aligned}$$

We would like to emphasize that (P10) is different from (P3) in that $Q^{-1}(P_f) > 0$ instead. In this case, (P10) can be transformed into the following convex problem by changing the optimization variable to $\mathbf{s} = \mathbf{\Sigma}_{\mathcal{H}_0}^{1/2}\mathbf{z}$, i.e.,

$$\begin{aligned} \min_{\mathbf{z}} \quad & Q^{-1}(P_f)\|\mathbf{s}\| - E_s\boldsymbol{g}^T\mathbf{\Sigma}_{\mathcal{H}_0}^{-1/2}\mathbf{s} \quad \text{(P11)} \\ \text{st.} \quad & \mathbf{s}^T\mathbf{\Sigma}_{\mathcal{H}_0}^{-T/2}\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{\Sigma}_{\mathcal{H}_0}^{-1/2}\mathbf{s} \le 1. \end{aligned}$$

Note that $Q^{-1}(P_f) > 0$, such that the objective function of (P11) consists of a convex function minus a linear function, and hence it is a convex function. It is obvious that $\mathbf{s}$ is limited in an ellipsoid because $\mathbf{\Sigma}_{\mathcal{H}_0} \succeq 0$ and $\mathbf{\Sigma}_{\mathcal{H}_1} \succeq 0$. Thus, (P11) is a convex optimization problem that can be easily solved.

Remark 5: Recall that the establishment of (P10) for aggressive systems is conditioned on (53). Hence, if the optimal value of (P10) satisfies $p_{10}^o < 0$ (or $p_{11}^o < 0$), then the system is aggressive. Otherwise, it is hostile since $P_d \le 1/2$ (i.e., $p_{10}^o \ge 0$) and we could not find a weight vector $\mathbf{w}$ such that (53) holds. It is hard to find an optimal solution for hostile systems. Fortunately, in cognitive radio networks, it is rarely allowed that the CR sets a low P_d to boost up spectrum efficiency, since this practice will cause unbearable interference to primary radios.

Once the optimal solution $\mathbf{s}^o$ of (P11) has been solved, the optimal detector at the fusion center can be determined, i.e., the optimal weight vector is given by

$$\mathbf{w}_{\text{opt}} = \frac{\mathbf{w}'_{\text{opt}}}{\|\mathbf{w}'_{\text{opt}}\|} \tag{54}$$

where

$$\mathbf{w}'_{\text{opt}} = \mathbf{\Sigma}_{\mathcal{H}_0}^{-1/2}\mathbf{s}^o\sqrt{\mathbf{s}^{oT}\mathbf{\Sigma}_{\mathcal{H}_0}^{-T/2}\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{\Sigma}_{\mathcal{H}_0}^{-1/2}\mathbf{s}^o} \tag{55}$$

and the test threshold at the fusion center is

$$\gamma_c = N\boldsymbol{\sigma}^T\mathbf{w}_{\text{opt}} + Q^{-1}(P_f)\sqrt{\mathbf{w}_{\text{opt}}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{w}_{\text{opt}}}. \tag{56}$$

Therefore, the strategy is to first assume that the system is aggressive if $Q^{-1}(P_f) > 0$. We then solve problem (P11) and check the optimal value p_{11}^o. If $p_{11}^o < 0$, then the assumption is valid and the solution corresponds to the optimal weight vector. Otherwise, the assumption is not true and the system is considered as a hostile system. The algorithm is summarized as follows in Algorithm 2.

Optimal Cooperation for Aggressive Systems

Solve (P11) for p_{11}^o and $\mathbf{s}^o$.

if $p_{11}^o < 0$ **then**

Calculate $\mathbf{w}_{\text{opt}}$ using **(54)** and γ_c using **(56)**.

Return $\mathbf{w}_{\text{opt}}$ and γ_c.

else

Return "It is a hostile system".

end if

Alternatively, one may formulate another optimization problem that minimizes P_f subject to a constraint on P_d. Mathematically, the alternative problem belongs to the same category as the original problem and can be solved using the algorithms developed in this subsection with trivial modifications.

