Management and Sensing of Spectrum in Cognitive Radio

by

Saeed Akhavan Astaneh

A thesis submitted to the
Department of Mathematics and Statistics
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
June 2013

Abstract

Under the contemporary spectrum usage regulations, radio frequency bands are allocated statically to licensed users in a large geographical area and over a long period of time. Recent investigations revealed that such static spectrum allocation has led to very poor spectrum utilization. Cognitive radio has emerged as a new communication paradigm to improve spectrum utilization. It is defined as an intelligent communication system that allows coexistence of the unlicensed users with the licensed ones. Additionally, the users in this system adopt efficient communication protocols to enhance spectral efficiency.

We employ cooperative mechanisms to accomplish two tasks:

- Cooperative spectrum sensing: Licensed users do not in this task. Instead, unlicensed users monitor the activity of the licensed users and transmit only during their absence.

- Cooperative spectrum management: Licensed and unlicensed users can benefit from cooperation with each other, e.g., can assist each other in transmission via relaying. Accordingly, they can save power or bandwidth and therefore, the whole network can accommodate more users.

In the first part of this thesis, we focus on cooperative spectrum sensing. We

first study the asymptotic performance of optimal distributed detectors and identify the conditions under which the highest or lowest asymptotic performance is achieved. For each condition, we study several detectors and investigate their performance-sensitivity-bandwidth tradeoffs. We then consider distributed detection of an OFDM signal. We propose several frequency-domain detectors that, despite their lower computational complexity, outperform state-of-the-art time-domain detectors. Finally, we consider spectrum sensing in fading environments and propose several novel detectors that significantly outperform traditional detectors. For all these detectors, we theoretically analyze their performance and prove that the proposed suboptimal detectors are asymptotically optimal.

We then focus on cooperative spectrum management and study the problem of cooperative relay selection and power allocation. We determine the conditions, in terms of channel gains and network geometry, under which such cooperation leads to an increase in rate, or a reduction in power and bandwidth usage. Lastly, we propose cooperative protocols that use these results to enhance spectrum efficiency.

Co-Authorship

List of publications as a result of the contributions of this thesis:

Chapter 2:

S. A. Astaneh and S. Gazor. Asymptotic performance analysis of distributed source detectors. *Submitted to IEEE Trans. Wireless Commun.*, 2013.

S. A. Astaneh and S. Gazor. Relay assited spectrum sensing. *Submitted to IET Commun.*, 2013.

Chapter 3:

S. A. Astaneh and S. Gazor. Noncoherent distributed OFDM source detection. *Submitted to IEEE Trans. Wireless Commun.*, 2013.

Chapter 4:

S. A. Astaneh and S. Gazor. Cooperative spectrum sensing over mixture-Nakagami channels. *IEEE Wireless Commun. Lett.*, 2013.

Chapter 5:

S. A. Astaneh and S. Gazor. Joint protocol and relay node selection in collaborative networks. *24th Biennial Symposium on Communications (QBSC)*, pages 162–165, 2008.

S. A. Astaneh and S. Gazor. Collaborative communications: Joint relay and protocol selection. *11th Canadian Workshop on Information Theory (CWIT)*, pages

25–28, 2009.

S. A. Astaneh and S. Gazor. Resource allocation and relay selection for collaborative communications. *IEEE Trans. Wireless Commun.*, 2009.

S. A. Astaneh, S. Gazor, and H. Behroozi. On the capacity of pairwise collaborative networks. *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1–5, 2008.

Acknowledgments

I had the honor of working with Dr. Gazor over the past few years. Through his help and guidance, we addressed many challenging topics. I would like to thank him for not just being a supervisor but also for being a great friend and mentor. He opened doors I could not see and helped me gain a new perspective of life.

I was privileged to have courses with Dr. Linder and and Dr. Kim. I like to thank them for teaching me the concepts that significantly contributed to this thesis. I wish to express my gratitude to Dr Alajaji, Dr. Blostein, Dr. Jiang, and Dr Beheshti from Ryerson University, for accepting to serve as the examiners for my defense, and for their valuable comments, that helped enhance the quality of the thesis.

I would also like to thank my father, mother, aunt, uncle, Sandra, Juergen, and Erika for their unconditional love and support.

I want to thank all my friends, who made the past few years a blast, especially, Behnam, Shervin, Ali, Renita, Mehrdad, Ahmad, Hosein, Somayyeh, and all my labmates. I want to thank Sousan and Ishrat for reading my dissertation and for their useful comments that helped me improve the quality of the work.

Finally, I would like to extend my gratitude to the Natural Sciences and Engineering Research Counsel (NSERC), Ontario Graduate Scholarship (OGS), and Queens Graduate Awards for funding my graduate studies during the past five years.

List of Abbreviations

AWGN Additive White Gaussian Noise

CD Cyclostationary Detector

CDF Cumulative Distribution Function

CO Communication Overhead

CP Cooperative Protocol

DFT Discrete Fourier Transform

FC Fusion Center

GLRD Generalized Likelihood Ratio Detector

IDFT Inverse Discreet Fourier Transform

i.i.d Independent and Identically Distributed

LOD Locally Optimum Detector

MGF Moment Generating Function

MLE Maximum Likelihood Estimate

NCP	Non-Cooperative Protocol
NPD	Neyman-Pearson Detector
OFDM	Orthogonal Frequency-Division Multiplexing
PDF	Probability Density Function
RFC	Rayleigh Fading Channel
ROC	Receiver Operating Characteristics
SC	Selection Combining
SF	Separating Function
SINR	Signal to Interference plus Noise Ratio
SNR	Signal to Noise Ratio
TERN	Transmit Energy to Received Noise Power
UMP	Uniformly Most Powerful
WSD	Weak Signal Detector
P_{d}	Probability of Detection
P_{md}	Probability of Mis-detection
P_{fa}	Probability of False Alarm

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The current wireless systems are regulated by fixed spectrum assignment policy where a single frequency band is assigned to a licensed user on long term basis and for a large geographical region. Figure 1.1 shows the spectrum distribution over a large portion of the spectrum for New York and Chicago. It can be observed that the spectrum is highly utilized over a certain portion of the spectrum while the rest is either under-utilized or un-utilized. A recent investigation by Shared Spectrum Company (SSC) and FCC has shown that on average less than 20 % of the allocated spectrum is utilized [69, 42]. On the other hand, in recent years, there has been a drastic increase in demand for accessing the scarce spectrum resources.

Limitation of the available spectrum resources and inefficiency of the past policy necessitate a new paradigm in wireless communication [3]. Cognitive radio is put forth as a new communication paradigm to improve the utilization of the radio spectrum. It is viewed as an intelligent wireless communication system that is aware of its environment and adapts to its variations for a more efficient use of the spectrum. As such, cognitive radio contemplates two different approaches [54]:

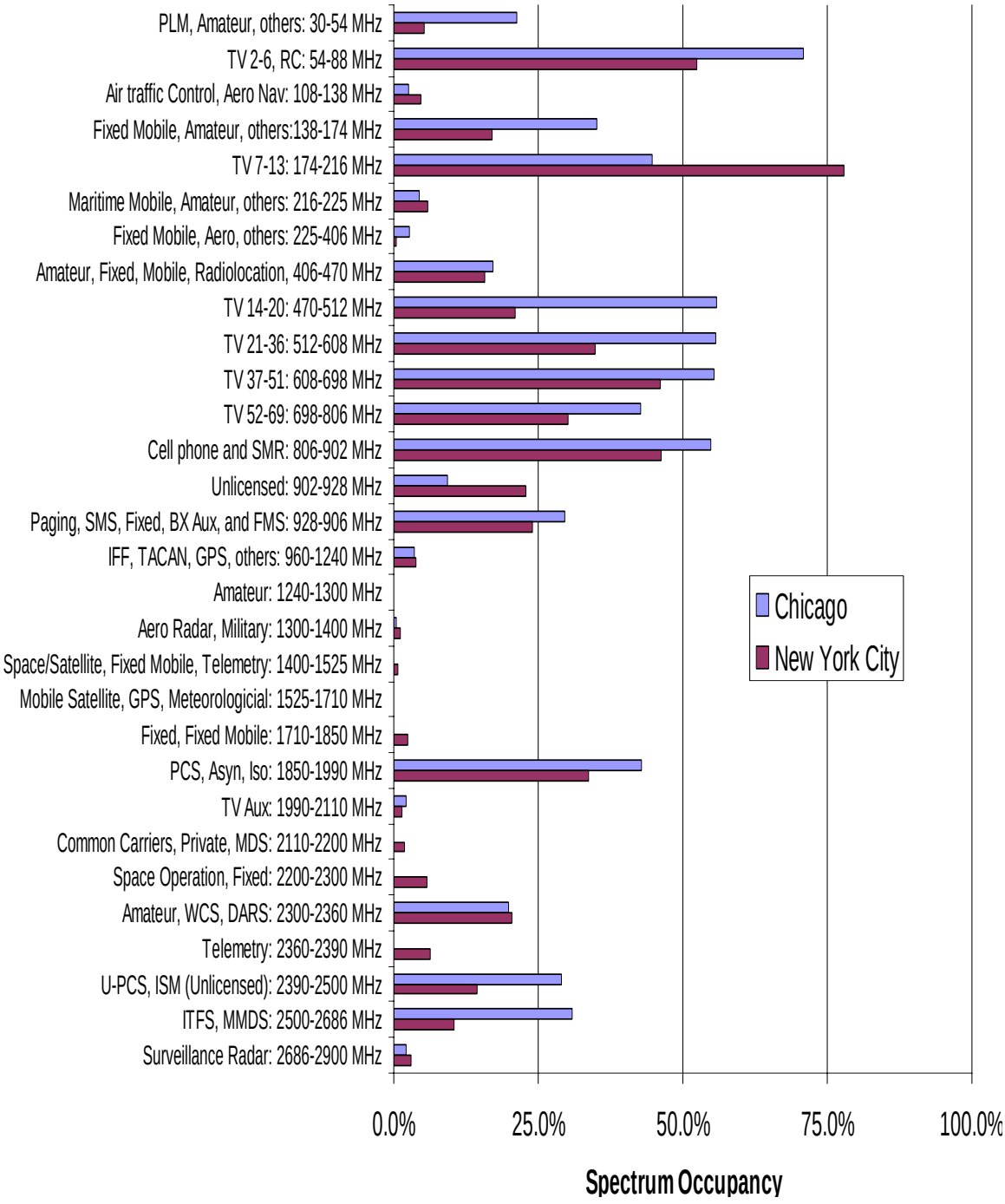


Figure 1.1: Measured spectrum utilization in Chicago and New York City (courtesy of the authors of [69]).

1. The unlicensed users can coexist with the licensed users in an interference free fashion. In other words, the licensed frequency bands that are not being fully utilized can be used by the unlicensed users [56].

2. The licensed and unlicensed users can cooperate with each other to manage their spectrum effectively in order to optimize a certain objective, e.g., to increase their rate or save power [60].

Employing the first strategy, the inefficient usage of the existing resources by the licensed users can be improved by opportunistic access to the spectrum by the unlicensed users [81]. More specifically, the unlicensed users first determine the frequency bands that are not being used. We refer to this stage as spectrum sensing. Having determined the availability of the spectrum opportunities, the unlicensed users may utilize these bands for data transmission. Since the spectrum sensing is prone to error, such opportunistic access may interfere with the transmission of the licensed users. This interference must be capped below some acceptable level.

Adopting the second strategy, the involved users (licensed and/or unlicensed users) monitor the radio environment in order to adjust their transmission or reception parameters. In addition, they can assist each other in transmission through relaying [24, 100]. Such cooperation can potentially result in more efficient spectrum utilization.

Cognitive radio, however, poses several research challenges. For instance, it calls for new distributed spectrum sensing techniques such that the information collected from different users are fused efficiently. In addition, it motivates the research on cooperative communications. In the following, we discuss these research challenges in more detail.

1.1 Spectrum Sensing

In order to improve spectrum utilization, cognitive radio allows unlicensed users to coexist with the licensed users. To this end, licensed users take a passive role while the unlicensed users have to comply with the interference regulation, which specifies the maximum allowed perceived interference [89, 44]. In this setting, the unlicensed users sense/detect the transmission of the licensed users and transmit only when the licensed users are not transmitting.

In order to cope with the environmental factors such as fading, cooperative spectrum sensing has been proposed [92, 91, 77, **?**]. In this way, the unlicensed users act as sensors and collect observations. Afterwards, they employ techniques such as distributed detection to decide on the presence or absence of the licensed users [97, 30]. Note that their inherent spatial distribution provides diversity and improves the detection performance.

Focusing on cooperative spectrum sensing, we aim to address these questions:

- **Optimal and Suboptimal Data Fusion**:
  - Optimal: How can we optimally fuse the collected observations from multiple sensors?
  - Suboptimal: How can we obtain practical suboptimal detectors when some parameters are unknown?
- **Performance Analysis**: What is the performance of these detectors?
- **Sensitivity-Performance Trade-off**: What is the impact of synchronization among the sensors and correlation of the observations on the performance and sensitivity of these detectors?

- **Performance Improvement**:
  - How can we exploit temporal-spectral features of the transmitted signal in these detectors for ?
  - How can we exploit fading and shadowing parameters in these detectors?

1.2 Spectrum Management

Cognitive radio is considered as a reconfigurable radio, which adapts its transmit and receive parameters based on the environment. The process of adaptation, primitively, starts with observation or monitoring. After acquiring the relevant information such as channel gains, bandwidth requirements, and rate requirements, the licensed and unlicensed users can jointly optimize their transmit/receive parameters or may decide to cooperate with each other for data transmission in order to optimize a certain objective [5, 7]. In this context, we usually encounter an optimization problem in a multi-user system. For instance, the problem of maximizing the achievable rate of the users with individual power constraints while guaranteeing a fair spectrum access for them is of great interest [101, 32, 74].

Cooperation among users gives the whole network an additional degree of freedom, which can provide an increase in rate or a reduction in energy or bandwidth usage. Therefore, the whole spectrum will be utilized more efficiently. Note that such cooperation is only an option since depending on the radio environment, cooperation may or may not be beneficial [6, 4]. We aim to study the problem of cooperation via relaying and joint power allocation in cognitive radio and determine the involved gains and losses in terms of rate increase and power and bandwidth usage reduction.

We also aim to identify the conditions, in terms of channel gains, network geometry, and other involved parameters, for a gainful cooperation.

1.3 Contributions

The contributions of the thesis presented in Chapters 2 to 5 are summarized as follows.

Considering the detection of a random source using multiple sensors, we prove that, as the number of samples L increases, the optimal Neyman-Pearson Detector (NPD) achieves its highest (or lowest) performance when the collected observations are correlated (or uncorrelated). We then obtain novel asymptotic expressions for the performance of several optimal and suboptimal detectors for correlated and uncorrelated observations and study the sensitivity-performance trade-off. We also propose several frequency-domain distributed detectors for an Orthogonal Frequency-Division Multiplexing (OFDM) random source that, despite their lower computational complexity, outperform state-of-the-art time-domain detectors. We also derive exact closed form expressions for their performance. Additionally, we propose novel optimal and suboptimal detectors for cooperative spectrum sensing in mixture-Nakagami fading channels that significantly outperform traditional detectors. Lastly, we propose several spectrum management protocols that greatly enhance spectrum efficiency.

More specifically, in Chapter 2, we consider detection of a random source using multiple sensors. Assuming partially correlated observations, we derive the optimal NPD. We prove that, as the number of samples L increases, the Probability of Misdetection (P_{md}) of this detector for a fixed Probability of False Alarm (P_{fa}), achieves the fastest exponential decay when the sensors observe an identical source signal in independent additive white Gaussian noise, namely, a coherent system. Whereas it

achieves the slowest exponential decay when the sensors observe independent pieces of the source signal, namely, in a noncoherent system. We then study/propose different optimal and suboptimal detectors in these systems. We also obtain novel asymptotic expressions for the performance of these detectors. We compare these detectors with three known detectors derived for the coherent system. We show that the coherent detectors require $2L$ times more bandwidth than the noncoherent ones. Additionally, in contrast to the coherent detector, the noncoherent detectors are robust to the signal coherence and sampling time/frequency errors. We also prove that all these detectors are either optimal or asymptotically optimal, i.e., their performance converges to that of the uniformly most powerful one as L increases. The results of this chapter are presented in [14]. Note that we have extended [14] in [17], where we studied spectrum sensing when a rate constraint is imposed on the communication link between each sensor/user and the Fusion Center (FC). However, the results in [17] are not presented in this thesis.

In Chapter 3, we consider distributed detection of an OFDM random source using multiple sensors and propose different frequency-domain optimal and suboptimal detectors. We also derive the exact closed form expressions for their performance. In addition, we prove that the suboptimal detectors are asymptotically optimal. We show that the proposed detectors have insignificant communication overhead and, despite their lower computational complexity, their performance surpasses that of the state-of-the-art time-domain detectors in practical cases. The results of this chapter are presented in [16].

In Chapter 4, we propose novel optimal and suboptimal detectors for cooperative spectrum sensing in mixture-Nakagami fading channels. They significantly outperform energy and cyclostationary detectors in practical scenarios. We study the performance of these detectors in special cases and show that in Reyleigh fading channels they simplify to a linear weighted-correlator or a weighted-energy detector. The results of this chapter are presented in [15].

In Chapter 5, we investigate spectrum management in a network in which users are willing to cooperate with each other; decode and forward the messages of the other along with their own messages to a destination. We study the problem of joint relay selection and power allocation and show that a considerable gain (i.e., increasing rate, saving energy, or reducing the time-bandwidth) and determine the conditions under which such cooperation offers gain in terms of 1) increasing the achievable rate, 2) saving the transmit energy, and 3) reducing the time-bandwidth requirement. We demonstrate that a rate or energy improvement by a factor of at most $\left(1+\sqrt[\eta]{\frac{k}{k+1}}\right)^{\eta}$ can be obtained, where η is the environment path loss exponent and k is the ratio of the rates of the involved users. In addition, we show that the cooperation is only beneficial for the middle range rate ratio. The results of this chapter are presented in [13, 12, 11].

1.4 Practical Impact

The results of Chapter 2 on the sensitivity-performance trade-off provides invaluable insights to the design of the distributed detectors. In addition, this chapter provides useful techniques for analyzing the asymptotic performance of these detectors, which is of great interest when the exact theoretical results are intractable.

The frequency-domain modeling of the OFDM signal in the presence of synchronization errors in Chapter 3 enables us to fully exploit the signal features. Subsequently, the proposed detectors outperform state-of-the-art time-domain detectors. In addition, the lower computational complexity of the proposed detectors make them attractive in practise. Similarly, effective incorporation of the channel statistic in spectrum sensing in Chapter 4 enables us to propose practical detectors with lower complexity and superior performance compared to the traditional detectors.

Lastly, Chapter 5 provides several practical cooperative protocols that can greatly enhance the spectrum utilization.

Chapter 2

Asymptotic Performance Analysis of Spectrum Sensing

We consider detection of a random source using multiple sensors. Assuming partially correlated observations, we derive the optimal NPD and prove that, for a fixed P_{fa} and as the number of samples L increases, its P_{md} achieves the fastest exponential decay, when the sensors observe an identical source signal in independent Additive White Gaussian Noise (AWGN), namely the coherent system. However, it achieves the slowest exponential decay when the sensors observe independent pieces of the source signal in AWGN, namely the noncoherent system. We then study different detectors in these systems. Assuming unknown signal to noise ratios or noise variance, we propose three novel detectors for the noncoherent system. We also use the Edgeworth expansion to obtain novel asymptotic expressions for the P_{fa} and P_{md} of these detectors and two other known detectors. We then compare these detectors with three known detectors derived for the coherent system in terms of bandwidth requirement and sensitivity to signal coherence and sampling time/frequency errors.

We also prove that these detectors are either optimal or asymptotically optimal, i.e., their performance converges to that of the uniformly most powerful one as L increases.

2.1 Introduction

In distributed source detection, a number of sensors locally process their observed data and transmit a summary of these observations to a FC. The FC then uses the reported summaries to make a global decision. This is in contrast to centralized detection where sensors transmit all the (raw) observations to the FC [55]. Distributed detection is of great practical interest in cognitive radio whereas its centralized counterpart is impractical due to the bandwidth scarcity in cognitive radio.

We refer to a system as noncoherent if 1) the bandwidth is significantly larger than the carrier frequency mismatch of the sensors, 2) the sensors have almost identical sampling frequency and 3) the sampling times of the sensors are such that the sensors observe independent pieces of the same source signal in independent AWGN and the observations of different sensors can be assumed uncorrelated. Assuming a noncoherent system, investigators proposed different techniques for distributed source detection. One approach is that the sensors report only a real-valued function of the local observations (as a summary of the observations) to the FC. For instance it is proposed in [94] to report the norm of the observed signals (or energy) to the FC. Employing suboptimal techniques in [94], the FC combines the reported energies to make a global decision. Reporting the likelihood ratio of the observations is also proposed, e.g., a censoring approach is presented in [82] in which sensors may choose to report the likelihood ratio of the local observations or keep silent. Another approach is to allow each sensor to report a single bit per observation (corresponding

to a local binary decision) [97, 55]. Using this scheme, suboptimal techniques such as counting rule, OR, and AND are investigated in [96]. The noncoherent assumption may fail to hold in some cases, e.g. when the sensors have overlapping sampling intervals. We refer to a system as coherent if the sampling time and frequency of the sensors are synchronized such that they observe identical pieces of the source signal in independent AWGN. This synchronization can be achieved by employing time and frequency synchronization methods provided in [88]. Considering such a partially or fully coherent system, [41, 65, 1] study optimal and suboptimal detectors where the local binary decisions are sent to the FC.

Most results on decentralized detection solely study detectors within a given system and, to the best of our knowledge, comparison of these systems with each other in terms of the performance, bandwidth requirement, sensitivity, and implementation aspects of their corresponding detectors, has never been considered. Here, we aim to analyze the asymptotic performance of distributed detectors in these systems and to understand the impact of bandwidth and synchronization among sensors on the performance. We consider distributed detection of a zero-mean Gaussian source signal in independent AWGN. Such a signal model is valid e.g. in OFDM systems where each sub-carrier is loaded with independent data streams and can be assumed Gaussian by virtue of the central limit theorem. Assuming that sensors may or may not have overlapping sampling intervals, we first consider distributed detection using partially correlated observations. For such partially correlated observations, we derive the NPD, denoted by Λ_1. We then prove that the P_{md} of Λ_1 achieves the fastest exponential rate of decay when the system is coherent and achieves the slowest exponential rate of decay when the system is noncoherent. We then move on to study

these two important systems. For the noncoherent system, we propose a novel Generalized Likelihood Ratio Detector (GLRD) for unknown Signal to Noise Ratio (SNR) and two novel heuristic detectors for unknown SNR or noise variance. In addition, we use the Edgeworth expansion to derive novel asymptotic expressions for the $P_{\rm md}$ and $P_{\rm fa}$ of these detectors as well as two previously known detectors (i.e., the NPD for the noncoherent system and a GLRD for unknown SNRs and noise variance). We then compare the detectors designed for the noncoherent system with several well-known detectors designed for the coherent system, and show that the coherent ones require $2L$ times more bandwidth than the noncoherent ones. In addition, we study the sensitivity of these detectors to the sampling time/frequency synchronization errors and show that, in contrast to the coherent detectors, the noncoherent ones are robust to such errors. We propose to use the detectors, designed for the noncoherent system, for the partially/fully coherent systems. This leads to a significant reduction in the bandwidth requirement at the expense of some performance loss, compared to the optimal detectors. We theoretically analyze the performance of the NPD for the noncoherent system, when used in the fully coherent system and show that the performance loss can be easily compensated by increasing the temporal sample size, without requiring extra bandwidth. Finally, we prove that all the aforementioned detectors are either optimal or asymptotically optimal and their performance converges to those of the Uniformly Most Powerful (UMP) tests as $L \to \infty$.

This chapter is organized as follows. We introduce the system model in Section 2.2. We study detection in the the noncoherent and coherent systems in Sections 2.3 and 2.4 respectively. In Section 2.5, we investigate the bandwidth requirement and sensitivity of these detectors to sampling time/frequency errors, propose two novel

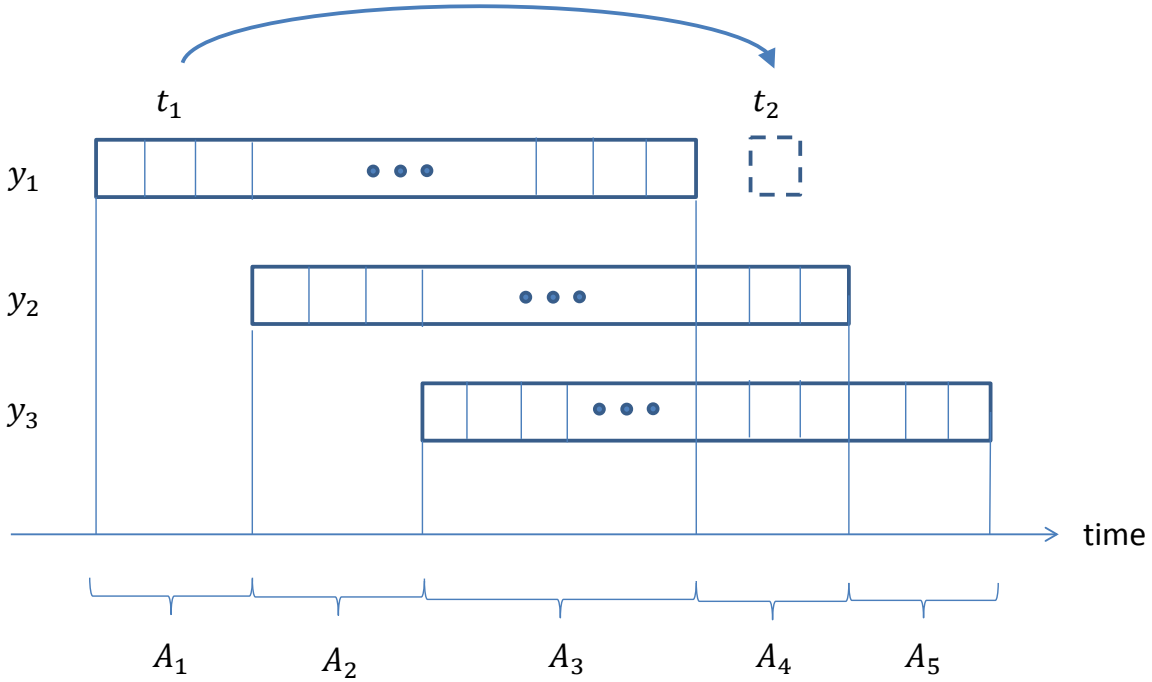


Figure 2.1: An example of overlapping sampling intervals.

asymptotically optimal detectors, study distributed detectors where sensors report the energies, and numerically evaluate the performance of these detectors. We give our concluding remarks in Section 2.6

2.2 System Model

We consider a network of M sensors, intending to detect the presence of a signal source $\mathbf{x}_i \in \mathbb{C}^L$, where L is the sample size, using the following observations from the sensors

$$\mathbf{y}_i = h_i \mathbf{x}_i + \mathbf{n}_i, \quad i = 1, 2, \dots M, \tag{2.1}$$

where $h_i \in \mathbb{C}$ denotes the channel gain between the source and the ith sensor and $\{\mathbf{x}_i\}$ denote the source components of the observed signals by different sensors. We assume $\{\mathbf{x}_i\}$ are complex white Gaussian random variables with zero mean and equal variance. Without loss of generality, let $\{h_i\}$ absorb the variance of the source signal and assume that $\{\mathbf{x}_i\}$ have unitary variance, i.e., $\mathbf{x}_i \sim \frac{1}{\pi^{LM}} \exp\left(-\|x_i\|^2\right)$, where $\|x_i\|$ denotes the norm of x_i. We also assume an AWGN at the ith sensor, denoted by $\mathbf{n}_i \in \mathbb{C}^L$, with zero mean and covariance matrix $\sigma^2 I_L$, where I_L denotes the identity matrix of size L, i.e.

$$\mathbf{N} = [\mathbf{n}_1, \ldots, \mathbf{n}_M] \sim \frac{1}{\pi^{LM}\sigma^{2LM}} \exp(-\frac{1}{\sigma^2}\mathrm{tr}\left(\mathbf{N}\mathbf{N}^H\right)). \tag{2.2}$$

Let also the signal vectors and noise vectors $\{\mathbf{x}_i, \mathbf{n}_i\}$ be mutually independent. In general, the observation interval of different sensors may or may not overlap since the sensors may start sampling with any desired delay. Figure 2.1 depicts a scenario where $M = 3$ sensors are collecting observations in overlapping intervals. Let $\{A_j\}$ denote the time intervals between the starts or ends of the sampling intervals (See Figure 2.1 for an example). During A_j, the source components of the observed signals of the participating sensors are identical. For instance, in Figure 2.1, the elements of $\mathbf{x}_1$ and $\mathbf{x}_2$ that are being observed during A_2 are identical. The joint Probability Density Function (PDF) of the observations $\mathbf{Y} = [\mathbf{y}_1, ..., \mathbf{y}_M]$ is given by

$$\mathbf{Y} = [\mathbf{y}_1, ..., \mathbf{y}_M] \sim \begin{cases} \prod_j \frac{\exp\left(-\mathrm{tr}\left(\mathbf{\Sigma}_{A_j}^{-1}\mathbf{Y}_{A_j}^H\mathbf{Y}_{A_j}\right)\right)}{|\pi\mathbf{\Sigma}_{A_j}|^{|A_j|}}, & \mathcal{H}_1 \\ \frac{\exp(-\sigma^{-2}\mathrm{tr}\left(\mathbf{Y}^H\mathbf{Y}\right))}{\pi^{LM}\sigma^{2LM}}, & \mathcal{H}_0 \end{cases} \tag{2.3}$$

where $\mathcal{H}_0$ and $\mathcal{H}_1$ represent the absence and the presence of the source respectively. In addition, $\mathbf{Y}_{A_j} \in \mathbb{C}^{|A_j| \times M_{A_j}}$ denotes the collected observations by M_{A_j} participating sensors during A_j, where $|A_j|$ denotes the number of samples collected by each participating sensor. We have $\mathbf{\Sigma}_{A_j} = E[\mathbf{Y}_{A_j}^H \mathbf{Y}_{A_j} | \mathcal{H}_1] = \boldsymbol{h}_{A_j}^H \boldsymbol{h}_{A_j} + \sigma^2 I_{M_{A_j}}$, where $\boldsymbol{h}_{A_j}$ is the vector of channel gains of the participating sensors (e.g., for A_2 in Figure 2.1, we have $\boldsymbol{h}_{A_2} = [h_1, h_2]$).

Before deriving the NPD for the hypothesis test in (2.3), we first briefly review its concept. The NPD, is used to perform a hypothesis test between two point hypotheses $\mathcal{H}_1 : \theta = \theta_1$ and $\mathcal{H}_0 : \theta = \theta_0$. As such, the NPD is characterized by the likelihood-ratio $\Lambda(y)$, which is given by the ratio of the PDFs of the observations y under these hypotheses as $\Lambda(y) = \frac{f(y|\mathcal{H}_1)}{f(y|\mathcal{H}_0)}$. This test favors $\mathcal{H}_1$ when $\Lambda(y) > \eta$, where η is the detection threshold. It is proven that [79, p. 23], for any given $P_{\text{fa}} = \alpha$, when η is such that $P(\Lambda(y) > \eta | \mathcal{H}_0) = \alpha$, this test has the highest Probability of Detection (P_{d}) (i.e., the lowest P_{md}) for this given P_{fa}.

As discussed above, assuming known parameters, the NPD for the hypothesis test in (2.3) is given by the ratio of the PDFs in (2.3), i.e.,

$$\frac{\sigma^{2LM}}{\prod_j |\mathbf{\Sigma}_{A_j}|^{|A_j|}} \frac{\prod_j \exp\left(-\text{tr}\left(\mathbf{\Sigma}_{A_j}^{-1} \boldsymbol{Y}_{A_j}^H \boldsymbol{Y}_{A_j}\right)\right)}{\exp(-\sigma^{-2}\text{tr}\left(\mathbf{Y}^H \mathbf{Y}\right))} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_1', \tag{2.4}$$

where η_1' is the detection threshold. Using the matrix inversion theorem, we have $\mathbf{\Sigma}_{A_j}^{-1} = \sigma^{-2} I_{M_{A_j}} - \frac{\sigma^{-2}}{\sigma^2 + \|\mathbf{h}_{A_j}\|^2} \mathbf{h}_{A_j}^H \mathbf{h}_{A_j}$. In addition, we have $\sum_j M_{A_j} |A_j| = LM$ and $\left|\mathbf{\Sigma}_{A_j}\right| = (\|\mathbf{h}_{A_j}\|^2 + \sigma^2)\sigma^{2(M_{A_j}-1)}$. Substituting the above in (2.4), we obtain the following NPD

$$\Lambda_1 \triangleq \sum_j \frac{\sigma^{-2} \|\mathbf{h}_{A_j} \mathbf{Y}_{A_j}^H\|^2}{\sigma^2 + \|\mathbf{h}_{A_j}\|^2} - |A_j| \log(1 + \frac{\|\mathbf{h}_{A_j}\|^2}{\sigma^2}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \log(\eta_1'). \tag{2.5}$$

The sampling intervals $\{A_j\}$ play a significant role in the performance of Λ_1. In the following, we find the conditions for $\{A_j\}$ under which $P_{\mathrm{md},\Lambda_1}$ achieves the fastest or slowest exponential decay rate.

Theorem 2.1. *As $L \to \infty$, P_{md,Λ_1} achieves the fastest exponential rate of decay when the system is coherent and achieves the slowest exponential rate of decay when the system is noncoherent, i.e.,*

$$\begin{aligned} \lim_{L\to\infty} \frac{\log P_{md,\Lambda_1}}{L} &\geq \sum_{i=1}^{M} \log(1+\tfrac{|h_i|^2}{\sigma^2}) - \tfrac{|h_i|^2}{\sigma^2+|h_i|^2}, \\ \lim_{L\to\infty} \frac{\log P_{md,\Lambda_1}}{L} &\leq \log(1+\tfrac{\|\mathbf{h}\|^2}{\sigma^2}) - \tfrac{\|\mathbf{h}\|^2}{\sigma^2+\|\mathbf{h}\|^2}, \end{aligned} \tag{2.6}$$

where $\mathbf{h} = [h_1, \cdots, h_M]$.

