Spectrum Sensing Techniques for Cognitive Radio Applications

A Thesis

Submitted for the Degree of

Doctor of Philosophy

in the Faculty of Engineering

Submitted by

Sanjeev G.

Electrical Communication Engineering
Indian Institute of Science, Bangalore
Bangalore – 560 012 (INDIA)

January 2015

TO

My parents

Smt. Shakunthala Aggithala Padmanabha
and
Sri. Gurugopinath Sanjeeva Rao

Acknowledgements

ಶ್ರೀಕಾಲತರಂಗಹಯ-
ಪ್ರಾಕಟಿ! ಗುರುಶಿಷ್ಯರನ್ವಯಾವನಲೀಲೇ! |
ಸಾಕಾರನಿರಾಕಾರಾ-
ಸ್ತೋಕಮಹಿಮೆ! ವಾಣಿ! ಕೃತಿಯ ಮೊದಲೊಳ್ ನಮಿಪೆಂ ||

ಪಿರಿಯನರುಹಿದೋಜಂ ಸಾ-
ಸಿರದೊಂದಾವರ್ತಿ ತಿರ್ದಿದರಿಮೆಯ ಗುರುಗಳ್ |
ಸರಸತಿಯಣುಗಂ ನಲ್ವಳಿ-
ಸಿರಿಯಂ ಪಡೆದಾರ್ಯಚಂದ್ರಮೂರ್ತಿಗೆ ನಮಿಪೆಂ ||

ಮನದ ಸೊಗಮನೊಂದನೆ ನಾಂ
ಮನನಿಸಿ ಗೆಯ್ದೆಲ್ಲಮಂ ಸಹಿಸಿ ಮುಳಿಯದೆಯೇ |
ಅನುನಯಿಸಿದರ್ಗೀದೊ ಮೌನದ
ವಿನತಿ ಜನಕಜನನಿಸೋದರಾರ್ಧಾಂಗಿಯರ್ಗಂ ||

ಎನಗೀ ಲೋಕಕಲಾವಿ-
ಜ್ಞನಯದತತ್ತ್ವದೊಳೆ ಸಾಮ್ಯಮರಿಪಿತು ಗೌಂಡ್ಯಂ |
ವಿನಯಂ ವಿವೇಕದುಕ್ತಿಯೊ-
ಳೆ ನಗೆಮೊಗಂ ತೋರ್ದ ಕೃಷ್ಣಮೂರ್ತಿಯನೊಲಿದೆಂ ||

ರಮ್ಯಂ ನಾಲ್ಕಯಾಮಂ
ಸಾಮ್ಯಮಿದರೊಳೆನಗೆ ದೊರೆತ ಭಾಗ್ಯಮಿವರ್ಗಳೊಳ್ |
ಸೌಮ್ಯಪ್ರೀತಿಯಿನಾಪ್ತಸು-
ಕಾಮ್ಯಂ ಬೇರೊಂದು ಪುಟ್ಟನೆಳಸೆಂ ಬುವಿಯೊಳ್ ||

ಮಿತಿಯಂ ಶಾನನಿಂ ಮಾ
ಹಿತಿಹಾದಿಯೊಳಿಪುರ್ದರಿತು ಕಲಿಕೆಯ ಮೊದಲೊಳ್|
ಸ್ಥಿತಿಯಾನಂದಾಶ್ರಿತಸ-
ದ್ಧತಿಯರಿತೆಂ ನಾಂ ಗಣೇಶನಿಂದರಿವೆಡೆಯೊಳ್ ||

ಖದ್ಯೋತಪುರಿಯ ಮಾರ್ಷೊ
ಉದ್ಯಮನಗರಿಯೊಳೆ ಕೂರ್ತ ಜಫರಲಿ ಎನ್ನೀ |
ಹೃದ್ಯಾಂಗಕೃತಿಯ ತಿರ್ದ-
ಲ್ಕುದ್ಯೋತಂ ಹೀರಹಾರಮೆನಿಸುತೆ ತೋರ್ಕುಂ ||

   In the above stanzas,[1] I acknowledge and salute *Saraswathi*, the Goddess of Learning, my mentor Chandra Murthy, my family: my parents, brother, and wife, my philosophical and spiritual guides D. V. Gundappa, K. Krishnamoorthy and M. Krishnamurti, Shannon, the Father of Information Theory, Shatavadhani R. Ganesh, my thesis examiners Marceau Coupechoux and Mohammed Zafar Ali Khan.