C. Optimization of the Modified Deflection Coefficient

From the previous discussions, we see that it is generally hard to find the optimal solution for all possible cognitive radio systems: conservative, aggressive, and hostile. As such, we now present a heuristic but general method to find the weight vector, which requires less computational complexity and incurs small performance degradation.

From (15) and (20) we observe that the weight vector $\mathbf{w}$ plays an important role in shaping the PDF of the global test statistic. To measure the effect of the PDF on the detection performance, we introduce a modified *deflection* coefficient (MDC)

$$d_m^2(\mathbf{w}) = \frac{[\mathrm{E}(y_c|\mathcal{H}_1) - \mathrm{E}(y_c|\mathcal{H}_0)]^2}{\mathrm{Var}(y_c|\mathcal{H}_1)} = \frac{\left(E_s\boldsymbol{g}^T\mathbf{w}\right)^2}{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{w}} \tag{57}$$

which provides a good measure of the detection performance since it characterizes the variance-normalized distance between the centers of two conditional PDFs. According to (18) and (20), the global test statistic y_c has different variances under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$. In particular, we have $\mathrm{Var}(y_c|\mathcal{H}_1) > \mathrm{Var}(y_c|\mathcal{H}_0)$. Moreover, if $\eta_i \gg 1$, we have $\mathrm{Var}(y_c|\mathcal{H}_1) \gg \mathrm{Var}(y_c|\mathcal{H}_0)$. Therefore, the PDF of y_c under $\mathcal{H}_1$ has a heavier tail than $\mathcal{H}_0$. This justifies using the measure d_m^2 for spectrum sensing in cognitive radio networks, where high detection sensitivity is desirable.

For accurate inference, we would like to maximize d_m^2 under the unit-norm constraint on the weight vector, i.e.,

$$\begin{aligned} \max_{\mathbf{w}} \quad & d_m^2(\mathbf{w}) \quad \text{(P12)} \\ \text{st.} \quad & \|\mathbf{w}\|_2^2 = 1. \end{aligned}$$

This problem can be solved as follows. Applying the linear transformation

$$\mathbf{w}' = \boldsymbol{\Sigma}_{\mathcal{H}_1}^{1/2}\mathbf{w} \tag{58}$$

we obtain

$$\begin{aligned} d_m^2(\mathbf{w}) &= \frac{E_s^2 \mathbf{w}'^T \boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2} \boldsymbol{g}\boldsymbol{g}^T \boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\mathbf{w}'}{\mathbf{w}'^T\mathbf{w}'} \\ &\overset{(a)}{\leq} E_s^2 \lambda_{\max}\left(\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\boldsymbol{g}^T\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right) \\ &= E_s^2 \left\|\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\right\|^2 \end{aligned} \tag{59}$$

where inequality (a) follows the *Rayleigh Ritz* inequality [19] and the equality is achieved if

$$\mathbf{w}' = \boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g} \tag{60}$$

which is the eigenvector of the positive semi-definite matrix $\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\boldsymbol{g}^T\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}$ corresponding to the maximum eigenvalue (the nonzero eigenvalue). Therefore, the optimal solution of (P12) is

$$\mathbf{w}_{12} = \frac{\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\mathbf{w}'}{\left\|\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\mathbf{w}'\right\|_2} \tag{61}$$

which maximizes d_m^2. To enforce $\mathrm{E}(y_c|\mathcal{H}_1) > \mathrm{E}(y_c|\mathcal{H}_0)$, we let $\mathbf{w}_{12}^o = \mathrm{sign}\left(\boldsymbol{g}^T\mathbf{w}_{12}\right)\mathbf{w}_{12}$. In the simulation results, we will show that the maximum d_m^2 leads to a large probability of detection. Such an approach performs similarly to the one maximizing P_d directly (as shown by simulation results in Section V) but with much less complexity and provides solution for general cognitive radio networks regardless of being conservative, aggressive, or hostile.