Proof. Since Λ_1 is the log-likelihood-ratio of the PDFs in (2.3), we can employ the Stein's theorem [9, 35] to compute the exponential decay rate of its P_{md} for a fixed P_{fa}. This theorem states that, assuming that the observations under each hypothesis are Independent and Identically Distributed (i.i.d) and for any given P_{fa}, the P_{md} is such that $\lim_{L\to\infty} \frac{\log P_{\mathrm{md}}}{L} = D(f(\text{observations}|\mathcal{H}_0)||f(\text{observations}|\mathcal{H}_1))$, where $D(.||.)$ denotes the Kullback-Leibler divergence between the PDFs of the observations under $\mathcal{H}_0$ and $\mathcal{H}_1$, i.e., $f(\text{observations}|\mathcal{H}_0)$ and $f(\text{observations}|\mathcal{H}_1)$ respectively. After simple manipulation, for Λ_1, we have $\lim_{L\to\infty} \frac{\log P_{\mathrm{md}}}{L} = -\lim_{L\to\infty} \frac{E[\Lambda_1|\mathcal{H}_0]}{L}$. Since the PDFs of $\mathbf{Y}_{A_j}$ under $\mathcal{H}_0$ is given by $f(\mathbf{Y}_{A_j}|\mathcal{H}_0) = \frac{\exp(-\frac{1}{\sigma^2}\mathrm{tr}\,(\mathbf{Y}_{A_j}^H \mathbf{Y}_{A_j}))}{(\sigma^2\pi)^{|A_j|M_{A_j}}}$, it can be shown that the Moment Generating Function (MGF) of $\|\mathbf{h}_{A_j}\mathbf{Y}_{A_j}^H\|^2$ is given by

$$E[e^{s\|\mathbf{h}_{A_j}\mathbf{Y}_{A_j}^H\|^2}|\mathcal{H}_0] = (1-s\|h\|^2\sigma^2)^{-|A_j|}. \tag{2.7}$$

Defining $\rho_j \triangleq \lim_{L\to\infty} \frac{|A_j|}{L}$, we have

$$-\lim_{L\to\infty} \frac{E[\Lambda_1|\mathcal{H}_0]}{L} = \sum_j \rho_j \big(\log(1+\frac{\|\mathbf{h}_{A_j}\|^2}{\sigma^2}) - \frac{\|\mathbf{h}_{A_j}\|^2}{\sigma^2+\|\mathbf{h}_{A_j}\|^2} \big). \tag{2.8}$$

In addition, it can be simply proved that

$$\log(1+x+y) - \frac{x+y}{1+x+y} \geq \log(1+x) - \frac{x}{1+x} + \log(1+y) - \frac{y}{1+y}, \tag{2.9}$$

for any $x, y > 0$. Consider two sampling instances $t_1 \in A$ and $t_2 \in B$ where A and B are two different sampling intervals such that the sets of the participating sensors in A and B are disjoint (e.g., $t_1 \in A = A_1$ and $t_2 \in B = A_4$ in Figure 2.1). Now we discuss that if these two sets of sensors sampled simultaneously at t_2, the expression on the right-hand-side of (2.8) would increase. This is verified using (2.9) as

$$\begin{aligned} \log(1+\tfrac{\|\mathbf{h}_A\|^2+\|\mathbf{h}_B\|^2}{\sigma^2}) - \tfrac{\|\mathbf{h}_A\|^2+\|\mathbf{h}_B\|^2}{\sigma^2+\|\mathbf{h}_A\|^2+\|\mathbf{h}_B\|^2} \geq \\ \log(1+\tfrac{\|\mathbf{h}_A\|^2}{\sigma^2}) - \tfrac{\|\mathbf{h}_A\|^2}{\sigma^2+\|\mathbf{h}_B\|^2} + \log(1+\tfrac{\|\mathbf{h}_B\|^2}{\sigma^2}) - \tfrac{\|\mathbf{h}_B\|^2}{\sigma^2+\|\mathbf{h}_B\|^2}. \end{aligned} \tag{2.10}$$

This implies that (2.8) achieves its maximum when any two sensors sample simultaneously (i.e., in a coherent system) and achieves its minimum when all the sensors sample in separate intervals (i.e., in a noncoherent system). Thus (2.6) can be obtained. □

Theorem 2.1 shows that the performance of the optimal detector for partially coherent observations for large values of L is bounded by the performance of the optimal detectors in the noncoherent (lower-bounded) and the coherent (upper-bounded) systems. In order to gain practical insights into the impact of the source, channel,

and noise parameters, we now study these two important cases.

2.3 Noncoherent Source Detection

Throughout this section, we consider a noncoherent system in which $\{\mathbf{x}_i\}$ are assumed independent. In this setting, it can be shown that (2.3) yields the following hypotheses test for detection of a noncoherence source:

$$\mathbf{Y} = [\mathbf{y}_1, ..., \mathbf{y}_M] \sim \begin{cases} \frac{\exp\left(-\sum_{i=1}^{M} \frac{\|\mathbf{y}_i\|^2}{p_i}\right)}{\pi^{LM} \prod_{i=1}^{M} p_i^L}, & \mathcal{H}_1 \\ \frac{\exp\left(-\sum_{i=1}^{M} \frac{\|\mathbf{y}_i\|^2}{\sigma^2}\right)}{\pi^{LM} \sigma^{2ML}}, & \mathcal{H}_0 \end{cases} \tag{2.11}$$

where $p_i = |h_i|^2 + \sigma^2$. We first investigate the NPD that requires the knowledge of the SNRs and noise power. Note that spectrum usage in cognitive radio has a fast changing dynamic, the availability of such knowledge is unreasonable. Nevertheless, the performance of the NPD is an upper-bound on the performance of any suboptimal detector (for noncoherent signals), thus can be used for performance assessment.

2.3.1 Noncoherent NPD: Known SNRs and noise power

For the noncoherent system in (2.11), it can be shown that the NPD in (2.5) is equivalent to the following weighted energy detector

$$\Lambda_2 = \sum_{i=1}^{M} w_i \frac{\|\mathbf{y}_i\|^2}{L} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_2, \tag{2.12}$$

where η_2 is the detection threshold, $w_i = \frac{q_i}{1+q_i}$, with $q_i = \frac{|h_i|^2}{\sigma^2}$ being the SNR of the ith sensor. Note that Λ_2 is a generalization of the detector in [62, p. 496], where

weights are assumed identical.

Remark 2.1. *Employing* (2.12)*, the FC needs only one real number from each sensor, i.e., the observed energies* $\|\mathbf{y}_i\|^2$*. This makes the noncoherent system very attractive. The FC then linearly combines the reported energies to make a global decision.*

The assigned weight w_i to the ith sensor is the ratio of the signal component of the energy to the total observed energy. These weights are larger for the observations with higher SNR. Furthermore, $\lim_{q_i\to\infty} w_i = 1$ and $\lim_{q_i\to 0} w_i = 0$. This suggests a suboptimal detector by discarding the observations from sensors with low SNR (assigning $w_i = 0$), and combining the observations with high SNR with equal weights (assigning $w_i = 1$). Note that $w_i = 1$ is the same as assuming that the SNR is very high. This suboptimal detector does not require the exact SNR values and only needs enough information about SNRs to select a subset of them, e.g., sensors can compare $\|\mathbf{y}_i\|^2$ with a threshold and assign $\{0, 1\}$ weights accordingly.

In the following, we obtain novel expressions for the P_{fa} and P_{md} of Λ_2 for large values of L.

Theorem 2.2. *The* P_{fa} *and* P_{md} *of* Λ_2 *for large* L *are approximately*

$$\begin{aligned} P_{\text{fa},\Lambda_2} &\approx 1 - F(\tfrac{\eta_2 - d_1}{\sqrt{d_2}}; [d_2, d_3, d_4]), \\ P_{\text{md},\Lambda_2} &\approx F(\tfrac{\eta_2 - c_1}{\sqrt{c_2}}; [c_2, c_3, c_4]), \end{aligned} \tag{2.13}$$

where $c_k = L^{1-k}\Gamma(k)\sum_i w_i^k p_i^k$*,* $d_k = L^{1-k}\Gamma(k)\sigma^{2k}\sum_i w_i^k$*, and*

$$\begin{aligned} &F(x, [\alpha_1, \alpha_2, \alpha_3]) = \\ &\Phi(x) - \tfrac{\alpha_2}{6\alpha_1^2}\Phi^{(2)}(x) + \tfrac{\alpha_3}{24\alpha_1^{2.5}}\Phi^{(3)}(x) + \tfrac{\alpha_2^2}{72\alpha_1^{3.5}}\Phi^{(5)}(x), \end{aligned} \tag{2.14}$$

with $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ *and* $\Phi^{(i)}(x) = \frac{\partial^i \Phi(x)}{\partial x^i}$.

Proof. We first approximate the Cumulative Distribution Function (CDF) of Λ_2 under $\mathcal{H}_1$. Under $\mathcal{H}_1$, the MGF of $L\Lambda_2$ is given by $\exp \zeta_1(s)$, where $\zeta_1(s) = -L \sum_i \log(1 - w_i p_i s)$. The kth cumulant of Λ_2 is $\frac{\partial^k \zeta_1(s)}{L^k \partial s^k}|_{s=0}$, which can be verified that is given by c_k (defined above). Using these cumulants, we now employ the Edgeworth expansion [83, 29, 26, 25] to compute the CDF of Λ_2 under $\mathcal{H}_1$. According to this expansion, the CDF of a random variable X can be expressed as $P(X \leq \eta) = F(\frac{\eta - \kappa_1}{\sqrt{\kappa_2}}; [\kappa_2, \kappa_3, \kappa_4]) + O(\frac{\kappa_5}{\sqrt{\kappa_2^5}})$, where the kth cumulant of X is given by κ_k. Noticing that $\kappa_k = c_k$, the CDF of Λ_2 under $\mathcal{H}_1$ can be approximated by (2.13), which is accurate to $O(\frac{\kappa_5}{\sqrt{\kappa_2^5}}) \sim O(\frac{1}{L^{3/2}})$.

Similarly, it can be shown that the cumulants of Λ_2 under $\mathcal{H}_0$ are given by d_k (defined above). Using these cumulants in the Edgeworth expansion, P_{fa,Λ_2} in (2.13) can be obtained, which is accurate to $O(\frac{1}{L^{3/2}})$. □

Note that the equation in (2.2) for the P_{fa} can be used to obtain an appropriate detection threshold for any given P_{fa}.

To draw simple conclusions from (2.13), we ignore the 3rd and 4th cumulants. This results in expressions for P_{fa} and P_{md} of Λ_2 that are accurate to $O(\frac{1}{L^{1/2}})$ [83, 29, 26, 25]. Accordingly, the following approximate Receiver Operating Characteristics (ROC) for Λ_2 can be obtained:

$$P_{\text{md},\Lambda_2} \approx \Phi\Big(\sqrt{\frac{\sum_{i=1}^{M} w_i^2}{\sum_{i=1}^{M} q_i^2}} \Phi^{-1}(1 - P_{\text{fa},\Lambda_2}) - \frac{\sum_{i=1}^{M} (q_i - w_i)}{\sqrt{\sum_{i=1}^{M} q_i^2 / L}}\Big). \tag{2.15}$$

The term $\frac{\sum_{i=1}^{M} (q_i - w_i)}{\sqrt{\sum_{i=1}^{M} q_i^2 / L}}$ in (2.15) plays an important role in the performance of this detector (see e.g. [61]). This term is positive and characterizes the normalized distance

between the centers of PDFs of the test statistic under two different hypotheses of $\mathcal{H}_0$ and $\mathcal{H}_1$. The larger it is, the higher performance can be achieved, e.g., this can be achieved by increasing the number of samples L or sensors M.

Assuming equal weights as $w_i = 1$, we obtain the conventional energy detector, which is given by

$$\Lambda_{\text{ED}} = \sum_{i=1}^{M} \frac{\|\mathbf{y}_i\|^2}{L} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{\text{ED}} \tag{2.16}$$

Using the same approach as in Theorem 2.2, it can be shown that the P_{fa} and P_{md} of Λ_{ED} for large L are given/approximated by

$$P_{\text{fa,ED}} \;\; = 1 - \gamma\big(LM, \tfrac{L\eta_5}{\sigma^2}\big), \qquad P_{\text{md,ED}} \;\; \approx F(\tfrac{\eta_2 - e_1}{\sqrt{e_2}}; [e_2, e_3, e_4]), \tag{2.17}$$

where $\gamma(s, x) = 1 - e^{-x} \sum_{q=0}^{s-1} \frac{x^q}{q!}$ and $e_k = L^{1-k}\Gamma(k) \sum_i p_i^k$. Note that $P_{\text{md,ED}}$ is accurate to $O(\frac{1}{L^{3/2}})$.

2.3.2 Noncoherent GLRD: Unknown SNRs and known noise power

The optimal detector derived in Section 2.3.1 requires the knowledge of SNRs and noise power, which may be unavailable in practice. Here, we assume that the SNRs are unknown and derive a novel GLRD that only uses the noise variance.

Before deriving the GLRD, we briefly describe its concept. The GLRD is a general procedure for composite testing problems. The basic idea is to compare the best model (most probable) under $\mathcal{H}_1$ to the best (most probable) under $\mathcal{H}_0$. Assume that we have two composite hypotheses of the form $\mathcal{H}_0 : \boldsymbol{\theta} \in \Theta_0$ and $\mathcal{H}_1 : \boldsymbol{\theta} \in \Theta_1$.

Also assume that the PDFs of the observations y is given by $f(y|\mathcal{H}_i)$ for $i = 0, 1$, the GLRD for this hypothesis test is given by $\Lambda(y) = \frac{\max\limits_{\theta\in\Theta_1} f(y|\mathcal{H}_1)}{\max\limits_{\theta\in\Theta_0} f(y|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta$. In fact, since θ is unknown, the Maximum Likelihood Estimate (MLE) of θ is replaced with the unknown value in the PDF under different hypothesis. Then their ratio is compared with some threshold [79, p. 38].

Now we derive the GLRD. To this end, we need to find the MLE of the unknown SNRs, i.e., $\widehat{\boldsymbol{q}} = [\widehat{q}_1, \cdots, \widehat{q}_M]$. The MLE of p_i, which can be obtained by maximizing (2.11) with respect to p_i, is given by $\widehat{p}_i = \max(\sigma^2, \frac{\|\mathbf{y}_i\|^2}{L})$. Since we have $\widehat{q}_i = \widehat{\frac{p_i - \sigma^2}{\sigma^2}}$, it can be verified that $\widehat{q}_i = \max(0, \frac{\|\mathbf{y}_i\|^2}{L\sigma^2} - 1)$. Substituting $\widehat{\boldsymbol{q}}$ in the ratio of PDFs in (2.11), we obtain the following GLRD

$$\Lambda_3 = \sum\nolimits_{\forall i: \frac{\|\mathbf{y}_i\|^2}{L\sigma^2} > 1} \left(\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} - 1 - \log \frac{\|\mathbf{y}_i\|^2}{L\sigma^2} \right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_3, \tag{2.18}$$

where η_3 is the detection threshold. This test first censors less informative observations (those for which $\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} \leqslant 1$) that have insignificant energy and only takes into account the observations that support the hypothesis $\mathcal{H}_1$. This is advantageous as each sensor needs to calculate its local energy $\|\mathbf{y}_i\|^2$ and report it to the FC only if the energy level is above the noise energy $L\sigma^2$.

Before we derive asymptotic expressions for the performance of Λ_3, we study its asymptotic optimality. Note that suboptimal detectors such as GLRDs are not always asymptotically optimal, i.e., their performance may not converge to that of the UMP tests as the number of samples increases. Note that if a test/detector for a composite hypothesis problem ($\mathcal{H}_0 : \boldsymbol{\theta} \in \Theta_0$ versus $\mathcal{H}_1 : \boldsymbol{\theta} \in \Theta_1$) is most powerful (i.e., has the highest P_{d} for a given P_{fa}) for all $\theta_1 \in \Theta_1$, it is said to be UMP. The following theorem expresses the conditions under which a GLRD is asymptotically optimal.

Theorem 2.3 ([49]). *Assume that Θ_0 and Θ_1 are disjoint subsets of $\mathbb{R}^M$ and the function $g : \mathbb{R}^M \to \mathbb{R}$ is a Separating Function (SF) for these subsets, i.e., it continuously maps Θ_0 and Θ_1 onto two disjoint intervals as $\Theta_0 \subseteq g^{-1}((-\infty, 0)]$ and $\Theta_1 \subseteq g^{-1}((0, \infty)]$. Then $g(\widehat{\boldsymbol{\theta}}) \gtrless_{\mathcal{H}_0}^{\mathcal{H}_1} \eta$ is asymptotically optimal for the hypothesis problem $\mathcal{H}_0 : \boldsymbol{\theta} \in \Theta_0$ versus $\mathcal{H}_1 : \boldsymbol{\theta} \in \Theta_1$, where $\boldsymbol{\theta}$ denotes the unknown parameters and $\widehat{\boldsymbol{\theta}}$ denotes the MLE of $\boldsymbol{\theta}$.*

Proof. According to [58] as $L \to \infty$, the PDF of $g(\widehat{\boldsymbol{\theta}})$ tends to a normal PDF where its mean and variance are given by $g(\boldsymbol{\theta})$ and $\frac{1}{I_g(\boldsymbol{\theta})}$ respectively. In addition the minimum variance and unbiased estimator of $g(\boldsymbol{\theta})$, i.e., $\widehat{g(\boldsymbol{\theta})}$, can be obtained from $\sum_{l=1}^{L} \frac{\partial}{\partial g(\boldsymbol{\theta})} \ln(f(\mathbf{x}_l; \boldsymbol{\theta})) = I_g(\boldsymbol{\theta})(\widehat{g(\boldsymbol{\theta})} - g(\boldsymbol{\theta}))$, where $f(\mathbf{x}_l; \boldsymbol{\theta})$ denotes the PDF of the observations $\mathbf{x}_l$. Employing the central limit theorem, we observe that the PDF of $\sum_{l=1}^{L} \frac{\partial}{\partial g(\boldsymbol{\theta})} \ln(f(\mathbf{x}_l; \boldsymbol{\theta}))$ and $\widehat{g(\boldsymbol{\theta})}$ tend to normal PDFs. Moreover, the mean and variance of $\widehat{g(\boldsymbol{\theta})}$ are $g(\boldsymbol{\theta})$ and $\frac{1}{I_g(\boldsymbol{\theta})}$ respectively. Thus, the PDF of $g(\widehat{\boldsymbol{\theta}})$ tends to that of $\widehat{g(\boldsymbol{\theta})}$ as $L \to \infty$. Therefore, $E_{\boldsymbol{\theta}}(u(g(\widehat{\boldsymbol{\theta}}) - \eta))$ approaches $E_{\boldsymbol{\theta}}(u(\widehat{g(\boldsymbol{\theta})} - \eta))$ as $L \to \infty$. The proof is complete as $E_{\boldsymbol{\theta}}(u(g(\widehat{\boldsymbol{\theta}}) - \eta))$ and $E_{\boldsymbol{\theta}}(u(\widehat{g(\boldsymbol{\theta})} - \eta))$ for $\boldsymbol{\theta} \in \Theta_0$ and $\boldsymbol{\theta} \in \Theta_1$ represent the P_{fa} and P_{d} using SF and UMP test, respectively. □

We now use this theorem to study the asymptotical optimality of Λ_3.

Theorem 2.4. *Λ_3 is asymptotically optimal, in the sense of Theorem 2.3, as $L \to \infty$.*

Proof. In order to prove the asymptotic optimality of Λ_3, we first need to find appropriate Θ_0 and Θ_1. We observe that the vector of unknown received SNRs, $\boldsymbol{q}$, is zero only under $\mathcal{H}_0$, i.e., $\Theta_0 \triangleq \{[0, \cdots, 0]\}$, and therefore, $\Theta_1 \triangleq \{[x_1, \cdots, x_M] \in \mathbb{R}^M | x_i \geq$

$0\} \setminus \Theta_0$. The proof is complete as it can be verified that

$$g_1(\boldsymbol{q}) \triangleq \sum_i (q_i - \log(1+q_i)), \tag{2.19}$$

satisfies the conditions in Theorem 2.3. □

In the following, we obtain novel expressions for the P_{fa} and P_{md} of Λ_3 for large values of L.

Theorem 2.5. *The P_{fa} and P_{md} of Λ_3 for large L are approximately*

$$\begin{aligned} P_{\text{fa},\Lambda_3} &\approx 1 - U(\eta_3)\epsilon_0^M - \textstyle\sum_{i=1}^{M} \binom{M}{i} \epsilon_0^{M-i}(1-\epsilon_0)^i \gamma(\tfrac{i}{2}, L\eta_3), \\ P_{\text{md},\Lambda_3} &\approx F(\tfrac{\eta_3 - e_1}{\sqrt{e_2}}, [e_2, e_3, e_4]), \end{aligned} \tag{2.20}$$

where $e_1 = g_1(\boldsymbol{\gamma}) + \frac{M}{2L}$, $e_2 = \frac{\sum_i q_i^2}{L} + \frac{M}{2L^2}$, $e_3 = \frac{\sum_i 2q_i^3 + 3q_i^2}{L^2} + \frac{M}{L^3}$, *and* $e_4 = \frac{\sum_i 6q_i^4 + 16q_i^3 + 12q_i^2}{L^3} + \frac{3M}{L^4}$, *and* $\epsilon_0 = \gamma(L, \frac{1}{L})$.

Proof. We first approximate the CDF of Λ_3 under $\mathcal{H}_1$. It is proven in [20, 73, p. 80] that the integral $\int_a^b e^{Lp(x)} q(x) dx$ can be expressed as $(1 + O(\frac{1}{L}))\sqrt{\frac{2\pi}{L|p''(x_0)|}} e^{Lp(x_0)} q(x_0)$, where $a < x_0 < b$ is such that $p'(x_0) = 0$, therefore, it can be approximated by $\sqrt{\frac{2\pi}{L|p''(x_0)|}} e^{Lp(x_0)} q(x_0)$ as $L \to \infty$. Then, it can be shown that

$$E[e^{sL(\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} - 1 - \log\frac{\|\mathbf{y}_i\|^2}{L\sigma^2})} | \mathcal{H}_1] \approx \big(\frac{1-s}{1-(1+q_i)s}\big)^{L(1-s)} \frac{(1+q_i)^{-Ls}}{\sqrt{1-s}}. \tag{2.21}$$

In addition since $q_i > 0$, we have $P(\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} > 1 | \mathcal{H}_1) \to 1$ as $L \to \infty$. Therefore,

$$E[e^{sL\Lambda_3} | \mathcal{H}_1] \approx \prod_i E[e^{sL(\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} - 1 - \log\frac{\|\mathbf{y}_i\|^2}{L\sigma^2})} | \mathcal{H}_1] \tag{2.22}$$

as $L \to \infty$, i.e., we have

$$E[e^{sL\Lambda_3}|\mathcal{H}_1] \approx \prod_{i=1}^{M} \left(\tfrac{1-s}{1-(1+q_i)s}\right)^{L(1-s)} \tfrac{(1+q_i)^{-Ls}}{\sqrt{1-s}}. \tag{2.23}$$

Using the Taylor series expansion of $\log(E[e^{sL\Lambda_3}|\mathcal{H}_1])$, the first four cumulants of Λ_3 can be approximated by $\{e_i\}_{i=1}^4$ (defined above), which are accurate to $\{O(\frac{1}{L^{i+1}})\}_{i=1}^4$ respectively. Using these approximate values, we can now use the Edgeworth expansion [83, 29, 26, 25] to obtain P_{md,Λ_3} in (2.20). Although we used approximate values for the cumulants of Λ_3, it is shown in [26, 25] that the accuracy of these cumulants (accuracy of $O(\frac{1}{L^{i+1}})$ for approximation of κ_i) is enough to guarantee that the final approximation for P_{md,Λ_3} is accurate to $O(\frac{1}{L^{3/2}})$.

Under $\mathcal{H}_0$, we can similarly show that the MGF of $\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} - 1 - \log\frac{\|\mathbf{y}_i\|^2}{L\sigma^2}$ approaches that of $\frac{1}{2}(\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} - 1)^2$ as $L \to \infty$. This implies that the latter converges in distribution to the former, i.e., $\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} - 1 - \log\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} \xrightarrow{d.} \frac{1}{2}(\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} - 1)^2$. Since for large L, the PDF of $\frac{1}{2}(\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} - 1)^2$ converges to the Gamma distribution $\Gamma(\frac{1}{2}, \frac{1}{L})$ (which is accurate to $O(\frac{1}{L^{1/2}})$) and $P(\frac{\|\mathbf{y}_i\|^2}{L\sigma^2} < 1|\mathcal{H}_0) = \gamma(L, \frac{1}{L})$, the PDFs of Λ_3 under $\mathcal{H}_0$ for large L can be approximated by

$$f(\Lambda_3|\mathcal{H}_0) \approx \delta(\Lambda_3)\epsilon_0^M + \sum_{i=1}^{M} \binom{M}{i} \epsilon_0^{M-i}(1-\epsilon_0)^i \frac{\Lambda_3^{\frac{i}{2}-1} e^{-L\Lambda_3}}{\Gamma(\frac{i}{2})L^{-\frac{i}{2}}}. \tag{2.24}$$

Using above, P_{fa,Λ_3} in (2.20) can be obtained, which is accurate to $O(\frac{1}{L^{1/2}})$. □

2.3.3 Noncoherent GLRD: Unknown SNRs and noise power

Thus far, we have assumed that the noise power is known. However, when the spectrum dynamic is fast, the noise power may vary for each sampling interval and therefore its value must be estimated based on the acquired samples. Here, we introduce a GLRD for the case where the noise power is also unknown. This test is obtained by maximizing the PDFs of the observations, with respect to noise power σ^2 and channel gains $|h_i|^2$, under both hypotheses. We observe that under $\mathcal{H}_0$, the MLE of the noise power is given by $\widehat{\sigma^2} = \frac{\sum_{i=1}^{M} \|y_i\|^2}{LM}$ while, the MLE of $\widehat{\sigma^2 + |h_i|^2}$ is given by $\widehat{\sigma^2 + |h_i|^2} = \frac{\|y_i\|^2}{L}$, substituting which in the ratio of PDFs in (2.11), we arrive at the following GLRD:

$$\Lambda_4 = \log\left(\tfrac{1}{M}\sum_{i=1}^{M} \|\mathbf{y}_i\|^2\right) - \tfrac{1}{M}\sum_{i=1}^{M} \log(\|\mathbf{y}_i\|^2) \gtrless_{\mathcal{H}_0}^{\mathcal{H}_1} \eta_4, \tag{2.25}$$

where η_4 is the detection threshold. This test compares the ratio of arithmetic to geometrical mean of the reported energies. This ratio is a well-known source enumeration criterion [98]. In order for this test to perform well, different sensors must have different channel gains, otherwise this detector misinterprets the observations as noise, and favors $\mathcal{H}_0$. This means that the availability of the knowledge of the noise power has significant impact on detection performance when the channel gains are not far apart from each other.

In the following, we study the asymptotic optimality of Λ_4.

Theorem 2.6. *Λ_4 is asymptotically optimal, in the sense of Theorem 2.3, as $L \to \infty$*

.

Proof. The vector of unknown parameters is $\boldsymbol{\theta} = [p_1, \cdots, p_M]$ which is the vector

of variances. We observe that the feasible values of $\boldsymbol{\theta}$ under $\mathcal{H}_0$ and $\mathcal{H}_1$ are $\Theta_0 \triangleq \{[x, \cdots, x] \in \mathbb{R}^M | x > 0\}$ and $\Theta_1 \triangleq \{[x_1, \cdots, x_M] \in \mathbb{R}^M | x_i > 0\} \setminus \Theta_0$ respectively. The proof is complete as it can be verified that

$$g_2(\boldsymbol{\theta}) = \log\left(\tfrac{1}{M}\sum_{i=1}^{M} p_i\right) - \tfrac{1}{M}\sum_{i=1}^{M}\log(p_i) \tag{2.26}$$

satisfies the conditions in Theorem 2.3. □

In the following, we obtain novel expressions for the P_{fa} and P_{md} of Λ_4 for large values of L.

Theorem 2.7. *The P_{fa} and P_{md} of Λ_4 for large L are approximately*

$$P_{\text{fa},\Lambda_4} \approx 1-\gamma(\frac{M-1}{2}, \frac{\eta_4}{ML}),\ P_{\text{md},\Lambda_4} \approx F(\frac{\eta_4 - k_1}{\sqrt{k_2}}, [k_2, k_3, k_4]), \tag{2.27}$$

where $k_i = \frac{\partial^i \zeta_3(s)}{L^i \partial s^i}|_{s=0}$, $\zeta_3(s) = LM\log(t) - L(1-\frac{s}{M})\sum_i \log(t-(1+q_i)s) - \frac{M}{2}\log(1-\frac{s}{M}) - \frac{1}{2}\log(1+\sum_i \frac{s(1+q_i)}{(t-(1+q_i)s)^2(1-\frac{s}{M})})$, *and t is such that* $\sum_i \frac{1+q_i}{t-(1+q_i)s} = 1$.

Proof. For $\mathcal{H}_1$, we can use the multi-dimensional Laplace's method [20] to show that as $L \to \infty$, the MGF of $L\Lambda_4$ can be approximated by $\exp\zeta_3(s)$, where ζ_3 is defined above. Hence, the first four cumulants of Λ_4 are approximately given by $k_i = \frac{\partial^i \zeta_3(s)}{L^i \partial s^i}|_{s=0}, i = 1, \cdots, 4$, which are accurate to $O(\frac{1}{L^{i+1}})$ respectively. We use these approximate values in the Edgeworth expansion to obtain P_{md,Λ_4} in (2.27). Similar to the discussion for Theorem 2.20, it can be shown that P_{md,Λ_3} is accurate to $O(\frac{1}{L^{3/2}})$. For $\mathcal{H}_0$, we substitute $q_i = 0$ in $\zeta_3(s)$ and observe that the MGF of $L\Lambda_4$ can be expressed as $(1-Ms)^{\frac{M-1}{2}}(1+O(\frac{1}{L}))$. Therefore, the PDFs of Λ_4 approaches $\Gamma(\frac{M-1}{2}, \frac{1}{ML})$ as $L \to \infty$. Hence, the approximation for P_{fa,Λ_4} in (2.27) can be obtained. □

2.4 Coherent Source Detection

In this section, we study the problem of source detection in the coherent system where the observation of the sensors are coherent, i.e., $\mathbf{x}_1 = \mathbf{x}_2 = \ldots = \mathbf{x}_M$. We consider detectors for the following hypotheses testing:

$$\mathbf{Y} \sim \begin{cases} \frac{\exp\left(-\operatorname{tr}\left(\mathbf{\Sigma}_1^{-1}\mathbf{Y}^H\mathbf{Y}\right)\right)}{\pi^{LM}|\mathbf{\Sigma}_1|^L}, & \mathcal{H}_1 \\ \frac{\exp(-\sigma^{-2}\operatorname{tr}\left(\mathbf{Y}^H\mathbf{Y}\right))}{\pi^{LM}\sigma^{2LM}}, & \mathcal{H}_0 \end{cases} \tag{2.28}$$

where $\mathbf{\Sigma_1} = \mathbf{h}^H\mathbf{h} + \sigma^2 I_M$ and $\mathbf{h} = [h_1, \ldots, h_M]$. We now study the NPD and GLRDs, where in the former case, the channel gains and noise power are known, and in the latter, they are assumed unknown.

2.4.1 Coherent NPD: Known channel gains and noise power

Assuming known channel gains and noise variance, it can be shown that the NPD in (2.5) for the coherent system in (2.28) is equivalent to the following matched-filter detector

$$\Lambda_5 = \frac{\|\mathbf{h}\mathbf{Y}^H\|^2}{\sigma^4} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_5, \tag{2.29}$$

where η_5 is the detection threshold. This detector needs accurate synchronization, otherwise its performance significantly deteriorates. In contrast to Λ_2, which only needs the amplitude of the channel gains $\{|h_i|^2\}$, Λ_5 requires both amplitude and phase information, i.e., $\{h_i\}$. In addition, all the observation vectors must be transmitted to the FC, which makes this detector impractical.

We now assess the performance of Λ_5. From (2.28), it is easy to show that the

MGF of Λ_5 is $E[e^{s\Lambda_5}|\mathcal{H}_1] = \big(1 - s\frac{\|h\|^2}{\sigma^2}(1 + \frac{\|h\|^2}{\sigma^2})\big)^{-L}$ and $E[e^{s\Lambda_5}|\mathcal{H}_0] = (1 - s\frac{\|h\|^2}{\sigma^2})^{-L}$. Thus the PDFs of Λ_5 is given by

$$f(\Lambda_5) = \begin{cases} \Gamma(L, \sum_{i=1}^{M} q_i), & \mathcal{H}_0 \\ \Gamma(L, (\sum_{i=1}^{M} q_i)(1 + \sum_{i=1}^{M} q_i)), & \mathcal{H}_1. \end{cases} \tag{2.30}$$

Hence, the P_{fa} and P_{md} of Λ_5 are given respectively by

$$P_{\text{fa},\Lambda_5} = 1 - \gamma\big(L, \frac{\eta_5}{\sum_i q_i}\big),\ P_{\text{md},\Lambda_5} = \gamma\big(L, \frac{\eta_5}{(\sum_i q_i)(1+\sum_i q_i)}\big). \tag{2.31}$$

From (2.31), we conclude that the performance of Λ_5 improves as the number of observations L or the SNR of observations q_i increases.

2.4.2 Coherent GLRD: Unknown channel gains and known noise power

Here, we assume that the channel gains are unknown while the noise power are known. It is shown in [10, 93] that the GLRD for such a case is given by

$$\Lambda_6 = \frac{\text{eigmax}(\frac{1}{L}\mathbf{Y}^H\mathbf{Y})}{\sigma^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_6, \tag{2.32}$$

where eigmax($\mathbf{\Sigma}$) is the largest eigenvalue of $\mathbf{\Sigma}$ and η_6 is the detection threshold. We now study the asymptotic optimality of Λ_6.

Theorem 2.8. *Λ_6 is asymptotically optimal, in the sense of Theorem 2.3, as $L \to \infty$*

.

Proof. Let $\mathbf{\Sigma}$ denote the unknown covariance matrix. We notice the set of feasible

covariance matrices under $\mathcal{H}_0$ and $\mathcal{H}_1$ are given by $\Theta_0 \triangleq \sigma^2 I_M$ and $\Theta_1 \triangleq \{x^H x + \sigma^2 I_M | x \in \mathcal{C}^M, \forall i\} \setminus \Theta_0$ respectively. It is easy to show that any $\mathbf{\Sigma} \in \Theta_1$ has two non-identical eigenvalues of σ^2 with a multiplicity of $M-1$ and $\sigma^2 + \|x\|^2$ with a multiplicity of 1. Hence, it can be verified that

$$g_3(\mathbf{\Sigma}) = \frac{\text{eigmax}(\mathbf{\Sigma})}{\sigma^2} - 1, \tag{2.33}$$

satisfies the conditions in Theorem 2.3. The proof is complete as the MLE of $\mathbf{\Sigma}$ is given by $\widehat{\mathbf{\Sigma}} = \frac{1}{L}\mathbf{Y}^H\mathbf{Y}$ [10]. □

2.4.3 Coherent GLRD: Unknown channel gains and noise power

It is shown in [10, 93] that the GLRD for a case when the channel gains and noise variance are all unknown is given by

$$\Lambda_7 = \frac{\text{eigmax}(\frac{1}{L}\mathbf{Y}^H\mathbf{Y})}{\text{tr}(\frac{1}{L}\mathbf{Y}^H\mathbf{Y})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_7, \tag{2.34}$$

where η_7 is the detection threshold. In the following, we study the asymptotic optimality of Λ_7 as $L \to \infty$.