---

[1]This piece of poetry, called as the *Kanda Padya*, is based on an ancient form of Chandas (metre), used in Kannada by 10th Century poets such as *Pampa*, *Ranna*, *Ponna* to the 21st Century poets such as *Basavappa Shastry*, *D. V. Gundappa* and *R. Ganesh*. Interested readers can refer to the following reference for more information on its aesthetic details.
S. Krishna Bhatta, "*Sediyapu Chandassamputa*," 1st ed., Rastrakavi Govinda Pai Samshodhana Kendra, 2006.

I would like to express my gratitude to my adviser, Prof. Chandra R. Murthy, without whom this work would not have been possible. During the kindergarten days of my research life, he taught me to walk with confidence, helped me stand up whenever I fell, and even carried me when I could not walk. It is seldom that one finds a man, who is punctual, diligent, a well-prepared teacher, does A-grade quality research work, and does it with both his feet on the ground. His dedication towards preparing for his lectures and presentations have simply stunned me over all these years. All the well-constructed sentences in this thesis are due to him; either directly or indirectly. I have learnt about the art of research, and a lot of presentation, teaching, and social skills from him. Thanks for everything, Chandra.

I thank the Director of the Indian Institute of Science, and the Chairman of the ECE department, for giving me an opportunity to work in this esteemed institution. Thanks to the funding from the Ministry of Human Resource Development, Government of India, for providing me rice, sambar and shelter throughout my stay here.

I thank the Govt. of India, Aerospace Network Research Consortium (ANRC), IEEE Globecom student travel grant, SPCOM student travel grant, and directorate of extra-mural research and intellectual property rights, Defence Research and Development Organization (DRDO), Govt. of India, for their financial support for my research.

I would like to thank all my co-authors and collaborators, especially Prof. Vinod Sharma, Prof. Chandra Sekhar Seelamantula, and Prof. Bharadwaj Amrutur.

I thank my teachers Profs. Chandra Murthy, Neelesh Mehta, Rajesh Sundaresan, Utpal Mukherjee, Anurag Kumar, A. Chokalingam, K. V. S. Hari, Chandra Sekhar Seelamantula, Vinod Sharma, M. K. Ghosh, and S. K. Iyer, for providing insights into the respective courses that they taught.

Heartfelt thanks to my dear friends Bharath, Nagananda, Deepa, Chandrashekar, Ganesan, J. Chandrasekhar, Krishna Chaythanya, Srinivas Reddy, Venugopalakrishna, Parthajit, Ranjitha, Abhay, Santhosh Kumar, Suma, Raghavendra, Anup, Reuben, Manesh, Sinchu, Karthik Velugoori, Sneha Latha and others for making my stay a memorable one. God bless everyone. Special thanks to pals Bharath and Nagananda for all the timely visits to Vidyarthi Bhavan, Ranjitha and Abhay for the memorable trips and treks, Venu and Parthajit for the philosophical discussions, and Venu for all the valuable time spent at GIPA. Also, sincere thanks to my friends from the other labs and

departments Karthik A., Gautam, Vijesh Joshi, Ashok, Harshan, Naveen K. P., Srinidhi, Harshavardhan, Naveen Deshpande, Shubha, Raghava, and Aditya Roychoudhury.

I thank various clubs in IISc that I was a part of. First off, I thank the Music Club for providing me with an opportunity to be the Convener for an year. Rhythmica, the official IISc band, has given me a lot of friends, and a rare opportunity to work with 50+ talented musicians all across the campus. Thanks, fellas! I would also like to thank the kung-fu, tennis and badminton clubs for giving me fun-filled days to cherish forever.

Interactions with my friends from the lab and Rhythmica at eat outlets in the campus such as the coffee board, tea kiosk, Prakruthi, Kabini canteen, Nesara, and gym-cafe have been very fruitful. Breakfast, lunch and dinner sessions in A-mess were awesome! I thank all the managing and helping hands in these places.