V. Numerical Results

In this section, the proposed cooperation schemes are evaluated numerically and compared with some existing methods. Consider a network of M cognitive radios, each of which independently senses the targeted spectrum. For simplicity, we assume that the transmitted primary signal is $s(k) = 1$. The proposed schemes are compared with the single CR spectrum sensing, the selection combining method [24] (denoted by SC, i.e., selecting the user with the maximum SNR), and the LRT-based fusion.

In Fig. 2, we plot the minimum probability of miss-detection $(1 - P_d)$ against the probability of false alarm (P_f), which indirectly measures the interference level to the primary radios for a given P_f. The result shows that the proposed optimal linear cooperation schemes, denoted as OPT. LIN for the maximum P_d method and OPT. MDC for the maximum modified deflection coefficient approach, lead to much less interference (with much higher P_d) to the primary radios than single CR and SC based approaches. In particular, the performance of optimal linear detectors is very close to that of the LRT-based detector. The lower bound of $(1 - P_d)$ corresponding to the upper bound of P_d given in (29) is fairly tight. In addition, we observe that the probability of detection given by the maximum MDC method closely approximates the exact maximum P_d value obtained from solving (P1) over a wide range of P_f. Therefore, the maximum MDC scheme can be used as an efficient suboptimal alternative for both conservative and aggressive opportunistic spectrum sensing.

In Fig. 3, we investigate the receiver operating characteristics ($1 - P_d$ versus P_f) for various numbers of cooperative CRs

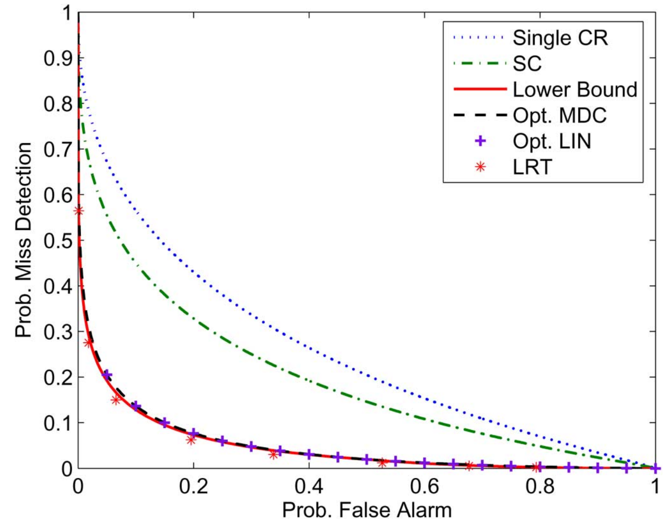

Fig. 2. Probability of miss-detection $(1 - P_d)$ versus the probability of false alarm (P_f), with $M = 6$, $N = 20$, $\sigma_i^2 = 1$, and $\delta_i^2 = 1$, $i = 1, 2, \ldots, M$. The local SNRs at individual CRs are $\{9.3, 7.8, 9.6, 7.6, 3.5, 9.2\}$ in dB, which are N times the single sample SNRs at individual CRs, i.e., $\{-3.7, -5.2, -3.4, -5.4, -9.5, -3.8\}$ in dB. The results are obtained from simulations over 1,000,000 noise realizations for the given set of channel gains and noise variances.