Theorem 2.9. *Λ_7 is asymptotically optimal, in the sense of Theorem 2.3, as $L \to \infty$*

.

Proof. We use g_3 in (2.33) for $\Theta_0 \triangleq \{tI_M | t > 0\}$ and $\Theta_1 \triangleq \{x^H x + tI_M | x \in \mathcal{C}^M, t > 0\} \setminus \Theta_0$. We must find the MLE of σ^2, which is given by $\widehat{\sigma^2} = \frac{\sum_{i=2}^{M} \widehat{\lambda}_i}{M}$, where $\widehat{\lambda}_i$ are the ith largest eigenvalue of $\frac{1}{L}\mathbf{Y}^H\mathbf{Y}$ [10]. The proof is complete since $\text{tr}(\widehat{\mathbf{\Sigma}}) = \sum_i \widehat{\lambda}_i$. □

Remark 2.2. *In contrast to the noncoherent detectors in* (2.5)*,* (2.12)*, and* (2.18)*, where the sensors only need to report the energies of the observed signals, i.e.,* $\{\|\mathbf{y}_i\|^2\}$*, coherent detectors in* (2.29)*,* (2.32)*, and* (2.34) *require the knowledge of all of the observation, i.e.,* $\{\mathbf{y}_i\}$*. This makes the coherent detectors impractical (particularly for large sample size) since they demand high bandwidth.*

2.5 Discussion

Following Remarks 2.1 and 2.2, we realize that the amount of data required by Λ_1 varies from one real number per sensor (for the noncoherent system) to $2L$ real numbers per sensor (for the coherent system). Note that each sensor quantizes its own data (the real numbers that must be transmitted to the FC) and then transmits the quantized values. We assume that the quantization error is small enough such that the quantization distortion in the received data by the FC can be neglected. In order to employ Λ_1, the participating sensors must report the raw correlated observations (two real numbers for each sample) and the energy of the noncoherent samples (one real number for all the samples) to the FC. As shown in Theorem 2.1, employing Λ_1, the lowest P_{md} for large values of L and for a given P_{fa} (i.e., the highest performance) is achieved when the signals are fully coherent and the highest P_{md} (i.e., the lowest performance) is obtained when the signals are fully noncoherent. Figure 2.2 shows the P_{md,Λ_1} versus the required number of real numbers (per sensor) for a similar scenario as in Figure 2.1 with $|A_i| \in \{0, \cdots, L\}$ for $i = 1, \cdots, M-1, L \in \{10, 20, 40\}, \sigma^2 = 50, M = 4, h_i = i$ and $P_{\text{fa}} = 0.1$. Such choice of channel gains (i.e., $h_i = i$) guarantees that the sensors experience distinct channel gains and also generates a reasonably wide range

of SNR $= 10\log_{10}(\frac{|h_i|^2}{\sigma^2}) =$ [-13, -7, -3, -1]. We observe that $P_{\mathrm{md},\Lambda_1}$ generally decreases (i.e., the performance improves) as L increases. Additionally, the performance of Λ_1 improves as the number of required real numbers increases for a given L. Figure 2.3 depicts $P_{\mathrm{md},\Lambda_1}$ versus $P_{\mathrm{fa},\Lambda_1}$ for a similar scenario as in Figure 2.1 with $L = 40, \sigma^2 = 50, M = 4$, $|A_i| \in \{40, 32, 16, 8, 0\}$ (i.e., number of required real numbers per sensor $\in \{1, 17, 33, 49, 65, 80\}$) for $i = 1, \cdots, M-1$. We observe that the performance of the (2.5) is bounded between those of (2.3) (requiring 1 real number) and (2.29) (requiring $2L$ real numbers).

The proposed detectors are also sensitive to sampling time/frequency synchronization errors, i.e., their performance degrades when exposed to such errors. Our simulation results show that the detectors designed for the coherent system (i.e., Λ_5, Λ_6, and Λ_7) are extremely sensitive while the ones for the noncoherent system (i.e., Λ_2, Λ_3, and Λ_4) are robust to such errors. To study the impact of the synchronization errors, we generated $\mathbf{x}_i = [x_{1,i}, \ldots, x_{L,i}]^T$ by $x_{n,i} = \exp(-jn\omega_i) \sum_{k=-5L}^{5L} \frac{\sin(\pi(n+\tau_i-k))}{\pi(n+\tau_i-k)} x_k$, where (τ_i, ω_i) is the pair of the sampling time offset and frequency offset of ith sensor and $\{x_k\}_{k=-5L}^{5L}$ is an i.i.d zero mean normal random process with unitary variance. Figures 2.4 and 2.5 show the average P_{md} of these detectors for an imperfectly synchronized system with P_{fa}= 0.1. In Figure 2.4, the sampling time offsets $\{\tau_i\}$ are independently and uniformly generated over $[-\frac{\Delta\tau}{2}, \frac{\Delta\tau}{2}]$ and $\omega_i = 0$ whereas in Figure 2.5, the sampling time offsets $\{\omega_i\}$ are independently and uniformly generated over $[-\frac{\Delta\omega}{2}, \frac{\Delta\omega}{2}]$ and $\tau_i = 0$. These figures show that the detectors Λ_2, Λ_3 and Λ_4 are quite robust to the time/frequency synchronization errors. In contrast, the detectors Λ_5, Λ_6, and Λ_7 are very sensitive to such errors and their performance dramatically degrades in imperfectly synchronized systems.

2.5.1 Detection Using The Energy Reports

From above, we conclude that the noncoherent system is more attractive and practical if the available bandwidth is limited, the synchronization is expensive or the synchronization errors are not negligible. Motivated by this, we propose that the sensors only report their observed energies and the FC uses (2.12), (2.18), or (2.25) to combine the reported energies.

Note that the performance results reported in (2.13) are invalid for the coherent signal model. In the following, we assess the asymptotic performance of Λ_2 when used for the detection of coherent signals.

Theorem 2.10. *The $P_{\rm fa}$ and $P_{\rm md}$ of Λ_2 when used in a coherent system for large L, can be approximated by*

$$\begin{aligned} P_{fa,\Lambda_2,\rm EN} &\approx 1 - F(\tfrac{\eta_{\rm EN}-d_1}{\sqrt{d_2}}; [d_2, d_3, d_4]), \\ P_{md,\Lambda_2,\rm EN} &\approx F(\tfrac{\eta_{\rm EN2}-l_1}{\sqrt{l_2}}; [l_2, l_3, l_4]), \end{aligned} \tag{2.35}$$

where $\{d_i\}$ are defined in Theorem 2.2, $l_i = \frac{\partial^i \zeta_{\rm EN}(s)}{L^i \partial s^i}|_{s=0}$, with $\zeta_{\rm EN}(s) = -L\log(1 - \sum_i \frac{s w_i |h_i|^2}{1-s\sigma^2 w_i}) - L\sum_i \log(1 - s\sigma^2 w_i)$.

Proof. Note that the $P_{\rm fa}$ of Λ_2 used in a coherent system is the same as $P_{\rm fa,\Lambda_2}$ in Theorem 2.13 and is accurate to $O(\frac{1}{L^{3/2}})$. Under $\mathcal{H}_1$, the conditional MGF of $L\Lambda_2$ (used for the coherent signal) given $r = \|\mathbf{x}\|^2$ is

$$E[e^{sL\Lambda_2}|r] = e^{r\sum_i \frac{s|h_i|^2 w_i}{1-s\sigma^2 w_i}} \prod_{i=1}^{M} (1 - s\sigma^2 w_i)^{-L}.$$

To find the unconditional distribution, we average the above MGF over $r = \|x\|^2 \sim$

$\Gamma(L,1)$, take logarithm and obtain $\log(E[e^{sL\Lambda_2}]) = \zeta_{\text{EN}}(s)$. The cumulants of Λ_2 (used for the coherent signal) are as $l_i = \frac{\partial^i \zeta_{\text{EN}}(s)}{L^i \partial s^i}|_{s=0}$, thus, $P_{\text{md},\Lambda_2,\text{EN}}$ can be directly obtained from the Edgeworth expansion of Λ_2 given $\{l_i\}_{i=1}^4$, which is accurate to $O(\frac{1}{L^{3/2}})$. □

To draw simple conclusions, we ignore the 3rd and 4th cumulants in (2.35). This results in expressions for the P_{fa} and P_{md} of Λ_2 (when used in a coherent system) that are accurate to $O(\frac{1}{L^{1/2}})$. It can be shown that $l_1 = \sum_i q_i$ and $l_2 = \sum_{i \neq j} w_i w_j q_i \gamma_j + \sum_i q_i^2$. Thus the following approximate ROC can be obtained:

$$P_{\text{md,EN}} \approx \Phi\left(\frac{\sqrt{\sum_i w_i^2}\,\Phi^{-1}(1-P_{\text{fa,EN}}) - L\sum_i (q_i - w_i)}{\sqrt{\sum_{i \neq j} w_i w_j q_i \gamma_j + \sum_i q_i^2}}\right). \tag{2.36}$$

Comparing (2.15) and (2.36), there is an extra term $\sum_{i \neq j} w_i w_j q_i \gamma_j$ in the denominator, which indicates that Λ_2 results in higher mis-detection probability (i.e., lower performance) for a given P_{fa} when used for detection of coherent signals.

To get the maximum performance out of the noncoherent signal model in (2.11), sensors may start sampling at random times to reduce the probability of concurrent sampling by different sensors. This means that not only these detectors do not require synchronization but also no coordination among sensors is required to ensure that the sensor observations are noncoherent. Our numerical analysis also shows that these detectors suffer solely from negligible performance loss when used for the coherent system. This loss can be compensated by a small increase in the number of samples L.

2.5.2 Novel Heuristic Detectors

One can use Laplace's method to approximate the MGF of $\left(\frac{\|\mathbf{y}_i\|^2}{L\sigma^2}-1\right)^2$ and $(\frac{\|\mathbf{y}_i\|^2}{\frac{1}{M}\sum_{i=1}^{M}\|\mathbf{y}_i\|^2}-1)^2$ for large values of L. As such, it can be observed that under both hypotheses (i.e., $\mathcal{H}_0$ and $\mathcal{H}_1$), expressions similar to the approximate expressions for the MGF of Λ_3 and Λ_4, respectively, in Theorems 2.5 and 2.7 can be obtained. This motivates us to propose the following detectors. These two heuristic detectors are asymptotically optimal and the proof for their asymptotic optimality is similar to Theorems 2.5 and 2.7. The following detector can be used for unknown channel gains:

$$\Lambda_8 = \sum\nolimits_{\forall i:\frac{\|\mathbf{y}_i\|^2}{L\sigma^2}>1}\left(\frac{\|\mathbf{y}_i\|^2}{L\sigma^2}-1\right)^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_8, \tag{2.37}$$

where η_8 denotes the detection threshold. The following detector can be used for unknown channel gains and noise variance:

$$\Lambda_9 = \sum\nolimits_{i=1}^{M}(\frac{\|\mathbf{y}_i\|^2}{\frac{1}{M}\sum_{i=1}^{M}\|\mathbf{y}_i\|^2}-1)^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_9 \tag{2.38}$$

where η_9 denotes detection threshold. As discussed above, similar to Theorems 2.5 and 2.7, it can be proved that the performance of Λ_8 and Λ_9 is similar to those of Λ_3 and Λ_4 respectively.

2.5.3 Numerical Results

Here, we evaluate the proposed distributed detectors numerically. Figure 2.6 shows the P_{md} versus P_{fa} of $\Lambda_2-\Lambda_5$, Λ_8, and Λ_9 for the noncoherent signal model, employing $M=4$ sensors, $L\in\{20,40\}$ samples, and assuming $h_i=i$ and $\sigma^2=20$. The theoretical performance (in (2.13), (2.20), and (2.27)) are given by dashed lines. We

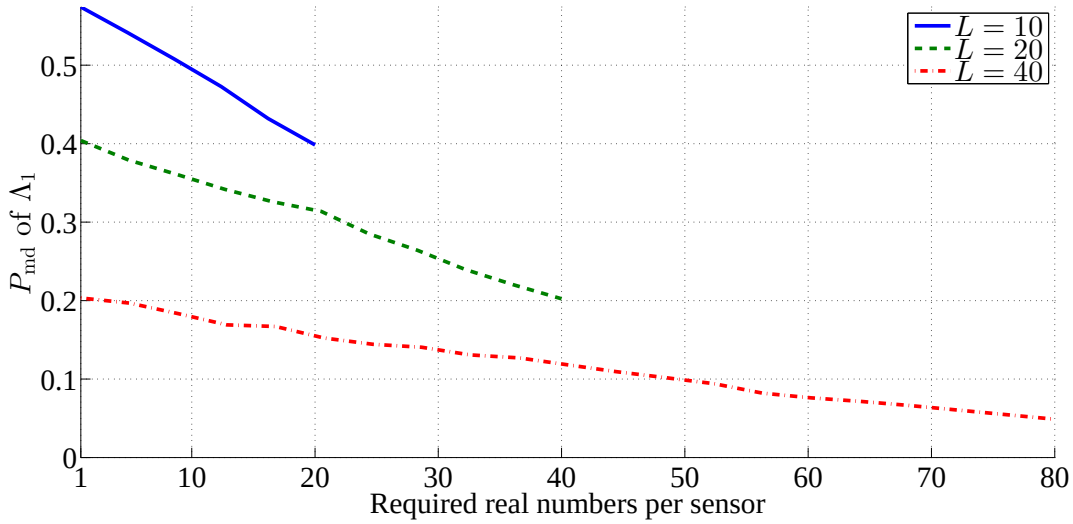


Figure 2.2: P_{md} of Λ_1 versus the number of required real numbers transmitted per sensor with $L \in \{10, 20, 40\}, \sigma^2 = 50, M = 4$ and $P_{\text{fa}} = 0.1$.

observe that Λ_2 (the optimal detector for the noncoherent system) outperforms all the other detectors. In addition, Λ_3 (which performs close to Λ_8) outperforms Λ_4 (which performs close to Λ_9). This is intuitively expected as Λ_3 and Λ_8 (in contrast to Λ_4 and Λ_9) make use of the knowledge of the channel gains. We also observe that as the number of samples increases, the performance of these detectors improves, i.e., the performance degradation of suboptimal detectors can be compensated by an increase in the sample size. For instance, our extended simulations show that the Λ_3 and Λ_8 with $L = 30$ outperform the optimal detector Λ_2 with $L = 20$ and Λ_4 and Λ_9 with $L = 80$ outperform Λ_3 and Λ_8 with $L = 20$.

Figure 2.7 compares the performance of the proposed detectors in the coherent signal model, with $h_i = i$, $M = 4$, $\sigma^2 = 50$ and $L = 20$. The theoretical performance (in (2.31) and (2.35)) are given by dashed lines. We observe that Λ_5, Λ_6, and Λ_7 outperform Λ_2, Λ_3 (and Λ_8), and Λ_4 (and Λ_9) respectively. This was expected as the detectors designed for the coherent system outperform those of the noncoherent

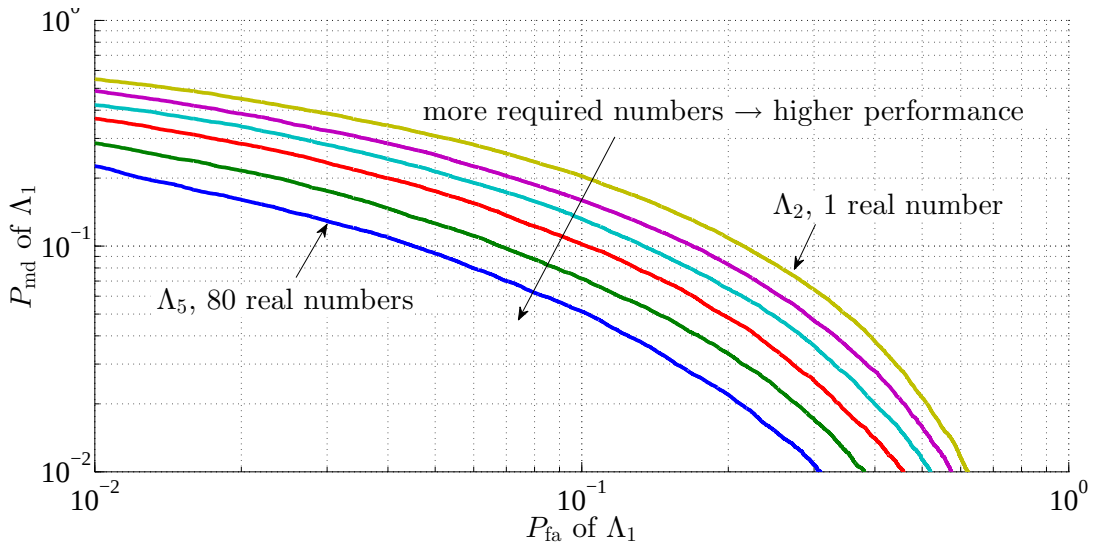


Figure 2.3: P_{md} versus P_{fa} of Λ_1 with $L = 40, \sigma^2 = 50, M = 4$, and required number of real numbers transmitted per sensor $\in [1, 17, 33, 49, 65, 80]$.

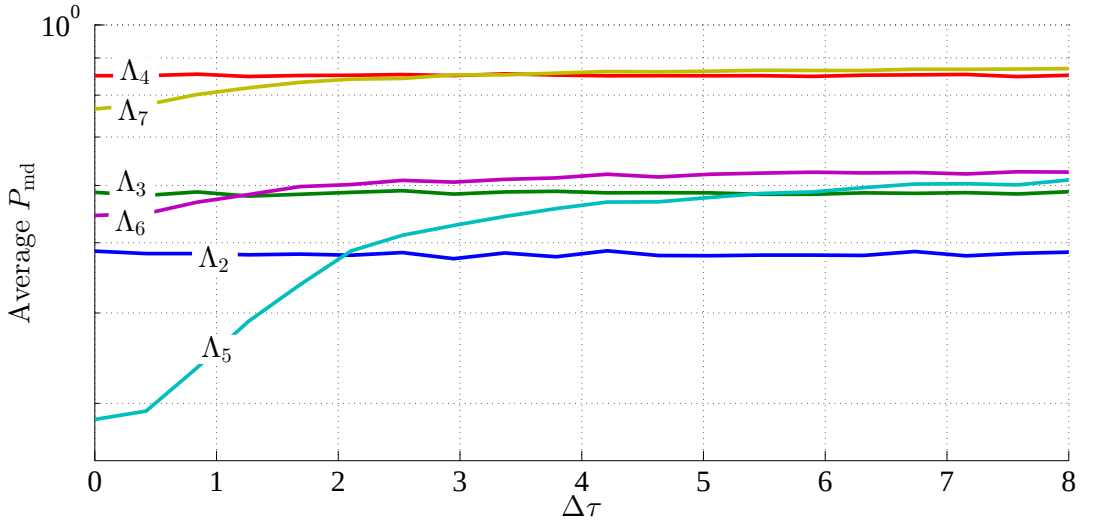


Figure 2.4: Impact of time synchronization error on the average P_{md} of the detectors $\Lambda_2, \ldots, \Lambda_7$ for a constant average $P_{\text{fa}} = 0.1$, $M = 4$, and $L = 15$, with sampling time and frequency offsets $\tau_i \sim U[\frac{-\Delta\tau}{2}, \frac{\Delta\tau}{2}]$ and $\omega_i = 0$.

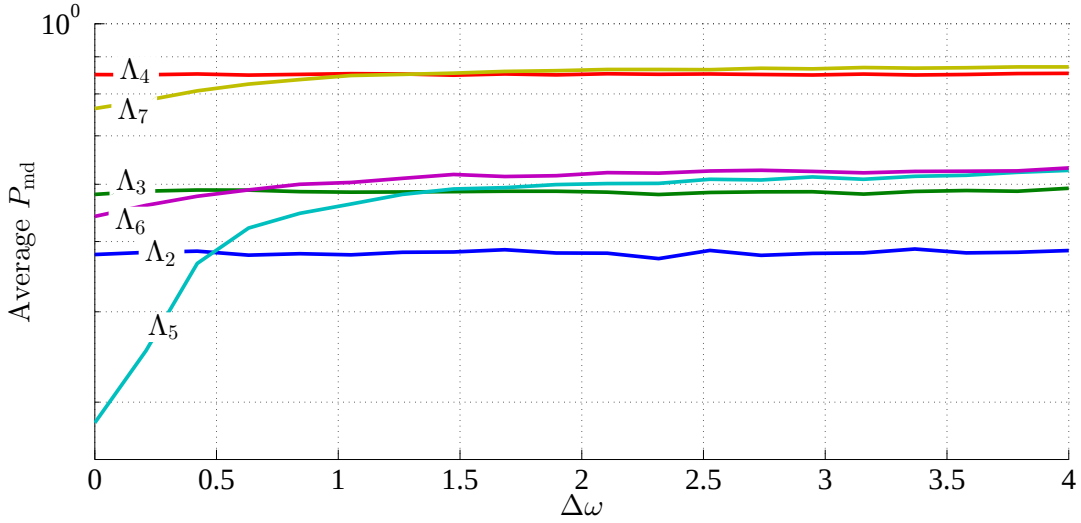


Figure 2.5: Impact of frequency synchronization error on the average P_{md} of the detectors $\Lambda_2, \ldots, \Lambda_7$ for a constant $P_{\text{fa}} = 0.1$, $M = 4$, and $L = 15$, with sampling time and frequency offsets $\tau_i = 0$ and $\omega_i \sim U[\frac{-\Delta\omega}{2}, \frac{\Delta\omega}{2}]$.

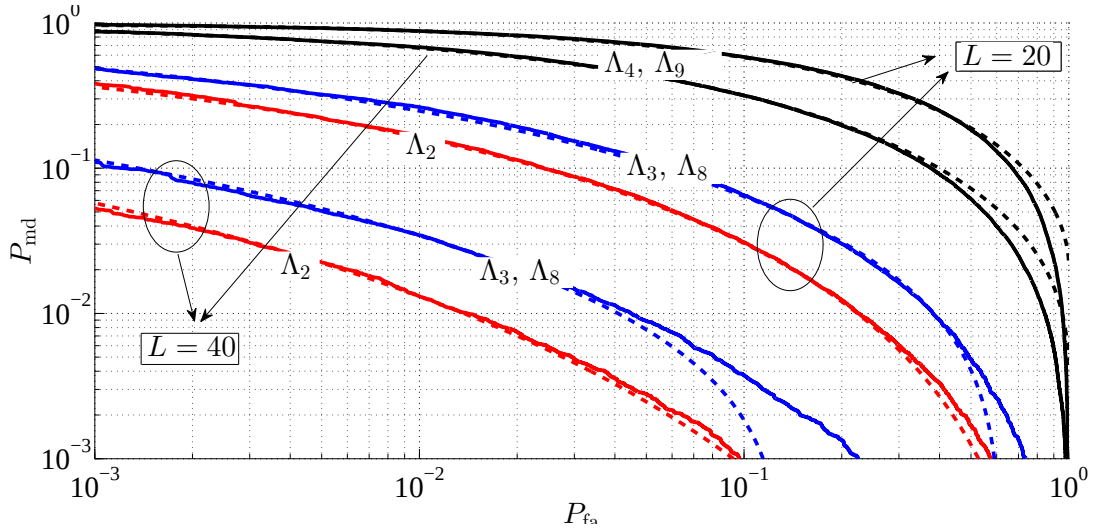


Figure 2.6: P_{md} versus P_{fa} of $\Lambda_2 - \Lambda_5$, Λ_8, and Λ_9 for the noncoherent signal model, employing $M = 4$ sensors, $L \in \{20, 40\}$ samples, and assuming $h_i = i$ and $\sigma^2 = 20$. The theoretical performance (in (2.13), (2.20), and (2.27)) are given by dashed lines.

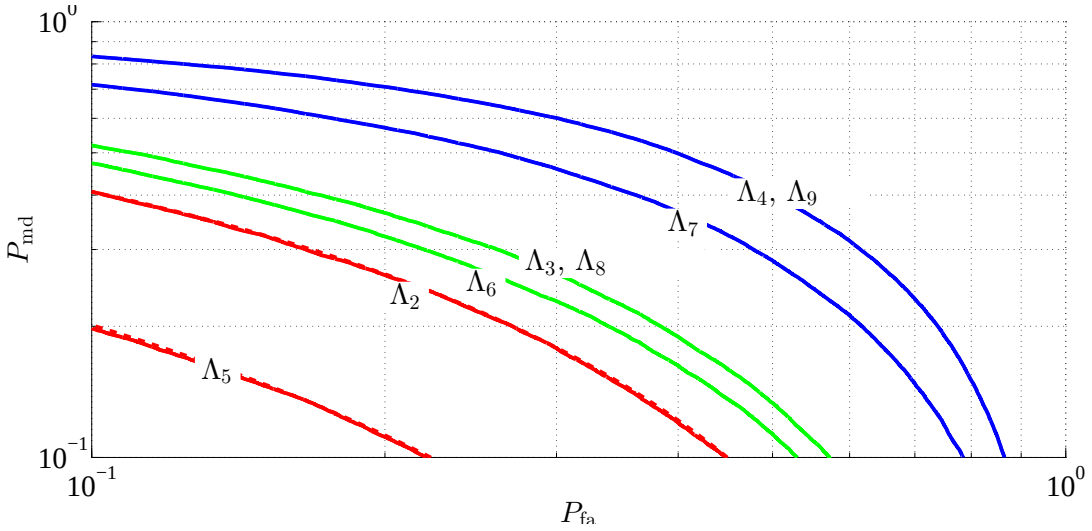


Figure 2.7: $P_{\rm md}$ versus $P_{\rm fa}$ of $\Lambda_1 - \Lambda_9$ for the coherent signal model, with $h_i = i$, $M = 4$, $\sigma^2 = 50$ and $L = 20$. The theoretical performance (in (2.31) and (2.35)) are given by dashed lines.

system, however, this comes at a very high cost as $\Lambda_5 - \Lambda_7$ require $2L$ times more bandwidth. The performance degradation of the suboptimal detectors can be compensated by an increase in the sample size, e.g., $\Lambda_2 - \Lambda_4$ with $L = 40$ outperform $\Lambda_5 - \Lambda_7$ with $L = 20$ respectively.

2.6 Conclusion

We have considered the problem of distributed detection where multiple sensors cooperatively detect presence or absence of a random source. We first considered partially correlated observations and obtained the optimal NPD. Assuming a noncoherent system, we obtained novel asymptotic expressions for the $P_{\rm fa}$ and $P_{\rm md}$ of several optimal/suboptimal detectors. We then compared these detectors with several detectors derived for the coherent system and showed that the coherent detectors require $2L$

times more bandwidth than the noncoherent ones. We also showed that the detectors for the noncoherent system are robust to the signal coherence and sampling time/frequency errors. Interestingly, the NPD for partially correlated observations reaches its highest performance when the observations are fully coherent and reaches its lowest performance when the observations are fully noncoherent. We also proposed one novel GLRD and two novel heuristic detectors for the noncoherent system. We also proved that all these detectors are either optimal or asymptotically optimal.

Chapter 3

Frequency Domain Spectrum Sensing

We consider distributed detection of an OFDM random source using a cooperative set of sensors assuming that the observations of different sensors are independent. The current OFDM detectors only make use of the signal correlation in time-domain and neglect the frequency-domain features. In order to exploit these features, we focus on the distributed detection of the OFDM signal in frequency-domain. We propose different frequency-domain detectors depending on the availability of knowledge of the SNRs and noise variances. We also derive the exact closed form expressions for the probabilities of false alarm and mis-detection for the NPD (which assumes SNRs and noise variances are known), and for a GLRD (which assumes they are unknown). Assuming known noise variances and unknown SNRs, we also propose two new GLRDs for two cases in which the transmit power is either uniformly allocated to all the subcarriers or not. We also prove that the GLRDs for unknown parameters and for uniform power allocation are asymptotically optimal. In addition, we propose a third

new Selection Combining (SC) based detector in which a select subset of observations is utilized for detection. Our theoretical analysis matches our simulation results and show that the proposed detectors have insignificant communication overhead. Despite their lower computational complexity, the proposed detectors outperform the state-of-the-art time-domain detectors in practical cases.

3.1 Introduction

In distributed detection, multiple sensors cooperatively detect presence or absence of a signal source. To this end, sensors transmit a summary of their observation to a FC, where the global decision is made. In contrast, sensors in the centralized counterpart, transmit all the (raw) observations to the FC [55]. We assume a noncoherent system and consider the distributed detection of an OFDM source signal in such a noncoherent system. Since it is crucial to efficiently detect the presence of the licensed users in cognitive radio, developing efficient detectors for OFDM systems using distributed noncoherent sensors is of great importance. As such, noncoherent detection of OFDM systems enables heterogeneous communication systems such as WiFi, WiMAX, LTE and DVB-T (which employ OFDM) to coexist.

Since the communication bandwidth in cognitive radio is restricted, distributed detection, wherein only a summary of the observations is transmitted by each sensor, is more practical. Several methods are proposed to obtain such summaries. For instance, it is proposed to transmit the energy [80, 94, 47], likelihood ratio [82], or correlation coefficient [33, 31] of the received signal to the FC. Employing sub-optimal techniques, the FC then combines the reported summaries to make a global decision. Another approach is to allow each sensor to transmit only a single bit to the FC.

In this setting, each sensor may employ matched filter detection [85], the energy detection [63], or the cyclo-stationary feature detection [43] to make a local decision, which will be transmitted to the FC. The FC may employ optimal or sub-optimal techniques (based on the counting rule, OR, and, AND detectors) to combine the local decisions to make a global decision [97, 55]. These detectors usually assume perfect OFDM symbol synchronization at the sensors. However, such synchronization is power consuming and adds more complexity and delay to the sensing task. The problem of OFDM signal detection with unsynchronized observations in time-domain is studied in [21, 31] and various time-domain detectors are proposed. These time-domain detectors only make use of the signal correlation, which is the most apparent feature of the OFDM signal in time-domain. However, they cannot exploit the frequency-domain features of the OFDM signal. We aim to design efficient OFDM signal detectors with low complexity and low communication overhead that can fully exploit its frequency-domain features. We also aim to evaluate the impact of the parameters of the OFDM source and sensors on the detection performance. To this end, we consider distributed detection of an OFDM source signal in frequency-domain with imperfect synchronization in two cases:

1. stationary channel: the channel responses remain unchanged over the entire detection interval,

2. non-stationary channel: the channel responses may vary over the detection interval.

The detectors in [21, 31] are computationally demanding since the distribution of unsynchronized OFDM symbols in time-domain is rather complex. We show that the distribution of such symbols in frequency-domain converges to a normal distribution

as the number of subcarriers increases. Using this result, we then derive detectors with lower computational complexity. Table 3.1 summarizes these frequency-domain detectors, denoted by $\Lambda_i, i = 1, \ldots, 5$, with the equation number, type, required parameters, and data summary collected from individual sensors. Two of these detectors, Λ_1 and Λ_4, have been proposed for different applications [62, 98]. We analyze the optimal NPD Λ_1 and derive the exact closed form expressions for the P_{fa} and P_{md}. In some situations the noise variance can be accurately estimated and is known to the sensors in advance. For such a situation, we propose two new GLRDs: Λ_2 and Λ_3; Λ_2 assumes that the unknown transmit power is uniformly distributed over all the subcarriers, whereas Λ_3 makes no such an assumption. When the SNRs and noise variances are unknown, the GLRD Λ_4 may be employed. Our novel asymptotical analysis accurately evaluates the performance of Λ_4 as the number of temporal samples is large enough in practice. We also propose a sub-optimal SC detector, Λ_5, where a subset of observations is selected and utilized for detection. This detector is a generalization of several existing detectors. We show that our proposed frequency-domain detectors have lower computational complexity compared to time-domain detectors. In addition, our simulation results show that the proposed frequency-domain detectors, despite their lower computational complexity, outperform some state-of-the-art time-domain detectors in practical cases.

Throughout this chapter, we denote the absence and the presence of the source signal by $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively. The Discrete Fourier Transform (DFT) of a vector $x \in \mathbb{C}^K$ is denoted by $X = \mathcal{F}x$, where $[X]_k = \sum_{n=0}^{K-1} x_n e^{-\frac{2\pi i}{K}kn}, k = 0, \ldots, K-1$ is the k^{th} element of X. The Inverse Discreet Fourier Transform (IDFT) of X is denoted by $x = \mathcal{F}^{-1}X$, where $[x]_n = \frac{1}{K}\sum_{k=0}^{K-1} X_n e^{\frac{2\pi i}{K}kn}, n = 0, \ldots, K-1$. We denote the MLE

Table 3.1: Summary of detectors. ✓: assumed to be known, ×: unknown parameter, ∗: unknown uniform transmit power, CO: $\frac{\#\text{ reported values}}{\#\text{ received samples}}$.

Decision statistic	Type	Subband SNRs	Noise variance	Local decision statistic to be reported to the FC	CO
Λ_1 (3.2)	NPD	✓	✓	$\sum_{n,k} w_{m,n,k} \frac{U_{m,n,k}}{\sigma_m^2}$	$\frac{1}{2KN}$
Λ_2 (3.9)	GLRD	×(∗)	✓	$\sum_{n,k} \frac{\lvert[\mathcal{F}\hat{\rho}_{m,n}(l)]_k\rvert}{\sigma_m^2+\lvert[\mathcal{F}\hat{\rho}_{m,n}(l)]_k\rvert} U_{m,n,k} - -\sigma_m^2 \log(1+\frac{\lvert[\mathcal{F}\hat{\rho}_{m,n}(l)]_k\rvert}{\sigma_m^2})$	$\frac{1}{2KN}$
Λ_3 (3.10)	GLR	×	✓	$\sum_{n,k} g\big(\frac{U_{m,n,k}}{\sigma_m^2}\big) u\big(\frac{U_{m,n,k}}{\sigma_m^2}-1\big)$	$\frac{1}{2KN}$
Λ_4 (3.12)	GLR	×	×	$\Omega(U_{m,1,1},\ldots,U_{m,N,K})$	$\frac{1}{2KN}$
Λ_5 (3.18)	SC	×	×	$\sum_{n,k} I_{m,n,k} \frac{U_{m,n,k}}{\bar{\sigma}_m^2}$	$\leq \frac{1}{2KN}$

of a parameter θ by $\hat{\theta}$.

This chapter is organized as follows. We introduce the system model in Section 3.2. We investigate the NPD in Section 3.3. In Section 3.4.1 assuming uniform transmit power, we derive the GLRD for unknown SNRs and known noise variance. We investigate GLRDs for two cases: 1) only the noise variances are known, and 2) the SNRs and noise variances are unknown in Section 3.4.2 and Section 3.4.3 respectively. We propose the SC detector in Section 3.5. In Section 3.6, we numerically evaluate the proposed detectors. Finally, we give our concluding remarks in Section 3.7.

