

Chapter 3

Spectrum Sensing Theories and Methods



Abstract Spectrum sensing is a critical step in cognitive radio based DSM to learn the radio environment. Despite its long history, in the past decade, the study on spectrum sensing has attracted substantial interests from the wireless communications community. In this chapter, we first provide the fundamental theories on spectrum sensing from the optimal likelihood ratio test perspective, then we review the classical methods including Bayesian method, robust hypothesis test, energy detection, matched filtering detection, and cyclostationary detection. After that, we discuss the robustness of the classical methods and review techniques that can enhance the sensing reliability under hostile environment. These methods include eigenvalue based sensing method, covariance based detection method. Finally, we discuss the cooperative sensing that uses data fusion or decision fusion from multiple sensors to enhance the sensing performance.

3.1 Introduction

As mentioned in Chap. 1, the basic idea of a cognitive radio is to support spectrum reuse or spectrum sharing, which allows the secondary networks/users to communicate over the spectrum allocated/licensed to the primary users (PUs) when they are not fully utilizing it. To do so, the secondary users (SUs) are required to frequently perform spectrum sensing, i.e., detecting the presence of the PUs. Whenever the PUs become active, the SUs have to detect the presence of them with a high probability, and vacate the channel or reduce transmit power within certain amount of time. For example, in the IEEE 802.22 standard [1–3], it is required for the SUs to detect the TV and wireless microphone signals and vacate the channel within two seconds once they become active. Furthermore, for TV signal detection, it is required to achieve 90% probability of detection and 10% probability of false alarm at signal-to-noise ratio (SNR) level as low as -20 dB [1].

Spectrum sensing has reborn as a very active research topic in the past decade despite its long history in signal detection field. Quite a few new sensing methods have been proposed which take into consideration practical requirements and constraints. In this chapter, we first provide the fundamental spectrum sensing theories

from the optimal likelihood ratio test perspective, then review the classical methods including Bayesian method, robust hypothesis test, energy detection, matched filtering detection, and cyclostationary detection. After that, we discuss the robustness of the classical methods and review techniques that can enhance the sensing reliability under hostile environment, including eigenvalue based sensing method, and covariance based detection. Finally, we discuss the cooperative spectrum sensing techniques that uses data fusion or decision fusion to combine the sensing data from multiple sensors.

3.1.1 System Model for Spectrum Sensing

We consider an SU spectrum sensor with $M \geq 1$ antennas. A similar scenario is the multi-node cooperative sensing, if all M distributed nodes are able to send their observed signals to a central node. There are two hypotheses: \mathcal{H}_0 , the PU is inactive; and \mathcal{H}_1 , the PU is active. The received signal at antenna/node i , $i = 1, \dots, M$, is given by

$$\mathcal{H}_0 : x_i(n) = \eta_i(n) \quad (3.1)$$

$$\mathcal{H}_1 : x_i(n) = s_i(n) + \eta_i(n) \quad (3.2)$$

where $\eta_i(n)$ is the received noise plus possible interference. At hypothesis \mathcal{H}_1 , $s_i(n)$ is the received primary signal at antenna/node i , which is the transmitted primary signal passing through the wireless channel to the sensing antenna/node. That is, $s_i(n)$ can be written as

$$s_i(n) = \sum_{l=0}^{q_i} h_i(l) \tilde{s}_{(n-l)} \quad (3.3)$$

where $\tilde{s}(n)$ stands for the transmitted primary signal, $h_i(l)$ and q_i denote the propagation channel coefficient and channel order from the PU to the i th antenna/node. For simplicity, it is assumed that the signal, noise, interference, and channel coefficients are all real numbers, though the theory and derivation can be directly extended to complex signals in most of the cases.

The objective of spectrum sensing is to choose one of the two hypotheses, \mathcal{H}_0 or \mathcal{H}_1 , based on the received signal samples at the SU sensor. The probability of detection, P_d , and probability of false alarm, P_{fa} , are defined as follows:

$$P_d = P(\mathcal{H}_1 | \mathcal{H}_1) \quad (3.4)$$

$$P_{fa} = P(\mathcal{H}_1 | \mathcal{H}_0) \quad (3.5)$$

where $P(\cdot | \cdot)$ defines the conditional probability. There are interesting physical meanings for the above two probabilities: P_d defines how well the PU is protected when

it is using the spectrum, while P_{fa} determines the opportunity the SU misses out when the PU is not using the spectrum. In general, a sensing algorithm is said to be “optimal” if it achieves the lowest P_{fa} for a given P_d with a fixed number of samples, though there could be other criteria to evaluate the performance of a sensing algorithm.

In order to apply both space and time processing, we stack the signals from the M antennas/nodes and L time samples to yield the following $ML \times 1$ vectors:

$$\mathbf{x}(n) = [x_1(n) \dots x_M(n) \ x_1(n-1) \dots x_M(n-1) \dots x_1(n-L+1) \dots x_M(n-L+1)]^T \quad (3.6)$$

$$\mathbf{s}(n) = [s_1(n) \dots s_M(n) \ s_1(n-1) \dots s_M(n-1) \dots s_1(n-L+1) \dots s_M(n-L+1)]^T \quad (3.7)$$

$$\boldsymbol{\eta}(n) = [\eta_1(n) \dots \eta_M(n) \ \eta_1(n-1) \dots \eta_M(n-1) \dots \eta_1(n-L+1) \dots \eta_M(n-L+1)]^T \quad (3.8)$$

Based on the above vector forms, the hypothesis testing problem can be reformulated as

$$\mathcal{H}_0 : \mathbf{x}(n) = \boldsymbol{\eta}(n), \quad n = 0, \dots, N-1. \quad (3.9)$$

$$\mathcal{H}_1 : \mathbf{x}(n) = \mathbf{s}(n) + \boldsymbol{\eta}(n), \quad n = 0, \dots, N-1. \quad (3.10)$$

Accurate knowledge on the noise power σ_η^2 is the key for many detection methods. Unfortunately, in practice, the noise uncertainty always presents. Due to the noise uncertainty [4–6], the estimated (or assumed) noise power may be different from the real noise power. Let the estimated noise power be $\hat{\sigma}_\eta^2 = \alpha \sigma_\eta^2$. It is assumed that α (in dB) is uniformly distributed in an interval $[-B, B]$, where B is called the noise uncertainty factor [5]. In practice, the noise uncertainty factor of a receiving device is typically in the range from 1 to 2 dB, but the environment/interference noise uncertainty can be much higher [5].

3.1.2 Design Challenges for Spectrum Sensing

The design of spectrum sensing methods in CR faces a few specific challenges including, among others, the following.

1. Low sensing SNR: A cognitive radio may need to sense the primary signal at very low SNR condition. This is to overcome the hidden node problem, i.e., a SU sensor hears very weak signal from the primary transmitter but can strongly interfere the primary receiver if it transmits (here the primary receiver looks like a hidden node). To avoid such interference, one solution is to require the SU sensor to have the capability to sense the presence of the primary signal at very low SNR.

For example, in the 802.22 standard, the sensing sensitivity requirement is as low as -20 dB.

2. **Channel uncertainty:** In wireless communications, the propagation channel is usually unknown and time variant. Such unknown channel makes coherent detection methods unreliable in practice.
3. **Non-synchronization:** It is hard to synchronize the received signal with the primary signal in time and frequency domains. This will cause traditional methods like preamble/pilot based detections less effective.
4. **Noise uncertainty:** The noise level may vary with time and location, which yields the noise power uncertainty issue for detection [4–7]. This makes methods relying on accurate noise power unreliable. Furthermore, the noise may not be white, which will further affect the effectiveness of many methods with white noise assumption.
5. **Interference:** There could be interferences from intentional or unintentional transmitters. Thus, the detector needs to have the capability to suppress the interference while identifying the primary signals.

While there have been a lot of spectrum sensing methods in the literature [8–15], many of them are based on ideal assumptions, and cannot work well in a hostile radio environment. We need the spectrum sensing to be robust to the unknown and maybe time-varying channel, noise and interference.

3.2 Classical Detection Theories and Methods

In this section, we first provide the fundamental theories on spectrum sensing from the optimal likelihood ratio test perspective, then we review the classical methods including Bayesian method, robust hypothesis test, energy detection, matched filtering detection, and cyclostationary detection.

3.2.1 Neyman–Pearson Theorem

The Neyman–Pearson (NP) theorem [16–18] states that, for a given probability of false alarm, the test statistic that maximizes the probability of detection is the likelihood ratio test (LRT) defined as

$$T_{LRT}(\mathbf{x}) = \frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_0)} \quad (3.11)$$

where $p(\cdot)$ denotes the probability density function (PDF), and \mathbf{x} denotes the received signal vector that is the aggregation of $\mathbf{x}(n)$, $n = 0, 1, \dots, N - 1$. Such a likelihood ratio test decides \mathcal{H}_1 when $T_{LRT}(\mathbf{x})$ exceeds a threshold γ , and \mathcal{H}_0 otherwise.

The major difficulty in using the LRT is its requirement on the exact distributions given in (3.11). Obviously, the distribution of random vector \mathbf{x} under \mathcal{H}_1 is related to the source signal distribution, the wireless channels, and the noise distribution, while the distribution of \mathbf{x} under \mathcal{H}_0 is related to the noise distribution. In order to use the LRT, we need to obtain the knowledge of the channels as well as the signal and noise distributions, which is practically difficult to realize.

If we assume that the channels are flat-fading, and the received source signal sample $s_i(n)$'s are independent over n , the PDFs in LRT are decoupled as

$$p(\mathbf{x}|\mathcal{H}_1) = \prod_{n=0}^{N-1} p(\mathbf{x}(n)|\mathcal{H}_1) \quad (3.12)$$

$$p(\mathbf{x}|\mathcal{H}_0) = \prod_{n=0}^{N-1} p(\mathbf{x}(n)|\mathcal{H}_0) \quad (3.13)$$

If we further assume that noise and signal samples are both Gaussian distributed, i.e., $\boldsymbol{\eta}(n) \sim \mathcal{N}(\mathbf{0}, \sigma_\eta^2 \mathbf{I})$ and $\mathbf{s}(n) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_s)$, the LRT becomes the estimator-correlator (EC) [16] detector for which the test statistic is given by

$$T_{EC}(\mathbf{x}) = \sum_{n=0}^{N-1} \mathbf{x}^T(n) \mathbf{R}_s (\mathbf{R}_s + \sigma_\eta^2 \mathbf{I})^{-1} \mathbf{x}(n) \quad (3.14)$$

From (3.10), we see that $\mathbf{R}_s (\mathbf{R}_s + 2\sigma_\eta^2 \mathbf{I})^{-1} \mathbf{x}(n)$ is actually the minimum-mean-squared-error (MMSE) estimation of the source signal $\mathbf{s}(n)$. Thus, $T_{EC}(\mathbf{x})$ in (3.14) can be seen as the correlation of the observed signal $\mathbf{x}(n)$ with the MMSE estimation of $\mathbf{s}(n)$.

The EC detector needs to know the source signal covariance matrix \mathbf{R}_s and noise power σ_η^2 . When the signal presence is unknown yet, it is unrealistic to have the knowledge on the source signal covariance matrix which is related to unknown channels.

3.2.2 Bayesian Method and the Generalized Likelihood Ratio Test

In practical scenarios, it is difficult to know the likelihood functions exactly. For instance, we may not know the noise power σ_η^2 and/or source signal covariance \mathbf{R}_s . Hypothesis testing in the presence of uncertain parameters is known as ‘‘composite’’ hypothesis testing. In classic detection theory, there are two main approaches to tackle this problem: the Bayesian method and the generalized likelihood ratio test (GLRT).

In the Bayesian method [16], the objective is to evaluate the likelihood functions needed in the LRT through marginalization, i.e.,

$$p(\mathbf{x}|\mathcal{H}_0) = \int p(\mathbf{x}|\mathcal{H}_0, \Theta_0)p(\Theta_0|\mathcal{H}_0)d\Theta_0 \quad (3.15)$$

where Θ_0 represents all the unknowns when \mathcal{H}_0 is true. Note that the integration operation in (3.15) should be replaced with a summation if the elements in Θ_0 are drawn from a discrete sample space. Critically, we have to assign a prior distribution $p(\Theta_0|\mathcal{H}_0)$ to the unknown parameters. In other words, we need to treat these unknowns as random variables, and use their known distributions to express our belief in their values. Similarly, $p(\mathbf{x}|\mathcal{H}_1)$ can be defined. The main drawbacks of the Bayesian approach are listed as follows:

1. The marginalization operation in (3.15) is often not tractable except for very simple cases;
2. The choice of prior distributions affects the detection performance dramatically and thus it is not a trivial task to choose them.

To make the LRT applicable, we may estimate the unknown parameters first and then use the estimated parameters in the LRT. Known estimation techniques could be used for this purpose. However, there is one major difference from the conventional estimation problem where we know that signal is present, while in the case of spectrum sensing we are not sure whether there is source signal or not (the first priority here is the detection of signal presence). At different hypothesis, \mathcal{H}_0 or \mathcal{H}_1 , the unknown parameters are also different.

The GLRT is one efficient method [16, 18] to resolve the above problem, which has been used in many applications, e.g., radar and sonar signal processing. For this method, the maximum likelihood estimation of the unknown parameters under \mathcal{H}_0 and \mathcal{H}_1 are first obtained as

$$\begin{aligned} \hat{\Theta}_0 &= \arg \max_{\Theta_0} p(\mathbf{x}|\mathcal{H}_0, \Theta_0) \\ \hat{\Theta}_1 &= \arg \max_{\Theta_1} p(\mathbf{x}|\mathcal{H}_1, \Theta_1) \end{aligned}$$

where Θ_0 and Θ_1 are the set of unknown parameters under \mathcal{H}_0 and \mathcal{H}_1 , respectively. Then, the GLRT statistic is formed as

$$T_{GLRT} = \frac{p(\mathbf{x}|\hat{\Theta}_1, \mathcal{H}_1)}{p(\mathbf{x}|\hat{\Theta}_0, \mathcal{H}_0)} \quad (3.16)$$

Finally, the GLRT decides \mathcal{H}_1 if $T_{GLRT}(\mathbf{x}) > \gamma$, where γ is a threshold, and \mathcal{H}_0 otherwise.

It is not guaranteed that the GLRT is optimal or closes to the optimum when the sample size goes to infinity. Since the unknown parameters in Θ_0 and Θ_1 are highly dependent on the noise and signal statistical models, the estimations of them could

be vulnerable to the modeling errors. Under the assumption of Gaussian distributed source signals and noises, and flat-fading channels, some efficient spectrum sensing methods based on the GLRT can be found in [19–21].