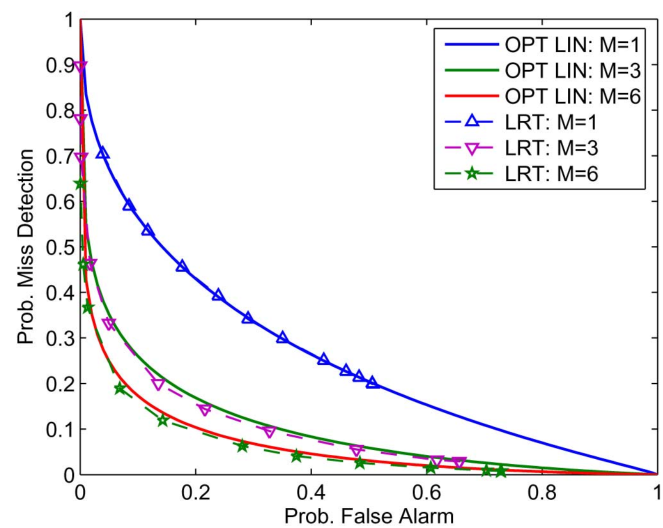

Fig. 3. Probability of miss-detection $(1 - P_d)$ versus the probability of false alarm (P_f) under various sensing noises, with fixed channel noise $\delta^2 = 0.5$, $N = 20$, and $M = 1$, 3, and 6. For $M = 1$, 3, and 6, the average local SNRs over individual CRs are respectively 8.3, 7.5, and 5.9 in dB. Specifically, for $M = 1$, the sensing noise level is $\sigma_1^2 = 1.9$ and the local SNR is 8.3 dB; for $M = 3$, the sensing noise levels are $\boldsymbol{\sigma} = \{0.7, 1.0, 0.9\}$ and the local SNRs are $\{10.4, 9.3, 2.6\}$ in dB; for $M = 6$, the noise levels are $\boldsymbol{\sigma} = \{0.9, 1.3, 1.0, 2.0, 1.8, 1.2\}$ and the local SNRs are $\{7.2, 5.1, 0.8, -1.2, 3.6, 9.7\}$ in dB. The results are obtained from simulations over 1,000,000 noise realizations for the given set of channel gains and noise variances.

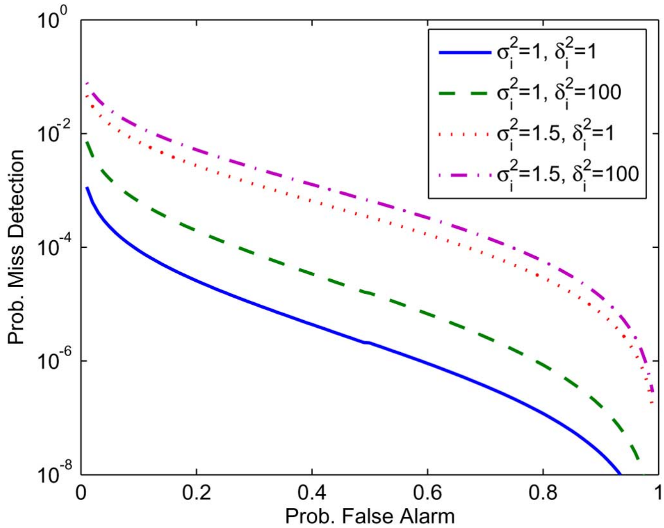

Fig. 4. Probability of miss-detection $(1 - P_d)$ versus the probability of false alarm (P_f) under different source and channel noises, with $M = 3$, $N = 100$, and uniform noises for different nodes. The local SNRs at individual CRs are $\{18.8, 15.2, 17.1\}$ in dB, corresponding to the single sample SNRs $\{5.8, 2.2, 4.1\}$ in dB. The results are obtained by numerically solving the optimization problem (P1) for the given set of channel gains and noise variances.

and different channel conditions. For each scenario, we randomly generate the channel gains and sensing noises that eventually influence the detection performance. The figure shows that with the same channel noise level, the sensing reliability improves as the number of cooperative nodes increases even if the average local SNR decreases. In addition, the performance of the optimal linear detector is very close to that of the optimal LRT-based detector in these three scenarios.

In Fig. 4, we draw $(1 - P_d)$ versus P_f under different noise conditions. As we can see, the detection performance degrades as the noise condition becomes severe. It can also be observed that the inference accuracy is more sensitive to the sensing noise change than to that of the communication channel noise. This is a good motivation for multi-CR cooperation to improve the spectrum sensing reliability.