3.2 System Model

We assume M sensors aim to cooperatively detect the presence or absence of an OFDM source employing K subcarriers. We assume a noncoherent system where sensors are not synchronized and observe different (independent) pieces of the same

source in independent additive white normal noise. We refer to a system as noncoherent if 1) the carrier frequency mismatch of the involved sensors are considerably less than the bandwidth of one subcarrier, i.e., they observe the same channel with no perfect synchronization, 2) the sampling times of the sensors are not synchronized, however, their sampling frequencies are (almost) identical, and 3) the sampling times of the sensors are such that the source observations from different sensors can be treated as independent.

Let $\{h_{m,n}(l)\}_{l=0}^{L_c-1}$ denote the channel impulse response between the source and $m \in \{0,\ldots,M-1\}$th sensor over the $n \in \{0,\ldots,N-1\}$th time interval, where L_c is the channel length. We assume that each sensor observes N discrete-time sequences with length K from the same OFDM source. Let $S_{n,k}$ denote the frequency-domain OFDM transmitted symbol at the k^{th} subcarrier during the n^{th} interval. The transmitter computes the IDFT of the symbols, i.e., $\{\mathcal{F}^{-1}S_{n,k}\}$, adds the cyclic prefix, and then transmits the result through the channel. We assume that the sensors are not synchronized and the m^{th} sensor records the n^{th} time interval of its received signal with some unknown delay, and some small offsets in carrier frequency and sampling rate. Therefore, the observed sequence may contain samples from the n^{th} and $n+1$st sets of OFDM symbols. The sensors then take DFT of the K samples of the recorded sequence to obtain $\{Y_{m,n,k}\}$. For a perfectly synchronized system it is known that $Y_{m,n,k} = H_{m,n,k}S_{n,k} + W_{m,n,k}$, where $W_{m,n,k}$ is a zero-mean complex normal noise with variance σ_m^2, i.e., $f(\{W_{m,n,k}\}) = \prod_{m,n,k} \frac{1}{\pi\sigma_m^2} \exp -\frac{|W_{m,n,k}|^2}{\sigma_m^2}$, and $H_{m,n,k} = \sum_{l=0}^{L_c-1} e^{-\frac{j2\pi kl}{K}} h_{m,n}(l)$ is the channel gain at frequency k.

However in an unsynchronized system, $Y_{m,n,k}$ is a linear transformation of the sequence $S_{n,k}$ and the relationship between $Y_{m,n,k}$ and $S_{n,k}$ is very complex. In this

chapter, we do not make use of this complex relationship as it requires the system to be fully synchronized, (i.e., the relationship depends on the unknown synchronization parameters). We argue that $\{Y_{m,n,k}\}_{k=0}^{K-1}$ are uncorrelated. We notice that the recorded data can be approximated as a wide sense stationary process and it is known that the DFT of such a process converges (with probability one) to the Karhunen-Loeve transform of the signal, as the DFT length increases [53, 95]. This is to say that as the DFT length increases (compared with the channel length), $\{Y_{m,n,k}\}_{k=0}^{K-1}$ become uncorrelated. Moreover, the variance of $Y_{m,n,k}$ converges (with probability one) to $P_k \left|H_{m,n,k}\right|^2 + \sigma_m^2$, where P_k represents the power spectral density of the input of the channel at k^{th} frequency index. This convergence is proven based on circular approximation of the symmetric Toeplitz autocovariance matrix of the recorded data. In addition, the symbols $S_{n,k}$ are usually taken from ν_k-QAM constellation points where the number of constellation points ν_k is random and unknown to the sensors. Thus the source symbols $S_{n,k}$ have a discrete mixture probability distribution with unknown parameters. In practice $Y_{m,n,k}$ is not only a linear combination of a number of source symbols $\{S_{n,k}\}$ but also depends on several other random parameters such as synchronization mismatches. Making use of the discrete mixture distribution of $S_{n,k}$ is impractical as it requires to deal with all the unknown parameters. Therefore by virtue of the central limit theorem, $\{Y_{m,n,k}\}$ converge in distribution to normal random variables as K increases. Thus throughout this chapter, we model the "unconditional distribution" of $\{Y_{m,n,k}\}_{k=0}^{K-1}$ as a set of independent normal random variables with zero mean and variances of $P_k \left|H_{m,n,k}\right|^2 + \sigma_m^2$. Obviously, it is not accurate to approximate the "conditional distribution" of $\{Y_{m,n,k}\}$ given the synchronization parameters and channel conditions with a normal PDFs. Figure 3.1 shows

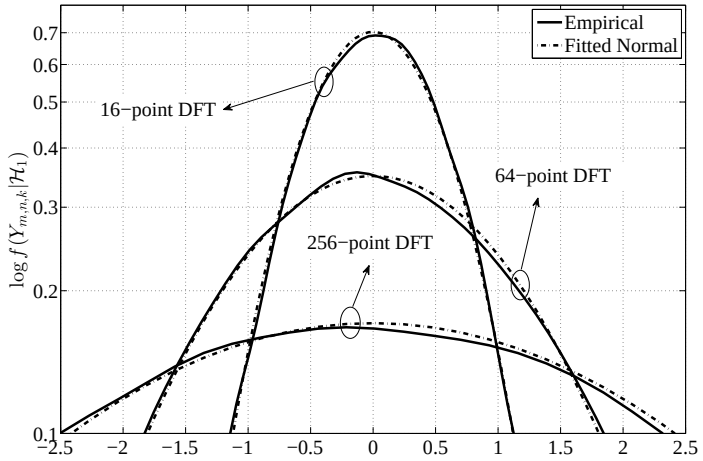


Figure 3.1: The empirical and the fitted normal PDFs of $Y_{m,n,k}$, i.e., $f(Y_{m,n,k}|\mathcal{H}_1)$, using DFT lengthes of 16, 64, and 256, where in each subcarrier either QPSK, 16QAM, or 64QAM constellations is randomly used with $[h_{m,n}(l)] = [1, \frac{1}{4}, \frac{1}{16}]$, $M = N = 1, K = 256$, uniform sampling time delay over $[0, K-1]$, and uniform frequency offsets over $[-\pi, \pi)$.

how closely the empirical PDFs of $Y_{m,n,k}$, i.e., $f(Y_{m,n,k}|\mathcal{H}_1)$, fit normal PDFs using DFT lengths of 16, 64, and 256, with $[h_{m,n}(l)] = [1, \frac{1}{4}, \frac{1}{16}]$, $M = N = 1$, and $K = 256$, where in each subcarrier either QPSK, 16QAM, or 64QAM is randomly used, the sampling time delay is uniformly distributed over $[0, K-1]$ and the frequency offsets is uniformly distributed over $[-\frac{\pi}{K}, \frac{\pi}{K})$.

Hence, the distribution of the recorded data by each sensor in the presence of time/frquency synchronization error and in frequency-domain converges to a normal PDF as $NK \to \infty$. Accordingly, we shall use this fact throughout this chapter to derive and study frequency-domain detectors.

3.3 Optimal Detector for Known Parameters

In this section, we derive the optimal NPD assuming that subband SNRs and noise variance are all known. This case is impractical, since availability of such knowledge in some applications where the spectrum dynamics is fast, is unreasonable.

The NPD compares the ratio of the PDFs of the observations $\{Y_{m,n,k}\}$ under two hypotheses with some threshold. As discussed in Section 3.2, we assume that $Y_{m,n,k}$ has zero mean complex normal distribution with variance σ_m^2 and $\theta_{m,n,k} = \sigma_m^2 + |H_{m,n,k}|^2 P_k$, under $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively, i.e.,

$$\begin{cases} f(\{Y_{m,n,k}\}|\mathcal{H}_1) = \prod_{m,n,k} \frac{\exp(-\frac{|Y_{m,n,k}|^2}{\theta_{m,n,k}})}{\pi\theta_{m,n,k}}, \\ f(\{Y_{m,n,k}\}|\mathcal{H}_0) = \prod_{m,n,k} \frac{\exp(-\frac{|Y_{m,n,k}|^2}{\sigma_m^2})}{\pi\sigma_m^2}. \end{cases} \tag{3.1}$$

It is easy to show that $\frac{f(\{Y_{m,n,k}\}|\mathcal{H}_1)}{f(\{Y_{m,n,k}\}|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \bar{\eta}_1$ can be rewritten as

$$\Lambda_1 \triangleq \sum_{m,n,k} w_{m,n,k} \frac{U_{m,n,k}}{\sigma_m^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_1, \tag{3.2}$$

where $U_{m,n,k} = |Y_{m,n,k}|^2$, $\eta_1 = \log(\bar{\eta}_1) + \sum_{m,n,k} \log(\frac{\theta_{m,n,k}}{\sigma_m^2})$, and $w_{m,n,k} = 1 - \frac{\sigma_m^2}{\theta_{m,n,k}}$.

Remark 3.1. *The detector Λ_1 in (3.2) first spatially normalizes the observations by dividing their norm by σ_m^2. The assigned weight $w_{m,n,k}$ to the normalized value is the ratio of the signal component of the subband energy, i.e., $|H_{m,n,k}|^2 P_k$, to the total subband energy, $\theta_{m,n,k}$. In addition, $w_{m,n,k} \in [0,1]$ is an increasing function of the subband SNR (defined as $\frac{|H_{m,n,k}|^2 P_k}{\sigma_m^2}$), i.e., $\lim_{SNR\to 0} w_{m,n,k} = 0$ and $\lim_{\mathrm{SNR}\to\infty} w_{m,n,k} = 1$. This suggests a sub-optimal detector by discarding the observations with lower received subband SNRs, and combining the other observations with higher channel*

gains with unit weight. This sub-optimal detector does not require the exact values of the received subband SNRs or noise variances and their approximate values can be used to select a subset of them.

In order to implement the fusion rule in (3.2), which is optimal for (3.1), sensors need to calculate two parameters, the noise variance σ_m^2 and total subband energy $\theta_{m,n,k}$. Each sensor computes

- the DFT of N observation vectors of size K,
- the subband observed energies $U_{m,n,k}$, and
- the weighted sum of the observed subband energies $\sum_{n,k} w_{m,n,k} \frac{U_{m,n,k}}{\sigma_m^2}$.

Since the number required operations is dominated by a DFT (requiring an order of $KN \log_2 K$ operations), Λ_1 requires an order of $KN \log_2 K$ operations at each participating sensor, which is fewer than that of the time-domain NPD proposed in [21] which requires an order of K^3N^2 operations. For every KN received observations (i.e., $2KN$ real numbers), the m^{th} sensor reports solely one real number, $\sum_{n,k} w_{m,n,k} \frac{U_{m,n,k}}{\sigma_m^2}$, to the FC. Obviously, the Communication Overhead (CO) defined as the ratio $\frac{\#\text{ reported values}}{\#\text{ received samples}}$ is $\frac{1}{2KN}$. The CO is very small for large values of K and N and is inversely proportional to the sensing time (see Table 3.1). The FC compares the summation of the reported values with η_1 to make the global decision.

The detector in (3.2) is investigated in [62, page 142] for a different application with a simpler model where the SNRs and noise variances are identical. However, to the best of our knowledge, the performance of (3.2) has not been analyzed to date. To this end, we introduce the following theorem that expresses the CDF of a linear combination of Gamma random variables.

Theorem 3.1 ([8]). *The CDF of $z = \sum_l \theta_l u_l$, where $\theta_l > 0$ and $\{u_l\}$ are independent random variables, each with Gamma distribution, i.e., $f(u_l) = u_l^{r_l - 1} \frac{e^{-u_l}}{\Gamma(r_l)}$, for $u_l > 0$, is given by*

$$F_{\mathbf{z}}(z; \boldsymbol{r}, \boldsymbol{\theta}) = \frac{1}{\prod_{l=1}^{L} \theta_l^{r_l}} \sum_{k=1}^{L} \sum_{l=1}^{r_k} \Psi_{k,l,\mathbf{r}} \theta_k^{r_l} \gamma(r_l, \frac{z}{\theta_k}) \quad \textit{for } x \geq 0, \tag{3.3a}$$

where $\gamma(s, x) = 1 - e^{-x} \sum_{q=0}^{s-1} \frac{x^q}{q!}$, $\mathbf{r} = [r_1, \ldots, r_L]$, $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_L]$, and,

$$\Psi_{k,l,\mathbf{r}} = \sum_{[i_1,\ldots,i_l] \in \mathbb{A}_{k,l}} \prod_{1 \leq j \leq l, j \neq k} \binom{i_j + r_j - 1}{i_j} \Big(\frac{1}{\theta_j} - \frac{1}{\theta_k}\Big)^{-(r_j + i_j)}, \tag{3.3b}$$

where the set $\mathbb{A}_{k,l}$ is defined as

$$\mathbb{A}_{k,l} = \Big\{ [i_1, \ldots, i_l] \in \mathbb{Z}^l \Big| \sum_{j=1}^{L} i_j = l - 1, i_k = 0, i_j \geq 0 \ \forall j \Big\}. \tag{3.3c}$$

In non-stationary environment when $w_{m,n,k}$ are distinct for different (m, n, k), employing Theorem 3.1 and some algebraic simplification, it can be shown that the CDF of $\Lambda_1 = \sum_l \theta_l u_l$ is given by

$$F_{\boldsymbol{\Lambda_1}}(\Lambda_1; \boldsymbol{\theta}) = \sum_{l=0}^{MNK-1} (1 - e^{-\frac{\Lambda_1}{\theta_l}}) \prod_{0 \leq \ell \leq MNK-1, \ell \neq l} \frac{\theta_\ell}{\theta_\ell - \theta_l}, \tag{3.4}$$

where $\boldsymbol{\theta} = [\theta_0, \ldots, \theta_{MNK-1}]$. Obviously, we have $\theta_{mNK+nK+k} = [\boldsymbol{\theta}_0]_{mNK+nK+k} = w_{m,n,k}$ under $\mathcal{H}_0$ and $\theta_{mNK+nK+k} = [\boldsymbol{\theta}_1]_{mNK+nK+k} = \frac{\theta_{m,n,k} - \sigma_m^2}{\sigma_m^2}$ under $\mathcal{H}_1$. As a result,

the P_{fa} and P_{md} of Λ_1 in (3.2) are expressed as

$$\text{non-stationary model:} \quad \begin{cases} P_{\text{fa},\Lambda_1} = 1 - F_{\mathbf{\Lambda_1}}\left(\eta_1; \boldsymbol{\theta}_0\right), \\ P_{\text{md},\Lambda_1} = F_{\mathbf{\Lambda_1}}\left(\eta_1; \boldsymbol{\theta}_1\right), \end{cases} \tag{3.5}$$

where $F_{\mathbf{\Lambda_1}}$ is defined in (3.4).

Remark 3.2. *Under the stationary channel model, the channel impulse response $h_{m,n}(l)$ remains static (invariant with respect to n) during the spectrum sensing interval. This is in contrast to the non-stationary channel model where the channel impulse response is constant only over the duration of one observation interval, but varies from interval to interval. Thus for the stationary channel model, the DFT of the channel impulse response in* (3.2) *is constant over the spectrum sensing interval and can be approximated as $H_{m,n,k} \approx H_{m,k}$, i.e., $w_{m,n,k} \approx w_{m,k}$, and $\theta_{m,n,k} \approx \theta_{m,k}$. In such condition for the case where $w_{m,k}$ are distinct values for different (m,k), Λ_1 becomes a linear combination of MK independent random variables $u_{k,m}$ with Gamma distribution, i.e., $f(u_{k,m}) = \frac{u_{k,m}^{N-1} e^{-u_{k,m}}}{\Gamma(N)}, u_{k,m} > 0$. Employing Theorem 3.1, the P_{fa} and P_{md} of* (3.2) *under stationary channel model are expressed as*

$$\textit{stationary model:} \quad \begin{cases} P_{\text{fa},\Lambda_1} = 1 - F_{\boldsymbol{z}}\left(\eta_1; \boldsymbol{r}, \boldsymbol{\theta}'_0\right), \\ P_{\text{md},\Lambda_1} = F_{\boldsymbol{z}}\left(\eta_1; \boldsymbol{r}, \boldsymbol{\theta}'_1\right), \end{cases} \tag{3.6}$$

where $\boldsymbol{r} = [N, \ldots, N]$, $[\boldsymbol{\theta}'_0]_{mK+k} = w_{m,k}$ and $[\boldsymbol{\theta}'_1]_{mK+k} = |H_{m,n,k}|^2 P_k / \sigma_m^2$, and $F_{\boldsymbol{z}}$ is defined in (3.3).

3.4 Generalized Likelihood Ratio Detectors

3.4.1 GLRD for Uniform Power Distribution and Unknown SNRs

We previously assumed that the subband SNRs and noise variance are known. However in some cases the environment has a fast changing dynamics and it is unreasonable to assume the availability of the channel knowledge at the sensors. On the other hand, due to the fast dynamics of the channel, the OFDM source cannot adapt to the channel variations and therefore allocates its transmit power uniformly to the different sub-channels. In this section, we address the detection of such an OFDM source, where the transmit power of subbands P_k are assumed equal yet unknown. Without loss of generality, let the channel gains $H_{m,n,k}$ absorb the subband transmit power, i.e., $P_k = 1$. Hence, to obtain the GLRD, we only need to find the MLE of $|H_{m,n,k}|^2$ by maximizing the PDFs of the observations $f(\{Y_{m,n,k}\}|\mathcal{H}_1)$ in (3.1) with respect to $|H_{m,n,k}|^2$. Note that since $h_m(n,l) = 0$ for $l \geq L_c$, the projection of the IDFT of $\{|H_{m,n,k}|^2\}$ is zero on some components, i.e., $[\mathcal{F}^{-1}|H_{m,n,k}|^2]_l = 0, L_c \leq l \leq K - L_c$. Hence, $\widehat{|H_{m,n,k}|^2}$ is the solution to the following optimization problem:

$$\begin{cases} \min\limits_{|H_{m,n,k}|^2} \sum\limits_{m,n,k} \frac{U_{m,n,k}}{\sigma_m^2 + |H_{m,n,k}|^2} + \log(\sigma_m^2 + |H_{m,n,k}|^2) \\ [\mathcal{F}^{-1}|H_{m,n,k}|^2]_l = 0, \quad L_c \leq l \leq K - L_c. \end{cases} \tag{3.7}$$

Since this is a non-convex optimization problem, finding the general optimum is computationally expensive when K is large. We notice the solution to the optimization problem (3.7), when the constraints are relaxed, is given by $\max(U_{m,n,k} - \sigma_m^2, 0)$. Using this solution, we propose to use the following method to obtain an approximate

solution:

1. Take IDFT of $\max(U_{m,n,k} - \sigma_m^2, 0)$ and retain the first L_c and last $L_c - 1$ elements, i.e.,

$$\hat{\rho}_{m,n}(l) = \begin{cases} 0, & L_c \leq l \leq K - L_c \\ [\mathcal{F}^{-1}\max(U_{m,n,k} - \sigma_m^2, 0)]_l, & \text{else.} \end{cases} \tag{3.8}$$

2. Take DFT of $\hat{\rho}_{m,n}(l)$ and use $|[\mathcal{F}\hat{\rho}_{m,n}(l)]|$ as a feasible initial estimate of $|H_{m,n,k}|^2$.

One may use exhaustive search to solve (3.7) and then obtain the GLRD. However to reduce the computational cost, we approximate the solution of (3.7) with $|\widehat{H_{m,n,k}}|^2 \approx |[\mathcal{F}\hat{\rho}_{m,n}(l)]_k|$.

Substituting this approximation in $\frac{f(\{Y_{m,n,k}\}|\mathcal{H}_1)}{f(\{Y_{m,n,k}\}|\mathcal{H}_0)} \gtrless_{\mathcal{H}_0}^{\mathcal{H}_1} \bar{\eta}_2$, with $\bar{\eta}_2$ being the detection threshold, we propose the following sub-optimal GLRD:

$$\Lambda_2 \triangleq \sum_{m,n,k} \frac{|[\mathcal{F}\hat{\rho}_{m,n}(l)]_k|}{\sigma_m^2 + |[\mathcal{F}\hat{\rho}_{m,n}(l)]_k|} \frac{U_{m,n,k}}{\sigma_m^2} - \log(1 + \frac{|[\mathcal{F}\hat{\rho}_{m,n}(l)]_k|}{\sigma_m^2}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_2, \tag{3.9}$$

where $\eta_2 = \log(\bar{\eta}_2)$.

Remark 3.3. *Under the stationary channel mode, relaxing the constraints, the solution to* (3.7) *is given by* $\max(\frac{1}{N}U_{m,k} - \sigma_m^2, 0)$*, where* $U_{m,k} = \sum_n U_{m,n,k}$*. In this condition, the same procedure, as described above, provides a feasible starting point when the term* $U_{m,n,k}$ *is replaced with* $\frac{U_{m,k}}{N}$ *in* (3.8)*.*

Using (3.9), each sensor only needs to know the noise variances σ_m^2 and performs an order of $NK\log_2 K$ operations (due to the dominant contribution of DFT in the complexity). This detector exploits the additional a-priori knowledge of uniform

transmit powers. However, this exploitation involves extra cost for the computation of $||[\mathcal{F}\hat{\rho}_{m,n}(l)]||$ using (3.8).

Since Λ_2 assumes uniform power allocation across all the sub-channels, its performance degrades substantially if this condition is not satisfied. For non-uniform transmit powers, this detector misinterprets the channel constraint and incorrectly imposes it on the received subband powers, i.e., $[\mathcal{F}^{-1}S_{m,n,k}]_l = 0$, for $L_c \leq l \leq K - L_c$ which is not necessarily correct if $\{P_k\}$ are non-identical.

3.4.2 GLRD for Non-uniform Power Distribution and Unknown SNRs

In contrast to the Section 3.4.1, here we assume the channel variations are slow and therefore the OFDM source employs bit-loading techniques (e.g. [99]) and adapt the transmit powers according to the channel. As such, we derive the GLRD assuming that $\{|H_{m,n,k}|^2 P_k\}$ are unknown while $\{\sigma_m^2\}$ are known. To this end, we need the MLE of $\theta_{m,n,k}$, i.e., $\widehat{\theta}_{m,n,k}$, which is found by maximizing $f(\{Y_{m,n,k}\}|\mathcal{H}_1)$ with respect to $\theta_{m,n,k}$ as $\widehat{\theta}_{m,n,k} = \max(U_{m,n,k}, \sigma_m^2)$. Substituting $\widehat{\theta}_{m,n,k}$ in the likelihood ratio $\frac{f(\{Y_{m,n,k}\}|\mathcal{H}_1)}{f(\{Y_{m,n,k}\}|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \bar{\eta}_3$, with $\bar{\eta}_3$ being the detection threshold, and simplifying the result, we obtain the following GLRD for the non-stationary channel model:

$$\Lambda_3 \triangleq \sum_{m,n,k} g\big(\frac{U_{m,n,k}}{\sigma_m^2}\big) u\big(\frac{U_{m,n,k}}{\sigma_m^2} - 1\big) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_3, \tag{3.10}$$

where $\eta_3 = \log \bar{\eta}_3$, $u(.)$ is the step function, and $g(x) = x - 1 - \log(x)$.

Remark 3.4. *Under the stationary channel model, the MLE of $\theta_{m,k}$ is given by $\widehat{\theta}_{m,k} = \max(\frac{U_{m,k}}{N}, \sigma_m^2)$. In this condition, the GLRD in* (3.10) *is still applicable.*

However, the term $U_{m,n,k}$ must be replaced with $\frac{U_{m,k}}{N}$.

For the GLRD in (3.10), each sensor needs to know its noise variance. In this setting, each sensor first censors less informative observations with insignificant energy, i.e., those for which $\frac{U_{m,n,k}}{\sigma_m^2} < 1$, (or $\frac{U_{m,k}}{N\sigma_m^2} < 1$ under the stationary model) and only takes into account the observations that favor the hypothesis $\mathcal{H}_1$. Then, using the function $g(x) = x - 1 - \log(x)$, sensors performs a transform on the local energies and forwards the summation $\sum_{n,k} g\big(\frac{U_{m,n,k}}{\sigma_m^2}\big) u\big(\frac{U_{m,n,k}}{\sigma_m^2} - 1\big)$. This detector requires each participating sensor to perform only an order of $KN \log_2 K$ operations (due to the dominant contribution of DFT in the complexity), which is fewer than that of the suboptimal time-domain detector proposed in [21] which requires an order of NK^2 operations.

3.4.3 GLRD for Unknown SNRs and Noise Variances

So far, the noise variance was assumed to be known. However in some situations, the noise spectrum may vary with time and therefore it must be estimated based on the acquired samples. Here, we treat the noise variance as unknown and derive a GLRD. To obtain this detector, we must find the MLE of the unknown parameters under each hypothesis. The obtained MLE then will be substituted in the corresponding PDFs, the ratio of which, gives the desired GLRD. The MLEs are computed by maximizing the PDFs of the observations in (3.1) with respect to $\theta_{m,n,k}$ and σ_m^2. Thus we have

$$\mathcal{H}_0 : \widehat{\sigma_m^2} = \frac{\sum_{n,k} U_{m,n,k}}{NK}, \quad \mathcal{H}_1 : \widehat{\theta}_{m,n,k} = U_{m,n,k}. \tag{3.11}$$

Substituting (3.11) respectively in $f(\{Y_{m,n,k}\}|\mathcal{H}_0)$ and $f(\{Y_{m,n,k}\}|\mathcal{H}_1)$, we obtain the following GLRD:

$$\Lambda_4 \triangleq \sum_m \Omega(U_{m,1,1},\ldots,U_{m,N,K}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_4. \tag{3.12}$$

where $\Omega(x_1,\ldots,x_p) = \log(\frac{x_1+\ldots+x_p}{p}) - \frac{\log x_1+\ldots+\log x_p}{p}$, referred to as the homogeneity index of $(x_1,\ldots,x_p)$, and η_4 is the detection threshold.

Remark 3.5. *For the stationary channel model and under $\mathcal{H}_1$, it can be shown that $\widehat{\theta}_{m,n,k} = \frac{U_{m,k}}{N}$. Thus in order to use Λ_4 for such condition, the term $U_{m,n,k}$ must be replaced by $\frac{U_{m,k}}{N}$.*

The detector Λ_4 is a constant false alarm detector and does not require a-priori knowledge about the noise variance and the SNRs. For this detector, each sensor must carry out an order of $KN\log_2 K$ computations (due to the dominant contribution of DFT in the complexity) to obtain the real number $\Omega(U_{m,1,1},\ldots,U_{m,N,K})$ and shall report it to the FC. The complexity of the time-domain GLR detector proposed in [21] is of order of NK^2 and is more than that of Λ_4. This detector compares the ratio of arithmetic to geometrical mean of the reported subband energies ($U_{m,n,k}$ or $\frac{U_{m,k}}{N}$) with a threshold. In fact, the ratio of geometrical mean to arithmetic mean is measure of spectral flatness, i.e., Λ_4 is always positive and quantifies how the observed signal resembles a white noise process. A smaller value for Λ_4 indicates that the energy is more uniformly distributed over all the subbands and the power spectrum is more flat. Otherwise, a large value of Λ_4 indicates that more power is concentrated in a number of bands than the other subbands [68]. Note that under $\mathcal{H}_0$, the signal is a white noise process and has a flat spectrum. As a drawback, this detector fails in

some cases where under $\mathcal{H}_1$, the spectrum is flat (see Remark 3.6). In such cases, Λ_4 misinterprets the observation as white noise and favors $\mathcal{H}_0$.

Interestingly, the ratio of the geometric mean to the arithmetic mean also appears as one of the well known source enumeration criterion in the literature (e.g. in MDL or AIC criteria) [98, 27]. In current practical standards NK is a large number whereas M can be small. To the best of our knowledge, the performance of (3.12) has not been analyzed for this case. We use the following theorems to find the asymptotic PDFs of Λ_4 as $KN \to \infty$ under $\mathcal{H}_0$ and $\mathcal{H}_1$. The analysis carried out in the following theorems is novel and has been recently cited and used in [77] for performance analysis of a similar detector.

Theorem 3.2. *For large NK, the distribution of Λ_4 in (3.12) under $\mathcal{H}_0$ is approximately normal with the following mean and variance*

$$E[\Lambda_4|\mathcal{H}_0] = M(\log(\alpha) - \psi(\alpha)),\ var[\Lambda_4|\mathcal{H}_0] = \frac{\alpha M \psi'(\alpha) - M}{NK}, \tag{3.13}$$

where $\psi(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha)$ is the digamma function and $\alpha = 1$ for non-stationary, and $\alpha = N$ for stationary model.

Proof. Similar to the proof in [50] assuming that $\{U_{m,n,k}\}$ are independent with identical Gamma distribution, i.e., $\{U_{m,n,k}\} \sim \exp(-\sum_{n,k} \frac{U_{m,n,k}}{\sigma_m^2}) \prod_{n,k} \frac{U_{m,n,k}^{\alpha_m - 1}}{\Gamma(\alpha_m)\sigma_m^{2\alpha_m}}$, it is easy to show that $\Omega(U_{m,1,1}, \ldots, U_{m,N,K})$ is the MLE of $\log(\alpha_m) - \psi(\alpha_m)$. It is proved that as $NK \to \infty$ the MLE gives an unbiased estimation and converges in distribution to a normal distribution (e.g., see [58]) with a variance given by Cramér–Rao bound. The proof is complete since Λ_4 is the summation of the MLEs obtained from different sensors. □

Theorem 3.3. *For large NK, the distribution of Λ_4 in (3.12) under $\mathcal{H}_1$ is approximately normal with the following mean and variance*

$$\begin{aligned} E[\Lambda_4|\mathcal{H}_1] &= \sum_m \Omega(\theta_{m,1,1},\ldots,\theta_{m,N,K}) + M(\log(\alpha) - \psi(\alpha)), \\ var[\Lambda_4|\mathcal{H}_1] &= \tfrac{\alpha M\psi'(\alpha)-M}{NK} + \alpha^{-1}\textstyle\sum_m \frac{\sum_{n,k}\theta^2_{m,n,k}}{(\sum_{n,k}\theta_{m,n,k})^2} - \frac{M}{NK}, \end{aligned} \tag{3.14}$$

Proof. Under $\mathcal{H}_1$, we rewrite Λ_4 as follows:

$$\begin{aligned} \Lambda_4 = \sum_m \Omega(\theta_{m,1,1},\ldots,\theta_{m,N,K}) + \Omega(Z_{m,1,1},\ldots,Z_{m,N,K}) + \\ \log\sum_{n,k}\left(\frac{Z_{m,n,k}}{\sum_{n,k}Z_{m,n,k}}\frac{NK\theta_{m,n,k}}{\sum_{n,k}\theta_{m,n,k}}\right), \end{aligned} \tag{3.15}$$

where $Z_{m,n,k} = \frac{U_{m,n,k}}{\theta_{m,n,k}}$. The first term in the above summation is constant which only appears in $E[\Lambda_4|\mathcal{H}_1]$. The distribution of the second term in (3.15) is given by Theorem 3.2 as $\{Z_{m,n,k}\}$ are independent with identical exponential distribution. For the last term in (3.15), we notice that it is a function of $\{x_{m,n,k} = \frac{Z_{m,n,k}}{\sum_{n,k}Z_{m,n,k}}\}$ which has a Dericlet distribution, i.e., $f(\{x_{m,n,k}\}) = \frac{\Gamma(NK)}{\Gamma^{KN/\alpha}(\alpha)}\prod_{n,k}x^{\alpha-1}_{m,n,k}$. It is easy to find the mean and the variance of the last term using a Taylor expansion, i.e.,

$$\log\frac{\sum_{n,k}NKx_{m,n,k}\theta_{m,n,k}}{\sum_{n,k}\theta_{m,n,k}} \approx$$

$$\left(\frac{\sum_{n,k}NKx_{m,n,k}\theta_{m,n,k}}{\sum_{n,k}\theta_{m,n,k}} - 1\right) - \frac{1}{2}\left(\frac{\sum_{n,k}NKx_{m,n,k}\theta_{m,n,k}}{\sum_{n,k}\theta_{m,n,k}} - 1\right)^2 + \ldots$$

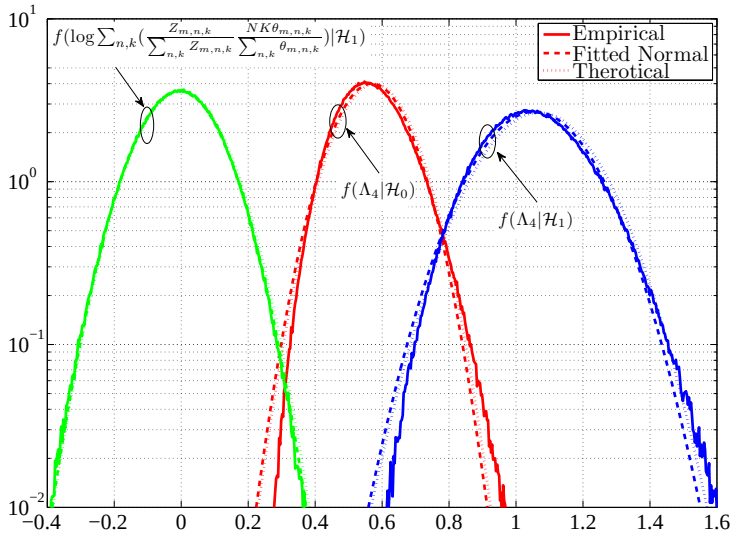


Figure 3.2: The empirical PDFs of the detector statistic Λ_4 under $\mathcal{H}_0$ and $\mathcal{H}_1$ and the approximated normal distributions in (3.13) and (3.3), with $M = P = N = 1, K = 512$.

as follows

$$E[\log \frac{\sum x_{m,n,k}\theta_{m,n,k}}{(NK)^{-1}\sum \theta_{m,n,k}}] \approx 0, \tag{3.16a}$$

$$\text{var}[\log \frac{\sum x_{m,n,k}\theta_{m,n,k}}{(NK)^{-1}\sum \theta_{m,n,k}}] \approx \alpha^{-1}\frac{\sum \theta^2_{m,n,k}}{(\sum \theta_{m,n,k})^2} - \frac{1}{NK}, \tag{3.16b}$$

$$\text{Cov}[\log(\frac{\sum \theta_{m,n,k}x_{m,n,k}}{(NK)^{-1}\sum \theta_{m,n,k}}), \sum \log(x_{m,n,k})] \approx 0, \tag{3.16c}$$

From (3.16c), we conclude that the 2nd and 3rd terms in (3.15) are uncorrelated for large NK (which is accurate to $O(\frac{1}{NK})$). In addition, our simulation results reveal that the last term is also asymptotically normal (see Figure 3.2). This implies that these terms are jointly normal and independent for large NK. Therefore, Λ_4 under $\mathcal{H}_1$ is a normal with the variance in (3.14) which is the sum of (3.16b) and the variance in (3.14). □

Figure 3.2 depicts the empirical PDFs, the normal PDFs obtained from the empirical mean and variance, and the theoretical normal PDFs in (3.13) and (3.14) for Λ_4 under $\mathcal{H}_0$ and $\mathcal{H}_1$, and $\log\sum(\frac{Z_{m,n,k}}{\sum Z_{m,n,k}}\frac{NK\theta_{m,n,k}}{\sum\theta_{m,n,k}})$ under $\mathcal{H}_1$, where $M=N=1$, $K=64$ and $\{\theta_{m,n,k}\}$ are such that $\frac{\theta_{m,n,k}-1}{50}\sim\exp(-x)$, $\Omega(\theta_{m,1,1},\ldots,\theta_{m,N,K})=0.5$, and $\frac{NK\sum\theta^2_{m,n,k}}{(\sum\theta_{m,n,k})^2}=1.8$. This reveals that our theoretical expressions are accurate. Note that a better accuracy can be obtained by considering higher moments and cumulants, which results in intractable expressions for the performance of Λ_4.