I would like to express my gratitude to the staff at the ECE dept. office for making my life simpler by directing me to the correct places to get all my administration-related work done smoothly. Thanks to the staff at the main hostel office, and the U-block office for their support, especially on the day when I had to break open the lock of my room! My sincere thanks and appreciation goes to the chief security officer Mr. Chandrasekhar, and staff at the security office, who work day and night to keep the campus safe for the residents.

Finally, my heartfelt thanks to my Father, Sri. Gurugopinath, my Mother, Smt. Shakunthala for providing me the opportunity to pursue my dreams while they toiled day and night. In the age when the entire race shifted its attention to the software industry, my Father believed in my teaching abilities and has taken all the financial responsibilities so that I could work in peace. My little brother, Sri. Rajiv, has taken a lot of responsibilities from my shoulders, and has done a great job, without complaining. My blessings are with him. The latest entry into my life, Shwetha, my wife has taken a very brave decision to marry a research student on stipend, and has sacrificed a lot of her small fun filled moments for my research. Thanks, everyone!

Abstract

Cognitive Radio (CR) has received tremendous research attention over the past decade, both in the academia and industry, as it is envisioned as a promising solution to the problem of spectrum scarcity. A CR is a device that senses the spectrum for occupancy by licensed users (also called as primary users), and transmits its data only when the spectrum is sensed to be available. For the efficient utilization of the spectrum while also guaranteeing adequate protection to the licensed user from harmful interference, the CR should be able to sense the spectrum for primary occupancy quickly as well as accurately. This makes Spectrum Sensing (SS) one of the fundamental blocks in the operation of a CR. At its core, SS is a hypothesis testing problem, where the goal is to test whether the primary user is inactive (the null or noise-only hypothesis), or not (the alternate or signal-present hypothesis). Computational simplicity, robustness to uncertainties in the knowledge of various noise, signal, and fading parameters, and ability to handle interference or other source of non-Gaussian noise are some of the desirable features of a SS unit in a CR.

In many practical applications, CR devices can exploit known structure in the primary signal. In the IEEE 802.22 CR standard, the primary signal is a wideband signal, but with a strong narrowband pilot component. In other applications, such as military communications, and bluetooth, the primary signal uses a Frequency Hopping (FH) transmission. These applications can significantly benefit from detection schemes that are tailored for detecting the corresponding primary signals. This thesis develops novel detection schemes and rigorous performance analysis of these primary signals in the presence of fading. For example, in the case of wideband primary signals with a strong narrowband pilot, this thesis answers the further question of whether to use the entire wideband for signal detection, or whether to filter out the pilot signal and

use narrowband signal detection. The question is interesting because the fading characteristics of wideband and narrowband signals are fundamentally different. Due to this, it is not obvious which detection scheme will perform better in practical fading environments.

At another end of the gamut of SS algorithms, when the CR has no knowledge of the structure or statistics of the primary signal, and when the noise variance is known, Energy Detection (ED) is known to be optimal for SS. However, the performance of the ED is not robust to uncertainties in the noise statistics or under different possible primary signal models. In this case, a natural way to pose the SS problem is as a Goodness-of-Fit Test (GoFT), where the idea is to either accept or reject the noise-only hypothesis. This thesis designs and studies the performance of GoFTs when the noise statistics can even be non-Gaussian, and with heavy tails. Also, the techniques are extended to the cooperative SS scenario where multiple CR nodes record observations using multiple antennas and perform decentralized detection.

In this thesis, we study all the issues listed above by considering both single and multiple CR nodes, and evaluating their performance in terms of (a) probability of detection error, (b) sensing-throughput tradeoff, and (c) probability of rejecting the null-hypothesis. We propose various SS strategies, compare their performance against existing techniques, and discuss their relative advantages and performance tradeoffs. The main contributions of this thesis are as follows:

- The question of whether to use pilot-based narrowband sensing or wideband sensing is answered using a novel, analytically tractable metric proposed in this thesis called the *error exponent with a confidence level*.

- Under a Bayesian framework, obtaining closed form expressions for the optimal detection threshold is difficult. Near-optimal detection thresholds are obtained for most of the commonly encountered fading models.

- For an FH primary, using the Fast Fourier Transform (FFT) Averaging Ratio (FAR) algorithm, the sensing-throughput tradeoff are derived in closed form.