3.2.3 Robust Hypothesis Testing

The searching for robust detection methods has been of great interest in the field of signal processing and many others. In this section, we start from a general paradigm that is called robust hypothesis testing, then we review a few methods that are robust to certain impairments. In Sects. 3.3–3.5, we will discuss a few new methods, including eigenvalue based detections, covariance based detections, and cooperative sensing.

A useful paradigm to design robust detectors is the maxmin approach, which maximizes the worst case detection performance. Among others, two techniques are very useful for robust spectrum sensing: the robust hypothesis testing [22, 23] and the robust matched filtering [24, 25]. In the following, we will give a brief overview on them.

Let the PDF of a received signal sample be f_1 at hypothesis \mathcal{H}_1 and f_0 at hypothesis \mathcal{H}_0 . If we know these two functions, the LRT-based detection described in Sect. 3.2.2 is optimal. However, in practice, due to the channel impairment, noise uncertainty, and interference, it is very hard to obtain these two functions exactly. One possible situation is when we only know that f_1 and f_0 belong to certain classes. One such class is called the ϵ -contamination class given by

$$\begin{aligned} \mathcal{H}_0 : f_0 \in F_0, F_0 &= \{(1 - \epsilon_0)f_0^0 + \epsilon_0g_0\} \\ \mathcal{H}_1 : f_1 \in F_1, F_1 &= \{(1 - \epsilon_1)f_1^0 + \epsilon_1g_1\} \end{aligned} \quad (3.17)$$

where f_j^0 ($j = 0, 1$) is the nominal PDF under hypothesis \mathcal{H}_j , ϵ_j in $[1, 0]$ is the maximum degree of contamination, and g_j is an arbitrary density function. Assume that we only know f_j^0 and ϵ_j (an upper bound for contamination), $j = 1, 2$. The problem is then to design a detection scheme to minimize the worst-case probability of error (e.g., probability of false alarm plus probability of mis-detection), i.e., finding a detector $\hat{\Psi}$ such that

$$\hat{\Psi} = \arg \min_{\Psi} \max_{(f_0, f_1) \in F_0 \times F_1} (P_{fa}(f_0, f_1, \Psi) + 1 - P_d(f_0, f_1, \Psi)) \quad (3.18)$$

Hubber [22] proved that the optimal test statistic is a “censored” version of the LRT given by

$$\hat{\Psi} = T_{CLRT}(\mathbf{x}) = \prod_{n=0}^{N-1} r(x(n)) \quad (3.19)$$

where

$$r(t) = \begin{cases} c_1, & c_1 \leq \frac{f_1^0(t)}{f_0^0(t)} \\ \frac{f_1^0(t)}{f_0^0(t)}, & c_0 < \frac{f_1^0(t)}{f_0^0(t)} < c_1 \\ c_0, & \frac{f_1^0(t)}{f_0^0(t)} \leq c_0 \end{cases} \quad (3.20)$$

and c_0, c_1 are nonnegative numbers related to $\epsilon_0, \epsilon_1, f_0^0$, and f_1^0 [22, 26]. Note that if choosing $c_0 = 0$ and $c_1 = +\infty$, the test is the conventional LRT with respect to nominal PDFs, f_0^0 and f_1^0 .

One special case is the robust matched filtering. We turn the model (3.10) into a vector form as

$$\mathcal{H}_0 : \mathbf{x} = \boldsymbol{\eta} \quad (3.21)$$

$$\mathcal{H}_1 : \mathbf{x} = \mathbf{s} + \boldsymbol{\eta} \quad (3.22)$$

where \mathbf{s} is the signal vector and $\boldsymbol{\eta}$ is the noise vector. Suppose that \mathbf{s} is known. In general, a matched-filtering detection is $T_{MF} = \mathbf{g}^T \mathbf{x}$. Let the covariance matrix of the noise be $\mathbf{R}_\eta = \mathbb{E}(\boldsymbol{\eta}\boldsymbol{\eta}^T)$. If $\mathbf{R}_\eta = \sigma_\eta^2 \mathbf{I}$, it is known that choosing $\mathbf{g} = \mathbf{s}$ is optimal. In general, it is easy to verify that the optimal \mathbf{g} to maximize the SNR is

$$\mathbf{g} = \mathbf{R}_\eta^{-1} \mathbf{s}. \quad (3.23)$$

In practice, the signal vector \mathbf{s} may not be known exactly. For example, \mathbf{s} may be only known to be around \mathbf{s}_0 with some errors modeled by

$$\|\mathbf{s} - \mathbf{s}_0\| \leq \Delta \quad (3.24)$$

where Δ is an upper bound on the Euclidean-norm of the error. In this case, we are interested in finding a proper value for \mathbf{g} such that the worst-case SNR is maximized, i.e.,

$$\hat{\mathbf{g}} = \arg \max_{\mathbf{g}} \min_{\mathbf{s}: \|\mathbf{s} - \mathbf{s}_0\| \leq \Delta} \text{SNR}(\mathbf{s}, \mathbf{g}) \quad (3.25)$$

It was proved in [24, 25] that the optimal solution for the above maxmin problem is

$$\hat{\mathbf{g}} = (\mathbf{R}_\eta + \delta \mathbf{I})^{-1} \mathbf{s}_0 \quad (3.26)$$

where δ is a nonnegative number such that $\delta^2 \|\hat{\mathbf{g}}\|^2 = \Delta$.

It is noted that there are also researches on the robust matched filtering detection when the signal has other types of uncertainty [26]. Moreover, if the noise has uncertainties, i.e., \mathbf{R}_η is not known exactly, or both noise and signal have uncertainties, the optimal robust matched-filtering detector was also found for some specific uncertainty models in [26].

3.2.4 Energy Detection

If we further assume that $\mathbf{R}_s = \sigma_s^2 \mathbf{I}$, the EC detection in (3.14) reduces to the well-known energy detector (ED) [4, 27, 28] for which the test statistic is given as follows (by discarding irrelevant constant terms).

$$T_{ED} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{x}^T(n) \mathbf{x}(n) \quad (3.27)$$

Note that for the multi-antenna/node case, T_{ED} is actually the summation of energies from all antennas/nodes, which is a straightforward cooperative sensing scheme [29–31].

The test statistic is compared with a threshold to make to decision. Obviously the threshold should be related to the noise power. Hence energy detection needs a priori information of the noise variance (power). It has been shown that energy detection is very sensitive to the inaccurate estimation of the noise power. We will give detailed discussion on this later.

From the derivation above, we know that *energy detection is the optimal detection if there is only one antenna, the signal and noise samples are independent and identically distributed (iid) Gaussian random variables, and the noise variance (power) is known*. Even if the signal and noise do not have Gaussian distribution, in most cases, energy detection is still approximately optimal for uncorrelated signal and noise at low SNR [32]. In general, the ED is not optimal if signals or noise samples are correlated.

The energy detection can be used in different ways and sometimes combined with other techniques.

(1) We can filter the signal before the energy detection is implemented. Let $f(l)$ ($l = 0, 1, \dots, L$) be a filter or the combining of a bank of filters. The received signal after filtering is

$$y_i(n) = \sum_{l=0}^L f(l)x_i(n-l) \quad (3.28)$$

The energy detection after the filtering is therefore

$$T_{ED, Filter} = \frac{1}{N} \sum_{n=0}^{N-1} \|\mathbf{y}(n)\|^2 \quad (3.29)$$

For practical applications, we can choose a narrowband filter or a bank of narrowband filters if we want to detect the signals in specific frequency bands.

(2) Energy detection can also be done in frequency domain. Let $S_i(k)$ be the power spectrum density (PSD) of the received signal $x_i(n)$. There are different methods to estimate the PSD including periodogram, multitaper method (MTM) [33, 34] and others. For the periodogram method, the received signal is divided into P non-overlapped

blocks: $x_{i,p}(n)$ with length N_f . Let $X_{i,p}(k)$ be the discrete Fourier transform (DFT) of $x_{i,p}(n)$. Note that DFT can be computed by the fast Fourier transform (FFT). The PSD is estimated as

$$S_i(k) = \frac{1}{P} \sum_{p=1}^P |X_{i,p}(k)|^2 \quad (3.30)$$

The test statistic of the frequency domain energy detection is:

$$T_{ED,F} = \frac{1}{NM} \sum_{i=1}^M \sum_{k=0}^{N_f-1} S_i(k) \quad (3.31)$$

Again the test statistic is compared with a threshold to make a decision and the threshold should be related to the noise power.

Among all the spectral estimation methods, the MTM is proved to achieve the performance close to the maximum likelihood PSD estimator [33, 34]. Thus it can be used to make a more accurate estimation of the PSD for spectrum sensing [35]. However, the computational complexity is also increased.

(3) The frequency domain energy detection can also be done in a more flexible way. Let ψ be a subset of set $\{0, 1, \dots, N_f - 1\}$. We can select the signal frequency within the bin ψ for detection and may also give different weights to different antennas and frequencies [36]. The test statistic is therefore

$$T_{ED,F} = \frac{1}{M|\psi|} \sum_{i=1}^M \sum_{k \in \psi} g_{i,k} S_i(k) \quad (3.32)$$

where $g_{i,k}$ is the weight for antenna i and frequency index k . This can have a better performance if we know that the interested signal power has peaks or is concentrated in the frequency bin ψ . For example, for ATSC signal detection, we know that the signal has a strong peak at the pilot. So we can choose the set ψ to be the frequency index around the pilot location. A special case is to choose just one frequency index nearest to the pilot location. In some OFDM based standards, the pilot subcarriers have higher power than other subcarriers. We can put more weights to the pilot subcarriers.

Another variation of the method is to change the averaging of the signal PSD to maximizing [37]. The test statistic becomes

$$T_{ED,Max} = \max_{k \in \psi} \frac{1}{M} \sum_{i=1}^M S_i(k) \quad (3.33)$$

The energy detection can also be done in other transform domains other than the Fourier transform domain. In general, let $X_{i,T}(k)$ ($k = 0, 1, \dots, K - 1$) be the transformed signal of the original received signal $x_i(n)$. The transform domain energy

detection is

$$T_{ED,T} = \frac{1}{MK} \sum_{i=1}^M \sum_{k=1}^K |X_{i,T}(k)|^2 \quad (3.34)$$

For example, the wavelet can be chosen as the transform, which turns to the wavelet multi-resolution detection [38]. In general, the transform should be chosen based on the signal and noise property or/and purpose of detection.

3.2.5 Sequential Energy Detection

In the discussions above, we assume that the sensing time (number of signal samples N) is a predefined fixed number. The detector is designed to have the optimal performance with the given sample size. In some applications, the sensing time may not be predefined. Our purpose is to design a detector that has the least average sensing time. A popular approach for this is the sequential detection [39–41].

In general, a sequential detector makes a decision whenever a new signal sample is available. For simplicity, here we consider single detector case. Let $\mathbf{x}_k = (x(0), x(1), \dots, x(k-1))^T$ be signal samples currently available at time k . The sequential detector calculates a test statistic $T(\mathbf{x}_k)$. It then makes a decision using two thresholds:

$$T(\mathbf{x}_k) \geq \gamma_1, \mathcal{H}_1 \quad (3.35)$$

$$T(\mathbf{x}_k) \leq \gamma_0, \mathcal{H}_0 \quad (3.36)$$

where $\gamma_1 > \gamma_0$ are some predefined thresholds. If the test statistic is between γ_1 and γ_0 , that is, $\gamma_0 < T(\mathbf{x}_k) < \gamma_1$, the detector does not make a decision on \mathcal{H}_1 or \mathcal{H}_0 yet; more samples are required. The detector will continue this process when new signal samples are available till a decision on \mathcal{H}_1 or \mathcal{H}_0 is reached.

Thus we do not know when the detection will finished to achieve the target sensing performance. The sensing time is a random number. It is proved that the average sensing time is shorter than the conventional methods (the Wald–Wolfowitz theorem [40]). However, the worst case sensing time could be much longer.

Using the energy detection for sequential detection is discussed in [40]. Some extensions and refinements can be found in [40].

3.2.6 Matched Filtering

If we assume that noise is Gaussian distributed and source signal $\mathbf{s}(n)$ is deterministic and known to the receiver, it is easy to show that the LRT in this case becomes the matched filtering based detector [16–18], for which the test statistic is

$$T_{MF} = R_e \left(\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{s}^\dagger(n) \mathbf{x}(n) \right) \quad (3.37)$$

The test statistic is compared with a threshold to make a decision. Obviously the threshold should be related to the noise power. Unlike energy detection, matched filtering (MF) is more robust to the inaccurate noise power estimation [6, 7]. MF is also widely used in other fields like radar signal processing.

Hence, theoretically *matched filtering is optimal if the signal is deterministic and known at the receiver*. The major difficulties in MF are the time delay, frequency offset and time dispersive channel.

For simplicity, we consider single antenna case here. In general, for a deterministic transmitted signal $s(n)$, the received signal $x(n)$ can be written as

$$x(n) = \frac{e^{j2\pi\epsilon n}}{\sqrt{N}} \sum_{l=0}^L h(l) s(n-1-\tau) + \eta(n) \quad (3.38)$$

where τ is the timing error, ϵ is the normalized frequency offset and $h(l)$ is the channel.

At ideal case of $\epsilon = 0$, $\tau = 0$, $L = 0$ and $h(0) > 0$,

$$T_{MF} = \frac{h(0)}{\sqrt{N}} \sum_{n=0}^{N-1} |s(n)|^2 + R_e \left(\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} s^*(n) \eta(n) \right) \quad (3.39)$$

In this case, MF is optimal.

In practical wireless communication applications, the CFO ϵ and timing error may not be zero, also the channel is most likely frequency selective: $L > 0$. So in general, the test statistic of MF should be expressed as

$$T_{MF} = R_e \left(\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{l=0}^L e^{j2\pi\epsilon n} s^*(n) s(n-l-\tau) \right) + R_e \left(\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} s^*(n) \eta(n) \right) \quad (3.40)$$

The CFO ϵ , timing error and frequency selective are three major obstacles for the MF. Anyone of them could reduce the performance of MF dramatically.

To deal with the timing error, a commonly used solution is to averaging or taking the maximum of the test statistic at different time delays of the received signal. Let

$$T_{MF}(v) = R_e \left(\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathbf{s}^\dagger(n) \mathbf{x}(n+v) \right) \quad (3.41)$$

be the test statistic of the signal with time delay $-v$. Obviously the best value is $v = \tau$. If we do not know the value of τ , we can average on different v or take the maximum, that is,

$$\hat{T}_{MF,A} = \frac{1}{2\Delta} \sum_{\nu=-\Delta}^{\Delta} T_{MF}(\nu) \quad (3.42)$$

or

$$\hat{T}_{MF,M} = \max_{\nu=-\Delta}^{\Delta} T_{MF}(\nu) \quad (3.43)$$

To tack the problem of CFO, we can modify the test of MF using the absolute value

$$T_{MF} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} |\mathbf{s}^\dagger(n)\mathbf{x}(n)| \quad (3.44)$$

This test is not affected by the carrier frequency offset.