Remark 3.6. *The last two theorems indicate that*

$$E[\Lambda_4|\mathcal{H}_1]-E[\Lambda_4|\mathcal{H}_0]=\sum_m\Omega(\theta_{m,1,1},\ldots,\theta_{m,N,K})$$

which is the homogeneity index of the subband energies. This is why a better performance is achieved when the center of the two PDFs are far apart, which happens only if $\theta_{m,n,k}$ are heterogenous, i.e., the source spectrum observed by sensors are non-flat. It is easy to show that the homogeneity index is a decreasing function of σ^2_m. In addition in high SNRs, the difference between the two means tends to the homogeneity index of the source component of the subband energies, i.e., $|H_{m,n,k}|^2P_k$.

The P_{fa} and P_{md} of Λ_4 can be expressed as follows:

$$\begin{cases} P_{\text{fa},\Lambda_4} & \approx Q(\frac{\overline{\eta_4}}{\sqrt{\frac{\alpha M\psi'(\alpha)-M}{NK}}}),\\ P_{\text{md},\Lambda_4} & \approx 1-Q(\frac{\overline{\eta_4}-\sum_m\Omega(\theta_{m,1,1},\ldots,\theta_{m,N,K})}{\sqrt{\frac{\alpha M\psi'(\alpha)-M}{NK}+\alpha^{-1}\frac{\sum\theta^2_{m,n,k}}{(\sum\theta_{m,n,k})^2}-\frac{1}{NK}}}), \end{cases} \tag{3.17}$$

where $Q(x)=\frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}dt$, and $\overline{\eta_4}=\eta_4-M(\log(P)-\psi(P))$.

It can happen that the power spectrum of the received source component is relatively flat, e.g., at low SNRs . In such a case, if all the sensors observe uniformly distributed $\theta_{m,n,k}$ over all the subbands, this detector misinterpret the observations as noise and favors $\mathcal{H}_0$. In addition under $\mathcal{H}_0$, the P_{fa} increases for highly colored noise where the noise energy is not uniformly distributed over subbands. This implies that Λ_4 performs well only if the received power spectrum from the source is non-flat and but that of noise is.

Theorem 3.4. *The detectors* (3.10) *and* (3.12) *are asymptotically optimal, in the sense of Theorem 2.3, as* $NK \to \infty$.

Proof. The functions (3.10) and (3.12) are two separating functions [49] respectively in terms of the unknown parameters $\{\theta_{m,n,k}\}$ and $\{\sigma_m^2\}$. Thus, they are separating function estimation tests and hence are asymptotically optimal [49, Th. 8]. □

3.5 Selection Combining Detector

The optimal NPD Λ_1 needs the exact values of $\theta_{m,n,k}$ and σ_m^2. However, in cognitive radio, sensors cannot obtain these parameters in isolation. Moreover, in most OFDM applications, the source occupies only a given subset of sub-channels [38]. However, such a-priori knowledge is not incorporated in the detectors Λ_2 in (3.9), Λ_3 in (3.10), and Λ_4 in (3.12). Here, motivated by these facts and Remark 3.1, we propose an SC detector that utilizes a selected subset of sensors and observations. We propose that each sensor approximates the noise variance σ_m^2 in (3.2) by $\overline{\sigma}_m^2 \in \mathbb{R}^+$ and the weight

factor $w_{m,n,k}$ in (3.2) by a binary value $I_{m,n,k} \in \{0,1\}$, as follows

$$\Lambda_5 \triangleq \sum_{m,n,k} I_{m,n,k} \frac{U_{m,n,k}}{\overline{\sigma}_m^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_5. \tag{3.18}$$

Here, the sensors do not need the exact values of the involved parameters and only require enough information to assign the weight of 0 or 1 and $\overline{\sigma}_m^2$ to normalize the observations. In this way, the m^{th} sensor reports $\sum_{n,k} I_{m,n,k} \frac{U_{m,n,k}}{\overline{\sigma}_m^2}$ to the FC only if this value is non-zero. For (3.18), the CO is smaller than other methods as the probability of zero weights for $I_{m,n,k}$ increases. Furthermore, the FC adds the reported values and makes the global decision. Similarly the computational complexity of Λ_5 is of order of $NK \log_2 K$ (due to the dominant contribution of DFT in the complexity), which is lower than that of the time-domain GLRD proposed in [21] which requires an order of NK^2 operations. The reason for the reduction in the complexity of the detectors in this chapter is that the signals are uncorrelated in frequency-domain, whereas time-domain algorithms deal with non-diagonal matrices. The following examples illustrate how $I_{m,n,k}$ and $\overline{\sigma}_m^2$ can be approximated in practise.

Example 3.1. *If we substitute $\overline{\sigma}_m^2 = 1, I_{m,n,k} = 1, \forall m,n,k$ in* (3.18)*, we obtain the energy detector [48]. The energy detector decides between $\mathcal{H}_0$ and $\mathcal{H}_1$ by comparing the sum of the measured subband energies to the threshold η_{EN}, i.e., $\sum_{m,n,k} U_{m,n,k} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{\text{EN}}$. The CDF of this detector is given by* (3.4)*. Thus its probabilities of false-alarm $P_{\text{fa,EN}}$ and mis-detection $P_{\text{md,EN}}$ can be expressed as follows:*

$$P_{\text{fa,EN}} = 1 - F_{\boldsymbol{z}}\left(\eta_{\text{EN}}; [NK, \ldots, NK], [\sigma_1^2, \ldots, \sigma_M^2]\right),$$
$$P_{\text{md,EN}} = F_{\boldsymbol{\Lambda_1}}\left(\eta_{\text{EN}}; \boldsymbol{\theta}_{\text{EN}}\right),$$

where $[\boldsymbol{\theta}_{\mathrm{EN}}]_{mNK+nK+k} = S_{m,n,k}$ *and* $F_{\boldsymbol{z}}$ *and* $F_{\boldsymbol{\Lambda_1}}$ *are defined in* (3.3) *and* (3.4)*, respectively.*

Example 3.2. *In applications such as LTE or WiMax, a number of subcarriers are not employed for signal transmission [38]. In this case, the sensors can use the average of the measured subband energies in those subbands as* $\overline{\sigma}_m^2$*. Or alternatively use the average of a number of smallest values of* $U_{m,n,k}$ *as* $\overline{\sigma}_m^2$*. The observations with lower energy provide a better approximation for the noise variance. In addition, for higher SNRs it is more likely that* $U_{m,k}$ *takes larger values. Thus we propose to sort data* $\{U_{m,k}\}$ *at each sensor and let* $\{\widetilde{U}_{m,l}\}_{l=1}^{K}$ *denote the sorted data and use the average of a number of smallest values as* $\overline{\sigma}_m^2$*, and use the following as decision statistic*

$$\sum_{m,l} \frac{\widetilde{U}_{m,l}}{\overline{\sigma}_m^2} u(\widetilde{U}_{m,l} - \widetilde{U}_{m,L}) \tag{3.19}$$

where $u(.)$ *is the step function,* $L \in \{1, \dots K-1\}$ *is the number subbands assumed to be vacant. For the stationary channel model, the same criteria can be used, however, the term* $U_{m,n,k}$ *must be replaced with* $\frac{U_{m,k}}{N}$*. This detector requires that each sensor performs an order of* $KN \log_2 K$ *operations. In addition, it has constant false alarm rate and does not need a-priori knowledge of the noise variance or SNRs.*

3.6 Simulation Results and Discussion

In this section, we numerically evaluate the proposed detectors. We generate a multi-path time varying channel with the correlation function of $E[h_{m,n}(l)h_{m,n'}^H(l)] = \sigma_{m,l}^2 J_0(2\pi f_d(n-n'))$, where σ_{ml}^2 represents the normalized power of the l^{th} path of the channel between the source and m^{th} sensor such that $\sum_l \sigma_{m,l}^2 = 1$, J_0 denotes

the zeroth-order Bessel function of the first kind, and f_d is Doppler frequency. In the simulations, a multi-path wireless channel with an exponentially decaying power delay profile is chosen, i.e., $L_c = 3$, and $\sigma^2_{m,0} = .448$, $\sigma^2_{m,1} = .322$, and $\sigma^2_{m,2} = .230$ [78].

Figure 3.3 depicts the average mis-detection probability versus the average false-alarm probability of the detectors Λ_1 to Λ_5 with $K = 512, N = 20, f_d = .0285, P_k = 1$ and a cyclic-prefix $L_p = 8$ for two scenarios: 1) $M = 4$ and $\sigma^2_m = 50m$ and 2) $M = 8$ and $\sigma^2_m = 50m$. We use the energy detector, as in Example 3.1 for the SC detector Λ_5. We observe that the performance of the detectors Λ_1 to Λ_4 improves as the number of observation intervals N increases. This means that the performance loss of the sub-optimal detectors can be compensated by increasing the number of observation intervals. For instance, our simulation results show that the sub-optimal detector Λ_2 for $N = 60$ outperforms the optimal detector Λ_1 with $N = 20$. Since the noise variances of the sensors $m \in \{4, \ldots, 8\}$ are large, the performance of Λ_1, Λ_2, Λ_3, and Λ_5 only slightly improves as the number of sensors increases. We also observe that the detector Λ_4 does not function in this setting, i.e., the outcome of the decision is independent of the observations, number of sensors M and observation intervals N. This is because the channels have a dominant path and the spectrum is almost flat. In addition a uniform power allocation at the OFDM transmitter is used, which results in a fairly flat power spectrum and therefore Λ_4 (which uses the homogeneity index) can not distinguish $\mathcal{H}_0$ from $\mathcal{H}_1$. We also observe that the performance of the energy detector Λ_5 is inferior to that of Λ_2 and Λ_3.

Figure 3.4 depicts the average probability of mis-detection versus the average probability of false alarm of $\{\Lambda_i\}_{i=2}^5$ for $K = 512, M = 4, N = 20, L_p = 8, f_d =$

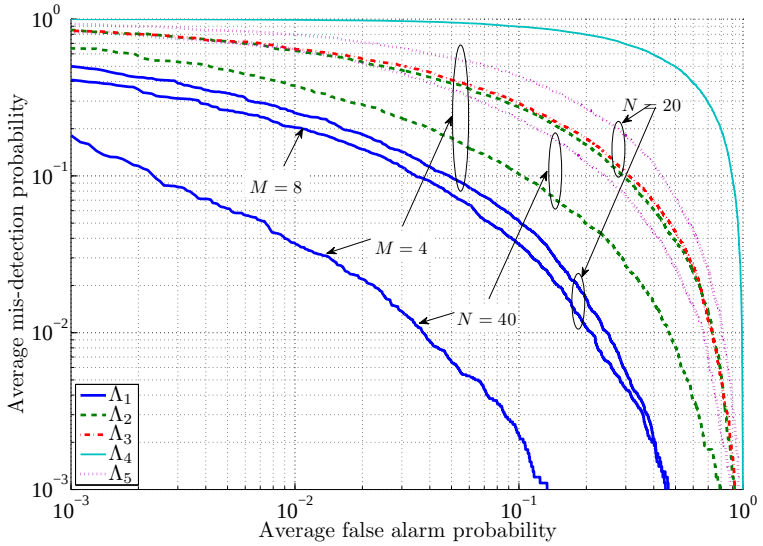


Figure 3.3: Average P_{md} versus average P_{fa} of Λ_1 to Λ_5 with $K = 512, N = 20, L_p = 8, f_d = .0285, P_k = 1$ for two scenarios: 1) $M = 4$ and $\sigma_m^2 = 50m$ and 2) $M = 8$ and $\sigma_m^2 = 50m$.

$.0285, \sigma_m^2 = 1$ and fixed total transmit power of $\sum P_k = 512$ for two cases where only 4 or 128 subcarriers are employed for transmission. The case where only 4 subcarriers are employed represents an extremely non-flat spectrum. Therefore, we expect that the performance of Λ_2, which assumes uniform power allocation, significantly degrades as the homogeneity index of the source spectrum increases. In addition, Λ_3 which estimates subband SNRs, outperforms other sub-optimal detectors. In contrast, the performance of Λ_4 is improved as the homogeneity in of the source spectrum increases. Similarly, for extremely non-flat spectrum, the performance of Λ_5 is improved as more subbands are occupied. All the aforementioned detectors provide considerably inferior performance compared to that of the optimal NPD that implies that the a-priori knowledge about the transmitter power allocation can enhance the performance of the sub-optimal detectors. Further simulation shows that Λ_5 presented in Example 3.2

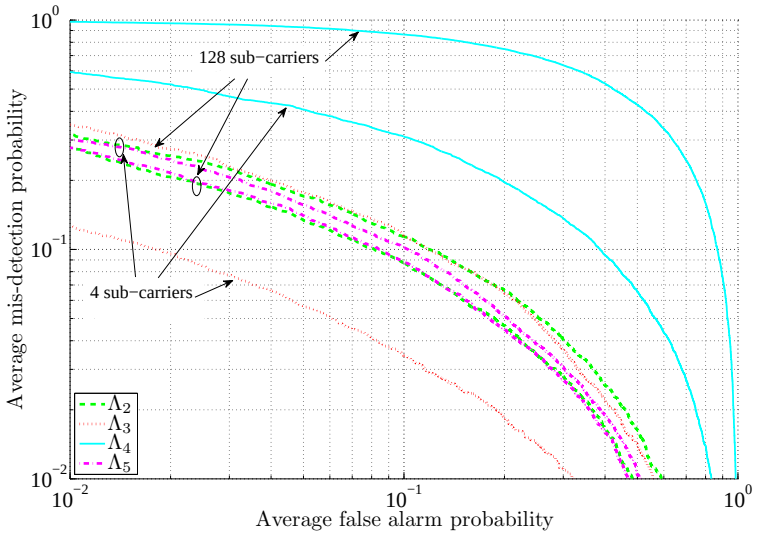


Figure 3.4: Impact of spectrum flatness on the performance of the detectors Λ_2 to Λ_5 with $K = 512, M = 4, N = 20, L_p = 8, = 100, f_d = .0285, \sigma_m^2 = 1$, and $\sum P_k = 512$ for two cases that only 4 or 128 subcarriers are employed for transmission.

with $L = 512 - 4$ or $L = 512 - 128$ provides a performance close to that of Λ_1. Note that Λ_5 in Example 3.2 requires the knowledge of the number of vacant subbands.

Figure 3.5 depicts the performance of the detectors Λ_1, Λ_2, and Λ_4 for $M = 4, N = 20, K = 512, L_p = 8, \sigma_m^2 = 50m, f_d = .0285, P_k = 1$ and using smaller (than K) DFT lengthes of $K' = 256$ (with $N' = 40$), and $K' = 32$ (with $N' = 320$). The performance of Λ_1 is also reported for the DFT length of 512 for the sake of comparison. Our simulation results show that the performance of the detectors Λ_3 and Λ_5 is not sensitive to the employed the DFT length from 32 to 512 (therefore, only the curve for 256 is plotted in this figure). In contrast, the performance using Λ_2 slightly degrades as the DFT length decreases. Using the smaller length K' allows to significantly reduce the computational cost of Λ_3 and Λ_5 to $NK \log_2(K')$ with almost no performance loss.

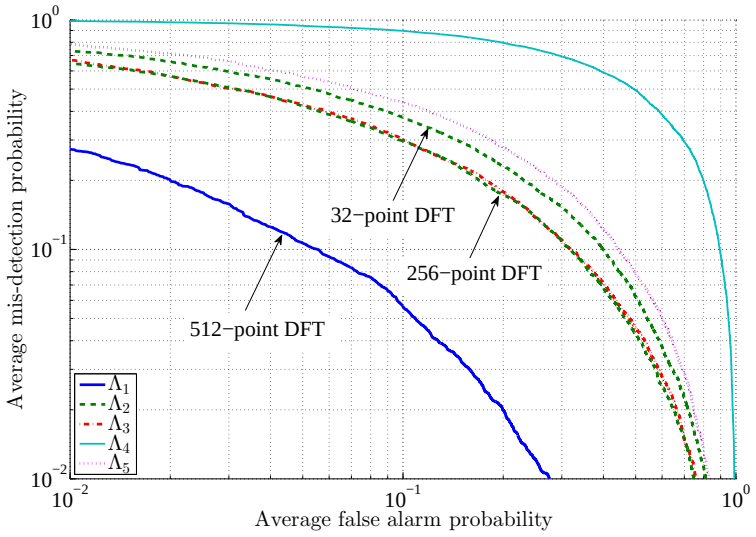


Figure 3.5: Impact of the DFT length on the performance of the detectors Λ_1 to Λ_5 with $K = 512, N = 20, L_p = 8, \sigma_m^2 = 100, f_d = .0285, P_k = 1$ for two number of sensors 1) $M = 4$, 2) $M = 8$.

Figures 3.6 and 3.7 show the achieved average mis-detection probability of the detectors Λ_1 to Λ_5 for a given false alarm of P_{fa}= 0.1 and $M = 4, N = 20, K = 512, L_p = 8, \sigma_m^2 = 50m, f_d = .0285, P_k = 1$ where there is no time and frequency synchronization, respectively. For imperfect time-synchronization sensors independently pick random times uniformly distributed over $[-\frac{\tau}{2}\frac{\tau}{2}]$ as the start of an OFDM frame. For imperfect frequency-synchronization sensors have independent random frequency offsets uniformly distributed over $[-\frac{\phi}{2}\frac{\phi}{2}]$. Interestingly, these detectors are robust against both time and frequency mismatches.

Figure 3.8 shows the achieved average mis-detection probability of the detectors Λ_1 to Λ_5 as a function of the transmitted signal variance $-20\text{dB} \leq P_k \leq 0\text{dB}$, for a given false alarm of $P_{\text{fa}} = 0.1$ and $M = 4, N = 20, K = 512, L_p = 8, \sigma_m^2 = 50m$, and $f_d = .0285$. We observe that the performance of the proposed detectors degrades for

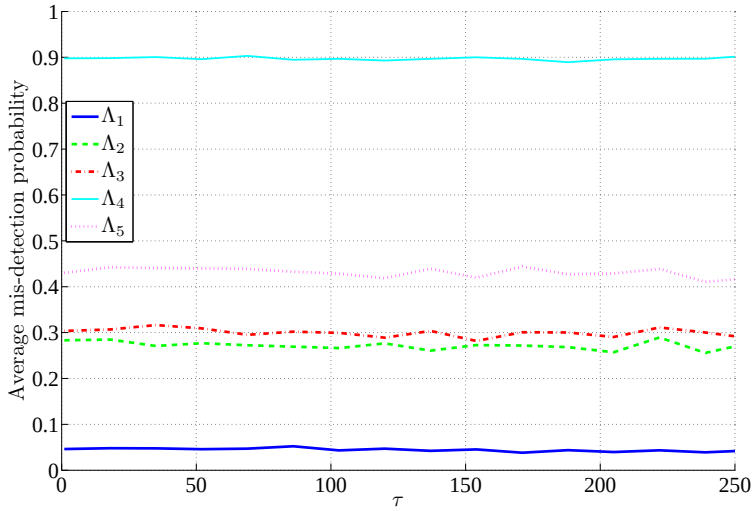


Figure 3.6: Impact of imperfect time synchronization error on the performance of the detectors Λ_1 to Λ_5 for a given false alarm of 0.1 with $M = 4, N = 20, K = 512, L_p = 8, \sigma_m^2 = 50m, f_d = .0285, P_k = 1$.

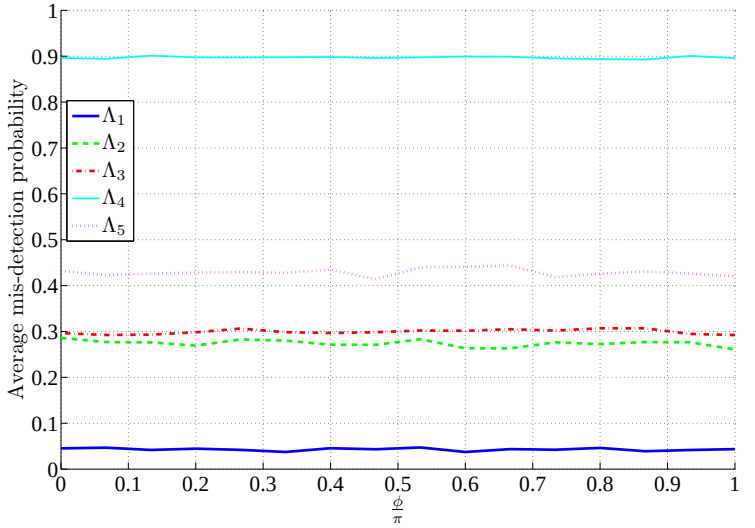


Figure 3.7: Impact of imperfect frequency synchronization error on the performance of the detectors Λ_1 to Λ_5 for a given false alarm of 0.1 with $M = 4, N = 20, K = 512, L_p = 8, \sigma_m^2 = 50m, f_d = .0285, P_k = 1$.

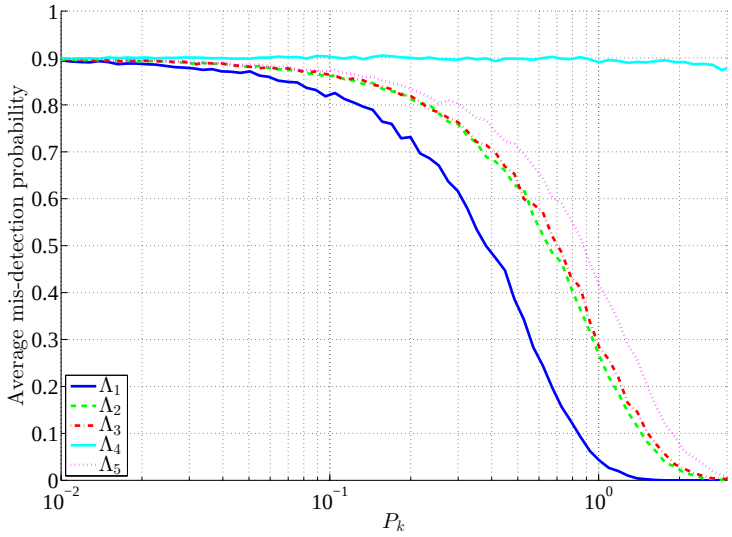


Figure 3.8: Impact of transmit power P_k on the performance of the detectors Λ_1 to Λ_5 for a given false alarm of $= 0.1$ with $M = 4, N = 20, K = 512, L_p = 8, \sigma_m^2 = 50m$, and $f_d = 0.0285$.

small values of P_k. Note that Λ_4 malfunctions as the subband received powers are equal and the spectrum is flat, i.e., the achieved mis-detection probability is 0.9 on average regardless of P_k. Other detectors are robust for small values of P_k and an average mis-detection probability of less than 0.9 is achievable for $P_k \geq -20$dB.

Figure 3.9 compares the performance of Λ_3 and Λ_4 with two time-domain detectors in [21, eq. (21)-(22)], which are designed for similar conditions and denoted by Λ_6 and Λ_7 respectively. We assume $M = 1, N = 10, K = 64, L_p = 8, \sigma_m^2 = 1, f_d = 0.0285$ for two scenarios: 1) fully-utilized spectrum with 64 employed subcarriers and 2) under-utilized spectrum with 32 employed subcarriers. The performance of Λ_4 (in contrast to Λ_3, Λ_6, and Λ_7) improves as the number of employed subcarrier decreases (See Remark 3.6). In addition, Λ_4 (despite lower computational complexity) outperforms Λ_7 when the spectrum is under-utilized. This makes Λ_4 attractive in for

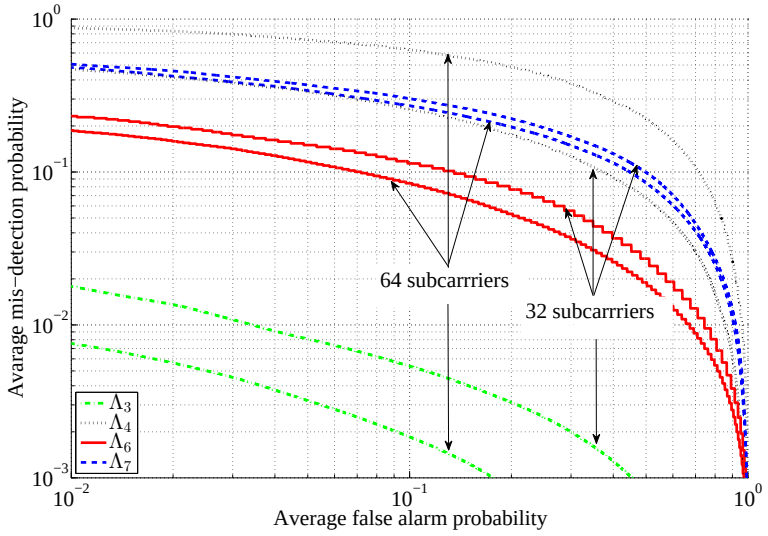


Figure 3.9: Comparison of Λ_3 and Λ_3 with Λ_6 and Λ_7 (proposed in [21, eq. (21)-(22)]) with $M = 1, N = 10, K = 64, L_p = 8, \sigma_m^2 = 1$, and $f_d = 0.0285$ when only 32 or 64 subcarriers are employed for transmission.

detecting under-utilized spectrum. In this case, Λ_4 may be used in conjunction with a second detector for fully-utilized spectrum (such as energy detector). The study of such multi-detector system is beyond the scope of this chapter. Additionally, Λ_3 (despite lower computational complexity) significantly outperforms Λ_6 since it takes into account the multichannel fading effect and non-uniform power allocation in frequency-domain. Note that Λ_6 and Λ_7 outperform previously known time-domain detectors, hence, this demonstrate that our proposed frequency-domain detectors outperform state-of-the-art time-domain detectors.

3.7 Conclusion

We have considered distributed detection of an OFDM random source using a cooperative set of sensors. Assuming a noncoherent system, we have shown that sensors

only need to transmit a real-valued function of their observations. We have analyzed the NPD (for known parameters) and a GLRD (for unknown SNRs and noise variance) and proposed two new GLRDs (for unknown SNRs or noise variances) and a new SC detector. For the SC detector, only a set of observation are selected by sensors and utilized by the FC for detection. We have derived the P_{md} and P_{fa} of these detectors. We have also proved that the GLRDs are asymptotically optimal. Our theoretical analysis and simulation results have shown that these frequency-domain detectors have negligible communication overhead and computational complexity and outperform the state-of-the-art time-domain detectors.

Chapter 4

Spectrum Sensing over Fading Channels

We propose novel detectors for cooperative spectrum sensing in mixture-Nakagami fading channels, namely 1) NPD, 2) a Locally Optimum Detector for weak signals that use the correlation between the observations and transmitted signals, 3) a weak signal detector for unknown parameters and 4) two GLRDs that use the received energies. They significantly outperform energy and cyclostationary detectors in practical scenarios. We also analyze the performance of the NPD and GLRD for unknown transmitted signal over Rayleigh channels, where they reduce to a linear weighted-correlator and a weighted-energy detector respectively.

4.1 Introduction

In cooperative spectrum sensing, a set of sensors aim to detect a signal source. The performance of such muti-sensor detectors deteriorates significantly in fading environments. Analogous to diversity techniques in communication systems, observations from multiple sensors with independent channels can be employed to combat fading. To this end, the sensors transmit a summary of their observations to a FC where a global decision is made.

Most of the existing cooperative spectrum sensing techniques do not make use of the channel fading parameters [47, 84, 71, 46, 90]. Assuming fixed channel gains between the source and sensors, the NPD becomes a linear weighted energy detector. We prove that this is not always true for random fading channels. Moreover in fading environments, the energy detector (ED) [47, 84] and Locally Optimum Detector (LOD) are suboptimal [71, 46]. Here, we aim to incorporate fading parameters into spectrum sensing. We consider mixture-Nakagami channels, which can accurately approximate most commonly used channel distributions (e.g., Rayleigh, Nakagami, and Rician, Nakagami-lognormal [19]), and derive novel detectors capable of exploiting the knowledge of the fading parameters. We show that the NPD and LOD are the sum of summaries reported by sensors, where each summary is a real-valued function of the correlation between the observations and the transmitted signal. Interestingly, the NPD for the Rayleigh Fading Channel (RFC) model simplifies to a linear combination of these correlations. The LOD provides a near-optimal performance and has a lower computational cost compared to the NPD. For unknown transmitted signals and known fading parameters, we derive a GLRD and propose a Weak Signal Detector (WSD) for which the sensors only need to report a real-valued function of

the energy of their observation. For RFCs, this GLRD simplifies to a weighted energy detector. The WSD has a lower computational cost and outperforms this GLRD even at high SNRs. We also derive a second GLRD that does not require the fading parameters and its decision statistic is computed numerically by solving some polynomial equations. This GLRD is attractive for Rayleigh-Lognormal environments [87, p.128] since without requiring the fading parameters performs close to the WSD.

We introduce the system model in Section 4.2. We derive the NPD, LOD, and GLRDs and assess their performance for some cases in Sections 4.3 and 4.4. We numerically evaluate these detectors in Section 4.5 and conclude in Section 4.6.

4.2 System Model

We use a network of M sensors to detect presence of a signal. The received signal by the ith sensor is denoted by

$$\begin{cases} y_i = n_i, & \mathcal{H}_0, \\ y_i = h_i s_i + n_i, & \mathcal{H}_1, \end{cases} \quad i = 1, \cdots, M, \tag{4.1}$$

where $h_i = r_i e^{j\phi_i} \in \mathbb{C}$ is the channel gain from the source to the ith sensor, $s_i \in \mathbb{C}^{1\times L}$ is the transmitted signal, L is number of samples and $n_i \in \mathbb{C}^{1\times L}$ is the received noise. The absence and presence of the source are denoted by $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively. We assume $\{n_i\}$ are independent and normal with zero mean and variance $\{\sigma_i^2\}$, i.e.,

$$f(\{n_i\}) = \prod_i f(n_i) = \prod_i \frac{1}{\pi^L \sigma_i^{2L}} \exp\left(-\frac{\|n_i\|^2}{\sigma_i^2}\right). \tag{4.2}$$

By virtue of maximum entropy principle, we assume that $\{\phi_i\}$ are independent and uniformly distributed over $[0, 2\pi)$ [22, 39, 45]. We use a mixture of Nakagami PDF, i.e.,

$$f(\{r_i\}) = \prod_i f(r_i) = \prod_i \sum_{j=1}^{N} \frac{2v_j m^m}{\Gamma(m)\Omega_{i,j}^m} r_i^{2m-1} \exp\left(-\frac{mr_i^2}{\Omega_{i,j}}\right), \tag{4.3}$$

where $\Omega_{i,j}$ is the average fading power, m is the fading parameter, $\Gamma(.)$ is the Gamma function, N is the number of terms in the mixture and $\sum_j v_j = 1$ with $v_j > 0$. Rayleigh, Nakagami, and Rician PDFs are accurately represented/approximated with $N = 1$. As N increases, the mixture $f(r_i)$ as a sum of Nakagami terms converges to any pdf. For Nakagami-lognormal distribution, the mixture parameters are obtained directly from the measurable Nakagami-lognormal parameters, i.e., as $\Omega_{i,j} = e^{\sqrt{2}\lambda_i t_j + \mu_i}$, where μ_i and λ_i are the mean and standard deviation of the lognormal pdf respectively, $v_j = \frac{w_j}{\sum_{k=1}^{N} w_k}$, t_j and w_j are given in [2, p. 890]. Using a sufficient number of terms this finite series accurately approximates Nakagami-lognormal distribution [19].

4.3 Optimal Detectors for Known Parameters

In some situations $\{s_i\}$ contains known pilots (intended for channel estimation and synchronization, e.g. in WiFi, WiMax, IEEE 802.22 beacon signals [38, 36]). We now derive the NPD for known $\{s_i\}$, $\{\Omega_{i,j}\}, m$, N, and $\{\sigma_i^2\}$. It optimally maximizes the probability of detection for a given probability of false alarm and favors $\mathcal{H}_1$ if $\frac{f(\{y_i\}|\mathcal{H}_1)}{f(\{y_i\}|\mathcal{H}_0)}$ is above some threshold.

Theorem 4.1. *The NPD for* (4.1) *is*

$$\Lambda_{NP} \triangleq \sum_{i=1}^{M} \log\Big(\sum_{j=1}^{N} v_j \frac{\exp(z_{i,j}^{NP}) P_{m-1}(-z_{i,j}^{NP})}{(1+\frac{\Omega_{i,j}\|s_i\|^2}{m\sigma_i^2})^m}\Big) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{NP}, \tag{4.4}$$

where $z_{i,j}^{NP} = \frac{|y_i s_i^H|^2}{\sigma_i^4 \frac{m}{\Omega_{i,j}} + \|s_i\|^2\sigma_i^2}$, η_{NP} *is the detection threshold and* $P_m(x)$ *is the Laguerre polynomial [2, p. 781].*

Proof. Under $\mathcal{H}_1$, the conditional pdf of the $\{y_i\}$ is Gaussian given the channel gains $\{h_i = r_i e^{j\phi_i}\}$. Integrating over the PDFs of r_i and ϕ_i, the pdf of y_i is given by

$$f(y_i|\mathcal{H}_1) = \int_0^\infty \int_0^{2\pi} e^{-\frac{\|y_i - r_i e^{j\phi_i} s_i\|^2}{\sigma_i^2}} \frac{f(r_i) d\phi_i dr_i}{2\pi^{L+1}\sigma_i^{2L}}. \tag{4.5}$$

Denoting $s_i y_i^H = |y_i s_i^H| e^{j\psi_i}$, we can rewrite the above as

$$f(y_i|\mathcal{H}_1) = \frac{e^{-\frac{\|y_i\|^2}{\sigma_i^2}}}{\pi^L \sigma_i^{2L}} \int_0^\infty \int_0^\pi e^{\frac{2|y_i s_i^H| r_i \cos(\theta_i) - r_i^2\|s_i\|^2}{\sigma_i^2}} f(r_i) \frac{d\theta_i}{\pi} dr_i, \tag{4.6}$$

where $\theta_i = \phi_i - \psi_i$. Using the equalities [52, p. 339, p. 707]

$$\begin{aligned} I_0(a) = J_0(a\sqrt{-1}) &= \int_0^\pi e^{a\cos(\theta)} \frac{d\theta}{\pi}, \\ \frac{\Gamma(m)}{2a^m} e^{\frac{b^2}{a}} P_{m-1}(-\tfrac{b^2}{a}) &= \int_0^\infty r^{2m-1} e^{-ar^2} I_0(2br) dr \end{aligned} \tag{4.7}$$

respectively, from (4.6), we obtain the pdf of $\{y_i\}$ under $\mathcal{H}_1$ as:

$$f(\{y_i\}|\mathcal{H}_1) = \prod_{i=1}^{M} \sum_{j=1}^{N} v_j \frac{\exp(z_{i,j}^{\mathrm{NP}} - \frac{\|y_i\|^2}{\sigma_i^2}) P_{m-1}(-z_{i,j}^{\mathrm{NP}})}{\pi^L \sigma_i^{2L} (1+\frac{\Omega_{i,j}\|s_i\|^2}{m\sigma_i^2})^m}. \tag{4.8}$$

From $f(\{y_i\}|\mathcal{H}_0) = \prod_{i=1}^{M} \frac{\exp -\frac{\|y_i\|^2}{\sigma_i^2}}{\pi^L \sigma_i^{2L}}$ and (4.8), we obtain (4.4). □

Λ_{NP} cannot be further simplified, thus it is the decision statistic for (4.1). Each sensor performs in the order of L operations to compute $\log(\sum_j v_j \frac{\exp(z_{i,j}^{\text{NP}}) P_{m-1}(-z_{i,j}^{\text{NP}})}{(1+\Omega_{i,j}\|s_i\|^2/m\sigma_i^2)^m})$. The FC compares the sum of these summaries/statistics with η_{NP} to make a decision.