- A GoFT technique based on the statistics of the number of zero-crossings in the observations is proposed, which is robust to uncertainties in the noise statistics,

and outperforms existing GoFT-based SS techniques.

- A multi-dimensional GoFT based on stochastic distances is studied, which provides better performance compared to some of the existing techniques. A special case, i.e., a test based on the Kullback-Leibler distance is shown to be robust to some uncertainties in the noise process.

All of the theoretical results are validated using Monte Carlo simulations. In the case of FH-SS, an implementation of SS using the FAR algorithm on a commercially off-the-shelf platform is presented, and the performance recorded using the hardware is shown to corroborate well with the theoretical and simulation-based results. The results in this thesis thus provide a bouquet of SS algorithms that could be useful under different CR-SS scenarios.

Glossary

ADC	:	Analog-to-Digital Converter
ADD	:	Anderson-Darling statistic based Detector
AR	:	Auto Regressive
ARMA	:	Auto Regressive Moving Average
AWGN	:	Additive White Gaussian Noise
B	:	Bhattacharyya Distance
BD	:	Blind Detector
BPF	:	Band Pass Filter
CCDF	:	Complementary CDF
CDF	:	Cumulative Distribution Function
CFAR	:	Constant False Alarm Rate
CLT	:	Central Limit Theorem
CR	:	Cognitive Radio
DAC	:	Digital-to-Analog Converter
DCM	:	Data Conversion Module
DP	:	Development Platform
DPM	:	Digital Processing Module
DSP	:	Digital Signal Processing
DTV	:	Digital Tele-Vision
ED	:	Energy Detection
EECL	:	Error Exponent with a Confidence Level
ER	:	Eigenvalue Ratio based Test
FAR	:	FFT Averaging Ratio
FC	:	Fusion Center
FDMA	:	Frequency Division Multiple Access
FFT	:	Fast Fourier Transform
FH	:	Frequency-Hopping
FPGA	:	Field-Programmable Gate Array

GoFT	:	Goodness-of-Fit Test
GUI	:	Graphical User Interface
H	:	Hellinger Distance
HOC	:	Higher Order Crossings
ID	:	Interpoint Distance
IEEE	:	Institute of Electrical and Electronics Engineers
IF	:	Intermediate Frequency
i.i.d.	:	Independent and Identically Distributed
KL	:	Kullback-Leibler Distance
LC	:	Level-Crossings
LR	:	Likelihood Ratio
MA	:	Moving Average
MAC	:	Medium Access Control Layer
MBDK	:	Model Based Design Kit
MDGoFT	:	Multi-Dimensional GoFT
NB	:	Narrow-Band
NCO	:	Numerically Controlled Oscillator
NP	:	Neyman-Pearson
NPU	:	Noise Parameter Uncertainty
NMU	:	Noise Model Uncertainty
NVU	:	Noise Variance Uncertainty
OSD	:	Ordered Statistics based Detector
PDF	:	Probability Density Function
PHY	:	Physical Layer
P-IV	:	Pearson type IV distribution
PN	:	Psuedo-random
Ψ_1SD	:	Ψ_wSD with uniform and equal weights
Ψ_eSD	:	Ψ_wSD with exponential weights
Ψ_wSD	:	Ψ^2 Statistic based Detector
PU	:	Primary User
QoS	:	Quality-of-Service
QPSK	:	Quadrature Phase Shift Keying
RF	:	Radio-Frequency
RFM	:	RF Module

$\mathcal{S}\alpha\mathcal{S}$	:	Symmetric-α Stable Distribution
SFF	:	Small Form Factor
SDR	:	Software Defined Radio
SNR	:	Signal-to-Noise Ratio
SS	:	Spectrum Sensing
ST	:	Sphericity Test
STFT	:	Short Time Fourier Transform
TDMA	:	Time Division Multiple Access
VHP	:	Virtual Hopping Period
VPBE	:	Video Processing Back End
VPFE	:	Video Processing Front End
VPSS	:	Video Processing Sub-System
WB	:	Wide-Band
WBX	:	Wide Bandwidth Transceiver
WRAN	:	Wireless Regional Area Network
WZCD	:	Weighted Zero-Crossings Detector
ZC	:	Zero-Crossings

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The term *Cognitive Radio* (CR) was coined by Joseph Mitola III in a series of papers in 1999 ([1–3]). In his Ph.D. thesis [4], Mitola explained the idea of CR from PHY, MAC and application layers' perspective. A CR transceiver is envisioned to possess the ability to adapt to its radio-environment, tuning its communication parameters, and matching the available resources to the network demand. Over the past decade, CR has received a significant research attention in signal processing for communications ([5–15]), sensor networks ([16–19]), information theory ([20–25]), game theory ([26,27]), machine learning ([28, 29]), and many other fields. Excellent overview articles on CR can be found in ([30–33]).