Similarly we can also average or take the maximum of the absolute value test statistics to deal with the timing error problem.

Like the energy detection, the MF can also be implemented in frequency domain or transformed domain. Using the MF for ATSC signal detection is discussed in [42].

3.2.7 Cyclostationary Detection

Practical communication signals may have special statistical features. For example, digitally modulated signals have non-random components such as double sidedness due to sine wave carrier and keying rate due to symbol period. Such signals have a special statistical feature called cyclostationarity, i.e., their statistical parameters vary periodically in time. This cyclostationarity can be extracted by the cyclic auto-correlation (CAC) or the spectral-correlation density (SCD) [43–45].

For simplicity, in this section we consider single antenna case, that is, $M = 1$. For notation simplicity, we omit the subscript for antenna. For a given α and time lag τ , the CAC of a signal $x(t)$ is defined as

$$R_x^\alpha(\tau) = \lim_{\Delta \rightarrow \infty} \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\alpha t} dt \quad (3.45)$$

where α is called a cyclic frequency. If there exists at least one non-zero α such that $\max_\tau |R_x^\alpha(\tau)| > 0$, we say that $x(t)$ exhibits cyclostationarity. The value of such α depends on the type of modulation, symbol duration, etc. For example, for a digitally modulated signal with symbol duration T_b , cyclostationary features exist at $\alpha = \frac{k}{T_b}$ and $\alpha = \pm 2f_c + \frac{k}{T_b}$, where f_c is the carrier frequency, and k is an integer. Equivalently, we can define the SCD, the Fourier transform of the CAC, as follows:

$$S_x^\alpha(f) = \int_{-\infty}^{\infty} R_x^\alpha(\tau) e^{-j2\pi f \tau} d\tau \quad (3.46)$$

In binary spectrum sensing or signal detection, there are two hypotheses: \mathcal{H}_0 , signal absent; and \mathcal{H}_1 , signal present. The received signal can be written as

$$\mathcal{H}_0 : y(t) = \eta(t) \quad (3.47)$$

$$\mathcal{H}_1 : y(t) = h(t) \otimes x(t) + \eta(t) \quad (3.48)$$

where $x(t)$ denotes the transmitted signal from the primary user, $h(t)$ is the channel response, and $\eta(t)$ is the additive noise.

When source signal $x(t)$ passes through a wireless channel $h(t)$, the received signal is impaired by the unknown propagation channel. It can be shown that the SCD function of $y(t)$ is

$$S_y(f) = H(f + \alpha/2)H^*(f - \alpha/2)S_x(f) \quad (3.49)$$

where $*$ denotes the conjugate, α denotes the cyclic frequency for $x(t)$, $H(f)$ is the Fourier transform of the channel $h(t)$, and $S_x(f)$ is the SCD function of $x(t)$. Thus, the unknown channel could have major impacts on the strength of SCD at certain cyclic frequencies.

Cyclostationary detection (CSD) is well studied when Nyquist rate signal samples are available. The rationale behind the CSD is that the signal $x(t)$ has cyclostationarity, that is, there exists at least one non-zero cyclic frequency α such that $R_x^\alpha(\tau) \neq 0$ for some τ , while the noise $\eta(t)$ is a pure stationary process, that is, for any non-zero α $R_\eta^\alpha(\tau) = 0$ for all τ 's, or equivalently $S_\eta^\alpha(f) = 0$ for all f 's. In the following, we list the cyclic frequencies for some signals with cyclostationarity in practical applications [44, 45].

1. Analog TV signal: It has cyclic frequencies at multiples of the TV-signal horizontal line-scan rate (15.75 KHz in USA, 15.625 KHz in Europe).
2. AM signal: $x(t) = a(t) \cos(2\pi f_c t + \phi_0)$. It has cyclic frequencies at $\pm 2f_c$.
3. PM and FM signal: $x(t) = \cos(2\pi f_c t + \phi(t))$. It usually has cyclic frequencies at $\pm 2f_c$. The characteristics of the SCD function at cyclic frequency $\pm 2f_c$ depend on $\phi(t)$.
4. Digital modulated signals:
 - a. Amplitude-Shift Keying: $x(t) = [\sum_{n=-\infty}^{\infty} a_n p(t - n\Delta_s - t_0)] \cos(2\pi f_c t + \phi_0)$. It has cyclic frequencies at k/Δ_s , $k \neq 0$ and $\pm 2f_c + k/\Delta_s$, $k = 0, \pm 1, \pm 2, \dots$
 - b. Phase-Shift Keying: $x(t) = \cos[2\pi f_c t + \sum_{n=-\infty}^{\infty} a_n p(t - n\Delta_s - t_0)]$. For BPSK, it has cyclic frequencies at k/Δ_s , $k \neq 0$, and $\pm 2f_c + k/\Delta_s$, $k = 0, \pm 1, \pm 2, \dots$. For QPSK, it has cyclic frequencies at k/Δ_s , $k \neq 0$.
5. OFDM modulated signal: It has cyclic features at cyclic frequency $k/(N_t T_b)$, where k is an integer number, N_t is the length of an OFDM block (FFT size plus CP size), and T_b is the symbol duration.

Let α_0 be a non-zero cyclic frequency such that $R_x^{\alpha_0}(\tau) \neq 0$ for some τ . Assume that the signal and noise are mutually independent. Then we have

$$\mathcal{H}_0 : R_y^{\alpha_0}(\tau) = 0 \quad (3.50)$$

$$\mathcal{H}_1 : R_y^{\alpha_0}(\tau) = R_x^{\alpha_0}(\tau) \neq 0, \text{ for some } \tau \quad (3.51)$$

In the frequency domain, this turns to

$$\mathcal{H}_0 : S_y^{\alpha_0}(f) = 0 \quad (3.52)$$

$$\mathcal{H}_1 : S_y^{\alpha_0}(f) = S_x^{\alpha_0}(f) \neq 0, \text{ for some } f \quad (3.53)$$

Therefore, \mathcal{H}_0 and \mathcal{H}_1 can be distinguished by generating a test statistic from the CAC/SCD of the received signal at cyclic frequency α_0 and comparing the test statistic with a threshold. A typical test statistic is $C_1 = \int |R_y^{\alpha_0}(\tau)|^2 d\tau$ or equivalently $C_1 = \int |S_y^{\alpha_0}(f)|^2 df$.

In practice, the received signal is sampled and only limited number of samples are available. Let T_s be the sampling period and N be the number of samples. The discrete version of the CAC is

$$R_y^{\alpha}(kT_s) = \frac{1}{N} \sum_{n=0}^{N-1} y((n+k)T_s) y^*(nT_s) e^{-j2\pi\alpha nT_s} \quad (3.54)$$

where the lag $k = 0, 1, \dots, M-1$ with $M \ll N$. Accordingly the discrete version of the test statistic is

$$C_1 = \sum_{k=0}^{M-1} |R_y^{\alpha_0}(kT_s)|^2 \quad (3.55)$$

In CSD, the test statistic is compared with a threshold to make a decision. Intuitively the threshold should be related to noise power. Due to the difficulty in acquiring the accurate noise power in practice [7, 46, 47], we can use the maximum likelihood estimation of the noise power. The maximum likelihood estimation of the noise power is

$$\hat{\sigma}_\eta^2 = \frac{1}{N} \sum_{n=0}^{N-1} |y(nT_s)|^2 \quad (3.56)$$

The threshold is thus chosen as $\beta \hat{\sigma}_\eta^4$, where β is a scalar to meet the pre-defined probability of false alarm.

There are other different test statistics and decision rules (thresholds) for the CSD. Especially, if the signal has cyclostationarity at multiple cyclic frequencies, how to use them to form a single test statistic is an interesting problem. In [48, 49], a general structure based on the GLRT principle is proposed to use the multiple cyclic frequencies. However, the method needs very high complexity and also some priori

information on the channel. Some simplified approaches have also been studied [50]. The use of the CSD for ATSC signal is proposed [51]. There are also researches on OFDM signal detections using the CSD [50, 52–54].

When interference exists, the CSD may still work well as long as the interference does not have the same cyclostationary feature as the primary signal. In general, the chance that the primary signal and the interference have the same cyclostationary feature is slim. That means CSD is robust to interference and noise uncertainty. Furthermore, it is possible to distinguish the signal type because different signal may have different non-zero cyclic frequencies.

Although cyclostationary detection has certain advantages, it also has some disadvantages:

1. The method needs a very high sampling rate;
2. The computation of SCD function requires large number of samples and thus high computational complexity;
3. The strength of SCD could be affected by the unknown channel [46];
4. The sampling time error and frequency offset could affect the cyclic frequencies [55, 56], which will be discussed further in the next section.

3.2.8 Detection Threshold and Test Statistic Distribution

To make a decision on whether signal is present, we need to set a threshold γ for each proposed test statistic, such that certain P_d and/or P_{fa} can be achieved. For a fixed sample size N , we cannot set the threshold to meet the targets for arbitrarily high P_d and low P_{fa} at the same time, as they are conflicting to each other. Since we have little or no prior information on the signal (actually we even do not know whether there is a signal or not), it is difficult to set the threshold based on P_d . Hence, a common practice is to choose the threshold based on P_{fa} under hypothesis \mathcal{H}_0 .

Without loss of generality, the test threshold can be decomposed into the following form: $\gamma = \gamma_1 T_0(\mathbf{x})$, where γ_1 is related to the sample size N and the target P_{fa} , and $T_0(\mathbf{x})$ is a statistic related to the noise distribution under \mathcal{H}_0 . For example, for the energy detection with known noise power, we have

$$T_0(\mathbf{x}) = \sigma_\eta^2 \quad (3.57)$$

For the matched-filtering detection with known noise power, we have

$$T_0(\mathbf{x}) = \sigma_\eta \quad (3.58)$$

In practice, the parameter γ_1 can be set either empirically based on the observations over a period of time when the signal is known to be absent, or analytically based on the distribution of the test statistic under \mathcal{H}_0 . In general, such distributions are difficult to find, while some known results are given as follows.

For energy detection defined in (3.27), it can be shown that for a sufficiently large values of N , its test statistic can be well approximated by the Gaussian distribution [14, 28], i.e.,

$$\frac{1}{NM} T_{ED}(\mathbf{x}) \sim \mathcal{N}\left(\sigma_\eta^2, \frac{2\sigma_\eta^4}{NM}\right) \quad \text{under } \mathcal{H}_0 \quad (3.59)$$

Accordingly, for given P_{fa} and N , the corresponding γ_1 can be found as

$$\gamma_1 = NM \left(\sqrt{\frac{2}{NM} Q^{-1}(P_{fa}) + 1} \right) \quad (3.60)$$

where

$$Q(t) = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-u^2/2} du \quad (3.61)$$

For the matched-filtering detection defined in (3.37), for a sufficiently large N , we have

$$\frac{1}{\sqrt{\sum_{n=0}^{N-1} \|\mathbf{s}(n)\|^2}} T_{MF}(\mathbf{x}) \sim \mathcal{N}(0, \sigma_\eta^2) \quad \text{under } \mathcal{H}_0 \quad (3.62)$$

Thereby, for given P_{fa} and N , it can be shown that

$$\gamma_1 = Q^{-1}(P_{fa}) \sqrt{\sum_{n=0}^{N-1} \|\mathbf{s}(n)\|^2} \quad (3.63)$$

For the GLRT-based detection, it can be shown that the asymptotic (as $N \rightarrow \infty$) log-likelihood ratio is central chi-square distributed [16]. More precisely,

$$2 \ln T_{GLRT}(\mathbf{x}) \sim \chi_r^2 \quad \text{under } \mathcal{H}_0 \quad (3.64)$$

where r is the number of independent scalar unknowns under \mathcal{H}_0 and \mathcal{H}_1 . For instance, if σ_η^2 is known while \mathbf{R}_s is not, r will be equal to the number of independent real-valued scalar variables in \mathbf{R}_s . However, there is no explicit expression for γ_1 in this case.

3.3 Eigenvalue Based Detections

Eigenvalue based detections (EBD) was first proposed in [47, 57–60]. The method was later studied and refined in [19, 61–64]. EBD can be derived from different approaches such as the GLRT principle or information theory. Some examples on the derivations can be found in [19, 20, 64]. The threshold setting of the EBD needs

random matrix theory [47, 57–60]. The EDB methods solve the noise uncertainty problem by using statistical covariance matrix to estimate the noise power. The method can detect signal without knowing explicit information of the signal. The method was also adopted by IEEE802.22 standard as a solution to detect TV and wireless microphone signals.

3.3.1 The Methods

We consider the same model as defined at the beginning of this chapter. Let $N_j \stackrel{\text{def}}{=} \max_i(q_{ij})$, zero-pad $h_{ij}(k)$ if necessary, and define

$$\mathbf{h}_j(n) \stackrel{\text{def}}{=} [h_{1j}(n), h_{2j}(n), \dots, h_{Mj}(n)]^T \quad (3.65)$$

We have [47]

$$\mathbf{x}(n) = \mathbb{H}\mathbf{s}(n) + \boldsymbol{\eta}(n) \quad (3.66)$$

where \mathbb{H} is a $ML \times (\hat{N} + PL)$ ($\hat{N} \stackrel{\text{def}}{=} \sum_{j=1}^P N_j$) matrix defined as

$$\mathbb{H} \stackrel{\text{def}}{=} [\mathbb{H}_1, \mathbb{H}_2, \dots, \mathbb{H}_P], \quad (3.67)$$

$$\mathbb{H}_j \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{h}_j(0) & \cdots & \cdots & \mathbf{h}_j(N_j) & 0 & \cdots & 0 \\ 0 & \mathbf{h}_j(0) & \cdots & \cdots & \mathbf{h}_j(N_j) & \cdots & 0 \\ & & \ddots & & & \ddots & \\ 0 & 0 & \cdots & \mathbf{h}_j(0) & \cdots & \cdots & \mathbf{h}_j(N_j) \end{bmatrix} \quad (3.68)$$

Note that the dimension of \mathbb{H}_j is $ML \times (N_j + L)$.

Define the statistical covariance matrices of the signals and noise as

$$\mathbf{R}_x = \mathbb{E}(\mathbf{x}(n)\mathbf{x}^\dagger(n)) \quad (3.69)$$

$$\mathbf{R}_s = \mathbb{E}(\mathbf{s}(n)\mathbf{s}^\dagger(n)) \quad (3.70)$$

$$\mathbf{R}_\eta = \mathbb{E}(\boldsymbol{\eta}(n)\boldsymbol{\eta}^\dagger(n)) \quad (3.71)$$

We can verify that

$$\mathbf{R}_x = \mathbb{H}\mathbf{R}_s\mathbb{H}^\dagger + \sigma_\eta^2\mathbf{I}_{ML} \quad (3.72)$$

where σ_η^2 is the variance of the noise, and \mathbf{I}_{ML} is the identity matrix of order ML .