We now consider weak signals as $\|s_i\|^2 \to 0$ and derive the LOD, which maximizes the slope of the average probability of mis-detection $\bar{P}_{\text{md}}$ for a fixed average probability of false alarm $\bar{P}_{\text{fa}}$, i.e., the LOD favors $\mathcal{H}_1$ when $\sum \frac{\partial f(\{y_i\}|\mathcal{H}_1)}{f(\{y_i\}|\mathcal{H}_0)\partial\|s_i\|^2}$ is above some threshold. From (4.4), the LOD is given by

$$\Lambda_{\text{LO}} \triangleq \sum_{i=1}^{M} \frac{u_{i,\text{LO}}}{\sigma_i^4} \frac{|y_i s_i^H|^2}{\|s_i\|^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{\text{LO}}, \tag{4.9}$$

where η_{LO} is the detection threshold, $u_{i,\text{LO}} = \sigma_i^2 \frac{\partial \text{SNR}_i}{\partial\|s_i\|^2} = \sum_j v_j \Omega_{i,j} > 0$ and $\text{SNR}_i = \frac{\|s_i\|^2}{\sigma_i^2}\sum_j v_j \Omega_{i,j}$. Each sensor performs in the order of L operations (yet fewer than that of Λ_{NP}). The FC linearly combines the sample correlations $|y_i s_i^H|^2$ with the weight $\frac{u_{i,\text{LO}}}{\sigma_i^4\|s_i\|^2}$, which are proportional to σ_i^{-4} and the channel quality $u_{i,\text{LO}}$. We now assess its performance in some cases.

4.3.1 Single-sensor case (SS)

Since $e^x P_{m-1}(-x)$ is monotonically increasing, (4.4) for $M = 1$ reduces to a correlator:

$$\Lambda_{\text{NP-SS}} \triangleq |y_i s_i^H|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{\text{NP-SS}}, \tag{4.10}$$

where $\eta_{\text{NP-SS}}$ is the detection threshold. A similar result for RFCs can be found in [62, p. 486]. Under $\mathcal{H}_0$, $\Lambda_{\text{NP-SS}}$ has an exponential pdf with mean $\|s_i\|^2\sigma_i^2$. Under $\mathcal{H}_1$ and

given h_i, it has a noncentral chi-squared pdf and its conditional MGF is

$$E[e^{t\Lambda_{\text{NP-SS}}}|\mathcal{H}_1, h_i] = e^{\frac{r_i^2\|s_i\|^4 t}{1-\|s_i\|^2\sigma_i^2 t}}(1-\|s_i\|^2\sigma_i^2 t)^{-1}. \tag{4.11}$$

Using the identity $\int_0^\infty x^{m-1}e^{-\frac{x}{a}}dx = a^m\Gamma(m)$, the MGF of $\Lambda_{\text{NP-SS}}$ is

$$E[e^{t\Lambda_{\text{NP-SS}}}|\mathcal{H}_1] = \sum_j v_j \frac{(1-\|s_i\|^2\sigma_i^2 t)^{m-1}}{(1-t_j t)^m}, \tag{4.12}$$

where $t_j = \|s_i\|^2(\sigma_i^2 + \frac{\Omega_{i,j}}{m}\|s_i\|^2)$. We use partial fraction expansion to get

$$E[e^{t\Lambda_{\text{NP-SS}}}|\mathcal{H}_1] = \sum_{j=1}^{N}\sum_{k=1}^{m} v_j\alpha_{k,j}(1-t_j t)^{-k}, \tag{4.13}$$

where $\alpha_{k,j}$ can be obtained by multiplying the both sides by $(1-t_j t)^m$ and evaluating the kth derivative of the result at $t = t_j^{-1}$, which amounts to

$$\alpha_{k,j} = \frac{\binom{m-1}{k-1}\|s_i\|^{2k-2}\Omega_{i,j}^{k-1}\sigma_i^{2m-2k}}{m^{k-1}(\sigma_i^2 + \frac{\Omega_{i,j}}{m}\|s_i\|^2)^{m-1}}. \tag{4.14}$$

Taking the inverse Laplace transform of $E[e^{t\Lambda_{\text{NP-SS}}}|\mathcal{H}_1]$, we obtain the $\bar{P}_{\text{fa}}$ and $\bar{P}_{\text{md}}$ of (4.10) as:

$$\begin{aligned} \bar{P}_{\text{NP-SS, fa}} &= \exp\left(-\eta_{\text{NP-SS}}\|s_i\|^{-2}\sigma_i^{-2}\right), \\ \bar{P}_{\text{NP-SS, md}} &= \textstyle\sum_{j=1}^{N}\sum_{k=1}^{m} v_j\alpha_{k,j}\gamma(k, \eta_{\text{NP-SS}}t_j^{-1}), \end{aligned} \tag{4.15}$$

where $\gamma(p,x) = 1 - e^{-x}\sum_{q=0}^{p-1}\frac{x^q}{q!}$ for any $p \in \mathbb{N}$.

4.3.2 Rayleigh fading channel (RFC)

For RFCs, i.e., $N = m = 1$, and $\Omega_{i,j} = \Omega_i$, we can simplify (4.4) to a linear weighted correlator as

$$\Lambda_{\text{NP-RFC}} \triangleq \sum_{i=1}^{M} \frac{u_{i,\text{RFC}}}{\sigma^2_{i,\text{RFC}}} \frac{|y_i s_i^H|^2}{\|s_i\|^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{\text{NP-RFC}}, \tag{4.16}$$

where $u_{i,\text{RFC}} = \frac{\text{SNR}_i}{1+\text{ SNR}_i}$, $\text{SNR}_i = \frac{\Omega_i \|s_i\|^2}{\sigma_i^2}$, and $\eta_{\text{NP-RFC}}$ is the detection threshold. $u_{i,\text{RFC}} \in [0, 1)$ is the ratio of the signal component of the received signal to the total observed energy and is an increasing function SNR_i.

We shall use the following theorem (See [8]) to obtain the CDF of $\Lambda_{\text{NP-RFC}}$.

Theorem 4.2. *If $\theta_l > 0$ and $u_l > 0$ are independent Gamma random variables, i.e., $f(u_l) = \frac{u_l^{r_l - 1} e^{-u_l}}{\Gamma(r_l)}$, then the CDF of $z = \sum \theta_l u_l$ is*

$$F_{\mathbf{z}}(z; [r_i]_{i=1}^L, [\theta_i]_{i=1}^L) = \prod_{l=1}^{L} \theta_l^{-r_l} \sum_{k=1}^{L} \sum_{l=1}^{r_k} \Psi_{k,l,\mathbf{r}} \theta_k^{r_l} \gamma(r_l, \frac{z}{\theta_k}), \tag{4.17a}$$

where $z > 0$ and $\Psi_{k,l,\mathbf{r}}$ is defined as

$$\Psi_{k,l,\mathbf{r}} = \sum_{[i_j]_{j=1}^l \in \mathbb{A}_{k,l}} \prod_{1 \le j \le l, j \ne k} \binom{i_j + r_j - 1}{i_j} \left(\frac{1}{\theta_j} - \frac{1}{\theta_k}\right)^{-(r_j + i_j)}, \tag{4.17b}$$

$$\mathbb{A}_{k,l} = \left\{ [i_j]_{j=1}^l \in \mathbb{Z}^l \,\middle|\, \sum_{j=1}^{L} i_j = l - 1, i_k = 0, i_j \ge 0 \right\}. \tag{4.17c}$$

To assess the performance of $\Lambda_{\text{NP-RFC}}$, we note that $|y_i s_i^H|^2$ has an exponential pdf under both hypotheses. Since $\Lambda_{\text{NP-RFC}}$ is a sum of Gamma random variables, we use

Theorem 4.2 with $r_l = 1$ and $\theta_i = \frac{\|s_i\|^2\Omega_i}{\sigma_i^2+\|s_i\|^2\Omega_i}$ under $\mathcal{H}_0$ and $\theta_i = \Omega_i$ under $\mathcal{H}_1$ to get

$$\begin{aligned} \bar{P}_{\text{NP-RFC, fa}} &= F_{\mathbf{z}}(\eta_{\text{NP-RFC}}; [1, \cdots, 1]_{1\times M}, [\tfrac{\|s_i\|^2\Omega_i}{\sigma_i^2+\|s_i\|^2\Omega_i}]_{i=1}^{M}), \\ \bar{P}_{\text{NP-RFC, md}} &= F_{\mathbf{z}}(\eta_{\text{NP-RFC}}; [1, \cdots, 1]_{1\times M}, [\Omega_i]_{i=1}^{M}). \end{aligned} \tag{4.18}$$

4.4 Detectors for Unknown Parameters

4.4.1 GLR1: Known Signal Energy and Fading Parameters

We previously assumed that $\{s_i\}$ are known. However in some cases, such knowledge is unavailable. We now derive a GLRD for known $\|s_i\|^2$ (instead of $\{s_i\}$) that utilizes the MLE of unknown parameters. This is motivated by sources with stationary power such as TV channels for which $\|s_i\|^2$ is almost constant if the observation duration is long enough. The MLE of s_i under $\mathcal{H}_1$ is $\hat{s}_i = \frac{\|s_i\|}{\|y_i\|} y_i^H$, substituting which in (4.4), we obtain the following:

$$\Lambda_{\text{GLR1}} \triangleq \sum_{i=1}^{M} \log \sum_{j=1}^{N} v_j \frac{e^{z_{i,j}^{\text{GLR1}}} P_{m-1}(-z_{i,j}^{\text{GLR1}})}{(1+\frac{\Omega_{i,j}\|s_i\|^2}{m\sigma_i^2})^m} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{\text{GLR1}}, \tag{4.19}$$

where $z_{i,j}^{\text{GLR1}} = \frac{\|y_i\|^2\|s_i\|^2}{\sigma_i^4 \frac{m}{\Omega_{i,j}} + \|s_i\|^2\sigma_i^2}$. Each sensor requires in the order of L operations to compute $\log \sum_j v_j \frac{\exp(z_{i,j}^{\text{GLR1}}) P_{m-1}(-z_{i,j}^{\text{GLR1}})}{(1+\Omega_{i,j}\|s_i\|^2/m\sigma_i^2)^m}$. The FC makes a global decision by comparing the sum of the reported values with η_{GLR1}. We now assess the performance of Λ_{GLR1} for some cases.

Single-sensor case (SS)

Since $e^x P_{m-1}(-x)$ is monotonically increasing, (4.19) reduces to the ED as

$$\Lambda_{\text{GLR1-SS}} \triangleq \|y_i\|^2 \gtrless_{\mathcal{H}_0}^{\mathcal{H}_1} \eta_{\text{GLR1-SS}}, \tag{4.20}$$

for $M = 1$, where $\eta_{\text{GLR1-SS}}$ is the detection threshold. A similar result for $m \to \infty$ and $N = 1$ can be found in [62, p. 250]. Under $\mathcal{H}_0$, $\Lambda_{\text{GLR1-SS}}$ has a Gamma pdf with parameters (L, σ_i^2). Similar to Section 4.3.1, we can show that

$$E[e^{t\Lambda_{\text{GLR1-SS}}}|\mathcal{H}_1] = \textstyle\sum_j v_j \frac{(1-\sigma_i^2 t)^{m-L}}{(1-(\sigma_i^2+\frac{\Omega_{i,j}}{m}\|s_i\|^2)t)^m}. \tag{4.21}$$

For $L > m$, the pdf corresponding to each term in $E[e^{t\Lambda_{\text{GLR1-SS}}}|\mathcal{H}_1]$, is a linear combination of two Gamma random variables. Thus using Theorem 4.2, $\bar{P}_{\text{fa}}$ and $\bar{P}_{\text{md}}$ of (4.20) are

$$\begin{aligned}
\bar{P}_{\text{GLR1-SS, fa}} &= 1 - \gamma(L, \sigma_i^{-2}\eta_{\text{GLR1-SS}}), \\
\bar{P}_{\text{GLR1-SS, md}} &= \textstyle\sum_j v_j F_{\mathbf{z}}(\eta_{\text{GLR1-SS}}; [L-m, m], [\sigma_i^2 + \frac{\Omega_{i,j}}{m}\|s_i\|^2, \sigma_i^2]).
\end{aligned}$$

Rayleigh Fading Channel (RFC)

For RFCs, we have $N = 1$, $m = 1$, and $\Omega_{i,j} = \Omega_i$. Thus (4.19) reduces to the following weighted energy detector, i.e.,

$$\Lambda_{\text{GLR1-RFC}} \triangleq \textstyle\sum_{i=1}^{M} \frac{u_{i,\text{RFC}}}{\sigma_i^2}\|y_i\|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{\text{GLR1-RFC}}. \tag{4.22}$$

Similar to Section 4.3.1, we can show that under $\mathcal{H}_1$ the MGF of $\Lambda_{\text{GLR1-SS}}$ is given by

$$E[e^{t\Lambda_{\text{GLR1-RFC}}}|\mathcal{H}_1] = \prod_{i=1}^{M}(1-u_{i,\text{RFC}}t)^{1-L}(1-\text{SNR}_i t)^{-1}. \tag{4.23}$$

Under $\mathcal{H}_0$, the MGF of $\Lambda_{\text{GLR1-RFC}}$ is similarly obtained as

$$E[e^{t\Lambda_{\text{GLR1-RFC}}}|\mathcal{H}_0] = \prod_{i=1}^{M}(1-u_{i,\text{RFC}}t)^{-L}.$$

Using these MGFs, we conclude that $\Lambda_{\text{GLR1-RFC}}$ is always a linear combination of Gamma random variables. Thus employing Theorem 4.2 yields

$$\begin{aligned}\bar{P}_{\text{GLR1-RFC, fa}} &= F_{\mathbf{z}}(\eta_{\text{GLR1-RFC}}; [L,\cdots,L]_{1\times M}, [u_{i,\text{RFC}}]_{i=1}^{M}),\\ \bar{P}_{\text{GLR1-RFC, md}} &= F_{\mathbf{z}}(\eta_{\text{GLR1-RFC}}; \boldsymbol{r}, \boldsymbol{\theta}),\end{aligned} \tag{4.24}$$

where $\boldsymbol{r} = [[L-1,\cdots,L-1]_{1\times M}, [1,\cdots,1]_{1\times M}]$, and $\boldsymbol{\theta} = [[u_{i,\text{RFC}}]_{i=1}^{M}, [\text{SNR}_i]_{i=1}^{M}] \in \mathbb{R}^{1\times 2M}$.

4.4.2 Weak Signal Detector (WSD)

To detect weak unknown transmitted signals, i.e., $\|s_i\|^2 \to 0$, we propose to substitute the MLE of s_i under $\mathcal{H}_1$ in (4.9), which results in the following weighted energy detector

$$\Lambda_{\text{WSD}} \triangleq \sum_{i=1}^{M} \frac{u_{i,\text{LO}}}{\sigma_i^4}\|y_i\|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{\text{WSD}}. \tag{4.25}$$

Note that each sensor performs in the order of L operations to compute its statistic (although fewer than that of Λ_{GLR1}). Although, this detector is heuristic, our

simulation results show that it outperforms Γ_{GLR1} in practical cases.

4.4.3 GLR2: Unknown Signal and Fading Parameters

Thus far, we assumed $\{\Omega_{i,j}\}$ and N are known, which may be impracticable in some cases. Here, we assume that $\{s_i\}$, $\|s_i\|^2$, $\{\Omega_{i,j}\}$, $\{v_i\}$ and N are unknown and derive a GLRD for such a case. From (4.19), we observe that the MLE of $\|s_i\|^2\Omega_{i,j}$ under $\mathcal{H}_1$ does not depend on j and is given by

$$\hat{\Omega}_i = \underset{\Omega_{i,j}\|s_i\|^2}{\operatorname{argmax}} \frac{e^{z_{i,j}^{\text{GLR1}}} P_{m-1}(-z_{i,j}^{\text{GLR1}})}{(1+\Omega_{i,j}\|s_i\|^2/m\sigma_i^2)^m}. \tag{4.26}$$

Additionally, N vanishes after substituting $\hat{\Omega}_i$ in (4.19). As $\hat{\Omega}_i$ is non-negative, it must be either zero or the largest positive solution to $\frac{\partial}{\partial\Omega_i} f(\{y_i\}|\mathcal{H}_1) = 0$, which yields $\hat{\Omega}_i = \max(\frac{m\sigma_i^2}{\|y_i\|^2+x\sigma_i^2}, 0)$ and x is the smallest positive root of the polynomial equation

$$\frac{\|y_i\|^2+x\sigma_i^2}{m\sigma_i^2}(P_{m-1}(x) - P'_{m-1}(x)) - P_{m-1}(x) = 0. \tag{4.27}$$

Substituting $\hat{\Omega}_i$ in (4.4), we obtain the following GLRD:

$$\Lambda_{\text{GLR2}} \triangleq \sum_{i=1}^{M} \frac{\hat{\Omega}_i\|y_i\|^2}{\sigma_i^2(\hat{\Omega}_i+m\sigma_i^2)} + \log \frac{P_{m-1}(-\frac{\hat{\Omega}_i\|y_i\|^2}{\sigma_i^2(\hat{\Omega}_i+m\sigma_i^2)})}{(1+\hat{\Omega}_i/m\sigma_i^2)^m} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{\text{GLR2}}. \tag{4.28}$$

Theorem 4.3. *The detectors* (4.19) *and* (4.28) *are asymptotically optimal, in the sense of Theorem 2.3, as* $L \to \infty$.

Proof. The functions (4.28) and (4.19) are two separating functions [49] respectively in terms of the unknown parameters $\{s_i\}$ and $\{\Omega_i\}$. Thus, they are separating function

estimation tests and hence are asymptotically optimal [49, Th. 8]. □

For RFCs, we have $m = 1$ and $\hat{\Omega}_i = \max(\|y_i\|^2 - \sigma_i^2, 0)$. Thus we obtain the following GLRD:

$$\Lambda_{\text{GLR2-RFC}} \triangleq \sum_{\forall i:\|y_i\|^2>\sigma_i^2} \frac{\|y_i\|^2}{\sigma_i^2} - 1 - \log \frac{\|y_i\|^2}{\sigma_i^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_{\text{GLR2-RFC}}. \tag{4.29}$$

Each sensor computes its test statistic with in the order of L operations. The test censors the uninformative observations with $\|y_i\|^2 < \sigma_i^2$ and only considers those that support $\mathcal{H}_1$. It reduces the communication overhead as the sensors report to the FC only if the energy level is above the noise energy.

4.5 Simulation Results

We now use Monte-Carlo simulations averaged out over 10^6 runs to evaluate the proposed detectors. We assume $\sigma_i^2 = i$, $\{s_i\}$ chosen from QPSK symbols with a period of 10 and $\|s_i\|^2 = L$. To capture the large-scale fading, different path-losses are assumed with $E[|h_i|^2] = .99$ for $1 \leq i \leq \frac{M}{2}$ and $E[|h_i|^2] = .01$ for $\frac{M}{2} < i \leq M$.

Figure 4.1 shows $\bar{P}_{\text{md}}$ versus $\bar{P}_{\text{fa}}$ for the proposed detectors and ED for $M=2$, L=10, N=1 and $m \in \{1, 3\}$. Λ_{NP} outperforms the rest and only slightly outperforms Λ_{LO}. Λ_{GLR1} outperforms Λ_{GLR2} as it uses the transmit power knowledge. Λ_{WSD} closely follows Λ_{GLR1} and the ED is significantly inferior to the rest as it does not use such information. Since the channel becomes less faded as m increases, the performance generally improves. Λ_{WSD} is attractive for unknown transmitted signal and known fading parameters since it outperforms Λ_{GLR1} and is computationally less expensive.

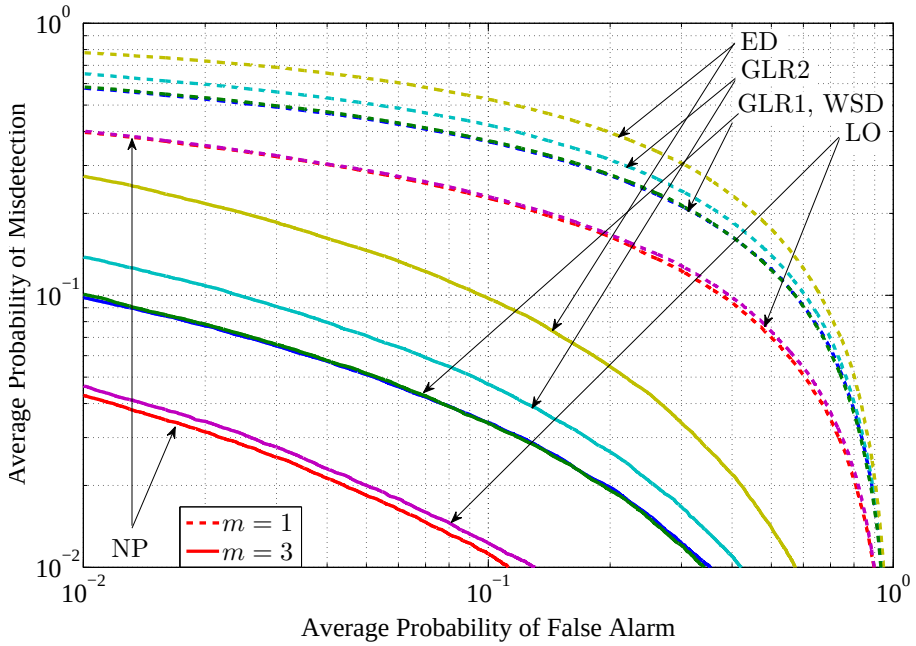


Figure 4.1: Impact of fading parameter $m \in \{1, 3\}$ on the performance of the detectors with $N = 1$, $M = 2$, and $L = 10$.

Figure 4.2 shows the same probabilities for $M \in \{2, 4\}$ and $L \in \{20, 40\}$ in RFCs. We observe a performance gain as the number of sensors or samples increases, i.e., the performance loss of the GLRDs (compared to the NPD) can be compensated by increasing M or L, e.g. in this setting, Λ_{GLR1} with $M = 2, L = 100$ outperforms Λ_{NP} with $M = 2, L = 20$ and Λ_{GLR2} with $M = 2, L = 30$ outperforms Λ_{GLR1} with $M = 2$, $L = 20$.

Figure 4.3 shows the same probabilities under Rayleigh-lognormal model with $m = 1, N = 10$ for $M = 2$ and $L \in \{20, 70\}$. The curves for Λ_{GLR1} and Λ_{GLR2} are very close. Thus Λ_{GLR2} is attractive in environments which involve both small and large scale fading, since it only needs the noise variance. Figures 4.2 and 4.3 depict the performance of the Cyclostationary Detector (CD). The CD performs close to the GLRDs and WSD only for large values of L. It comes at a high computational

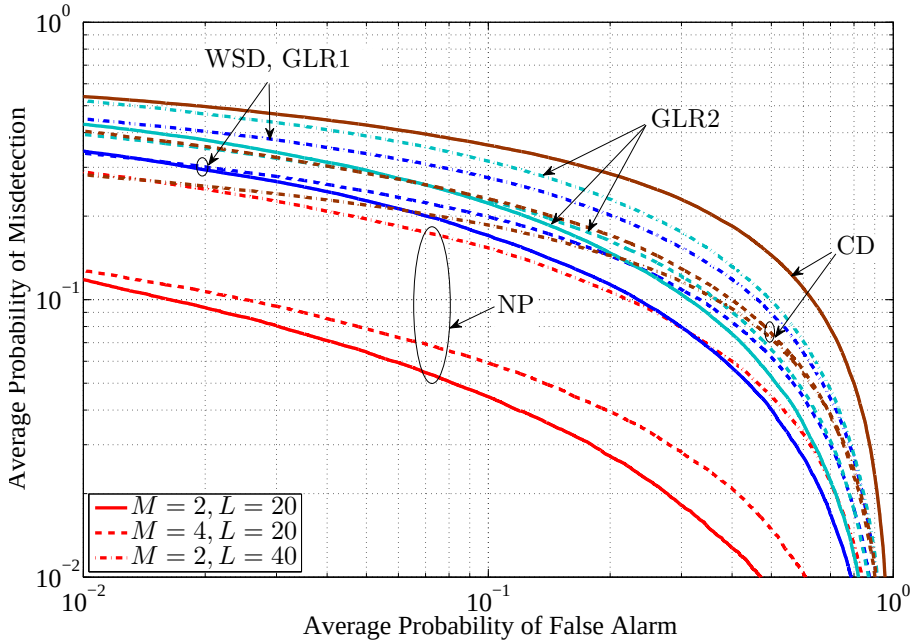


Figure 4.2: Impact of the number sensors $M \in \{2, 4\}$ and samples $L \in \{20, 40\}$ on the performance of the detector in Rayleigh channels.

cost as it needs about L^3 operations, whereas the proposed detectors require about L operations.

4.6 Conclusion

We derived several detectors for cooperative spectrum sensing in a mixture-Nakagami environment that make use of channel statistic for performance improvement. These detectors significantly outperform the ED. The LOD and WSD are attractive since they are computationally less expensive than the others, while the LOD provides a near optimal performance and the WSD outperforms the GLRDs. Under Rayleigh-lognormal model, the GLRD for unknown fading parameters is attractive since its performance is very close to the WSD, which uses the fading parameters.

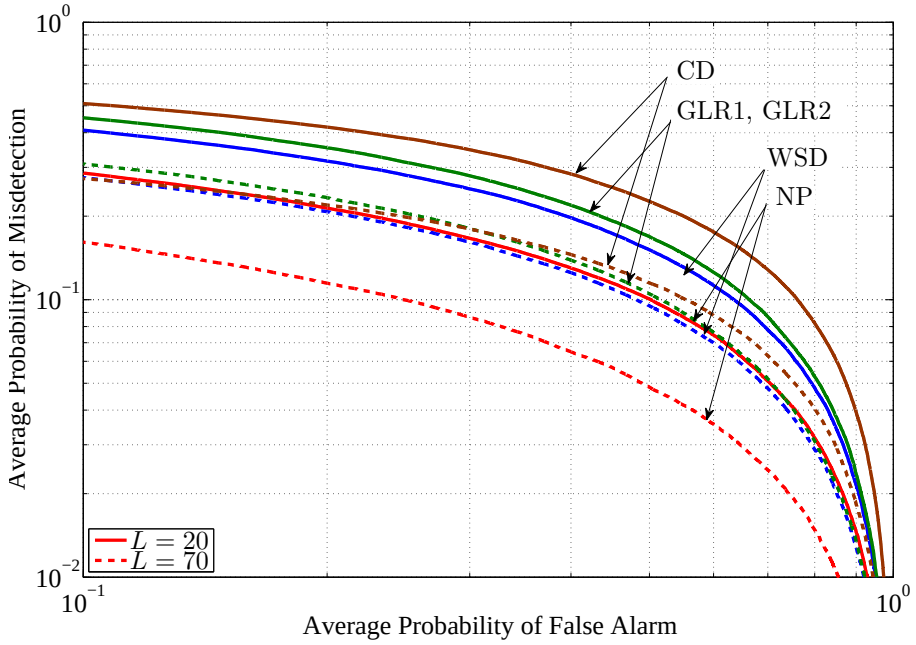


Figure 4.3: Performance of the proposed detector in Rayleigh-lognormal channels with $L \in \{20, 70\}$ and $M = 2$, $N = 10$, $m = 1$.

Chapter 5

Cooperative Spectrum Management

We investigate the joint resource allocation and relay selection problem in a cognitive radio network where (licensed and/or unlicensed) users are willing to cooperate with each other; decode and forward the messages of the other along with their own messages to the destination. We study the performance obtained from cooperation in terms of 1) increasing the achievable rate, 2) saving the transmit energy, and 3) reducing the resource (time-bandwidth) requirement. To ensure fairness, we assume that the ratio of the transmit energy to the rate is fixed for all the users. We optimally allocate the radio resources for the proposed Cooperative Protocol (CP) and compare the results with the Non-Cooperative Protocol (NCP) where users transmits their messages directly to the destination. Such comparison allows us 1) to decide whether to cooperate or not and 2) to select one relay among the possible relay users. We show that a considerable gain (in terms of increasing rate, saving energy, or reducing time-bandwidth) can be obtained if the direct source-destination channel gain is

significantly smaller than those of alternative links. We demonstrate that a rate and energy improvement of up to $\left(1+\sqrt[\eta]{\frac{k}{k+1}}\right)^{\eta}$ can be obtained, where η is the environment path loss exponent and k is the ratio of the required rates (i.e., rate demands) of the involved users. We also show that the cooperation is only beneficial for the middle range rate ratio.

5.1 Introduction

In wireless networks, the main interrelated quantities are achievable rate, transmit or receive energy and efficiency of resource. Many recent results, e.g.,[86, 76, 64], show that cooperation among users in wireless networks, depending on channel condition and available energy, may increase the rate or reduce the energy or resource (time-bandwidth) requirement. Here, we ask the questions:

- Depending on channel condition and available energy, when is cooperation beneficial?
- What are the involved gain or loss from the possible cooperation?
- How can we select a relay among the possible candidates?

In order to answer these questions, we consider a network of licensed and/or unlicensed users intending to send independent information to their corresponding destinations (see Figure 5.1 for different scenarios). Since the role of the licensed and unlicensed users is assumed interchangeable, we do not distinguish between them and throughout this chapter simply refer to the both as users. We propose that the users assist each other only if, in a fair way, the cooperation offers benefit in terms

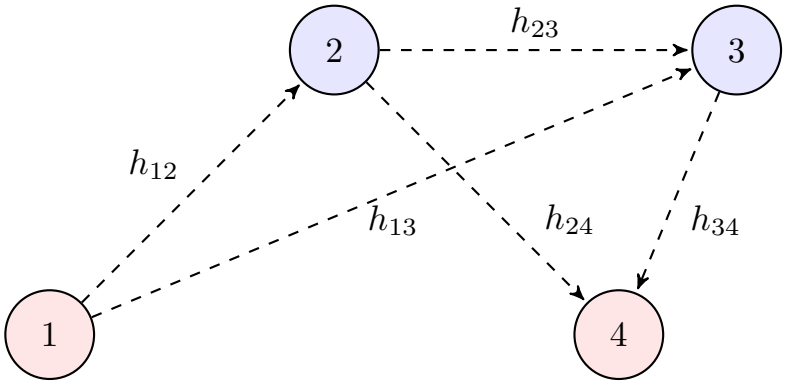


Figure 5.1: A cooperative network, the channel energy gain between i^{th} and j^{th} user is denoted by h_{ij}. Consider three scenarios: 1) the 1^{st} and 2^{nd} users use time-frequency division and transmit to the 3^{rd} user, 2) the 1^{st} user first transmits to the 3^{rd} user and then the 2^{nd} user broadcasts to the 3^{rd} and 4^{th} users, 3) the 1^{st} user transmits to the 3^{rd} user and then the 2^{nd} and 3^{rd} users use time-frequency division to transmit to the 4^{th} user.

of rate, energy, or resource. Here, the notion of fairness means that the achievable rates of different users would be proportional to their energy levels. We evaluate the effect of such cooperation on the performance of the system (in terms of gains in rate, energy, or time-bandwidth) and then, based on the achieved gains or losses, present our cooperative protocols.

Most of the existing CPs, e.g.,[86, 76, 64], implicitly assume that a relay is already chosen, although, selection schemes have been investigated recently and several interesting methods have been proposed to choose the best relay among the potential relay users using different optimization criteria. For example the error rate in [102, 59], energy consumption and network lifetime in [34, 103, 57], diversity gain and outage probability in [28, 23, 75], the pricing technique in [72], and convex optimization in [66] are used for user selection. In all these references [86, 76, 64, 102, 59, 34, 103, 57, 28, 23, 75, 72, 66], it is assumed that the relay node

provides free service to the source which is obviously beneficial to the source. Following [12, 11, 18], we study the problem of relay and protocol selection using three criteria; rate, energy and resource. In addition, to capture fairness among users, we assume that users will assist each other in relaying only if they gain from such a cooperation. Thus, those users having no data to transmit will not engage in such a cooperation. In contrast to [12, 11, 18], as it is important to take into consideration the different rate demands of various users, we introduce a new priority parameter. This parameter is imposed by an upper layer in order to determine the ratio of rate demands of involved users in the network. The motivation of this approach is to provide differentiated/prioritized services (see [40]). In this chapter assuming that a rate ratio is provided by the upper layer, we either maximize the achievable rate, minimize the energy consumption, or minimize the resource utilization.

The remainder of the chapter is organized as follows. Next we present the system model and present the protocols in Section 5.2. In Section 5.3 we study single relay networks and investigate the rate, energy and resource improvement from the possible cooperation. We then provide conditions on the location of the relay user for cooperation to be beneficial. In Section 5.4, we present our relay selection protocols. Extensions to the general network with multiple source and relay topology are discussed in Section 5.5. Finally, in Section 5.6 we give our concluding remarks.

5.2 System Model

Consider the first scenario in Figure 5.1, where we assume that the 1^{st} and 2^{nd} users wish to transmit independent messages respectively with rates R_1 and R_2 to the 3^{rd} user over an additive white Gaussian noise channel and the 2^{nd} user may also assist

the 1st user to transmit its messages to the 3rd user. The scenarios with multiple relays or with different destinations will be addressed later in this chapter. Let h_{ij} denote the channel gain of the communication link between the i^{th} and j^{th} user. We assume that the gain of all the channel links are perfectly known to all the users. Thus, users need to acquire their channel gains via efficient channel estimation algorithms (see e.g. [67]) and make it available to all other users. To this end, we assume that the involved users could initially exchange messages to establish a cooperation protocol before the main data streams are transmitted. Obviously, such exchanges consume a fraction of the available energy and resource. In this chapter, we ignore the extra cost incurred by this communication overhead.