In communications engineering, CR is a promising solution to the ever-increasing demand for RF spectrum, and to the apparent scarcity of the bandwidth caused by fixed frequency allocations [34]. The idea of CR has been formalized for access over the digital TV bands in the IEEE 802.22 standard for the secondary communication in a wireless regional area network [35].

In its most commonly envisioned mode of operation, a CR continuously monitors the

spectrum usage activity of a primary user (or the licensed user) in a given frequency band, and opportunistically utilizes it, whenever it is found to be unoccupied. Therefore, reliable and fast detection of the presence/absence of a primary user is the first, key step in enabling CR. This problem is referred to as *spectrum sensing*, and is discussed in detail in the next section.

1.1 Spectrum Sensing

Spectrum Sensing (SS), or the detection of the presence or absence of a primary signal in a given frequency band of interest, is a well-studied topic in cognitive radios. At its core, spectrum sensing is a binary hypothesis testing problem between the *noise-only* (or the *signal-absent* or the *null*) hypothesis (denoted by $\mathcal{H}_0$) and the *signal-present* (or the *alternative*) hypothesis (denoted by $\mathcal{H}_1$) [36]. If Y_i, n_i, s_i, and h_i denote the received observation, noise sample, primary signal sample and the frequency-flat channel between the primary transmitter and a CR node at a time instant i, respectively, then the SS problem can be modeled as testing $\mathcal{H}_0$ versus $\mathcal{H}_1$, where

$$
\begin{aligned}
\mathcal{H}_0 \quad &: \quad Y_i = n_i, \\
\mathcal{H}_1 \quad &: \quad Y_i = h_i s_i + n_i, \quad i = 1, 2, \cdots, M.
\end{aligned}
\tag{1.1}
$$

In the above, M is the number of observations used for detection. In such problems, a test-statistic (denoted by $T(\cdot)$) calculated as a function of the recorded observations is compared with a suitably chosen threshold (denoted by τ), and a decision is made in

favor of one of the two hypotheses. Mathematically, the detector is represented as

$$T(Y_1, \cdots, Y_M) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau. \tag{1.2}$$

The key design choices that need to be made in order to solve the hypothesis testing problem are a) how to choose the test statistic, and b) how to set the detection threshold. These choices depend on a variety of factors such as the performance metric, available knowledge about the primary signal, computational complexity constraints, and whether the detection is based on observations at a single sensor, or whether multiple nodes collaboratively sense for the presence or absence of the primary signal. In particular, multi-sensor based detection or *decentralized detection* [37] offers resilience against the so-called *hidden node problem* ([5], [10] [38], [30]). In the next section, we discuss some of the issues underlying the aforementioned design choices in greater detail.