Let the eigenvalues of \mathbf{R}_x and $\mathbb{H}\mathbf{R}_s\mathbb{H}^\dagger$ be $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{ML}$ and $\rho_1 \geq \rho_2 \geq \dots \geq \rho_{ML}$, respectively. Obviously, $\lambda_n = \rho_n + \sigma_\eta^2$. When there is no signal, that is, $\mathbf{s}(n) = 0$ (then $\mathbf{R}_s = 0$), we have $\lambda_1 = \lambda_2 = \dots = \lambda_{ML} = \sigma_\eta^2$. Hence, $\lambda_1/\lambda_{ML} = 1$. When there is a signal, if $\rho_1 > \rho_{ML}$, we have $\lambda_1/\lambda_{ML} > 1$. Hence, we can detect if signal exists by checking the ratio λ_1/λ_{ML} . Obviously, $\rho_1 = \rho_{ML}$ if and only if $\mathbb{H}\mathbf{R}_s\mathbb{H}^\dagger = \lambda\mathbf{I}_{ML}$, where λ is a positive number. From the definition of the matrix \mathbb{H} and \mathbf{R}_s , it is highly probable that $\mathbb{H}\mathbf{R}_s\mathbb{H}^\dagger \neq \lambda\mathbf{I}_{ML}$. In fact, the worst case is $\mathbf{R}_s = \sigma_s^2\mathbf{I}$, that is, the source signal samples are iid. At this case, $\mathbb{H}\mathbf{R}_s\mathbb{H}^\dagger = \sigma_s^2\mathbb{H}\mathbb{H}^\dagger$. Obviously, $\sigma_s^2\mathbb{H}\mathbb{H}^\dagger = \lambda\mathbf{I}_{ML}$ if and only if all the rows of \mathbb{H} have the same power and they are co-orthogonal. This is only possible when $N_j = 0$, $j = 1, \dots, P$ and $M = 1$, that is, the source signal samples are iid, all the channels are flat-fading and there is only one receiver.

Thus, if $M > 1$ (multiple antennas) or the channel has multiple paths or the source signal itself is correlated, the eigenvalues of the \mathbf{R}_x are not identical, while at pure noise case, the \mathbf{R}_x should have identical eigenvalues. Hence, we can check the eigenvalues of \mathbf{R}_x to see if signal presents.

In practice, we only have finite number of samples. Hence, we can only obtain the sample covariance matrix other than the statistic covariance matrix. The sample covariance matrix is defined as

$$\mathbf{R}_x(N) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=L-1}^{L-2+N} \mathbf{x}(n)\mathbf{x}^\dagger(n) \quad (3.73)$$

where N is the number of collected samples. Based on the sample covariance matrix and its eigenvalues, a few methods have been proposed based on different perspectives [19, 47, 57–64]. Such methods are called eigenvalue based detections (EBD). Here we summarize the methods as follows.

Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{ML}$ be the eigenvalues of the sample covariance matrix.

Algorithm Eigenvalue based detections

Step 1. Compute the sample covariance matrix as defined in (3.73).

Step 2. Calculate the eigenvalues of the sample covariance matrix.

Step 3. Compute a test statistic from the eigenvalues. There are different approaches to construct the test statistic. A few simple but effective method are as follows:

1. Maximum eigenvalue to trace detection (MET). The test statistic is

$$T_{MET} = \lambda_1/t_r(\mathbf{R}_x(N)) \quad (3.74)$$

where $t_r(\cdot)$ is the trace of a matrix, $t_r(\mathbf{R}_x(N)) = \sum_{i=1}^{ML} \lambda_i$. This method is also called blindly combined energy detection (BCED) in [60].

2. Maximum to minimum eigenvalue detection (MME) [47]. The test statistic is

$$T_{MME} = \lambda_1/\lambda_{ML} \quad (3.75)$$

3. **Arithmetic to geometric mean (AGM)** [19]. The test statistic is

$$T_{AGM} = \frac{1}{ML} \sum_{i=1}^{ML} \lambda_i / \left(\prod_{i=1}^{ML} \lambda_i \right)^{1/ML} \quad (3.76)$$

Step 4. Compare the test statistic with a threshold to make a decision.

All these methods do not use the information of the signal, channel and noise power as well. The methods are robust to synchronization error, channel impairment, and noise uncertainty.

3.3.2 Threshold Setting

To find a formula for the threshold is mathematically involved. In general we need to find the theoretical distribution of some combination of the eigenvalues of a random matrix. There have been some exciting works on this by using the random matrix theory [47, 61–63, 65–68]. For simplicity, in the following, we provide an example for the maximum eigenvalue detection (MED) with known noise power [59]. At this case, we actually compare the ratio of the maximum eigenvalue of the sample covariance matrix $\mathbf{R}_x(N)$ to the noise power σ_η^2 with a threshold γ_1 . To set the value for γ_1 , we need to know the distribution of $\lambda_1(N)/\sigma_\eta^2$ for any finite N . Fortunately, the random matrix theory has laid the foundation to derive the distributions.

When there is no signal, $\mathbf{R}_x(N)$ reduces to $\mathbf{R}_\eta(N)$, which is the sample covariance matrix of the noise only. It is known that $\mathbf{R}_\eta(N)$ is a Wishart random matrix [69]. The study of the eigenvalue distributions for random matrices is a very hot research topic over recent years in mathematics, communications engineering, and physics [69–72]. The joint PDF of the ordered eigenvalues of a Wishart random matrix has been known for many years [69]. However, since the expression of the joint PDF is very complicated, no simple closed-form expressions have been found for the marginal PDFs of the ordered eigenvalues, although some computable expressions have been found in [73]. Recently, I. M. Johnstone and K. Johansson have found the distribution of the largest eigenvalue [70, 71] of a Wishart random matrix as described in the following theorem.

Theorem 3.1 Let $\mathbf{A}(N) = \frac{N}{\sigma_\eta^2} \mathbf{R}_\eta(N)$, $\mu = (\sqrt{N-1} + \sqrt{M})^2$, and $\nu = (\sqrt{N-1} + \sqrt{M})(\frac{1}{\sqrt{N-1}} + \frac{1}{\sqrt{M}})^{1/3}$. Assume that $\lim_{N \rightarrow \infty} \frac{M}{N} = y$ ($0 < y < 1$). Then, $\frac{\lambda_{\max}(\mathbf{A}(N)) - \mu}{\nu}$ converges (with probability one) to the Tracy–Widom distribution of order 1 [74, 75].

Table 3.1 Numerical table for the Tracy–Widom distribution of order 1

t	-3.90	-3.18	-2.78	-1.91	-1.27	-0.59	0.45	0.98	2.02
$F_1(t)$	0.01	0.05	0.10	0.30	0.50	0.70	0.90	0.95	0.99

The Tracy–Widom distribution provides the limiting law for the largest eigenvalue of certain random matrices [74, 75]. Let F_1 be the cumulative distribution function (CDF) of the Tracy–Widom distribution of order 1. We have

$$F_1(t) = \exp\left(-\frac{1}{2} \int_t^\infty (q(u) + (u-t)q^2(u)) du\right) \quad (3.77)$$

where $q(u)$ is the solution of the nonlinear Painlevé II differential equation given by

$$q''(u) = uq(u) + 2q^3(u) \quad (3.78)$$

Accordingly, numerical solutions can be found for function $F_1(t)$ at different values of t . Also, there have been tables for values of $F_1(t)$ [70] as shown in Table 3.1.

Using the above results, we can derive the probability of false alarm as

$$\begin{aligned} P_{fa} &= P(\lambda_1(N) > \gamma_1 \sigma_\eta^2) \\ &= P\left(\frac{\lambda_{\max}(\mathbf{A}(N)) - \mu}{\nu} > \frac{\gamma_1 N - \mu}{\nu}\right) \approx 1 - F_1\left(\frac{\gamma_1 N - \mu}{\nu}\right) \end{aligned} \quad (3.79)$$

Thus we have

$$F_1\left(\frac{\gamma_1 N - \mu}{\nu}\right) \approx 1 - P_{fa} \quad (3.80)$$

or equivalently,

$$\frac{\gamma_1 N - \mu}{\nu} \approx F_1^{-1}(1 - P_{fa}) \quad (3.81)$$

From the definitions of μ and ν in Theorem 3.1, we finally obtain the value for γ_1 as

$$\gamma_1 \approx \frac{(\sqrt{N} + \sqrt{M})^2}{N} \left(1 + \frac{(\sqrt{N} + \sqrt{M})^{-2/3}}{(NM)^{1/6}} F_1^{-1}(1 - P_{fa})\right) \quad (3.82)$$

Note that γ_1 depends only on N and P_{fa} . A similar approach like the above can be used for the case of MME detection, as shown in [47, 68].

Figure 3.1 shows the expected (theoretical) and actual (by simulation) probability of false alarm values based on the theoretical threshold in (3.82) for $N = 5000$,

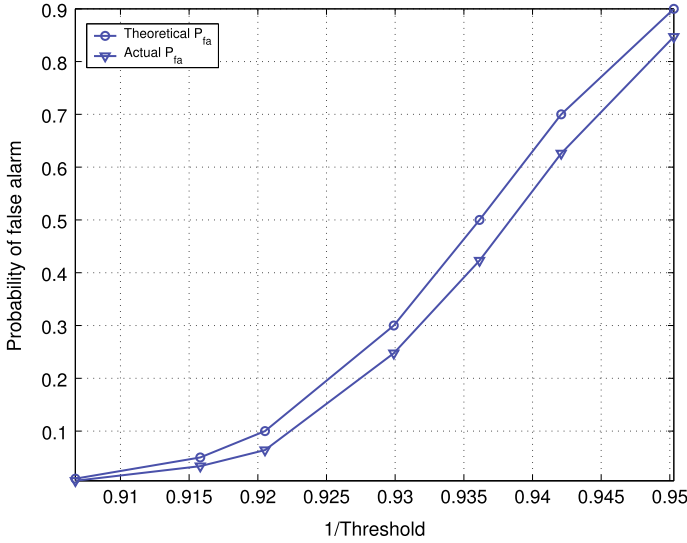


Fig. 3.1 Comparison of theoretical and actual P_{fa}

$M = 8$, and $L = 1$. It is observed that the differences between these two sets of values are reasonably small, suggesting that the choice of the theoretical threshold is quite accurate.

3.3.3 Performances of the Methods

To show the performance and the robustness of the methods, here we give some simulation results for the EBDs. Comparison with the energy detection (ED) is also included. We consider two cases here: the signal is time uncorrelated and the signal is time correlated. The Receiver Operating Characteristics (ROC) curves (P_d versus P_{fa}) at $\text{SNR} = -15$ dB, $N = 5000$, and $M = 4$ are plotted at the two cases. The performance at first case is shown in Fig. 3.2 with $L = 1$ and that at the second case is shown in Fig. 3.3 with $L = 6$, where “ED- u dB” means energy detection with u dB noise uncertainty. In Fig. 3.3, the source signal is the wireless microphone signal [76] and a multipath fading channel (with eight independent taps of equal power) is assumed. For both cases, MET, MME and AGM perform better than ED. MET, MME and AGM are totally immune to noise uncertainty. However, the ED is very vulnerable to noise power uncertainty [4–6].

Obviously the eigenvalue based detections do not use the information of the signal, channel and noise power as well. The methods are robust to synchronization error, channel impairment, and noise uncertainty. However, like other blind detections, the methods are vulnerable to unknown narrowband interferences.

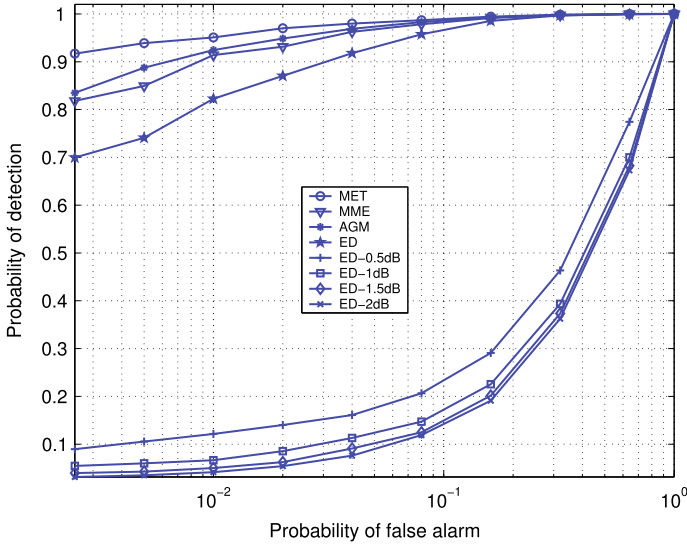


Fig. 3.2 ROC curve: i.i.d source signal

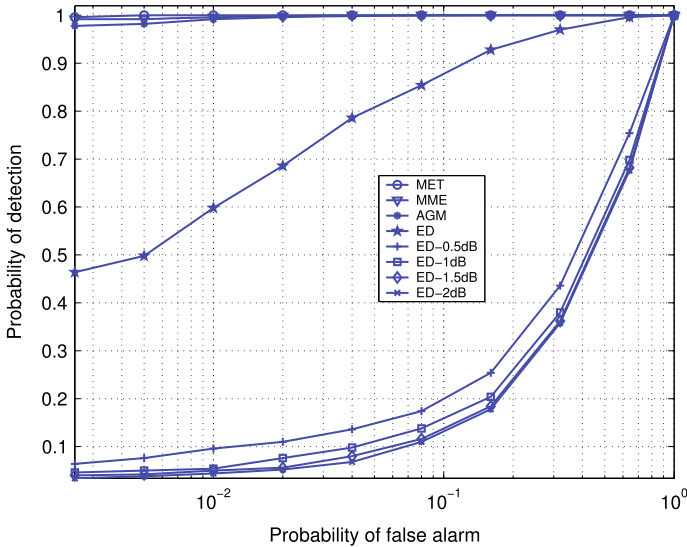


Fig. 3.3 ROC curve: wireless microphone source signal

3.4 Covariance Based Detections

Covariance based detections (CBD) was first proposed in [65, 77]. The method solved the noise uncertainty problem by using the online estimated noise power. The method can detect signal without knowing explicit information of the signal. The

method was also adopted by IEEE802.22 standard for detecting TV signal and as the first choice for sensing the wireless microphone signals.

3.4.1 The Methods

As shown in the last section, the covariance matrix of the received signal can be written as

$$\mathbf{R}_x = \mathbb{H}\mathbf{R}_s\mathbb{H}^\dagger + \sigma_\eta^2\mathbf{I}_{ML} \quad (3.83)$$

If the signal $s(n)$ is not present, $\mathbf{R}_s = 0$. Hence the off-diagonal elements of \mathbf{R}_x are all zeros. If there is signal and the signal samples are correlated, \mathbf{R}_s is not a diagonal matrix. Hence, some of the off-diagonal elements of \mathbf{R}_x should be non-zeros.