When the users cooperate, the network is a multi-hopping network where one user receives the messages of another user and forwards the decoded messages to the intended receiver as well as its own messages. Otherwise, they form a multiple access channel, i.e., they transmit directly to the receiver via a resource sharing method.

We assume that users transmit via a resource division protocol where the i^{th} user can transmit over a portion β_i of available resource (by resource we mean the product of used time and bandwidth, i.e., $B \times T$.) Using time/frequency division requires perfect time/frequency synchronization. In this chapter, we ignore the required overhead to achieve this synchronization and assume perfect synchronization. The received energy to noise ratio within the resource slot $\beta_i BT$ can be expressed as $\frac{h_{ij}E_i}{N\beta_i BT}$, where E_i denotes the transmit energy of the i^{th} user and N denotes the received noise power. Unless otherwise stated, we consider the case where the available resource BT is unity, i.e., $BT = 1$. Let us define the ratio of Transmit Energy to Received Noise

Power (TERN) as $\epsilon_i = \frac{E_i}{N}$. Thus, the achievable rate for the user i is given by [37]

$$R_i = \beta_i \log\left(1 + \frac{h_{ij}\epsilon_i}{\beta_i}\right). \tag{5.1}$$

To pose the fairness issue in a multi-user communication network, we first need to define a fairness constraint. Most data applications are sensitive to error but tolerant to delay. It is clear that a higher Signal to Interference plus Noise Ratio (SINR) level at the output of the receiver will generally result in a lower bit-error rate, and hence higher throughput. However, achieving a high SINR level requires the user terminals to transmit at a high power, which in turn results in low battery life. Similar to [70, 51], we impose the constraint $\frac{\epsilon_i}{R_i}$ = Const. in order to maintain the fairness in the network for all users. This constraint ensures fairness among users as the energy spent by users is proportional to their demand for rate. This constraint captures the tradeoff between the throughput and energy consumption and is particularly suitable for applications in which energy efficiency is critical [70]. Note that this constraint penalizes the nodes with higher channel gain. This approach can be deemed as fair since, although users have no control over their channel gains, they can control their rate demand versus the required transmit energy. The fairness constraint also can be recast as

$$\frac{R_2}{R_1} = \frac{\epsilon_2}{\epsilon_1} \stackrel{\text{def}}{=} k. \tag{5.2}$$

which hereafter we denote k as the rate ratio and as a design parameter imposed by upper layers of the network such as the MAC layer.

We consider a half-duplex communication network where each user can either

transmit or receive (but not both) at any time and any frequency band. Throughout this chapter, we consider two following communication protocols:

- Non-Cooperative protocol where users transmit directly to the destination via a resource (time and frequency) division method.
- Cooperative protocol where over the 1st resource slot, the 1^{st} user transmits its message and the 2^{nd} user decodes the message of the 1^{st} user. Then, over the 2^{nd} resource slot, the 2^{nd} user re-encodes the decoded message of the 1^{st} user in conjunction with its own message, the 2^{nd} message, and broadcasts the encoded message.

5.3 Cooperation in Single Relay Networks

In the following we study some properties of proposed protocols and investigate upper and lower bounds for achievable rates. We use the Shannon's capacity as the performance metric. In order to justify this assumption, we assume that the users employ a capacity-achieving channel coding.

5.3.1 Non-Cooperative Protocol (NCP)

In this protocol, during the 1^{st} portion of resource slot, i.e., β_1, the 1^{st} user transmits its message. The receiver, the 3^{rd} user, decodes this message correctly for a maximum rate of $R_1(\beta_1) = \beta_1 \log\left(1 + \frac{h_{13}\epsilon_1}{\beta_1}\right)$. In a similar manner, the maximum rate of the 2^{nd} user which could be decoded reliably at the 3^{rd} user is $R_2(\beta_2) = \beta_2 \log\left(1 + \frac{h_{23}\epsilon_2}{\beta_2}\right)$. Since, we assume that one unit of resource is available, i.e., $\beta_1 + \beta_2 = 1$, hereafter, we denote $\epsilon_1 \stackrel{\text{def}}{=} \epsilon$, $\epsilon_2 = k\epsilon$, $\beta_1 \stackrel{\text{def}}{=} \beta$ and $\beta_2 = 1 - \beta$. Hence, we get the following

optimization problem for the NCP:

$$\begin{aligned} R_{\text{NCP}} &= \max_{\beta} \left(R_1(\beta) + R_2(1-\beta) \right) \\ \text{s.t.} \quad & \tfrac{R_2}{R_1} = k \end{aligned} \tag{5.3}$$

where R_{NCP} is the achievable sum rate of the users $R_1(\beta) = \beta \log\left(1 + \frac{h_{13}\epsilon}{\beta}\right)$ and $R_2(1-\beta) = (1-\beta)\log\left(1 + \frac{h_{23}k\epsilon}{1-\beta}\right)$. Since $R_1(\beta)$ and $R_2(1-\beta)$ are increasing and decreasing function of β, respectively, the solution of the above optimization is the unique solution of the following

$$R_{\text{NCP}} = (k+1)\beta \log\left(1 + \frac{h_{13}\epsilon}{\beta}\right) = \frac{k+1}{k}(1-\beta)\log\left(1 + \frac{h_{23}k\epsilon}{1-\beta}\right). \tag{5.4}$$

The optimal resource β_i as a solution of (5.4) is a function of h_{ij} and ϵ. However for ease of notation and abbreviation, we denote the optimal resource only by β_i.

5.3.2 Cooperative protocol (CP)

In this protocol, over the 1st portion of the resource slot, i.e., β, the 1st user transmits its messages at rate R_1. During this time, The 3rd user is switched off and thus ignores the received signal from the 1st user. The 2nd user attempts to decode the messages of the 1st user. Hence, the maximum achievable rate for the 1st user is expressed as $R_1(\beta) = \beta \log\left(1 + \frac{h_{12}\epsilon}{\beta}\right)$. Over the remaining portion of resource slot, i.e., $1-\beta$, the 2nd user re-encodes the decoded messages of the 1st user and transmits the messages of the 1st user as well as its own messages to the intended destination. In fact, during this time, the 2nd user must transmit at rate of $\frac{k+1}{k}R_2$ to accommodate both data. The maximum achievable rate which may be decoded reliably at the 3rd

user is $R_2(1-\beta) = \frac{k(1-\beta)}{k+1} \log\left(1 + \frac{h_{23}k\epsilon}{1-\beta}\right)$. This yields the following max-min resource allocation problem:

$$\begin{aligned} R_{\text{CP}} &= \max_{\beta} \left(R_1(\beta) + R_2(1-\beta)\right) \\ \text{s.t.} \quad & \frac{R_2}{R_1} = k \end{aligned} \tag{5.5}$$

where R_{CP} is the achievable sum rate of the users which will be compared with R_{NCP}. In a similar way, the optimal solution is the unique solution of the following equation with respect to β:

$$R_{\text{CP}} = (k+1)\beta \log\left(1 + \frac{h_{12}\epsilon}{\beta}\right) = (1-\beta)\log\left(1 + \frac{h_{23}k\epsilon}{1-\beta}\right). \tag{5.6}$$

Similar to the NCP, for ease of notation, we denote the optimal resource by β_i.

5.3.3 Rate Improvement for Given Resource and Energy

Here, we define the cooperation gain as the ratio of achievable sum rate of the CP to that of the NCP, i.e., $\frac{R_{\text{CP}}}{R_{\text{NCP}}}$. This ratio represents the achievable sum rate improvement of these protocols. We derive tight upper and lower bounds and study the asymptotic behavior of the cooperation gain at a low and high TERN and rate ratio.

Since $R_1(\beta)$ and $R_2(1-\beta)$ are increasing and decreasing convex and continuous functions of β, respectively, the maximization (5.4) is guaranteed to have a unique solution. Unfortunately, this solution has no closed form expression. Nevertheless, the following bounds can be obtained.

Theorem 5.1. *For the NCP, we have*

$$R_{\text{NCP}} > \frac{\frac{1}{k}\log\left(1+kh_{23}\epsilon\right)\log\left(1+h_{13}\epsilon\right)}{\frac{1}{k}\log\left(1+kh_{23}\epsilon\right)+\log\left(1+h_{13}\epsilon\right)}, \tag{5.7a}$$

$$R_{\text{NCP}} < \frac{\frac{\log(1+h_{13}(k+1)\epsilon)}{\left(1-\frac{1}{1+h_{13}(k+1)\epsilon}-\log(1+h_{13}(k+1)\epsilon)\right)}+\frac{k\log(1+h_{23}(k+1)\epsilon)}{\left(1-\frac{1}{1+h_{23}(k+1)\epsilon}-\log(1+h_{23}(k+1)\epsilon)\right)}}{\frac{(k+1)}{\left(1-\frac{1}{1+h_{13}(k+1)\epsilon}-\log(1+h_{13}(k+1)\epsilon)\right)}+\frac{k(k+1)}{\left(1-\frac{1}{1+h_{23}(k+1)\epsilon}-\log(1+h_{23}(k+1)\epsilon)\right)}} \tag{5.7b}$$

Proof. We use the first-order Taylor series approximation at point $\frac{1}{k+1}$ for $R_1(\beta)$ and $R_2(1-\beta)$ which is accurate for a high TERN regime. The intersection point of the approximate lines gives an upper bound for achievable capacity for the NCP. The coordinates of this intersection point are given by

$$\beta = \frac{1}{k+1}+\frac{\frac{1}{k+1}\log\left(\frac{1+(k+1)h_{23}\epsilon}{1+(k+1)h_{13}\epsilon}\right)}{\log\left(\left(1+\left(k+1\right)h_{23}\epsilon\right)\left(1+\left(k+1\right)h_{13}\epsilon\right)\right)-\frac{(k+1)h_{13}\epsilon}{1+(k+1)h_{13}\epsilon}-\frac{(k+1)h_{23}\epsilon}{1+(k+1)h_{23}\epsilon}} \tag{5.8}$$

and (5.7b), which as noted before is tight for a high TERN. The lower bound in (5.7a) is obtained from the intersection point of the two lines connecting end points of the rate curves. □

These bound are tight for a high TERN $\epsilon \to \infty$; this is the case where the noise power is negligible compared with the received signal powers. For a high TERN, the available resource is allocated to the users' receive in proportion with their rate demands, i.e., $\lim_{\epsilon\to\infty}\beta = \frac{1}{k+1}$.

Using the same approach, we can find the bounds in (5.9a) and (5.9b) for the achievable sum-rate of the CP

Theorem 5.2. *For the CP, we have*

$$R_{\text{CP}} > \frac{\frac{1}{k+1}\log\left(1+kh_{23}\epsilon\right)\log\left(1+h_{12}\epsilon\right)}{\frac{1}{k+1}\log\left(1+kh_{23}\epsilon\right)+\log\left(1+h_{12}\epsilon\right)}, \tag{5.9a}$$

$$R_{\text{CP}} < \frac{\frac{\log(1+h_{12}(k+1)\epsilon)}{\left(1-\frac{1}{1+h_{12}(k+1)\epsilon}-\log(1+h_{12}(k+1)\epsilon)\right)}+\frac{(k+1)\log\left(1+h_{23}\frac{k(k+2)}{k+1}\epsilon\right)}{\left(1-\frac{1}{1+h_{23}(k+1)\epsilon}-\log\left(1+h_{23}\frac{k(k+2)}{k+1}\epsilon\right)\right)}}{\frac{k+2}{\left(1-\frac{1}{1+h_{12}(k+1)\epsilon}-\log(1+h_{12}(k+1)\epsilon)\right)}+\frac{(k+1)(k+2)}{\left(1-\frac{1}{1+h_{23}\frac{k(k+2)}{k+1}\epsilon}-\log\left(1+h_{23}\frac{k(k+2)}{k+1}\epsilon\right)\right)}}, \tag{5.9b}$$

which are tight in the high TERN regime.

From (5.7b) and (5.9a), it is easy to see that $\lim_{\epsilon\to\infty}\frac{R_{\text{CP}}}{R_{\text{NCP}}} \geq \frac{k+1}{k+2}$. In addition, from (5.7a), we can see that $\lim_{\epsilon\to\infty}\frac{R_{\text{CP}}}{R_{\text{NCP}}} \leq \frac{k+1}{k+2}$. Thus $\lim_{\epsilon\to\infty}\frac{R_{\text{CP}}}{R_{\text{NCP}}} = \frac{k+1}{k+2}$. We conclude that the cooperative protocols are not attractive in the high TERN regime since the sum rate gain $\frac{k+1}{k+2}$ is smaller than one.

We also can derive tight bounds for the low TERN regime (small values of ϵ).

Theorem 5.3. *For the NCP, we have*

$$R_{\text{NCP}} > \epsilon\frac{2h_{23}+2h_{13}-h_{13}^2\epsilon-kh_{23}^2\epsilon-\sqrt{A}}{4}, \tag{5.10a}$$

$$R_{\text{NCP}} < \min\left\{\log\left(1+h_{13}\epsilon\right), \frac{1}{k}\log\left(1+kh_{23}\epsilon\right)\right\} \leq \epsilon\min\left\{h_{13}, h_{23}\right\}, \tag{5.10b}$$

where $A = 4\left(h_{23}-h_{13}\right)^2+\epsilon^2\left(h_{13}^2+kh_{23}^2\right)^2+4\epsilon\left(h_{23}-h_{13}\right)\left(h_{13}^2-kh_{23}^2\right)$.

Proof. To find a lower bound, we can approximate the functions in (5.4) by their second order Taylor series versus ϵ and obtain $R_{\text{NC}} \geq \max\{h_{13}\epsilon-\frac{h_{13}^2\epsilon^2}{2\beta}, h_{23}\epsilon-\frac{kh_{23}^2\epsilon^2}{2(1-\beta)}\}$.

To find a tight bound we solve $(h_{23}-h_{13})\,\beta^2+\left(\frac{h_{13}^2\epsilon}{2}+\frac{kh_{23}^2\epsilon}{2}+h_{13}-h_{23}\right)\beta-\frac{h_{13}^2\epsilon}{2}=0$. This quadratic equation has only one feasible solution in the interval $[0,1]$. This bound is described by (5.10a) and

$$\beta=\frac{(h_{23}-h_{13}-\frac{\epsilon kh_{23}^2}{2}-\frac{\epsilon h_{13}^2}{2})}{2\,(h_{23}-h_{13})}+\frac{\sqrt{\left(h_{23}-h_{13}-\frac{\epsilon kh_{23}^2}{2}-\frac{\epsilon h_{13}^2}{2}\right)^2+2\,(h_{23}-h_{13})\,\epsilon h_{13}^2}}{2\,(h_{23}-h_{13})}. \tag{5.11}$$

To obtain (5.10b), we notice that the achievable rate is also upper bounded by two end points of the curves. This upper bound is tight in the low TERN regime, i.e., where the received signal is dominated by noise power.

□

From the above, we conclude that

$$\lim_{\epsilon\to 0^+}\frac{R_{\text{NCP}}}{\epsilon}=\min\{h_{13},h_{23}\}. \tag{5.12}$$

Similar to the non-cooperative case, we derive the following upper and lower bounds for the CP:

Theorem 5.4. *For the CP, we have*

$$R_{\text{CP}} \;<\min\left\{\log(1+h_{12}\epsilon),\tfrac{1}{k+1}\log(1+kh_{23}\epsilon)\right\}\le\epsilon\min\left\{h_{12},\tfrac{k}{k+1}h_{23}\right\}, \tag{5.13a}$$

$$R_{\text{CP}} \;>\tfrac{\epsilon}{4}(\tfrac{2kh_{23}}{k+1}+2h_{12}-h_{12}^2\epsilon-\tfrac{h_{23}^2k^2\epsilon}{k+1}-B), \tag{5.13b}$$

where $B=\sqrt{4\left(\frac{kh_{23}}{k+1}-h_{12}\right)^2+\epsilon^2\left(h_{12}^2+\left(\frac{kh_{23}}{k+1}\right)^2\right)^2+4\epsilon\left(\frac{kh_{23}}{k+1}-h_{12}\right)\left(h_{12}^2-(\frac{kh_{23}}{k+1})^2\right)}$.

Thus, we conclude that

$$\lim_{\epsilon \to 0^+} \frac{R_{\text{CP}}}{\epsilon} = \min\{h_{12}, \frac{k}{k+1} h_{23}\}. \tag{5.14}$$

By combining (5.12) and (5.14), we get the following result

$$\lim_{\epsilon \to 0^+} \frac{R_{\text{CP}}}{R_{\text{NCP}}} = \frac{\min\left\{h_{12}, \frac{k}{k+1} h_{23}\right\}}{\min\left\{h_{13}, h_{23}\right\}}. \tag{5.15}$$

In addition, It is easy to show that $\frac{R_{\text{CP}}}{R_{\text{NCP}}}$ is always smaller than $\frac{\min\left\{h_{12}, \frac{k}{k+1} h_{23}\right\}}{\min\{h_{13}, h_{23}\}}$, i.e., $\frac{R_{\text{CP}}}{R_{\text{NCP}}} \leq \frac{\min\left\{h_{12}, \frac{k}{k+1} h_{23}\right\}}{\min\{h_{13}, h_{23}\}}$. This means that the rate gain can be greater than unity only if $h_{13} \leq \min\{h_{12}, h_{23}\frac{k}{k+1}\}$. In this case, the maximum rate gain ($\min\{\frac{h_{12}}{h_{13}}, \frac{h_{23}}{\frac{k+1}{k} h_{13}}\}$) is only achievable in the low TERN regime.

Now, we examine the cooperative gain when the rate ratio is large. It is easy to see that for large k, the optimal β, which is either the solution of (5.4) or (5.6), tends to zero, i.e., $\beta \to 0$. This implies that more resources should be allocated to the higher demanding user. Hence, it is easy to show that $\lim_{k \to \infty} \frac{\log(k)}{k} R_{\text{NCP}} = \lim_{k \to \infty} \frac{\log(k)}{k} R_{\text{CP}} = 1$. Then, it follows that

$$\lim_{k \to \infty} \frac{R_{\text{CP}}}{R_{\text{NCP}}} = 1 \tag{5.16}$$

On the other hand, if k tends to zero (where the rates of the 1$^{\text{st}}$ user is larger than the rate of the 2$^{\text{nd}}$ user), the optimal β for the NCP tends to unity, while for the CP tends to zero. Thus, the cooperative gain for small values of k, i.e., $k \to 0$, is

$$\lim_{k \to 0^+} \frac{1}{k} \frac{R_{\text{CP}}}{R_{\text{NCP}}} = \frac{h_{23}\epsilon}{\log\left(1 + h_{13}\epsilon\right)}. \tag{5.17}$$

It follows that for small enough rate ratio the achievable rate of the NCP is strictly greater than that of the CP, i.e, $R_{\mathrm{NCP}} > R_{\mathrm{CP}}$.

5.3.4 Energy Saving for Given Capacity and Resource

Here, we are interested in quantifying the advantage of the cooperation in terms of energy saving. This is in contrast to the previous section where the rate is maximized providing a fixed amount of available energy. Here, we assume that each user requires some specified rate R_i and has to allocate TERN proportional to R_i. In order to meet these rate requirements, users may cooperate (or not) to use available resource efficiently. Given a unit of the shared resource, we minimize the TERN as follows

$$\mathrm{CP}: \begin{cases} \min \epsilon_{\mathrm{CP}}, \\ \text{s.t. } R &= \beta \log\left(1+\frac{h_{12}\epsilon_{\mathrm{CP}}}{\beta}\right) \\ &= \frac{1-\beta}{k}\log\left(1+\frac{h_{23}k\epsilon_{\mathrm{CP}}}{1-\beta}\right) \end{cases} \tag{5.18a}$$

$$\mathrm{NCP}: \begin{cases} \min \epsilon_{\mathrm{NCP}}, \\ \text{s.t. } R &= \beta \log\left(1+\frac{h_{13}\epsilon_{\mathrm{CP}}}{\beta}\right) \\ &= \frac{1-\beta}{k+1}\log\left(1+\frac{h_{23}k\epsilon_{\mathrm{CP}}}{1-\beta}\right) \end{cases} \tag{5.18b}$$

Since the rates in (5.3) and (5.5) are monotonically increasing functions of TERN, it is easy to show that the optimization problem (5.18) is the dual of (5.3) and (5.5). This means that under similar channel gains, the TERN cooperation gain (i.e., the ratio of TERN in the NCP to that of cooperative one $\frac{\epsilon_{\mathrm{NCP}}}{\epsilon_{\mathrm{CP}}}$) obtained from (5.18) is the same as the rate cooperation gain from (5.3) and (5.5). More specifically from

this duality, we conclude that

$$\frac{\epsilon_{\mathrm{NCP}}}{\epsilon_{\mathrm{CP}}} \leq \frac{\min\left\{h_{12}, \frac{k}{k+1}h_{23}\right\}}{\min\left\{h_{13}, h_{23}\right\}}. \tag{5.19}$$

Similarly, the maximum gain is obtained for small rate demands, i.e., when $R \to 0$.

5.3.5 Resource Efficiency for Given Capacity and Energy

Let $\beta_{i,\mathrm{NCP}}$ ($\beta_{i,\mathrm{CP}}$), denote the required resource for the i^{th}, $i = 1, 2$, user to transmit its own information, R and kR (R and $(k+1)R$), under TERN constraints of ϵ and $k\epsilon$, respectively, in the NCP (CP). We also define resource efficiency as $\frac{\beta_{\mathrm{NCP}}}{\beta_{\mathrm{CP}}}$, where β_{NCP} and β_{CP} are solution of the following equations:

$$\begin{cases} R = \beta_{1,\mathrm{NCP}} \log\left(1 + \frac{h_{13}\epsilon}{\beta_{1,\mathrm{NCP}}}\right) = \frac{\beta_{2,\mathrm{NCP}}}{k} \log\left(1 + \frac{h_{23}\epsilon}{\beta_{2,\mathrm{NCP}}}\right), \\ \beta_{\mathrm{NCP}} = \beta_{1,\mathrm{NCP}} + \beta_{2,\mathrm{NCP}}, \end{cases} \tag{5.20a}$$

$$\begin{cases} R = \beta_{1,\mathrm{CP}} \log\left(1 + \frac{h_{12}\epsilon}{\beta_{1,\mathrm{CP}}}\right) = \frac{\beta_{2,\mathrm{CP}}}{k+1} \log\left(1 + \frac{h_{23}\epsilon}{\beta_{2,\mathrm{CP}}}\right), \\ \beta_{\mathrm{CP}} = \beta_{1,\mathrm{CP}} + \beta_{2,\mathrm{CP}}. \end{cases} \tag{5.20b}$$

Note that we have a feasible solution only if $R \leq \epsilon \min\{h_{13}, h_{23}\}$ for the NCP and $R \leq \epsilon \min\{h_{12}, h_{23}\frac{k}{k+1}\}$ for the CP. As the required rates approach these upper bounds the resource usage tends to infinity. In both protocols, due to the fairness constraint, the user with the worst channel obtains a larger amount of resource.

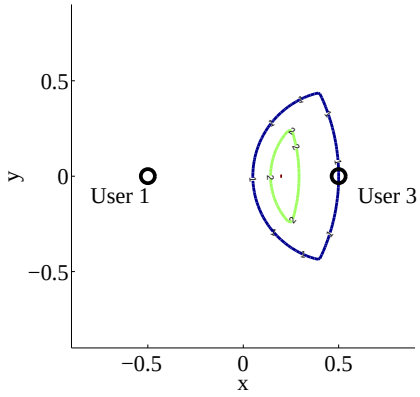
5.3.6 Effect of Network Geometry

In the following, we investigate the impact of the location of the relay user on the cooperation gain. In particular, we assume that the signal attenuation is governed

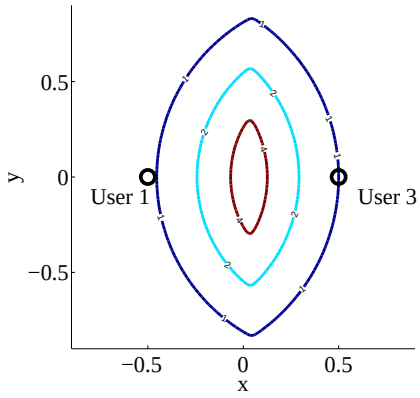
by geometry of users as $h_{ij} = \frac{1}{{d_{ij}}^{\eta}}$ on two dimensional plane, where d_{ij} denotes the distance between the i^{th} and j^{th} users. In order to understand the impact of users relative locations on the cooperation gain, we investigate the region where transmission via cooperation provides more gain. We assume that in the two dimensional plane, the source, relay and destination are located on $(-\frac{1}{2}, 0)$, (x, y) and $(\frac{1}{2}, 0)$, respectively. Plugging the channel gains as $\frac{1}{d^{\eta}}$ and $\frac{1}{(1-d)^{\eta}}$ into the equations (5.4) and (5.6), we obtain the rate improvement of both protocols as a function of geometry of relay user. Figure 5.2 depicts the region where cooperation provides more benefit, i.e., the rate of the CP is more than that of the NCP. This figure also depicts the contours of rate gain, where the ratio of achievable rate of protocols is fixed numbers (we plotted for the rate gains of 1, 2 and 4). We observe that as the rate ratio k increases the cooperation contours enlarge. Further increase in the rate ratio would result in shrinkage in the area of the gain contours. It implies that only the users with middle rate demand have incentive to cooperate with other users.

Since the channel gains are symmetric in two dimensional space, it is clear that the optimal relay user lies on the line connecting the source to the destination. We observe that the gain contours are approximately the intersections of two arcs with the radii $(\text{gc})^{1/\eta}$ and $(\frac{k+1}{k}\text{gc})^{1/\eta}$ with $\text{gc} = \frac{R_{\text{CP}}}{R_{\text{NCP}}}$. In order to find the optimal placement of the relay user we examine the equation (5.15). It is easy to see that the optimal location is

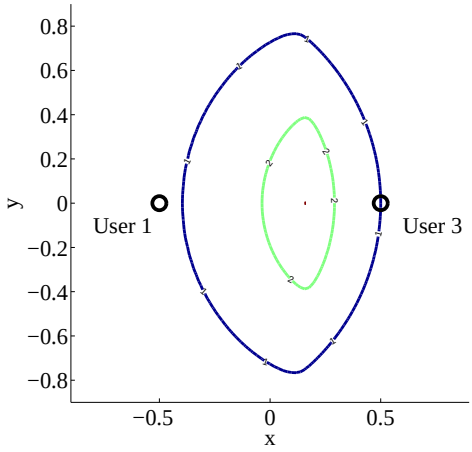
$$d = \frac{1}{1 + \left(\frac{k}{k+1}\right)^{1/\eta}} \tag{5.21}$$



(a)



(b)



(c)

Figure 5.2: Contours of the rate gain $\frac{R_{\mathrm{CP}}}{R_{\mathrm{NCP}}}$ (5.4), (5.6) versus relay (2^{nd} user) location (x, y) for $\epsilon = 0.01$, $h_{ij} = \frac{1}{d_{ij}^{\eta}}$ and $\eta = 3$, (a) $k = 0.1$, (b) $k = 10$, (c) $k = 100$

.

where at that point the maximum rate gain of

$$\frac{R_{\mathrm{CP}}}{R_{\mathrm{NCP}}} \leq \left(1+\sqrt[\eta]{\frac{k}{k+1}}\right)^{\eta} \tag{5.22}$$

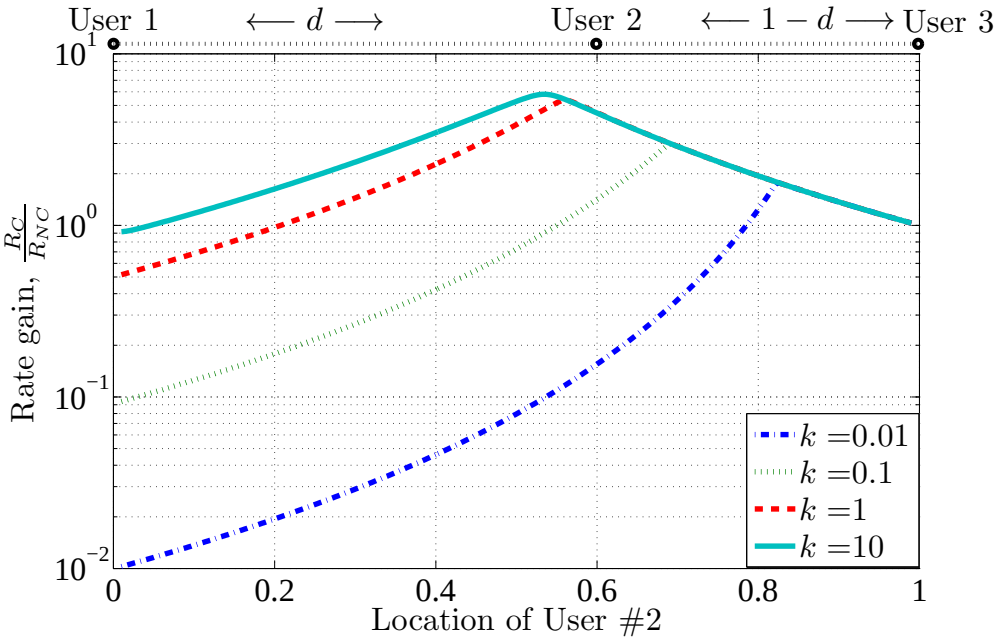
is achievable.

Clearly the optimal location lies on the line connecting nodes 1 and 3. We now assume that all three nodes are on a one dimensional line and are located at 0, d and 1, respectively, i.e. $h_{12}=\frac{1}{d^{\eta}}$, $h_{13}=1$ and $h_{23}=\frac{1}{(1-d)^{\eta}}$. Figure 5.3 shows the resulting ratio of the maximum achievable rates using the CP and NCP versus the location of the relay node $d\in[0,1]$ for $\eta=3$. As intuited from the upper bound in (5.22), for higher rate ratio k, more gain is expected (see in Figure 5.3(a)), however, as depicted in Figure 5.3(b), for large enough available energy of users ϵ, the cooperation gain degrades as rate ratio k increases. Figure 5.4 presents the rate improvement from the CP and NCP protocols versus the rate ratio of users k. We observe that for small rate ratio, the rate improvement is zero and for large values of k, the rate improvement tends to unity which also confirms that for high rate ratio k, the cooperation is not beneficial. Figure 5.5 depicts the resource gain of the CP compared with the NCP, i.e., $\frac{\beta_{\mathrm{NCP}}}{\beta_{\mathrm{CP}}}$ (5.20), for a required rate of $0.5h_{13}\epsilon$ versus location of the relay node. We observe that for a given required rate, depending on the relay channel condition, the resource gain is greater than unity. We have noticed that for small rate ratio k, the CP provides more gain in terms of resource usage. In addition, for small rate ratio, the best location for relay user is almost in the middle of the source and destination users. Figure 5.6 shows the energy gain of the CP compared with the NCP, i.e., $\frac{\epsilon_{\mathrm{NCP}}}{\epsilon_{\mathrm{CP}}}$ (5.18), for a given required rate of $R=0.09h_{13}$ versus the location of the relay node. Employing the CP, we obtain significant energy savings even for $\eta=3$, provided

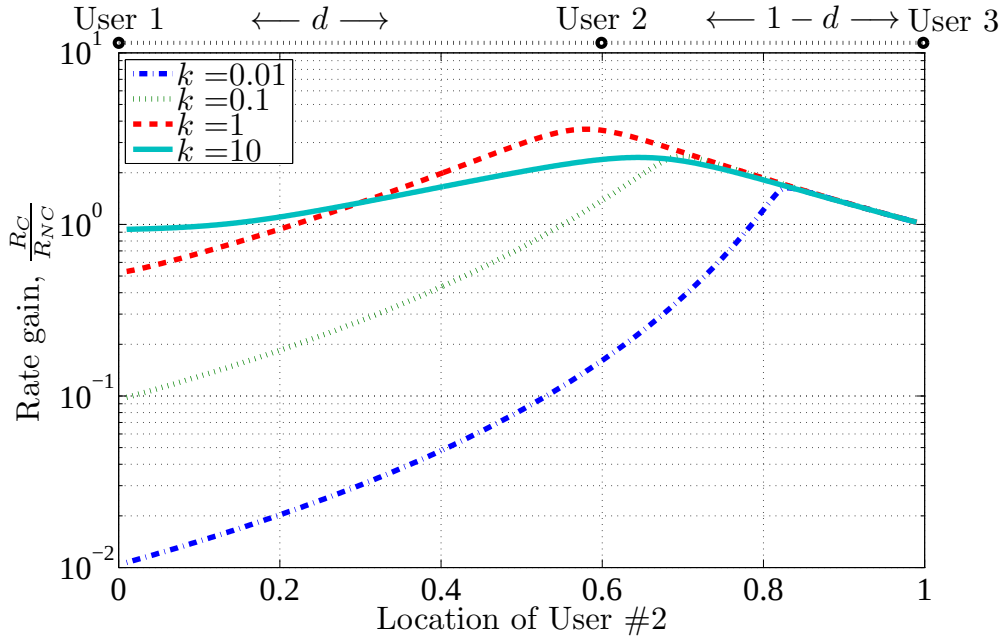
that the relay is located appropriately. In addition, we observe that for higher rate ratio (see Figure 5.5), the users benefit less in terms of resource efficiency. We deduce that only users which are interested in resource efficiency, with less rate requirement, can gain from possible cooperation. It is interesting to note that the CP provides rate/energy gain even for $\eta = 2$, by contrast, for such a small η there is no gain in rate/energy if the relay has no information to transmit (traditional multi-hopping) [64, 12, 11, 18].

5.4 Cooperation in Multiple Relay Networks

In the following, we propose our relay selection protocols based on the cooperation gain which is introduced in previous section. We use the channel gains to select one relay among the available relay users to participate in cooperation. We note that if the NCP outperforms the cooperative one, we fall back on the NCP, i.e., no relay user would be selected and the source sends its information to the destination directly. Otherwise, the source employs one relay in forwarding its information to the destination. The main objective of the proposed protocols are to achieve higher cooperation gain, higher rate improvement, energy saving or resource efficiency while guaranteeing fairness for all users. However, in large networks, the cost incurred by communication overhead must be considered in future works.



(a)



(b)

Figure 5.3: Impact of the relay location d on rate improvement $\frac{R_{\text{CP}}}{R_{\text{NCP}}}$ (5.4), (5.6) for $h_{12} = \frac{1}{d^{\eta}}$, $h_{13} = 1$, $h_{23} = \frac{1}{(1-d)^{\eta}}$, for $\eta = 3$, $k = 0.01, 0.1, 1, 10$, respectively, and different TERN values (a) $\epsilon = 0.01$, and (b) $\epsilon = 0.1$.