VI. Conclusion

We have developed an optimal linear framework for cooperative spectrum sensing in cognitive radio networks. The proposed methods optimize the detection performance by operating over a linear combination of local test statistics from individual secondary users, which combats the destructive channel effects between the target primary radio and the opportunistic CRs. Within this framework, we have given exact solutions for finding the optimal weight vector for aggressive and conservative cognitive systems. Furthermore, we proposed an MDC-based optimization method, which would approximate the maximum-P_d approach for any given probability of false alarm and is applicable for general cognitive radio systems. The proposed novel numerical algorithms can be applied to similar non-convex problems in other applications. Some interesting extensions of this work may include studying constrained communications between the secondary users and the fusion center over fading wireless channels.

Appendix A
Proof of Lemma 1: Bounds for Aggressive and Hostile Systems

Proof: Equation (40) can be rewritten as

$$f(\mathbf{w}) = Q^{-1}(P_f)\sqrt{\frac{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{w}}{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{w}}} - \frac{E_s\boldsymbol{g}^T\mathbf{w}}{\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{w}}} \tag{62}$$

where each of the two terms containing $\mathbf{w}$ can be bounded respectively as follows:

$$\lambda_{\min}^{1/2}\left(\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{\Sigma}_{\mathcal{H}_1}^{-1/2}\right) \le \sqrt{\frac{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{w}}{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{w}}} \le \lambda_{\max}^{1/2}\left(\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\mathbf{\Sigma}_{\mathcal{H}_0}\mathbf{\Sigma}_{\mathcal{H}_1}^{-1/2}\right) \tag{63}$$

and

$$\lambda_{\min}^{1/2}\left(\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\boldsymbol{g}^T\mathbf{\Sigma}_{\mathcal{H}_1}^{-1/2}\right) \le \frac{\boldsymbol{g}^T\mathbf{w}}{\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{w}}} \le \lambda_{\max}^{1/2}\left(\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\boldsymbol{g}^T\mathbf{\Sigma}_{\mathcal{H}_1}^{-1/2}\right). \tag{64}$$

Note that $\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\boldsymbol{g}^T\mathbf{\Sigma}_{\mathcal{H}_1}^{-1/2}$ is a rank-one matrix. Thus, we have

$$\lambda_{\max}\left(\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\boldsymbol{g}^T\mathbf{\Sigma}_{\mathcal{H}_1}^{-1/2}\right) = \left\|\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\right\|^2 \tag{65}$$

and

$$\lambda_{\min}\left(\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\boldsymbol{g}^T\mathbf{\Sigma}_{\mathcal{H}_1}^{-1/2}\right) = 0. \tag{66}$$

Substituting (65) and (66) into (64) gives

$$0 \le \frac{\boldsymbol{g}^T\mathbf{w}}{\sqrt{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{w}}} \le \left\|\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\boldsymbol{g}\right\|. \tag{67}$$

Considering $P_f < 1/2$, we have $Q^{-1}(P_f) > 0$. According to (30) and (31), it follows that

$$\psi_a(P_f) \le f(\mathbf{w}) \le \phi_a(P_f). \tag{68}$$

Accordingly, Lemma 1 is established since the Q-function is a monotonically decreasing function. ■

Appendix B
Proof of Lemma 2: Bounds for Conservative Systems

Proof: For a conservative system where $P_f \ge 1/2$, we have $Q^{-1}(P_f) \le 0$. Similar to the proof in Appendix A, we can obtain the following bounds on $f(\mathbf{w})$ from (37) and (39):

$$\psi_c(P_f) \le f(\mathbf{w}) \le \phi_c(P_f). \tag{69}$$

On the other hand, $f(\mathbf{w})$ can be lower bounded as follows:

$$\begin{aligned} f(\mathbf{w}) &= -\left(\frac{\left[E_s \boldsymbol{g}^T\mathbf{w} - Q^{-1}(P_f)\sqrt{\mathbf{w}^T\boldsymbol{\Sigma}_{\mathcal{H}_0}\mathbf{w}}\right]^2}{\mathbf{w}^T\boldsymbol{\Sigma}_{\mathcal{H}_1}\mathbf{w}}\right)^{1/2} \\ &\overset{(b)}{\geq} -\left(\frac{\left[E_s^2 + Q^{-2}(P_f)\right]\mathbf{w}^T\left(\boldsymbol{g}\boldsymbol{g}^T + \boldsymbol{\Sigma}_{\mathcal{H}_0}\right)\mathbf{w}}{\mathbf{w}^T\boldsymbol{\Sigma}_{\mathcal{H}_1}\mathbf{w}}\right)^{1/2} \\ &\overset{(c)}{\geq} -\sqrt{E_s^2 + Q^{-2}(P_f)}\lambda_{\max} \\ &\quad \times \left[\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-T/2}\left(\boldsymbol{g}\boldsymbol{g}^T + \boldsymbol{\Sigma}_{\mathcal{H}_0}\right)\boldsymbol{\Sigma}_{\mathcal{H}_1}^{-1/2}\right] \end{aligned}$$

where (b) follows from the Cauchy-Schwartz inequality and (c) is due to the Rayleigh Ritz inequality. Therefore, we have

$$\max\{\psi_c(P_f), \omega_c(P_f)\} \leq f(\mathbf{w}) \leq \phi_c(P_f) \tag{70}$$

which leads to Lemma 2. ■

ACKNOWLEDGMENT

The authors would like to thank Prof. L. Vandenberghe at UCLA for helpful discussions on the dual problem (P5).

REFERENCES

[1] Z. Quan, S. Cui, and A. H. Sayed, "An optimal strategy for cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE GLOBECOM*, Washington, DC, Nov. 2007.
[2] J. Mitola, III and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun*, vol. 6, pp. 13–18, 1999.
[3] Fed. Commun. Comm. 2003, Et docket-322.
[4] D. Cabric, S. M. Mishra, and R. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proc. 38th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2004, pp. 772–776.
[5] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
[6] D. Cabric, A. Tkachenko, and R. W. Brodersen, "Experimental study of spectrum sensing based on energy detection and network cooperation," in *Proc. ACM 1st Int. Workshop on Technology and Policy for Accessing Spectrum (TAPAS)*, Aug. 2006.
[7] S. Enserink and D. Cochran, "A cyclostationary feature detector," in *Proc. 28th Asilomar Conference on Signals, Systems, and Computers*, Monterey, CA, Oct. 1994, pp. 806–810.
[8] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Proc. Allerton Conf. Communication, Control, and Computing*, Oct. 2004, pp. 131–136.
[9] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Select. Areas Commun*, vol. 23, pp. 201–220, Feb. 2005.
[10] P. K. Varshney, *Distributed Detection and Data Fusion*. New York: Springer-Verlag, 1997.
[11] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors: Part II—Advanced topics," *Proc. IEEE*, vol. 85, pp. 64–79, Jan. 1997.
[12] 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, Feb. 1992.
[13] A. Ghasemi and E. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. IEEE Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Baltimore, MD, Nov. 2005, pp. 131–136.
[14] E. Vistotsky, S. Kuffner, and R. Peterson, "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," in *Proc. IEEE Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Baltimore, MD, Nov. 2005, pp. 338–345.
[15] G. Ghurumuruhan and Y. Li, "Agility improvement through cooperative diversity in cognitive radio," in *Proc. IEEE GLOBECOM*, St. Louis, MO, Nov. 2005, pp. 2505–2509.
[16] B. Chen and P. K. Willett, "On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 693–699, Feb. 2005.
[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[18] B. V. Gendenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*. Reading, MA: Addison-Wesley, 1954.
[19] A. H. Sayed, *Fundamentals of Adaptive Filtering*. NJ: Wiley, 2003.
[20] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer-Verlag, 1998.
[21] P. W. Glynn and W. Whitt, "Logarithmic asymptotics for steady-state tail probabilities in a single-server queue," *J. Appl. Prob*, pp. 131–155, 1994.
[22] C.-S. Chang and J. A. Thomas, "Effective bandwidth in high-speed digital networks," *IEEE J. Select. Areas Commun*, vol. 13, pp. 1091–1100, Aug. 1995.
[23] R. J. Stern and H. Wolkowicz, "Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations," *SIAM J. Optim.*, vol. 5, no. 2, pp. 286–313, 1995.
[24] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