In practice, the statistical covariance matrix can only be calculated using a limited number of signal samples. For notation simplicity, here we consider the case of single antenna/sensor $M = 1$, and drop the indices for antenna/sensor. Define the sample auto-correlations of the received signal as

$$r(l) = \frac{1}{N_s} \sum_{m=0}^{N_s-1} x(m)x(m-l), \quad l = 0, 1, \dots, L-1 \quad (3.84)$$

where $x(m)$ is the received signal samples, and N_s is the number of available samples. The statistical covariance matrix \mathbf{R}_x can be approximated by the sample covariance matrix $\mathbf{R}_x(N_s)$ as defined in the last section. At $M = 1$, $\mathbf{R}_x(N_s)$ can be formed by the auto-correlations $r(l)$. Note that the sample covariance matrix is symmetric and Toeplitz.

Based on the generalized likelihood ratio test (GLRT) or information/signal processing theory, there have been a few methods proposed based on the sample covariance matrix. One class of such methods is called covariance based detections (CBD) [1, 65, 76, 77]. Some methods that directly use the auto-correlations of the signal can also be included in this class [78]. The covariance based detections directly use the elements of the covariance matrix to construct detection methods, which can reduce computational complexity. The methods are summarized in the following.

Let the entries of the matrix $\mathbf{R}_x(N_s)$ be c_{mn} ($m, n = 1, 2, \dots, ML$).

Algorithm Covariance based detections

Step 1. Compute the sample covariance matrix $\mathbf{R}_x(N_s)$ as defined in (3.73).

Step 2. Construct a test statistic directly from the entries of the sample covariance matrix. In general, the test statistic of the CBD is

$$T_{CBD} = F_1(c_{mn})/F_2(c_{mm}) \quad (3.85)$$

where F_1 and F_2 are two functions. At single antenna/sensor case, it can be written equivalently as

$$T_{CBD} = F_1(r(0), \dots, r(L-1))/F_2(r(0), \dots, r(L-1)) \quad (3.86)$$

There are many ways to choose the two functions. Some special cases are shown in the following.

1. **Covariance absolute value detection (CAVD).** The test statistic is

$$T_{CAVD} = \frac{\sum_{m=1}^{ML} \sum_{n=1}^{ML} |c_{mn}|}{\sum_{m=1}^{ML} |c_{mm}|} \quad (3.87)$$

2. **Maximum auto-correlation detection (MACD).** The test statistic is

$$T_{MACD} = \max_{m \neq n} |c_{mn}| / \sum_{m=1}^{ML} |c_{mm}| \quad (3.88)$$

3. **Fixed auto-correlation detection (FACD):** The test statistic is

$$T_{FACD} = |c_{m_0 n_0}| / \sum_{m=1}^{ML} |c_{mm}| \quad (3.89)$$

where m_0 and n_0 are fixed numbers between 1 and ML . At single antenna case, the detection can be written equivalently as

$$T_{FACD} = |r(l_0)|/r(0) \quad (3.90)$$

This detection is especially useful when we have some prior information on the source signal correlation and knows the lag that produces the maximum auto-correlation. For example, it can be used for detect the OFDM signal by using the CP or pilot property [52]. \square

Step 3. Compare the test statistic with a threshold to make a decision.

All these methods do not use the information of the signal, channel and noise power as well. The methods are robust to synchronization error, channel impairment, and noise uncertainty.

The test statistic is compared with a threshold γ to make a decision. The threshold γ is determined based on the given P_{fa} . To find a formula for the thresholds is mathematically involved [65, 77]. We will show an example for $M = 1$ in the following subsection.

The computational complexity of the algorithm is as follows (for $M = 1$). Computing the auto-correlations of the received signal requires about LN_s multiplications and additions. Computing $T_1(N_s)$ and $T_2(N_s)$ requires about L^2 additions. Therefore, the total number of multiplications and additions is about $LN_s + L^2$.

3.4.2 Detection Probability and Threshold Determination

It is generally difficult to find closed-form detection probabilities. For this purpose, we need to find the distribution of test statistics. In [65, 76, 77], approximations for the distribution of the test statistics has been found by using central limit theorem for $M = 1$. Furthermore, the theoretical estimations for the two probabilities, P_d , P_{fa} , as well as the threshold associated with these probabilities, were also discussed. Here we summarize the results as follows.

In the following, we consider the case of $M = 1$. Denote c_{nm} as the element of sample covariance matrix $\mathbf{R}_x(N_s)$ at the n th row and m th column, and let

$$T_1(N_s) = \frac{1}{L} \sum_{n=1}^L \sum_{m=1}^L |c_{nm}| \quad (3.91)$$

$$T_2(N_s) = \frac{1}{L} \sum_{n=1}^L |c_{nn}| \quad (3.92)$$

The test statistic of the CAVD is then $T_{CAVD} = T_1(N_s)/T_2(N_s)$.

It is shown in [65, 76, 77] that

$$\lim_{N_s \rightarrow \infty} \mathbb{E}(T_1(N_s)) = \sigma_s^2 + \sigma_\eta^2 + \frac{2\sigma_s^2}{L} \sum_{l=1}^{L-1} (L-l)|\alpha_l| \quad (3.93)$$

where

$$\alpha_l = \mathbb{E}[s(n)s(n-l)]/\sigma_s^2 \quad (3.94)$$

σ_s^2 is the signal power, $\sigma_s^2 = \mathbb{E}[s^2(n)]$. $|\alpha_l|$ defines the correlation strength among the signal samples, here $0 \leq |\alpha_l| \leq 1$. For simplicity, we denote

$$\Upsilon_L \triangleq \frac{2}{L} \sum_{l=1}^{L-1} (L-l)|\alpha_l| \quad (3.95)$$

which is the overall correlation strength among the consecutive L samples. When there is no signal, we have

$$T_1(N_s)/T_2(N_s) \approx E(T_1(N_s))/E(T_2(N_s)) = 1 + (L-1)\sqrt{\frac{2}{\pi N_s}} \quad (3.96)$$

Note that this ratio approaches to 1 as N_s approaches to infinite. Also note that the ratio is not related to the noise power (variance). On the other hand, when there is signal (signal plus noise case), we have

$$\begin{aligned} T_1(N_s)/T_2(N_s) &\approx E(T_1(N_s))/E(T_2(N_s)) \\ &\approx 1 + \frac{\sigma_s^2}{\sigma_s^2 + \sigma_\eta^2} \Upsilon_L = 1 + \frac{\text{SNR}}{\text{SNR} + 1} \Upsilon_L \end{aligned} \quad (3.97)$$

Here the ratio approaches to a number larger than 1 as N_s approaches to infinite. The number is determined by the correlation strength among the signal samples and the SNR. Hence, for any fixed SNR, if there are sufficiently large number of samples, we can always differentiate if there is signal or not based on the ratio.

However, in practice we have only limited number of samples. So, we need to evaluate the performance at fixed N_s .

First we analyze the P_{fa} at hypothesis \mathcal{H}_0 . For given threshold γ_1 , the probability of false alarm for the CAVD algorithm is

$$\begin{aligned} P_{fa} &= P(T_1(N_s) > \gamma_1 T_2(N_s)) \approx P\left(T_2(N_s) < \frac{1}{\gamma_1} \left(1 + (L-1)\sqrt{\frac{2}{N_s\pi}}\right) \sigma_\eta^2\right) \\ &= P\left(\frac{T_2(N_s) - \sigma_\eta^2}{\sqrt{\frac{2}{N_s}\sigma_\eta^2}} < \frac{\frac{1}{\gamma_1} \left(1 + (L-1)\sqrt{\frac{2}{N_s\pi}}\right) - 1}{\sqrt{2/N_s}}\right) \\ &\approx 1 - Q\left(\frac{\frac{1}{\gamma_1} \left(1 + (L-1)\sqrt{\frac{2}{N_s\pi}}\right) - 1}{\sqrt{2/N_s}}\right) \end{aligned} \quad (3.98)$$

where

$$Q(t) = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-u^2/2} du \quad (3.99)$$

For a given P_{fa} , the associated threshold should be chosen such that

$$\frac{\frac{1}{\gamma_1} \left(1 + (L-1)\sqrt{\frac{2}{N_s\pi}}\right) - 1}{\sqrt{2/N_s}} = -Q^{-1}(P_{fa}) \quad (3.100)$$

That is,

$$\gamma_1 = \frac{1 + (L-1)\sqrt{\frac{2}{N_s\pi}}}{1 - Q^{-1}(P_{fa})\sqrt{\frac{2}{N_s}}} \quad (3.101)$$

Note that here the threshold is not related to noise power and SNR. After the threshold is set, we now calculate the probability of detection at various SNR. For the given threshold γ_1 , when signal presents,

$$\begin{aligned}
P_d &= P(T_1(N_s) > \gamma_1 T_2(N_s)) = P\left(T_2(N_s) < \frac{1}{\gamma_1} T_1(N_s)\right) \\
&\approx P\left(T_2(N_s) < \frac{1}{\gamma_1} E(T_1(N_s))\right) \\
&= P\left(\frac{T_2(N_s) - \sigma_s^2 - \sigma_\eta^2}{\sqrt{\text{Var}(T_2(N_s))}} < \frac{\frac{1}{\gamma_1} E(T_1(N_s)) - \sigma_s^2 - \sigma_\eta^2}{\sqrt{\text{Var}(T_2(N_s))}}\right) \\
&= 1 - Q\left(\frac{\frac{1}{\gamma_1} E(T_1(N_s)) - \sigma_s^2 - \sigma_\eta^2}{\sqrt{\text{Var}(T_2(N_s))}}\right)
\end{aligned} \tag{3.102}$$

For very large N_s and low SNR, we have

$$\text{Var}(T_2(N_s)) \approx \frac{2\sigma_\eta^2}{N_s} (2\sigma_s^2 + \sigma_\eta^2) \approx \frac{2(\sigma_s^2 + \sigma_\eta^2)^2}{N_s} \tag{3.103}$$

and

$$E(T_1(N_s)) \approx \sigma_s^2 + \sigma_\eta^2 + \sigma_s^2 \Upsilon_L \tag{3.104}$$

Hence, we have a further approximation

$$P_d \approx 1 - Q\left(\frac{\frac{1}{\gamma_1} + \frac{\Upsilon_L \sigma_s^2}{\gamma_1(\sigma_s^2 + \sigma_\eta^2)} - 1}{\sqrt{2/N_s}}\right) = 1 - Q\left(\frac{\frac{1}{\gamma_1} + \frac{\Upsilon_L \text{SNR}}{\gamma_1(\text{SNR}+1)} - 1}{\sqrt{2/N_s}}\right) \tag{3.105}$$

Obviously, the P_d increases with the number of samples, N_s , the SNR and the correlation strength among the signal samples. Note that γ_1 is also related to N_s as shown above, and $\lim_{N_s \rightarrow \infty} \gamma_1 = 1$. Hence, for fixed SNR, P_d approaches to 1 when N_s approaches to infinite.

3.4.3 Performance Analysis and Comparison

To compare the performances of any methods, first we need a criterion. By properly choosing the thresholds, many methods can achieve any given P_d and $P_{fa} > 0$ if

sufficiently large number of samples are available. The key point is how many samples are needed to achieve the given P_d and $P_{fa} > 0$. Hence, we choose this as the criterion to compare the two algorithms.

For a target pair of P_d and P_{fa} , based on (3.105) and (3.101), we can find the required number of samples for the CAVD as

$$N_c \approx \frac{2(Q^{-1}(P_{fa}) - Q^{-1}(P_d) + (L-1)/\sqrt{\pi})^2}{(\Upsilon_L \text{SNR})^2} \quad (3.106)$$

For fixed P_d and P_{fa} , N_c is only related to the smoothing factor L and the overall correlation strength Υ_L . Hence, the best smoothing factor is

$$L_{best} = \min_L \{N_c\} \quad (3.107)$$

which is related to the correlation strength among the signal samples.

Here we give a comparison of the CBD with the energy detection. Energy detection simply compares the average power of the received signal with the noise power to make a decision. To guarantee a reliable detection, the threshold must be set according to the noise power and the number of samples [4–6]. On the other hand, the proposed methods do not rely on the noise power to set the threshold (see Eq. (3.101)), while keeping other advantages of the energy detection. Simulations have shown that the proposed method is much better than the energy detection when noise uncertainty is present [65, 76, 77]. Hence, here we only compare the proposed method with the ideal energy detection (assume that noise power is known exactly).

For energy detection, the required number of samples is approximately [5]

$$N_e = \frac{2(Q^{-1}(P_{fa}) - Q^{-1}(P_d))^2}{\text{SNR}^2} \quad (3.108)$$

Comparing (3.106) and (3.108), if we want $N_c < N_e$, we need

$$\Upsilon_L > 1 + \frac{L-1}{\sqrt{\pi}(Q^{-1}(P_{fa}) - Q^{-1}(P_d))} \quad (3.109)$$

For example, if $P_d = 0.9$ and $P_{fa} = 0.1$, we need $\Upsilon_L > 1 + \frac{L-1}{4.54}$. In conclusion, if the signal samples are highly correlated such that (3.109) holds, the CAVD is better than the ideal energy detection; otherwise, the ideal energy detection is better.

In terms of the computational complexity, the energy detection needs about N_s multiplications and additions. Hence, the computational complexity of the proposed methods is about L times that of the energy detection.

3.5 Cooperative Spectrum Sensing

When there are multiple secondary users/receivers distributed at different locations, it is possible for them to cooperate to achieve higher sensing reliability. There are various sensing cooperation schemes in the current literature [28, 29, 41, 79–92]. In general, these schemes can be classified into two categories: (A) Data fusion: each user sends its raw data or processed data to a specific user, which processes the data collected and then makes the final decision; and (B) Decision fusion: multiple users process their data independently and send their decisions to a specific user, which then makes the final decision.

3.5.1 Data Fusion

Theoretically, the LRT based on the multiple sensors is the best. However, there are two major difficulties in using the optimal LRT based method: (1) it needs the exact distribution of \mathbf{x} , which is related to the source signal distribution, the wireless channels, and the noise distribution; (2) it may need the raw data from all sensors, which is very expensive for practical applications.