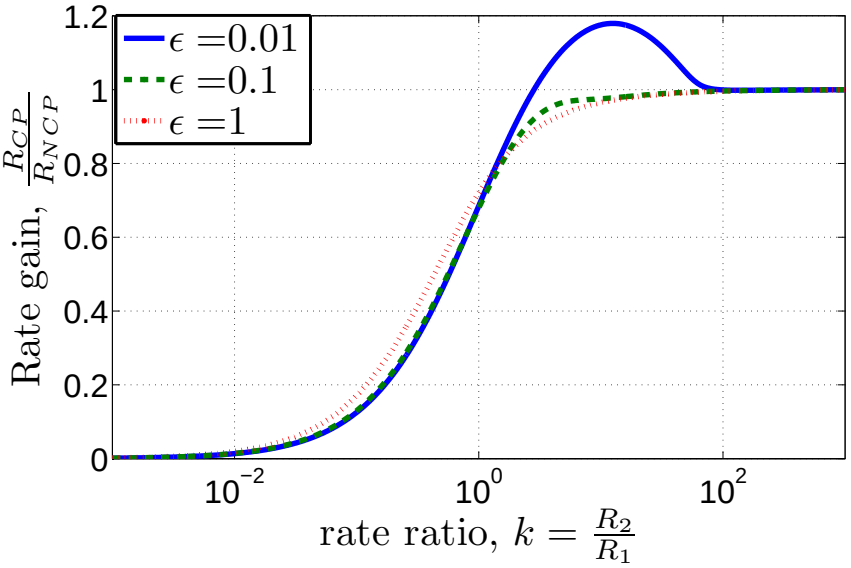


Figure 5.4: Impact of rate ratio k on rate improvement, $\frac{R_{\text{CP}}}{R_{\text{NCP}}}$, (5.4), (5.6), for $h_{12} = \frac{1}{d^\eta}$, $h_{13} = 1$, $h_{23} = \frac{1}{(1-d)^\eta}$ for a fixed relay location $d = 0.5$ for $\eta = 3$ and different TERN values $\epsilon = 0.01, 0.1, 1$.

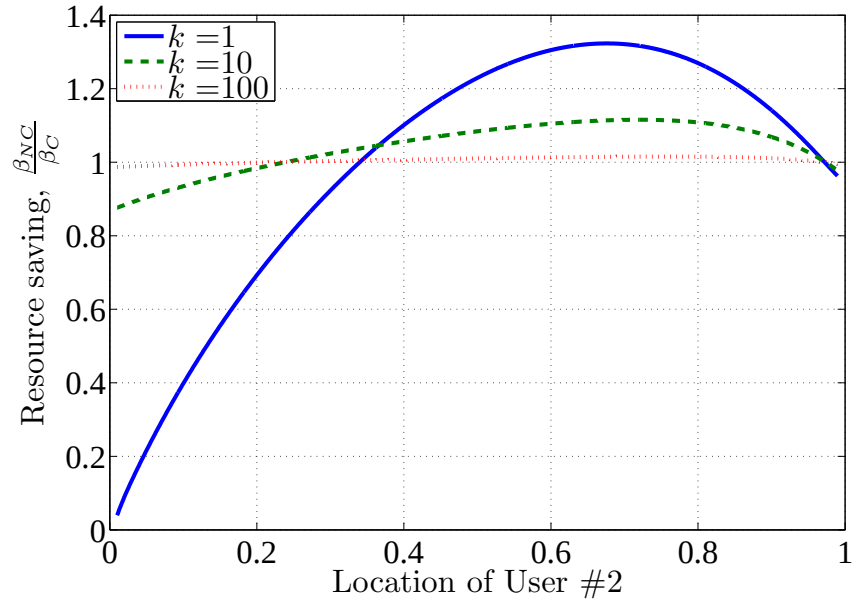


Figure 5.5: Ratio of resource usage in the CP and NCP $\frac{\beta_{\text{NCP}}}{\beta\text{CP}}$ (5.20) for $h_{12} = \frac{1}{d^\eta}$, $h_{13} = 1$ and $h_{23} = \frac{1}{(1-d)^\eta}$ versus relay location d for a required rate of $R = 0.5 h_{13}\epsilon$, $\eta = 3$ and $h_{13}\epsilon = 0.01$, and $k = 1, 10, 100$.

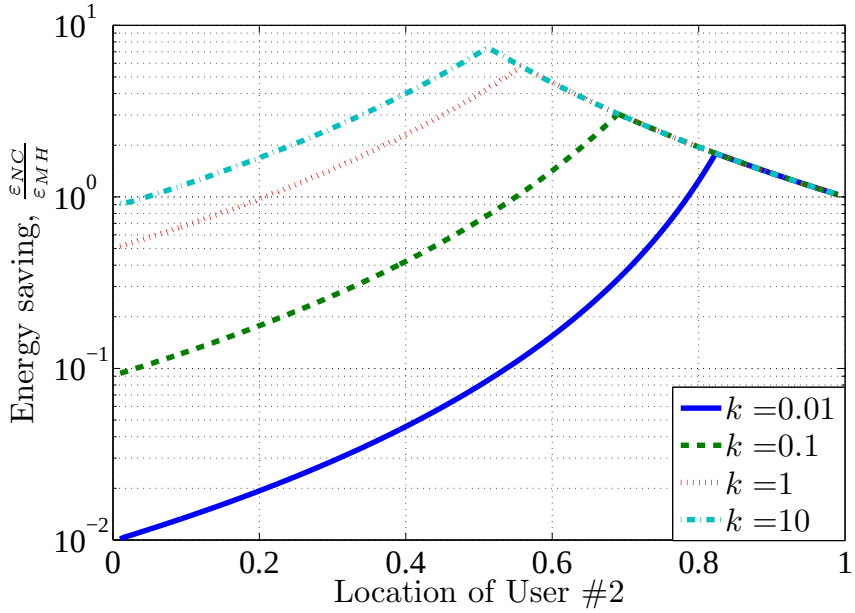


Figure 5.6: Ratio of energy usage in the CP and NCP $\frac{\epsilon_{\text{NCP}}}{\epsilon\text{CP}}$ (5.18) for $h_{12} = \frac{1}{d^\eta}$, $h_{13} = 1$, $h_{23} = \frac{1}{(1-d)^\eta}$ and $\eta = 3$ versus relay location d for unit resource and a given required rate of $R = h_{13}/100$, (a) $k = 0.01, 0.1, 1, 10$.

5.4.1 Relay Selection: Rate Improvement and Energy Saving

First, we consider the rate improvement as a criterion to select the best relay. As shown in previous section, the energy minimization problem is dual of the rate maximization problem, hence the relay selection protocol holds for the energy saving as well.

The results in (5.15) are very intuitive and suggests a strategy in deciding to use cooperation and to choose a relay user among the potential candidates. Given the full CSI, cooperation protocol is preferred if $\epsilon \ll 1$ and $h_{13} \ll \min\{h_{12}, h_{23}\frac{k}{k+1}\}$. In order to maximize the rate gain, the best relay user is the one that maximizes the $\frac{\min\{h_{12},h_{23}\frac{k}{k+1}\}}{h_{13}}$.

The results in (5.16) and (5.17) also provide an attractive guideline that for low and high rate ratio, the NCP is preferred. The equation (5.21) implies that the best relay user, in order to maximize the rate gain, is located almost in the middle of the source and destination users. We observe that under severe path loss, users benefit more from the proposed cooperation relative to direct transmission. Authors in [76] showed the same result in the context of diversity gain which is not in the scope of this chapter. This result also appear attractive that, in contrast to traditional multi-hopping, appropriately designed cooperation can provide a significant rate gain (see e.g., [64, 12, 11, 18]). Figures 5.3(a) and 5.3(b) confirm the above results. This indicates that the best location for the relay user is in the vicinity of the midpoint between the transmitter and the receiver pair. This means that by appropriately selecting the relay user, we can efficiently take advantage of the geometrical distribution of users. A relay with optimal location almost achieves (5.15), which serves for relay selection. Note that by selecting one relay, the multiple relay network becomes

a single relay network. Thus, the exact rate improvement or energy saving can be examined as in (5.4), (5.6), and (5.18).

5.4.2 Relay Selection: Resource Efficiency

Now, we address resource efficiency and the objective is to select a relay user among the potential candidates and to decide whether to cooperate or not. We propose the following procedure:

- Feasibility check: We compare R with $\epsilon \min\{h_{13}, h_{23}\}$ for the NCP and with $\epsilon \min\{h_{12}, h_{23}\frac{k}{k+1}\}$ for the CP. Then, we ignore the protocol which is infeasible.

- Resource usage: If both are feasible, we must choose the protocols with the least resource usage. The resource usages β_{NCP} and β_{CP} are the solutions of (5.20). It is worth noting that this criterion is different from the rate and energy criteria since here we are willing to minimize resource usage for a given amount of energy and rate which does not necessarily yields the same result as maximizing rate or energy ratios (compare Figures 5.5 with Figure 5.3 and Figure 5.6 for difference between obtained rate, energy and resource gains). However, in order to maximize the resource efficiency, the simulation result (Figures 5.5) shows that the best location for the relay user is almost in the middle of the source and the destination user.

- User selection: Similarly, we can use the resource usages as the criterion to select users (to cooperate with) among multiple feasible candidates.

5.5 Cooperation in General Networks

We can extend the proposed protocols to the general networks where more than one user is available to relay the messages of multiple users (as source users) toward the different destinations. As we have shown here, we focus on one relay system and look for the best user to serve as relay to maximize the achievable rate, minimize the energy consumption or utilize the available resource more efficiently. To this end, we provide a rough guideline that if direct link channel gain is smaller than the other links, often the CP outperforms the NCP. Otherwise, if a fixed rate is required, the feasibility of different scenarios must be verified. Among feasible solutions, we must choose the protocol and relays which provide maximum rate, or maximize savings on resource (5.20) or on energy (5.18). For the CP, a relay among possible candidates must be selected which maximizes $\min\{h_{23}k/(k+1), h_{12}\} \gg h_{13}$.

For example, suppose that in Figure 5.1 the 1^{st} user wishes to send data to the 3^{rd} user, while the 2^{nd} user wishes to broadcast independent messages to the 3^{rd} and 4^{th} users. Using this guideline, the 2^{nd} user can cooperate with the 1^{st} user via acting as relay (the more information, the more incentive to cooperate). In this example the 3^{rd} user has no data to send and thus, ironically, has no incentive to cooperate. So the 2^{nd} user should send his data directly to the 4^{th} user.

So far, we have assumed the same destination for both transmissions. We might relax this constraint. For example in Figure 5.1, suppose that the 1^{st} user wishes to send messages to the 3^{rd} user and the 2^{nd} and 3^{rd} users wish to send messages to the 4^{th} user. Using the CP, the 2^{nd} user can act as the relay between the 1^{st} and 3^{rd} users and the 3^{rd} user acts as the relay between the 2^{nd} and 4^{th} users.

We have shown that cooperation have the potential to increase the rate gain of

the users by a factor of at most $\left(1+\sqrt[\eta]{\frac{k}{k+1}}\right)^{\eta}$. This result shows that appropriately choosing the relay user and cooperation protocol considerably save the transmit energy, and also reduce interference amongst the users. Our proposed protocols not only improves rate, energy or resource utilization of the involved users, but also have the potential to increase the spectral utilization of the whole network. We have shown that cooperation can mitigate the effects of path loss, thus, users can save transmit energy. This saving reduces interference among users which allows to increase density of users in the network through resource reusing.

5.6 Conclusion

We used rate, energy and resource usage as criteria for cooperation and relay user selection. We found the conditions under which the cooperation is preferred for all the users. Interestingly, the gain of the users from cooperation in various terms (increase their achievable rate, reduce their transmit energy or use resources more efficiently) can be more significant at a low TERN, where the background noise is strong. Clearly, if the background noise is very weak, the cooperation is less attractive. The relative geometrical location of users (i.e., channel responses) must be considered in the relay selection. Very simple criteria are proposed for relay selection. If the relay is in the vicinity of the midpoint between the transmitter and the receiver pair, cooperation can offer good performance. A maximum rate gain (as well as energy saving gain) of up to $\left(1+\sqrt[\eta]{\frac{k}{k+1}}\right)^{\eta}$ can be obtained provided that a cooperation is established with an appropriately located relay, where η is the environment path loss exponent. Furthermore, we present several protocols on how to select the best relay among the possible candidates to maximize the cooperation gain.

Chapter 6

Conclusions and Future Works

6.1 Conclusions

The contemporary spectrum allocation policies have led to very poor usage of the overall spectrum. Cognitive radio is a promising paradigm in communication that has the potential to improve the spectrum utilization. Cognitive radio is distinguished by its two main features:

- cognitive users monitor the radio environment and respond to the changes by adapting their transmit-receive parameters, and
- peaceful existence of licensed and unlicensed users is allowed as long as the inflicted interference is below some certain threshold.

In this thesis, we addressed two main functionalities of every cognitive users, i.e.,

- Spectrum sensing: To sense/detect the presence of the licensed users to avoid collision with their transmission

- Spectrum management: To allocate the available radio resources more effectively between cognitive users (either licensed or unlicensed)

For the spectrum sensing task to mitigate the effect of fading, we used cooperation between the sensors/users. We started by answering fundamental questions on the interplay between the synchronization and performance of the distributed detectors. We then successfully incorporated channel characteristics and time-frequency features of the signal transmitted by the licensed users into the system. For the spectrum management task, we added one degree of freedom to the system by considering cooperative communications between the users. This led to some optimization problems that allowed us to re-allocate radio resources more efficiently, i.e., save power and bandwidth or gain rate.

In Chapter 2, we considered the problem of distributed detection where multiple sensors cooperatively detect the presence or absence of a random source. We first considered partially correlated observations and obtained the optimal detector. Assuming a noncoherent system, we obtained novel asymptotic expressions for the P_{fa} and P_{md} of several optimal/suboptimal detectors. We then compared these detectors with several detectors derived for the coherent system and showed that the coherent detectors require $2L$ times more bandwidth than the noncoherent ones. We also showed that the detectors for the noncoherent system are robust to the signal coherence and sampling time/frequency errors. Interestingly, the NPD for partially correlated observations reaches its highest performance when the observations are fully coherent and reaches its lowest performance when the observations are fully noncoherent. We also proposed one novel GLRD and two novel heuristic detectors for the noncoherent system. We also proved that all these detectors are either optimal

or asymptotically optimal.

In Chapter 3, we considered the distributed detection of an OFDM random source using a cooperative set of sensors. Assuming a noncoherent system, we have shown that sensors only need to transmit a real-valued function of their observations. We have analyzed several optimal and suboptimal detectors. We derived the P_{fa} and P_{md} of some of the proposed detectors. We also proved that the suboptimal detectors are asymptotically optimal. Our theoretical analysis and simulation results showed that using these distributed detectors, the sensors require negligible communication overhead. Our results also showed that the proposed frequency-domain detectors outperform the state-of-the-art time-domain detectors.

In Chapter 4, we derived several detectors for cooperative spectrum sensing in a mixture-Nakagami environment that exploit channel statistic for performance improvement. These detectors significantly outperform the energy detector. In addition, they are attractive in Rayleigh-lognormal fading channels, since, despite their lower computational complexity, they provides a near optimal performance.

In Chapter 5, we used rate, energy and resource usage as criteria for cooperation and relay-user selection. We found the conditions under which the cooperation is preferred by all the users. We studied the gains and losses involved with the possible cooperations. Interestingly, the gain of the users from cooperation is more significant at low TERN. We proposed very simple criteria for relay selection. A maximum rate gain (as well as energy saving gain) of up to $\left(1+\sqrt[\eta]{\frac{k}{k+1}}\right)^{\eta}$ can be obtained. Furthermore, we presented several protocols on how to select the best relay among the possible candidates to maximize the cooperation gain.

6.2 Future Works

The work in this thesis may be extended in many different ways.

- In Chapter 2, we used the Edgeworth expansion to derive the performance of the proposed tests. This can be further generalized and a framework for the asymptotic performance analysis of detectors can be investigated.

- In Chapter 3, we did not incorporate the channel fading statistic and the correlation between subchannels into the model. These can be potential avenues for future works.

- In Chapter 4, we considered mixture-Nakagami fading model. A potential extension is to derive detectors for other fading models such as $\mathbb{K}$-channel.

- In Chapters 2, 3, and 4, we assumed Gaussian noise, which may not hold in some cases. Studying detectors for non-Gaussian noise is also of great interest.

- In Chapter 5, we only considered users that cooperate by decoding and forwarding. The disadvantage of this method is that users are able to decode each other's message. The concept of secrecy can be introduced to such networks, where the private message of the users must not be decoded by the others. In addition, other cooperative protocols such as amplify/compress and forward, or beamforming may be considered.

Bibliography

[1] V. Aalo and R. Viswanathou. On distributed detection with correlated sensors: Two examples. *IEEE Trans. Aerosp. Electron. Syst.*, 25(3):414–421, 1989.

[2] M. Abramowitz and I.A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover publications, 1964.

[3] I.F. Akyildiz, Y. Altunbasak, F. Fekri, and Sivakumar R. AdaptNet: Adaptive protocol suite for next generation wireless internet. *IEEE Commun. Mag.*, 42(3):128–138, 2004.

[4] I.F. Akyildiz, W. Lee, M.C. Vuran, and S. Mohanty. A survey on spectrum management in cognitive radio networks. *IEEE Commun. Mag.*, 46(4):40–48, 2008.

[5] I.F. Akyildiz, W.Y. Lee, M.C. Vuran, and S. Mohanty. NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey. *Computer Networks*, 50(13):2127–2159, 2006.

[6] S. Alrabaee, A. Agarwal, N. Goel, M. Zaman, and M. Khasawneh. Comparison of spectrum management without game theory (smwg) and with game theory (smg) for network performance in cognitive radio network. In *7th International*

Conference onBroadband, Wireless Computing, Communication and Applications (BWCCA), pages 348–355. IEEE, 2012.

[7] S. Alrabaee, A. Agarwal, N. Goel, M. Zaman, and M. Khasawneh. Routing management algorithm based on spectrum trading and spectrum competition in cognitive radio networks. In *7th International Conference on Broadband, Wireless Computing, Communication and Applications (BWCCA)*, pages 614–619. IEEE, 2012.

[8] S.V. Amari and R.B. Misra. Closed-form expressions for distribution of sum of exponential random variables. *IEEE Trans. Rel.*, 46(4):519–522, 1997.

[9] A. Anandkumar, A. Willsky, and L. Tong. Detection error exponent for spatially dependent samples in random networks. In *IEEE International Symposium on Information Theory (ISIT)*, pages 2882–2886, 2009.

[10] T.W. Anderson. Asymptotic theory for principal component analysis. *The Annals of Mathematical Statistics*, 34(1):122–148, 1963.

[11] S. A. Astaneh and S. Gazor. Joint protocol and relay node selection in collaborative networks. *24th Biennial Symposium on Communications (QBSC)*, pages 162–165, 2008.

[12] S. A. Astaneh and S. Gazor. Collaborative communications: Joint relay and protocol selection. *11th Canadian Workshop on Information Theory (CWIT)*, pages 25–28, 2009.

[13] S. A. Astaneh and S. Gazor. Resource allocation and relay selection for collaborative communications. *IEEE Trans. Wireless Commun.*, 2009.

[14] S. A. Astaneh and S. Gazor. Asymptotic performance analysis of distributed source detectors. *Submitted to IEEE Trans. Wireless Commun.*, 2013.

[15] S. A. Astaneh and S. Gazor. Cooperative spectrum sensing over mixture-Nakagami channels. *IEEE Wireless Commun. Lett.*, 2013.

[16] S. A. Astaneh and S. Gazor. Noncoherent distributed OFDM source detection. *Submitted to IEEE Trans. Wireless Commun.*, 2013.

[17] S. A. Astaneh and S. Gazor. Relay assited spectrum sensing. *Submitted to IET Commun.*, 2013.

[18] S. A. Astaneh, S. Gazor, and H. Behroozi. On the capacity of pairwise collaborative networks. *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1–5, 2008.

[19] S. Atapattu, C. Tellambura, and H. Jiang. Energy detection based cooperative spectrum sensing in cognitive radio networks. *IEEE Trans. Wireless Commun.*, 10(4):1232–1241, 2011.

[20] I. Avramidi. Lecture notes on asymptotic expansion, 2000.

[21] E. Axell and E.G. Larsson. Optimal and sub-optimal spectrum sensing of OFDM signals in known and unknown noise variance. *IEEE J. Sel. Areas Commun.*, 29(2):290–304, 2011.

[22] N.C. Beaulieu and C. Cheng. Efficient Nakagami-m fading channel simulation. *IEEE Trans. Veh. Commun.*, 54(2):413–424, 2005.

[23] E. Beres and R. Adve. Selection cooperation in multi-source cooperative networks. *IEEE Trans. Wireless Commun.*, 7(1):118–127, 2008.

[24] L. Berlemann, S. Mangold, and B.H. Walke. Policy-based reasoning for spectrum sharing in cognitive radio networks. *Proc. IEEE DySPAN*, pages 1–10, Nov. 2005.

[25] R.N. Bhattacharya and J.K. Ghosh. On the validity of the formal Edgeworth expansion. *The Annals of statistics*, 6(2):434–451, 1978.

[26] R.N. Bhattacharya and J.K. Ghosh. On moment conditions for valid formal Edgeworth expansions. *Journal Of Multivariate Analysis*, 27:68–79, 1988.

[27] P. Bianchi, M. Debbah, M. Maïda, and J. Najim. Performance of statistical tests for single-source detection using random matrix theory. *IEEE Trans. Inf. Theory*, 57(4):2400–2419, 2011.

[28] A. Bletsas, H. Shin, and MZ Win. Cooperative communications with outage-optimal opportunistic relaying. *IEEE Trans. Wireless Commun.*, 6(9):3450–3460, 2007.

[29] S. Blinnikov and R. Moessner. Expansions for nearly Gaussian distributions. *Astronomy And Astrophysics*, 1997.

[30] R.S. Blum, S.A. Kassam, and H.V. Poor. Distributed detection with multiple sensors I. Advanced topics. *Proc. IEEE*, 85(1):64–79, 2002.

[31] S. Bokharaiee, H.H. Nguyen, and E. Shwedyk. Blind spectrum sensing for OFDM-based cognitive radio systems. *IEEE Trans. Veh. Commun.*, 60(3), 2011.

[32] L. Cao and H. Zheng. Distributed spectrum allocation via local bargaining. *Proc. IEEE Sensor and Ad Hoc Communications and Networks (SECON)*, pages 475–486, Sept. 2005.

[33] S. Chaudhari, V. Koivunen, and H.V. Poor. Autocorrelation-based decentralized sequential detection of OFDM signals in cognitive radios. *IEEE Trans. Signal Process.*, 57(7):2690–2700, 2009.

[34] M. Chen, S. Serbetli, and A. Yener. Distributed power allocation strategies for parallel relay networks. *IEEE Trans. Wireless Commun.*, 7(2):552–561, 2008.

[35] P.N. Chen. General formulas for the Neyman-Pearson type-II error exponent subject to fixed and exponential type-I error bounds. *IEEE Trans. Inf. Theory*, 42(1):316–323, 1996.

[36] C. Cordeiro, K. Challapali, D. Birru, and S. Shankar. IEEE 802.22: The first worldwide wireless standard based on cognitive radios. *Proc. IEEE DySPAN*, pages 328–337, Nov. 2005.

[37] T.M. Cover and J.A. Thomas. *Elements of information theory*. Wiley-Interscience, 2006.

[38] E. Dahlman. *3G evolution: HSPA and LTE for mobile broadband*. Academic Pr., 2008.

[39] M. Debbah and R.R. Muller. MIMO channel modeling and the principle of maximum entropy. *IEEE Trans. Inf. Theory*, 51(5):1667–1690, 2005.

[40] C. Dovrolis, D. Stiliadis, and P. Ramanathan. Proportional differentiated services: Delay differentiation and packet scheduling. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, pages 109–120, 1999.

[41] E. Drakopoulos and C.C. Lee. Optimum multisensor fusion of correlated local decisions. *IEEE Trans. Aerosp. Electron. Syst.*, 27(4):593–606, 1991.

[42] FCC. ET Docket No 03-222 Notice of proposed rule making and order. Dec. 2003.

[43] A. Fehske, J. Gaeddert, and J.H. Reed. A new approach to signal classification using spectral correlation and neural networks. In *1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, pages 144–150, 2005.

[44] M. Gastpar. On capacity under receive and spatial spectrum-sharing constraints. *IEEE Trans. Inf. Theory*, 53(2):471–487, 2007.

[45] S. Gazor and H.S. Rad. Space-time-frequency characterization of mimo wireless channels. *IEEE Trans. Wireless Commun.*, 5(9):2369–2375, 2006.

[46] S. Gezici, Z. Sahinoglu, and H.V. Poor. On the optimality of equal gain combining for energy detection of unknown signals. *IEEE Commun. Lett.*, 10(11):772–774, 2006.

[47] A. Ghasemi and E.S. Sousa. Collaborative spectrum sensing for opportunistic access in fading environments. In *1st IEEE International Symposium on New*

Frontiers in Dynamic Spectrum Access Networks (DySPAN), pages 131–136, 2005.

[48] A. Ghasemi and E.S. Sousa. Opportunistic spectrum access in fading channels through collaborative sensing. *J. Commun.*, 2(2):71, 2007.

[49] A. Ghobadzadeh, S. Gazor, M. R. Taban, Tadaion A. A., and M. Gazor. Separating function estimation tests: A new perspective on binary composite hypothesis testing. *IEEE Trans. Signal Process.*, 2012.

[50] R.E. Glaser. The ratio of the geometric mean to the arithmetic mean for a random sample from a gamma distribution. *J. Am. Stat. Assoc.*, 71(354):480–487, 1976.

[51] D. Goodman and N. Mandayam. Power control for wireless data. *IEEE Trans. Wireless Commun.*, 7(2):48–54, 2000.

[52] I.S. Gradshteyn, I.M. Ryzhik, and A. Jeffrey. *Table of integrals, series, and products*. Academic Pr, 2000.

[53] R.M. Gray. *Toeplitz and circulant matrices: A review*. Now Pub, 2006.

[54] S. Haykin. Cognitive radio: brain-empowered wireless communications. *IEEE J. Sel. Areas Commun.*, 23(2):201–220, 2005.

[55] I.Y. Hoballah and P.K. Varshney. Distributed bayesian signal detection. *IEEE Trans. Inf. Theory*, 35(5):995–1000, 2002.

[56] W.D. Horne. Adaptive spectrum access: Using the full spectrum space. *Proc. Telecommun. Policy Research Conference (TPRC)*, Sept. 2003.

[57] W.J. Huang, Y.W. Hong, and C.C.J. Kuo. Lifetime maximization for amplify-and-forward cooperative networks. *IEEE Trans. Wireless Commun.*, 7(5 Part 2):1800–1805, 2008.

[58] P.J. Huber. The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 221–33, 1967.

[59] A.S. Ibrahim, A.K. Sadek, W. Su, and K.J.R. Liu. Cooperative communications with relay-selection: When to cooperate and whom to cooperate with? *IEEE Trans. Wireless Commun.*, 7(7), 2008.

[60] F.K. Jondral. Software-defined radio-basic and evolution to cognitive radio. 2005.

[61] S. Kay. Asymptotically optimal detection in unknown colored noise via autoregressive modeling. *IEEE Trans. Acoust., Speech, Signal Process.*, 31(4):927–940, 1983.

[62] S.M. Kay. *Fundamentals of Statistical signal processing, Volume 2: Detection theory*. Prentice Hall PTR, 1998.

[63] V.I. Kostylev. Energy detection of a signal with random amplitude. In *IEEE International Conference on Communications (ICC)*, volume 3.

[64] G. Kramer, M. Gastpar, and P. Gupta. Cooperative strategies and capacity theorems for relay networks. *IEEE Trans. Inf. Theory*, 51(9):3037–3063, 2005.

[65] G.S. Lauer and N.R. Sandell. Distributed detection with waveform observations: correlated observation processes. In *American Control Conference (ACC)*, pages 812–819, 1982.

[66] L. Le and E. Hossain. Cross-layer optimization frameworks for multihop wireless networks using cooperative diversity. *IEEE Trans. Wireless Commun.*, 7(7), 2008.

[67] Y. Li, N. Seshadri, and S. Ariyavisitakul. Channel estimation for OFDM systems with transmitter diversity inmobile wireless channels. *IEEE J. Sel. Areas Commun.*, 17(3):461–471, 1999.

[68] N. Madhu. Note on measures for spectral flatness. *Electronics letters*, 45(23):1195–1196, 2009.

[69] M.A. McHenry, P.A. Tenhula, D. McCloskey, D.A. Roberson, and C.S. Hood. Chicago spectrum occupancy measurements & analysis and a long-term studies proposal. In *Proceedings of the 1st International Workshop on Technology and Policy for Accessing Spectrum*, page 1, 2006.

[70] F. Meshkati, D. Guo, H.V. Poor, and S.C. Schwart. A unified approach to power control in large energy-constrained cdms systems. *IEEE Trans. Wireless Commun.*, 7(4):1208–1216, 2008.

[71] F. Moghimi, R. Schober, and R.K. Mallik. Hybrid coherent/energy detection for cognitive radio networks. *IEEE Trans. Wireless Commun.*, 10(5):1594–1605, 2011.

[72] T.C. Ng and W. Yu. Joint optimization of relay strategies and resource allocations in cooperative cellular networks. *IEEE J. Sel. Areas Commun.*, 25(2):328–339, 2007.

[73] A.F. Nikiforov, V.B. Uvarov, and R.P. Boas. *Special functions of mathematical physics*. Birkhäuser BaselBoston, 1988.

[74] D. Niyato and E. Hossain. Competitive spectrum sharing in cognitive radio networks: a dynamic game approach. *IEEE Trans. Wireless Commun.*, 7(7):2651–2660, 2008.

[75] A. Nosratinia and T.E. Hunter. Grouping and partner selection in cooperative wireless networks. *IEEE J. Sel. Areas Commun.*, 25(2):369, 2007.

[76] H. Ochiai, P. Mitran, and V. Tarokh. Variable-rate two-phase collaborative communication protocols for wireless networks. *IEEE Trans. Inf. Theory*, 52(9):4299–4313, 2006.

[77] M.R. Oularbi, S. Gazor, S. Houcke, and A. Aissa-El-Bey. GLR test for OFDM system identification using pilot tones pattern. *EURASIP J. Wireless Commun. and Networking*, 2013(1):72, 2013.

[78] E. Panayirci, H. Senol, and H.V. Poor. Joint channel estimation, equalization, and data detection for OFDM systems in the presence of very high mobility. *IEEE Trans. Signal Process.*, 58(8):4225–4238, 2010.

[79] H.V. Poor. *An Introduction to Signal Detection and Estimation*, volume 1. 1988.

[80] Z. Quan., S. Cui, and A.H. Sayed. Optimal linear cooperation for spectrum sensing in cognitive radio networks. *IEEE J. Sel. Topics Signal Process.*, 2(1):28–40, 2008.

[81] Z. Quan, S. Cui, A.H. Sayed, and H.V. Poor. Optimal multiband joint detection for spectrum sensing in cognitive radio networks. *IEEE Trans. Signal Process.*, 57(3):1128–1140, 2009.

[82] C. Rago, P. Willett, and Y. Bar-Shalom. Censoring sensors: A low-communication-rate scheme for distributed detection. *IEEE Trans. Aerosp. Electron. Syst.*, 32(2):554–568, 2002.

[83] N. Reid. Saddlepoint methods and statistical inference. *Statistical Science*, 3(2):213–227, 1988.

[84] K. Ruttik, K. Koufos, and R. Jantti. Detection of unknown signals in a fading environment. *IEEE Commun. Lett.*, 13(7):498–500, 2009.

[85] A. Sahai, N. Hoven, and R. Tandra. Some fundamental limits on cognitive radio. In *Allerton Conference on Communication, Control, and Computing*, pages 1662–1671, 2004.

[86] A. Sendonaris, E. Erkip, and B. Aazhang. User cooperation diversity. part I. system description. *IEEE Trans. Commun.*, 51(11):1927–1938, 2003.

[87] Marvin K Simon and Mohamed-Slim Alouini. *Digital communication over fading channels*, volume 95. Wiley-Interscience, 2005.

[88] F. Sivrikaya and B. Yener. Time synchronization in sensor networks: a survey. *IEEE Network*, 18(4):45–50, 2004.

[89] J.A. Stine. Spectrum management: The killer application of ad hoc and mesh networking. *Proc. IEEE DySPAN*, pages 184–193, Nov. 2005.

[90] A.A. Tadaion, M. Derakhtian, S. Gazor, and M.R. Aref. A fast multiple-source detection and localization array signal processing algorithm using the spatial filtering and ml approach. *IEEE Trans. Signal Process.*, 55(5):1815–1827, 2007.

[91] A Taherpour, S Gazor, and M Nasiri-Kenari. Wideband spectrum sensing in unknown white gaussian noise. *IET commun.*, 2(6):763–771, 2008.

[92] A. Taherpour, M. Nasiri-Kenari, and S. Gazor. Invariant wideband spectrum sensing under unknown variances. *IEEE Trans. Wireless Commun.*, 8(5):2182–2186, 2009.

[93] A. Taherpour, M. Nasiri-Kenari, and S. Gazor. Multiple antenna spectrum sensing in cognitive radios. *IEEE Trans. Wireless Commun.*, 9(2):814 –823, 2010.

[94] J. Unnikrishnan and V.V. Veeravalli. Cooperative sensing for primary detection in cognitive radio. *IEEE J. Sel. Topics Signal Process.*, 2(1):18–27, 2008.

[95] M. Unser. On the approximation of the discrete karhunen-loeve transform for stationary processes. *Signal Processing*, 7(3):231–249, 1984.

[96] R. Viswanathan and V. Aalo. On counting rules in distributed detection. *IEEE Trans. Acoust., Speech, Signal Process.*, 37(5):772–775, 2002.

[97] R. Viswanathan and P.K. Varshney. Distributed detection with multiple sensors I. Fundamentals. *Proc. IEEE*, 85(1):54–63, 2002.

[98] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Trans. Acoust., Speech, Signal Process.*, 33(2):387–392, 1985.

[99] H. Yin and H. Liu. An efficient multiuser loading algorithm for OFDM-based broadband wireless systems. In *IEEE Global Telecommunications Conference (GLOBECOM)*, volume 1, pages 103–107, 2000.

[100] Q. Zhang, J. Jia, and J. Zhang. Cooperative relay to improve diversity in cognitive radio networks. *IEEE Commun. Mag.*, 47(2):111–117, 2009.

[101] Q. Zhao, L. Tong, and A. Swami. Decentralized cognitive MAC for dynamic spectrum access. *Proc. IEEE DySPAN*, pages 224–232, Nov. 2005.

[102] Y. Zhao, R. Adve, and T.J. Lim. Improving amplify-and-forward relay networks: optimal power allocation versus selection. *IEEE Trans. Wireless Commun.*, 6(8):3114–3123, 2007.

[103] Z. Zhou, S. Zhou, J. Cui, and S. Cui. Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks. *IEEE Trans. Wireless Commun.*, 7(8):3066–3078, 2008.