Zhi Quan (S'06) received the B.E. degree in communication engineering from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, and the M.S. degree in electrical engineering from Oklahoma State University (OSU), Stillwater. He is currently pursuing the Ph.D. degree in electrical engineering, University of California, Los Angeles.

He was a Visiting Researcher with Princeton University, Princeton, NJ, during the summer of 2007. His current research interests include statistical signal processing, linear and nonlinear optimization, wireless communications, and cognitive radios.

Shuguang Cui (S'99–M'05) received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 2005, the M.Eng. in electrical engineering from McMaster University, Hamilton, ON, Canada, in 2000, and the B.Eng. in radio engineering with the highest distinction from Beijing University of Posts and Telecommunications, Beijing, China, in 1997.

He is now working as an Assistant Professor in electrical and computer engineering with Texas A&M University, College Station, TX. From 1997 to 1998, he was with Hewlett-Packard, Beijing, China, as a System Engineer. In the summer of 2003, he was with National Semiconductor, Santa Clara, CA, on the ZigBee project. From 2005 to 2007, he worked as an Assistant Professor in the Department of Electrical and Computer Engineering, University of Arizona, Tucson. His current research interests include cross-layer energy minimization for low-power sensor networks, hardware and system synergies for high-performance wireless radios, statistical signal processing, and general communication theories.

Dr. Cui was a recipient of the NSERC graduate fellowship from the National Science and Engineering Research Council of Canada and the Canadian Wireless Telecommunications Association (CWTA) graduate scholarship. He has been serving as the TPC Co-chair for the 2007 IEEE Communication Theory Workshop and the ICC'08 Communication Theory Symposium. He is currently serving as an associate editor for the IEEE COMMUNICATION LETTERS and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

Ali H. Sayed (S'90–M'92–SM'99–F'01) received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1992.

He is Professor and Chairman of the Electrical Engineering Department at the University of California, Los Angeles (UCLA). He is also the Principal Investigator of the UCLA Adaptive Systems Laboratory (http://www.ee.ucla.edu/asl). He is the author or coauthor over 290 journal and conference publications and is the author of the textbook *Fundamentals of Adaptive Filtering* (New York: Wiley, 2003), the coauthor of the research monograph *Indefinite Quadratic Estimation and Control* (Philadelphia, PA: SIAM, 1999), and of the graduate-level textbook *Linear Estimation* (Englewood Cliffs, NJ: Prentice-Hall, 2000). He is also coeditor of *Fast Reliable Algorithms for Matrices with Structure* (Philadelphia, PA: SIAM, 1999). He has contributed several articles to engineering and mathematical encyclopedias and handbooks and has served on the program committees of several international meetings. His research interests include adaptive and statistical signal processing, distributed processing, filtering and estimation theories, signal processing for communications, interplays between signal processing and control methodologies, system theory, and fast algorithms for large-scale problems.

Dr. Sayed is a recipient of the 1996 IEEE Donald G. Fink Award, a 2002 Best Paper Award from the IEEE Signal Processing Society, the 2003 Kuwait Prize in Basic Science, the 2005 Frederick E. Terman Award, the 2005 Young Author Best Paper Award from the IEEE Signal Processing Society, and two Best Student Paper awards at international meetings. He is also a member of the technical committees on Signal Processing Theory and Methods (SPTM) and on Signal Processing for Communications (SPCOM), both of the IEEE Signal Processing Society. He has served as Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2003–2005) and the *EURASIP Journal on Advances in Signal Processing* (2006-2007). He is the General Chairman of International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2008 and sits on the Board of Governors of the IEEE Signal Processing Society.