In some situations, the signal samples are independent in time, that is, $E(s_i(n)s_j(m)) = 0$, for $n \neq m$. If we further assume that the noise and signal samples have Gaussian distribution, i.e., $\boldsymbol{\eta}(n) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_\eta)$ and $\mathbf{s}(n) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_s)$, where

$$\mathbf{R}_s = E(\mathbf{s}(n)\mathbf{s}^T(n)), \quad \mathbf{R}_\eta = E(\boldsymbol{\eta}(n)\boldsymbol{\eta}^T(n)) \quad (3.110)$$

the LRT can be obtained explicitly as [89]

$$\log T_{LRT} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{x}^T(n) \mathbf{R}_\eta^{-1} \mathbf{R}_s (\mathbf{R}_s + \mathbf{R}_\eta)^{-1} \mathbf{x}(n) \quad (3.111)$$

Note that in general the cross-correlations among the signals from different sensors are used in the detection here. It means that the fusion center needs the raw data from all sensors, if the signals from different sensors are correlated in space. The reporting of the raw data is very expensive for practical applications.

If the sensors are distributed at different locations and far apart, the primary signal will very likely arrive at different sensors at different times. *That is, in (3.3) τ_{ik} may be different for different i .* For example, assuming that we are sensing a channel with 6 MHz bandwidth with sampling rate 6 MHz, delay of one data sample approximately equals to 50 m distance. In a large size network like a 802.22 cell (typically with radius 30 km), the distance differences of different sensors to the primary user could be as large as several kilo-meters. Therefore, the relative time delays τ_{ik} can be as large as 20 samples or more. If the delays are different, the signals at the sensors will be independent in space.

For distributed sensors, their noises are independent in space. If we aim for sensing at very low SNR, the received signal at a sensor will be dominated by noise. Hence even if the primary signals at different sensors may be weakly correlated, the whole signals (primary signals plus noises) can be treated approximately as independent in space at low SNR. So, in the following, we further assume that $E(s_i(n)s_j(n)) = 0$, for $i \neq j$.

Under the assumptions we have

$$\mathbf{R}_\eta = \text{diag}(\sigma_{\eta,1}^2, \dots, \sigma_{\eta,M}^2) \quad (3.112)$$

$$\mathbf{R}_s = \text{diag}(\sigma_{s,1}^2, \dots, \sigma_{s,M}^2) \quad (3.113)$$

where $\sigma_{\eta,i}^2 = E(|\eta_i(n)|^2)$ and $\sigma_{s,i}^2 = E(|s_i(n)|^2)$. Under the assumptions, we can express the LRT equivalently as

$$\log T_{LRT} = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{i=1}^M \frac{\sigma_{s,i}^2}{\sigma_{\eta,i}^2(\sigma_{s,i}^2 + \sigma_{\eta,i}^2)} |x_i(n)|^2 = \sum_{i=1}^M \frac{\gamma_i}{1 + \gamma_i} T_{ED,i} \quad (3.114)$$

where

$$T_{ED,i} = \frac{1}{N\sigma_{\eta,i}^2} \sum_{n=0}^{N-1} |x_i(n)|^2 \quad (3.115)$$

and $\gamma_i = \sigma_{s,i}^2 / \sigma_{\eta,i}^2$.

Note that $T_{ED,i}$ is the normalized energy at sensor i . The LRT is simply a linearly combined (LC) cooperative sensing. This method is also called cooperative energy detection (CED), which combines the energy from different sensors to make a decision. Thus there are three assertions for cooperative sensing by distributed sensors with time independent signals:

1. the optimal cooperative sensing is the linearly combined energy detection;
2. the combining coefficient is a simple function of the SNR at the sensor;
3. a sensor only needs to report its normalized energy and SNR to the fusion center, and no raw data transmission is necessary.

If the signals are time dependent, the derivation of the LRT becomes much more difficult. Furthermore, the information of correlation among the signal samples is required. There have been methods to exploit the time and space correlations of the signals in a multi-antenna system [14]. If the raw data from all sensors are sent to the fusion center, the sensor network may be treated as a single multi-antenna system (virtual multi-antenna system). If the fusion center does not have the raw data, how to fully use the time and space correlations is still an open question, though there have been some sub-optimal methods. For example, a fusion scheme based on the CAVD is given in [87], which has the capability to mitigate interference and noise uncertainty.

A major difficulty in implementing the method is that the fusion center needs to know the SNR at each user. Also the decision and threshold are related to the SNR's, which means that the detection process changes dynamically with the signal strength and noise power.

If $P = 1$, the propagation channels are flat-fading ($q_{ik} = 0, \forall i, k$), and $\tau_{ik} = 0, \forall i, k$, the signal at different antennas can be coherently combined first and then the energy detection is used [28, 31, 93]. The method is called maximum ratio combined (MRC) cooperative energy detection.

$$T_{MRC} = \frac{1}{N} \sum_{n=0}^{N-1} \left| \sum_{i=1}^M h_i x_i(n) \right|^2 \quad (3.116)$$

It is optimal if the noise powers at different sensors are equal. Note that the MRC needs the *raw data* from all sensors and also the channel information.

We have proved that the LRT is actually a LC scheme. It is natural to also consider other LC schemes. In general, a LC scheme simply sums the weighted energy values to obtain the following test statistic

$$T_{LC} = \sum_{i=1}^M g_i T_{ED,i} \quad (3.117)$$

where g_i is the combining coefficient with $g_i \geq 0$. If we allow the combining coefficients to depend on the SNRs of sensors, we know that the *optimal sensing* should choose $g_i = \gamma_i / (1 + \gamma_i)$. So the problem is how we design a LC scheme that does not need the SNR information or only uses partially available SNR information, while its performance does not degrade much.

One such scheme is the equal gain combine (EGC) [14, 28, 83, 84, 93, 94], i.e., $g_i = 1/M$ for all i :

$$T_{EGC} = \frac{1}{M} \sum_{i=1}^M T_{ED,i} \quad (3.118)$$

EGC totally ignores the differences of sensors.

If the normalized signal energies at different sensors have large differences, a natural way is to choose the largest normalized energy for detection. We call this maximum normalized energy (MNE) cooperative sensing. The test statistic is

$$T_{MNE} = \max_{1 \leq i \leq M} T_{ED,i} \quad (3.119)$$

Note that this is different from the method that uses the known sensor with the largest normalized signal energies. The largest normalized energy may not always be at the same sensor due to the dynamic changes of wireless channels. The method is equivalent to the ‘‘OR decision rule’’ [79, 86].

There have been many researches for the “selective energy detection”. Such methods select the “optimal” sensor to do the sensing based on different criterias [41, 90–92, 95–98].

3.5.2 Decision Fusion

In decision fusion, each sensor sends its one-bit (hard decision) or multiple-bit decision (soft-decision) to a central processor that deploys a fusion rule to make the final decision.

Let us consider the case of hard decision: sensor i sends its decision bit u_i (“1” for signal present and “0” for signal absent) to the fusion center. Let u be the vector formed from u_i . The test statistic of the optimal fusion rule is thus the LRT [79]:

$$T_{DFLRT} = \frac{p(u|\mathcal{H}_1)}{p(u|\mathcal{H}_0)} \quad (3.120)$$

Assuming that the sensors are independent, we have

$$T_{DFLRT} = \prod_{i=1}^M \frac{p(u_i|\mathcal{H}_1)}{p(u_i|\mathcal{H}_0)} \quad (3.121)$$

Let A_1 be the set of i such that $u_i = 1$ and A_0 be the set of i such that $u_i = 0$. The above expression can be rewritten as

$$T_{DFLRT} = \prod_{i \in A_1}^M \frac{P_{d,i}}{P_{fa,i}} \prod_{A_0}^M \frac{1 - P_{d,i}}{1 - P_{fa,i}} \quad (3.122)$$

where $P_{d,i}$ and $P_{fa,i}$ are the probability of detection and probability of false alarm for user i , respectively. Taking logarithm, we obtain

$$\log T_{DFLRT} = \sum_{i \in A_1}^M \log \frac{P_{d,i}}{P_{fa,i}} \sum_{A_0}^M \log \frac{1 - P_{d,i}}{1 - P_{fa,i}} \quad (3.123)$$

By ignoring some constants not related to u_i , the expression can be rewritten as

$$\log T_{DFLRT} = \sum_{i=1}^M u_i \log \frac{P_{d,i}(1 - P_{fa,i})}{P_{fa,i}(1 - P_{d,i})} \quad (3.124)$$

The test statistic is a weighted linear combination of the decisions from all sensors. The weight for a particular sensor reflects its reliability, which is related to the

status of the sensor (for example, signal strength, noise power, channel response, and threshold).

If all sensors have the same status and choose the same threshold, the weights are equal and therefore the LRT is equivalent to the popular “ K out of M ” rule: if and only if K decisions or more are “1”, the final decision is “1”. This includes “Logical-OR (LO)” ($K = 1$), “Logical-AND (LA)” ($K = M$) and “Majority” ($K = \lceil \frac{M}{2} \rceil$) as special cases [79]. Let the probability of detection and probability of false alarm of the method are respectively

$$P_d = \sum_{i=K}^M \binom{M}{i} (1 - P_{d,i})^{M-i} P_{d,i}^i \quad (3.125)$$

and

$$P_{fa} = \sum_{i=K}^M \binom{M}{i} (1 - P_{fa,i})^{M-i} P_{fa,i}^i \quad (3.126)$$

While the Neyman–Pearson Theorem tells us that “ K out of M ” rule is optimal (for equal sensors case), it does not stipulate how to choose the threshold t_h and K . In general, to get the best threshold t_h and K we need to solve some optimization problems for different purpose.

If each user can send multiple-bit decision to the fusion center, a more reliable decision can be made. A fusion scheme based on multiple-bit decisions is shown in [29]. In general, there is a tradeoff between the number of decision bits and the fusion reliability. There are also other fusion rules that may require additional information [79, 99].

3.5.3 Robustness of Cooperative Sensing

Let the noise uncertainty factor of sensor i be α_i . Assume that all sensors have the same noise uncertainty bound. For the linear combination, the expectation of noise power in T_{LC} is therefore

$$\sigma_{LC}^2 = \sum_{i=1}^M g_i \hat{\sigma}_\eta^2 / \alpha_i = \hat{\sigma}_\eta^2 \sum_{i=1}^M g_i / \alpha_i \quad (3.127)$$

Hence, the noise uncertainty factor for LC fusion is $\alpha_{LC} = 1 / \sum_{i=1}^M (g_i / \alpha_i)$. Note that α_i and $1/\alpha_i$ are limited in $[10^{-B/10}, 10^{B/10}]$ and have the same distribution. Hence α_{LC} is also limited in $[10^{-B/10}, 10^{B/10}]$. EGC is a special case of LC. Based on the well-known central limit theorem (CLT), it is easy to verify the following theorem for EGC [56].

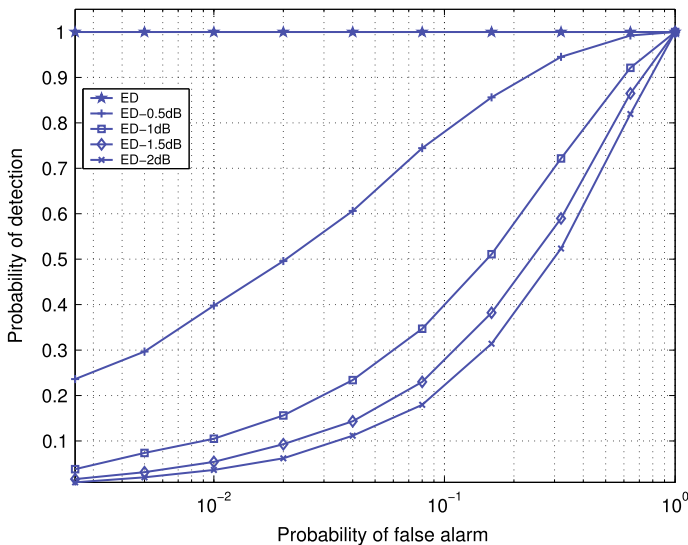


Fig. 3.4 ROC curve for data fusion: $N = 5000$, $\mu = -15$ dB, 20 sensors

Theorem 3.2 Assume that all sensors have the same noise uncertainty bound B and their noise uncertainty factors are independent. As M goes to infinite, the noise uncertainty factor of EGC α_{LC} converges in probability to a deterministic number $1/E(\alpha_i) = \frac{\log(10)B}{5(10^{B/10} - 10^{-B/10})}$, that is, for any $\epsilon > 0$,

$$P(|\alpha_{LC} - 1/E(\alpha_i)| > \epsilon) = 0 \tag{3.128}$$

It means that, as M approaches to infinite, there is no noise uncertainty for EGC fusion rule. We can prove similar result for some other data fusion rules. Hence, data fusion does reduce the noise uncertainty impact. For example, at $N = 5000$ and SNR $\mu = -15$ dB, the ROC curve for 20 sensors is shown in Fig. 3.4.

Although cooperative sensing can achieve better robustness and performance, there are some issues associated with it. First, additional bandwidth is required to exchange information among the cooperating users. In an ad-hoc network, this is by no means a simple task. Second, the information exchange may induce errors, which may have a major impact on fusion performance.

3.5.4 Cooperative CBD and EBD

As shown in the last sections, CBD and EBD are robust sensing methods that are immune to noise uncertainty. Thus it is interesting to use it for cooperative sensing

as well. In [87], methods were proposed to use the CBD and EBD for cooperative sensing. Here we give a brief review of the methods.

It is assumed that there are $M \geq 1$ sensors/receivers in a network. The sensors are distributed in different locations so that their local environments are different and independent. Each sensor has only one antenna. Other than the previous model, here we consider that the received signal may be contaminated by interference. There are two hypotheses: \mathcal{H}_0 and \mathcal{H}_1 , which corresponds to signal absent or present, respectively. The received signal at sensor/receiver i and time n is given as

$$\mathcal{H}_0 : x_i(n) = \rho_i(n) + \eta_i(n) \quad (3.129)$$

$$\mathcal{H}_1 : x_i(n) = h_i s(n - \tau_i) + \rho_i(n) + \eta_i(n) \quad (3.130)$$

Here $\rho_i(n)$ is the interference (like spurious signals) to sensor i , which may be emitted from other electronic devices due to non-linear Analog-to-Digital Converters (ADC) or from other intentional/un-intentional transmitters. Note that interferences to different sensors could be different due to their location differences. $\eta_i(n)$ is the Gaussian white noise to receiver i . $s(n)$ is the primary user's signal and h_i is the propagation channel from the primary to receiver i . τ_i is the relative time delay of the primary signal reaching sensor i . Note that primary signal may reach different sensors at different times due to their location differences. In the following we consider baseband processing and assume that the signal, noise and channel coefficients are complex numbers.

3.5.4.1 The Methods

Let the auto-correlation of the signal be

$$\hat{r}_i(l) = E(x_i(n)x_i^*(n-l)), \quad l = 0, 1, \dots, L-1 \quad (3.131)$$

where L is the number of lags. Then, at hypothesis \mathcal{H}_0 ,

$$\hat{r}_i(l) = \hat{r}_{\rho,i}(l) + \hat{r}_{\eta,i}(l) \quad (3.132)$$

where

$$\hat{r}_{\rho,i}(l) = E(\rho_i(n)\rho_i^*(n-l)) \quad (3.133)$$

$$\hat{r}_{\eta,i}(l) = E(\eta_i(n)\eta_i^*(n-l)) \quad (3.134)$$

Since $\eta_i(n)$ are white noise samples, we have

$$\hat{r}_{\eta,i}(0) = \sigma_{\eta,i}^2, \quad \hat{r}_{\eta,i}(l) = 0, \quad l > 0 \quad (3.135)$$

where $\sigma_{\eta,i}^2$ is the expected noise power at sensor i . At hypothesis \mathcal{H}_1 , we have

$$\hat{r}_i(l) = |h_i|^2 \hat{r}_s(l) + \hat{r}_{\rho,i}(l) + \hat{r}_{\eta,i}(l) \quad (3.136)$$

where

$$\hat{r}_s(l) = E(s(n)s^*(n-l)) \quad (3.137)$$

In practice, there are only limited number of samples at each sensor. Let N be the number of samples. Then the auto-correlations can only be estimated by the sample auto-correlations defined as

$$r_i(l) = \frac{1}{N} \sum_{n=0}^{N-1} x_i(n)x_i^*(n-l), \quad l = 0, 1, \dots, L-1 \quad (3.138)$$

It is known that $r_i(l)$ approaches to $\hat{r}_i(l)$ if N is large. Each sensor computes its sample auto-correlations $r_i(l)$ and then sends them to a fusion center (the fusion center could be one of the sensor). The fusion center first averages the received auto-correlations, that is, compute

$$r(l) = \frac{1}{M} \sum_{i=1}^M r_i(l) \quad (3.139)$$

Then the covariance based detection (CBD) in [65] is used for the detection. Let

$$T_1 = \sum_{l=0}^{L-1} g(l)|r(l)|, \quad T_2 = r(0) \quad (3.140)$$

where $g(l)$ are positive weight coefficients and $g(0) = 1$. The decision statistic of the cooperative covariance based detection (CCBD) is

$$T_{CCBD} = T_1/T_2 \quad (3.141)$$

Let

$$\hat{r}(l) = \frac{1}{M} \sum_{i=1}^M \hat{r}_i(l) \quad (3.142)$$

Then $r(l)$ approaches to $\hat{r}(l)$ for large simple size. At hypothesis \mathcal{H}_0 ,

$$\hat{r}(l) = \frac{1}{M} \sum_{i=1}^M \hat{r}_{\rho,i}(l) + \frac{1}{M} \sum_{i=1}^M \hat{r}_{\eta,i}(l) \quad (3.143)$$

At hypothesis \mathcal{H}_1 ,

$$\hat{r}(l) = \left\{ \frac{1}{M} \sum_{i=1}^M |h_i|^2 \right\} \hat{r}_s(l) + \frac{1}{M} \sum_{i=1}^M \hat{r}_{\rho,i}(l) + \frac{1}{M} \sum_{i=1}^M \hat{r}_{\eta,i}(l) \quad (3.144)$$

Therefore, at hypothesis \mathcal{H}_0 ,

$$T_1/T_2 \approx \frac{\sum_{l=0}^{L-1} g(l) \left| \frac{1}{M} \sum_{i=1}^M \hat{r}_{\rho,i}(l) \right| + \frac{1}{M} \sum_{i=1}^M \sigma_{\eta,i}^2}{\frac{1}{M} \sum_{i=1}^M \left(\hat{r}_{\rho,i}(0) + \sigma_{\eta,i}^2 \right)} \quad (3.145)$$

while at hypothesis \mathcal{H}_1 ,

$$T_1/T_2 \approx \frac{\sum_{l=0}^{L-1} g(l) \left| \frac{1}{M} \sum_{i=1}^M (|h_i|^2 \hat{r}_s(l) + \hat{r}_{\rho,i}(l)) \right| + \frac{1}{M} \sum_{i=1}^M \sigma_{\eta,i}^2}{\frac{1}{M} \sum_{i=1}^M \left(|h_i|^2 \hat{r}_s(0) + \hat{r}_{\rho,i}(0) + \sigma_{\eta,i}^2 \right)} \quad (3.146)$$

Unlike white noise, the interference may be correlated in time. Hence it is possible that $\hat{r}_{\rho,i}(l) \neq 0$ for $l > 0$. However, if we assume that the interferences at different sensors are different and independently distributed, it is highly possible that $\frac{1}{M} \sum_{i=1}^M \hat{r}_{\rho,i}(l)$ ($l > 0$) will be small. This is proved in [87] for some special cases. Thus CCBD does improve the robustness to interference.

As long as the primary signal samples are time correlated, we have $T_1/T_2 > 1$ at hypothesis \mathcal{H}_1 . Hence, we can use T_1/T_2 to differentiate hypothesis \mathcal{H}_0 and \mathcal{H}_1 . We summarize the cooperative covariance based detection (CCBD) as follows.

Algorithm Cooperative Covariance Based Detection

- Step 1. Each sensor computes its sample auto-correlations $r_i(l)$, $l = 0, 1, \dots, L - 1$.
- Step 2. Every sensor sends its sample auto-correlations to the fusion center.
- Step 3. The fusion center computes the average of the sample auto-correlations of all sensors as described in (3.139).
- Step 4. The fusion center computes two statistics T_1 and T_2 as described in (3.140).
- Step 5. Determine the presence of the signal based on T_1 , T_2 and a threshold γ . That is, if $T_1/T_2 > \gamma$, signal exists; otherwise, signal does not exist. \square

In Algorithm CCBD, a special choice for the weights is: $g(0) = 1$, $g(l) = 2(L - l)/L$ ($l = 1, \dots, L - 1$). For this choice, it is equivalent to choose T_1 as the summation of absolute values of all the entries of matrix \mathbf{R}_x , and T_2 as the summation of absolute values of all the diagonal entries of the matrix.

We can form the sample covariance matrix defined as

$$\mathbf{R}_x = \begin{bmatrix} r(0) & \cdots & r(L-1) \\ \vdots & \vdots & \vdots \\ r^*(L-1) & \cdots & r(0) \end{bmatrix} \quad (3.147)$$

Based on the analysis above, at hypothesis \mathcal{H}_0 , \mathbf{R}_x is approximately an diagonal matrix, while at hypothesis \mathcal{H}_1 , \mathbf{R}_x is far from diagonal if the primary signal samples are time correlated.

Based on the sample covariance matrix, the eigenvalue based detections (EBD) discussed in the last sections can also be used here. We summarize the cooperative eigenvalue based detection (CEBD) as follows.

Algorithm Cooperative Eigenvalue Based Detection

Step 1–Step 3. Same as Algorithm CCBD.

Step 4. Form the sample covariance matrix and compute the maximum eigenvalue ζ_{max} , and trace of the matrix \mathbf{R}_x , denoted as T_r .

Step 5. Determine the presence of the signal based on ζ_{max} and T_r and a threshold γ . That is, if $\zeta_{max}/T_r > \gamma$, signal exists; otherwise, signal does not exist. \square

3.5.4.2 Comparisons with Other Methods

There have been extensive studies on cooperative sensing. Some of the methods have been discussed in Sect. 3.5.1. Among them, the cooperative energy detection (CED) is the most popular method. Here we choose the CED for comparison.

In general, ED needs to know the noise power. A wrong estimation of noise power will greatly degrade its performance [7, 47]. CED improves somewhat but still vulnerable to the noise power uncertainty as shown above. Furthermore, when unexpected interference presents, CED will treat it as signal and hence gives high probability of false alarm.

Compared with CED, advantages of CCBD/CEBD are: (1) as an inherent property of covariance and eigenvalue based detection [47, 65], CCBD/CEBD is robust to noise uncertainty; (2) due to the cancellation of auto-correlations at non-zeros lags, CCBD/CEBD is not sensitive to interferences; (3) it is naturally immune to wide-band interference, since such interferences have very weak time correlations; (4) there is no need for noise power estimation at all which reduces implementation complexity.

Compared to single sensor covariance and eigenvalue based detections [47, 65], which may be affected by correlated interferences, CCBD/CEBD overcomes this drawback by cancelation of the adversary impact in the data fusion.

3.6 Summary

In this chapter, spectrum sensing techniques, including classical and newly-developed robust methods, have been reviewed in a systematic way. We start with the fundamental sensing theories from the optimal likelihood ratio test perspective, then review the classical methods including Bayesian method, robust hypothesis test, energy detection, matched filtering detection, and cyclostationary detection. After that, robust

sensing methods, including eigenvalue based sensing method, and covariance based detection, are discussed in detail, which enhance the sensing reliability under hostile environment, Finally, cooperative spectrum sensing techniques are reviewed which improve the sensing performance through combining the test statistics or decision data from multiple sensors. It is pointed out that this chapter only covers the basics of spectrum sensing, but there are many topics are not covered here, such as wideband spectrum sensing [100–103] and compressive sensing [104–107], interested readers are encouraged to refer to the relevant literatures.

References

1. IEEE 802.22 Working Group, IEEE 802.22-2011(TM) standard for cognitive wireless regional area networks (RAN) for operation in TV bands (IEEE, 2011)
2. C. Stevenson, G. Chouinard, Z.D. Lei, W.D. Hu, S. Shellhammer, W. Caldwell, IEEE 802.22: the first cognitive radio wireless regional area network standard. *IEEE Commun. Mag.* **47**(1), 130–138 (2009)
3. Y.H. Zeng, Y.-C. Liang, Z.D. Lei, S.W. Oh, F. Chin, S.M. Sun, Worldwide regulatory and standardization activities on cognitive radio, in *Proceedings of the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Singapore (2010)
4. A. Sonnenschein, P.M. Fishman, Radiometric detection of spread-spectrum signals in noise of uncertainty power. *IEEE Trans. Aerosp. Electron. Syst.* **28**(3), 654–660 (1992)
5. A. Sahai, D. Cabric, Spectrum sensing: fundamental limits and practical challenges, in *Proceedings of the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Baltimore, MD (2005)
6. R. Tandra, A. Sahai, Fundamental limits on detection in low SNR under noise uncertainty, in *WirelessCom 2005, Maui, HI* (2005)
7. R. Tandra, A. Sahai, SNR walls for signal detection. *IEEE J. Sel. Top. Signal Process.* **2**, 4–17 (2008)
8. I.F. Akyildiz, W.Y. Lee, M.C. Vuran, S. Mohanty, Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey. *Comput. Netw.* 2127–2159 (2006)
9. Q. Zhao, B.M. Sadler, A survey of dynamic spectrum access: signal processing, networking, and regulatory policy. *IEEE Signal Process. Mag.* **24**, 79–89 (2007)
10. T. Yucek, H. Arslan, A survey of spectrum sensing algorithms for cognitive radio applications. *IEEE Commun. Surv. Tutor.* **11**(1), 116–130 (2009)
11. A. Sahai, M. Mishra, R. Tandra, K.A. Woyach, Cognitive radios for spectrum sharing. *IEEE Signal Process. Mag.* **26**(1), 140–145 (2009)
12. A. Ghasemi, E.S. Sousa, Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs. *IEEE Commun. Mag.* **46**(4), 32–39 (2008)
13. Sensing techniques for cognitive radio-state of the art and trends, in *IEEE SCC41-P1900.6 White Paper* (2009)
14. Y.H. Zeng, Y.-C. Liang, A.T. Hoang, R. Zhang, A review on spectrum sensing for cognitive radio: challenges and solutions. *EURASIP J. Adv. Signal Process.* **2010**, Article ID 381465, 1–15 (2010)
15. E. Axell, G. Leus, E.G. Larsson, H.V. Poor, Spectrum sensing for cognitive radio: state-of-the-art and recent advances. *IEEE Signal Process. Mag.* **29**(3), 101–116 (2012)
16. S.M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. 2 (Prentice Hall, Englewood Cliffs, 1998)
17. H.V. Poor, *An Introduction to Signal Detection and Estimation* (Springer, Berlin, 1988)
18. H.L. Van-Trees, *Detection, Estimation and Modulation Theory* (Wiley, New York, 2001)

19. T.J. Lim, R. Zhang, Y.-C. Liang, Y.H. Zeng, GLRT-based spectrum sensing for cognitive radio, in *Proceedings of the IEEE Global Communications Conference (GLOBECOM), New Orleans, USA* (2008)
20. R. Zhang, T.J. Lim, Y.-C. Liang, Y.H. Zeng, Multi-antenna based spectrum sensing for cognitive radio: a GLRT approach. *IEEE Trans. Commun.* **58**(1), 84–88 (2010)
21. S.K. Zheng, P.Y. Kam, Y.-C. Liang, Y.H. Zeng, Spectrum sensing for digital primary signals in cognitive radio: a Bayesian approach for maximizing spectrum utilization. *IEEE Trans. Wirel. Commun.* **12**(4), 1774–1782 (2013)
22. P.J. Huber, A robust version of the probability ratio test. *Ann. Math. Stat.* **36**, 1753–1758 (1965)
23. P.J. Huber, Robust estimation of a location parameter. *Ann. Math. Stat.* **35**, 73–104 (1964)
24. L.-H. Zetterberg, Signal detection under noise interference in a game situation. *IEEE Trans. Inf. Theory* **8**, 47–57 (1962)
25. H.V. Poor, Robust matched filters. *IEEE Trans. Inf. Theory* **29**(5), 677–687 (1983)
26. S.A. Kassam, H.V. Poor, Robust techniques for signal processing: a survey. *Proc. IEEE* **73**(3), 433–482 (1985)
27. H. Urkowitz, Energy detection of unknown deterministic signals. *Proc. IEEE* **55**(4), 523–531 (1967)
28. Y.-C. Liang, Y.H. Zeng, E. Peh, A.T. Hoang, Sensing-throughput tradeoff for cognitive radio networks. *IEEE Trans. Wirel. Commun.* **7**(4), 1326–1337 (2008)
29. J. Ma, Y. Li, Soft combination and detection for cooperative spectrum sensing in cognitive radio networks, in *Proceedings of the IEEE Global Communications Conference (GLOBECOM), Washington, USA* (2007)
30. H. Uchiyama, K. Umehayashi, Y. Kamiya, Y. Suzuki, T. Fujii, Study on cooperative sensing in cognitive radio based ad-hoc network, in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Athens, Greece* (2007)
31. A. Pandharipande, J.P.M.G. Linnartz, Performance analysis of primary user detection in multiple antenna cognitive radio, in *Proceedings of the IEEE International Conference on Communications (ICC), Glasgow, Scotland* (2007)
32. A. Sahai, N. Hoven, R. Tandra, Some fundamental limits on cognitive radio, in *Proceedings of the Allerton Conference on Communication, Control, and Computing* (2004)
33. D.J. Thomson, Spectrum estimation and harmonic analysis. *Proc. IEEE* **70**(9), 1055–1096 (1982)
34. D.J. Thomson, Jackknifing multitaper spectrum estimates. *IEEE Signal Process. Mag.* **24**(4), 20–30 (2007)
35. S. Haykin, D.J. Thomson, J.H. Reed, Spectrum sensing for cognitive radio. *Proc. IEEE* **97**(5), 849–877 (2009)
36. Y.H. Zeng, T.H. Pham, Pilot subcarrier feature detection for multi-carrier signals, in *Proceedings of the Asia-Pacific Conference on Communications (APCC)* (2017)
37. C. Liu, Y.H. Zeng, S. Attllah, Max-to-mean ratio detection for cognitive radio, in *IEEE VTC Spring* (2008)
38. N.M. Neihart, S. Roy, D.J. Allstot, A parallel, multi-resolution sensing technique for multiple antenna cognitive radios, in *Proceedings of the IEEE International Symposium on Circuits and System (ISCAS)* (2007)
39. A. Wald, Sequential tests of statistical hypotheses. *Ann. Math. Stat.* **16**, 117–186 (1945)
40. H.V. Poor, O. Hadjiladis, *Quickest Detection*, vol. 2 (Cambridge University Press, Cambridge, 2009)
41. J. Zhao, Q. Liu, X. Wang, S. Mao, Scheduling of collaborative sequential compressed sensing over wide spectrum band. *IEEE/ACM Trans. Netw.* **26**(1), 492–505 (2018)
42. H.-S. Chen, W. Gao, D.G. Daut, Signature based spectrum sensing algorithms for IEEE 802.22 WRAN, in *Proceedings of the IEEE International Conference on Communications (ICC)* (2007)
43. W.A. Gardner, Exploitation of spectral redundancy in cyclostationary signals. *IEEE Signal Process. Mag.* **8**, 14–36 (1991)

44. W.A. Gardner, Spectral correlation of modulated signals: part i-analog modulation. *IEEE Trans. Commun.* **35**(6), 584–595 (1987)
45. W.A. Gardner, W.A. Brown III, C.-K. Chen, Spectral correlation of modulated signals: part ii-digital modulation. *IEEE Trans. Commun.* **35**(6), 595–601 (1987)
46. R. Tandra, A. Sahai, SNR walls for feature detectors, in *Proceedings of the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), Dublin, Ireland* (2007)
47. Y.H. Zeng, Y.-C. Liang, Eigenvalue based spectrum sensing algorithms for cognitive radio. *IEEE Trans. Commun.* **57**(6), 1784–1793 (2009)
48. A.V. Dandawate, G.B. Giannakis, Statistical tests for presence of cyclostationarity. *IEEE Trans. Signal Process.* **42**(9), 2355–2369 (1994)
49. J. Lunden, V. Koivunen, A. Huttunen, H.V. Poor, Spectrum sensing in cognitive radio based on multiple cyclic frequencies, in *Proceedings of CROWNCOM, Orlando, USA* (2007)
50. M. Kim, K. Po, J.-I. Takada, Performance enhancement of cyclostationarity detector by utilizing multiple cyclic frequencies of OFDM signals, in *Proceedings of the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), Singapore* (2010)
51. H.-S. Chen, W. Gao, D.G. Daut, Spectrum sensing using cyclostationary properties and application to IEEE 802.22 WRAN, in *IEEE GlobeCOM* (2007), pp. 3133–3138
52. H.-S. Chen, W. Gao, D.G. Daut, Spectrum sensing for OFDM systems employing pilot tones. *IEEE Trans. Wirel. Commun.* **8**(12), 5862–5870 (2009)
53. K. Maeda, A. Benjebbour, T. Asai, T. Furuno, T. Ohya, Cyclostationarity-inducing transmission methods for recognition among OFDM-based systems. *EURASIP J. Wirel. Commun. Netw.* **2008**, Article ID 586172 (14 p) (2008)
54. H. Zhang, D.L. Ruyet, M. Terre, Spectral correlation of multicarrier modulated signals and its application for signal detection. *EURASIP J. Adv. Signal Process.* **2010**, Article ID 794246 (14 p) (2010)
55. Y.H. Zeng, Y.-C. Liang, Robustness of the cyclostationary detection to cyclic frequency mismatch, in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Istanbul, Turkey* (2010)
56. Y.H. Zeng, Y.-C. Liang, Robust spectrum sensing in cognitive radio (invited paper), in *IEEE PIMRC Workshop, Istanbul, Turkey* (2010)
57. Y.H. Zeng, Y.-C. Liang, Eigenvalue based sensing algorithms, in *IEEE 802.22-06/0118r0* (2006)
58. Y.H. Zeng, Y.-C. Liang, Maximum-minimum eigenvalue detection for cognitive radio, in *The 18th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication Proceedings, Athens, Greece* (2007)
59. Y.H. Zeng, C.L. Koh, Y.-C. Liang, Maximum eigenvalue detection: theory and application, in *IEEE International Conference on Communications (ICC), Beijing, China* (2008)
60. Y.H. Zeng, Y.-C. Liang, R. Zhang, Blindly combined energy detection for spectrum sensing in cognitive radio. *IEEE Signal Process. Lett.* **15**, 649–652 (2008)
61. P. Bianchi, J.N.G. Alfano, M. Debbah, Asymptotics of eigenbased collaborative sensing, in *IEEE Information Theory Workshop (ITW 2009), Sicily, Italy* (2009)
62. M. Maida, J. Najim, P. Bianchi, M. Debbah, Performance analysis of some eigen-based hypothesis tests for collaborative sensing, in *IEEE Workshop on Statistical Signal Processing, Cardiff, UK* (2009)
63. F. Penna, R. Garello, M.A. Spiritog, Cooperative spectrum sensing based on the limiting eigenvalue ratio distribution in wishart matrices. *IEEE Commun. Lett.* **13**(7), 507–509 (2009)
64. P. Wang, J. Fang, N. Han, H. Li, Multiantenna-assisted spectrum sensing for cognitive radio. *IEEE Trans. Veh. Technol.* **59**(4), 1791–1800 (2010)
65. Y.H. Zeng, Y.-C. Liang, Spectrum sensing for cognitive radio based on statistical covariance. *IEEE Trans. Veh. Technol.* **58**, 1804–1815 (2009)
66. A. Kortun, T. Ratnarajah, M. Sellathurai, Exact performance analysis of blindly combined energy detection for spectrum sensing, in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Istanbul, Turkey* (2010)

67. A. Kortun, T. Ratnarajah, M. Sellathurai, C. Zhong, On the performance of eigenvalue-based spectrum sensing for cognitive radio, in *Proceedings of the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), Singapore* (2010)
68. Y.-C. Liang, G. Pan, Y.H. Zeng, On the performance of spectrum sensing algorithms using multiple antenna, in *Proceedings of the IEEE Global Communications Conference (GLOBECOM), Miami, USA* (2010)
69. A.M. Tulino, S. Verdú, *Random Matrix Theory and Wireless Communications* (Now Publishers Inc., Hanover, 2004)
70. I.M. Johnstone, On the distribution of the largest eigenvalue in principle components analysis. *Ann. Stat.* **29**(2), 295–327 (2001)
71. K. Johansson, Shape fluctuations and random matrices. *Commun. Math. Phys.* **209**, 437–476 (2000)
72. Z.D. Bai, Methodologies in spectral analysis of large dimensional random matrices, a review. *Stat. Sin.* **9**, 611–677 (1999)
73. A. Zanella, M. Chiani, M.Z. Win, On the marginal distribution of the eigenvalues of Wishart matrices. *IEEE Trans. Commun.* **57**, 1050–1060 (2009)
74. C.A. Tracy, H. Widom, On orthogonal and symplectic matrix ensembles. *Commun. Math. Phys.* **177**, 727–754 (1996)
75. C.A. Tracy, H. Widom, The distribution of the largest eigenvalue in the Gaussian ensembles, in *Calogero-Moser-Sutherland Models*, ed. by J. van Diejen, L. Vinet (Springer, New York, 2000), pp. 461–472
76. Y.H. Zeng, Y.-C. Liang, Simulations for wireless microphone detection by eigenvalue and covariance based methods, in *IEEE 802.22-07/0325r0* (2007)
77. Y.H. Zeng, Y.-C. Liang, Covariance based detections for cognitive radio, in *Proceedings of the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), Dublin, Ireland* (2007)
78. Y.H. Zeng, Y.-C. Liang, T.H. Pham, Spectrum sensing for OFDM signals using pilot induced auto-correlations. *IEEE J. Sel. Areas Commun.* **31**(3), 353–363 (2013)
79. P.K. Varshney, *Distributed Detection and Data Fusion* (Springer, New York, 1996)
80. J. Unnikrishnan, V.V. Veeravalli, Cooperative sensing for primary detection in cognitive radio. *IEEE J. Sel. Top. Signal Process.* **2**(1), 18–27 (2008)
81. G. Ganesan, Y. Li, Cooperative spectrum sensing in cognitive radio: part I: two user networks. *IEEE Trans. Wirel. Commun.* **6**, 2204–2213 (2007)
82. G. Ganesan, Y. Li, Cooperative spectrum sensing in cognitive radio: part II: multiuser networks. *IEEE Trans. Wirel. Commun.* **6**, 2214–2222 (2007)
83. Z. Quan, S. Cui, A.H. Sayed, Optimal linear cooperation for spectrum sensing in cognitive radio networks. *IEEE J. Sel. Top. Signal Process.* **2**, 28–40 (2007)
84. S.M. Mishra, A. Sahai, R.W. Brodensen, Cooperative sensing among cognitive radios, in *Proceedings of the IEEE International Conference on Communications (ICC), Istanbul, Turkey* (2006)
85. C. Sun, W. Wang, K.B. Letaief, Cluster-based cooperative spectrum sensing for cognitive radio systems, in *Proceedings of the IEEE International Conference on Communications, Glasgow, UK* (2007)
86. E. Peh, Y.-C. Liang, Optimization of cooperative sensing for cognitive radio networks, in *Proceedings of the IEEE WCNC, Hong Kong* (2007)
87. Y.H. Zeng, Y.-C. Liang, E. Peh, A.T. Hoang, Cooperative covariance and eigenvalue based detections for robust sensing, in *Proceedings of the IEEE Global Communications Conference (GLOBECOM), Hawaii, USA* (2009)
88. E. Peh, Y.-C. Liang, Y.L. Guan, Y.H. Zeng, Optimization of cooperative sensing in cognitive radio networks: a sensing-throughput trade off view. *IEEE Trans. Veh. Technol.* **58**, 5294–5299 (2009)
89. Y.H. Zeng, Y.-C. Liang, S.K. Zheng, C.Y. Peh, Optimal cooperative sensing and its robustness to decoding errors, in *Proceedings of the IEEE International Conference on Communications (ICC), Kyoto, Japan* (2011)

90. W. Yuan, H. Leung, S. Chen, W. Cheng, A distributed sensor selection mechanism for cooperative spectrum sensing. *IEEE Trans. Signal Process.* **59**(12), 6033–6044 (2011)
91. M. Monemian, M. Mahdavi, M.J. Omid, Optimum sensor selection based on energy constraints in cooperative spectrum sensing for cognitive radio sensor networks. *IEEE Sens. J.* **16**(6), 1829–1841 (2016)
92. A. Ebrahimzadeh, M. Najimi, S.M.H. Andargoli, A. Fallahi, Sensor selection and optimal energy detection threshold for efficient cooperative spectrum sensing. *IEEE Trans. Veh. Technol.* **64**(4), 1565–1577 (2015)
93. J. Ma, G. Zhao, Y. Li, Soft combination and detection for cooperative spectrum sensing in cognitive radio networks. *IEEE Trans. Wirel. Commun.* **7**(11), 4502–4507 (2008)
94. K.B. Letaief, W. Zhang, Cooperative communications for cognitive radio networks. *Proc. IEEE* **97**(5), 878–893 (2009)
95. Z. Dai, J. Liu, K. Long, Selective-reporting-based cooperative spectrum sensing strategies for cognitive radio networks. *IEEE Trans. Veh. Technol.* **64**(7), 3043–3055 (2015)
96. R. Kishore, C.K. Ramesha, S. Gurugopinath, K.R. Anupama, Performance analysis of superior selective reporting-based energy efficient cooperative spectrum sensing in cognitive radio networks. *Ad Hoc Netw.* **65**, 99–116 (2017)
97. S. Althunibat, M. Di Renzo, F. Granelli, Cooperative spectrum sensing for cognitive radio networks under limited time constraints. *Comput. Commun.* **43**, 55–63 (2014)
98. W. Ejaz, G.A. Shah, N.U. Hasan, H.S. Kim, Energy and throughput efficient cooperative spectrum sensing in cognitive radio sensor networks. *Trans. Emerg. Telecommun. Technol.* **26**(7), 1019–1030 (2015)
99. Z. Chair, P.K. Varshney, Optimal data fusion in multiple sensor detection systems. *IEEE Trans. Aerosp. Electron. Syst.* **22**(1), 98–101 (1986)
100. Y.H. Zeng, Y.-C. Liang, M.W. Chia, Edge based wideband sensing for cognitive radio: algorithm and performance evaluation, in *Proceedings of the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), Germany* (2011)
101. H. Sun, A. Nallanathan, C.-X. Wang, Y. Chen, Wideband spectrum sensing for cognitive radio networks: a survey. *IEEE Wirel. Commun.* **20**(2), 74–81 (2013)
102. F. Wang, J. Fang, H. Duan, H. Li, Phased-array-based sub-Nyquist sampling for joint wideband spectrum sensing and direction-of-arrival estimation. *IEEE Trans. Signal Process.* **66**(23), 6110–6123 (2018)
103. Z. Qin, Y. Gao, M.D. Plumbley, C.G. Parini, Wideband spectrum sensing on real-time signals at sub-Nyquist sampling rates in single and cooperative multiple nodes. *IEEE Trans. Signal Process.* **64**(12), 3106–3117 (2016)
104. Z. Tian, Y. Tafesse, B.M. Sadler, Cyclic feature detection with sub-Nyquist sampling for wideband spectrum sensing. *IEEE J. Sel. Top. Signal Process.* **6**(1), 58–69 (2012)
105. D. Cohen, Y.C. Eldar, Sub-Nyquist sampling for power spectrum sensing in cognitive radios: a unified approach. *IEEE Trans. Signal Process.* **62**(15), 3897–3910 (2014)
106. D. Romero, D.D. Ariananda, Z. Tian, G. Leus, Compressive covariance sensing: structure-based compressive sensing beyond sparsity. *IEEE Signal Process. Mag.* **33**(1), 78–93 (2016)
107. D. Romero, G. Leus, Compressive covariance sampling, in *Information Theory and Applications Workshop (ITA), San Diego, CA, USA*, 10–15 Feb 2013

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