1.2 Scenarios for Spectrum Sensing

As mentioned earlier, several scenarios for SS have been investigated in the CR literature. These depend on the problem framework, the number or type of observations at hand, the possibility of cooperation among different CR nodes, and the knowledge about the primary signal. Some of the approaches that have been explored in the literature are pictorially shown in Fig.1.1.
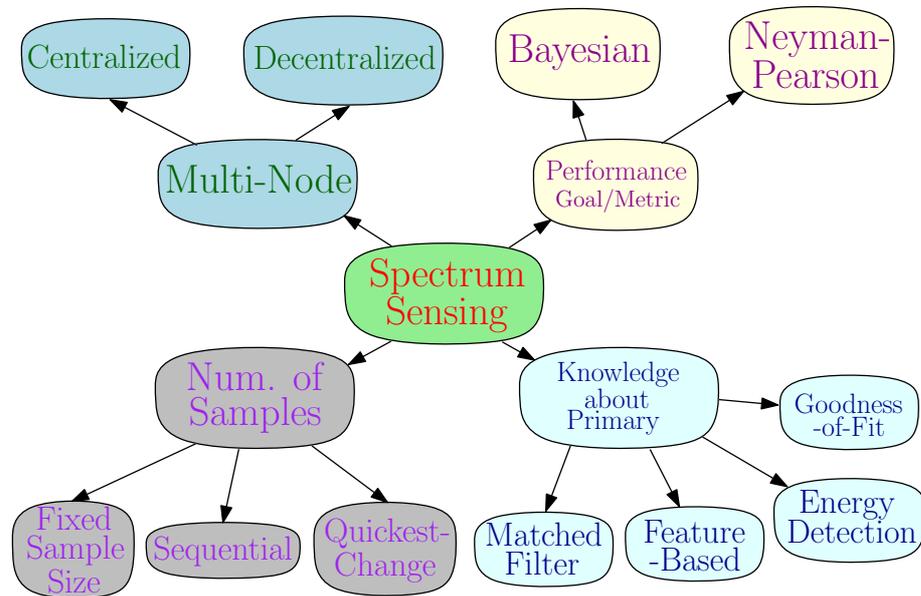
Figure 1.1: Different Scenarios for Spectrum Sensing.

1.2.1 Available Knowledge About the Primary Signal

1. *Matched-Filter Detection*: When the primary signal, e.g., packet headers, training signals, etc., are known at the CR node, matched-filter based detection is a computationally efficient, high-performing detector. Matched filtering maximizes the Signal-to-Noise Ratio (SNR) at the output of the filter, in turn improving signal detection. However, a limitation of this approach is that it requires the CR node to know the primary signal, and have accurate timing and carrier-frequency synchronization. Another disadvantage with this approach is that in co-existence of CR with primary users following different standards, or signaling schemes, the CR node needs to have dedicated receivers for each type of primary. This increases the complexity in the secondary system.

2. *Feature based Detection*: In this approach, a particular *feature* of the primary signal is utilized for increasing the accuracy of signal detection. For instance, the *Cyclostationarity Based Detection* (CBD), ([5], [39] [40]) offers benefits such as the Constant False Alarm Rate (CFAR) property even with inaccurate knowledge of the noise variance [41]. Since the modulated signals are coupled with sinusoidal carriers, they exhibit a natural, inherent periodicity. The CBD takes advantage of this structure, and offers good performance even at very low SNRs ([5], [42]).

3. *Energy Detection*: Energy Detector (ED) is a non-coherent detector which uses the average energy in the observations as the decision statistic. ED is very simple to construct and implement. The threshold chosen for ED is dependent on the noise power. This makes the performance of the ED sensitive to uncertainty in the noise variance, especially at low SNRs ([5], [38]).[1] Another drawback is that the ED does not have the ability to differentiate between the signal, noise and interference. The ED does not work well for spread spectrum signals, where the SNR is very low. Despite these disadvantages, ED has received tremendous attention in spectrum sensing due to its simplicity and ease of implementation ([5, 10, 13, 38, 43]). Additionally, the ED is known to be optimal when the primary signal is unknown but i.i.d. and the noise-only samples are i.i.d., with known distributions [38].

---

[1]This is the SNR wall problem, where, due to the noise variance uncertainty, reliable detection is not possible when the SNR is below a certain threshold, even if the number of samples used for detection is made arbitrarily large.

1.2.2 Signal Acquisition Scenarios

Another way to view the SS problems is in terms of the rates at which the samples are acquired and processed. When the sampling rate is significantly faster than the rate at which each sample can be processed, or when the decision can only be made using a block of samples, one employs *fixed-sample size detection*. When each sample can be processed before the arrival of the next sample, it is pertinent to consider *sequential detection*. Here, each time a new sample arrives, a decision statistic is computed, based on the samples collected till that time. Based on the statistic, the detector either stops and declares in the favor of one of the two hypotheses, or decides to continue taking observations [44]. Thus, the detector consists of both a stopping criterion and a detection rule. Generally speaking, at a given performance level (e.g., as measured through the probability of error of the detector), sequential detectors result in a lower average detection delay compared to the fixed sample size detectors, albeit with higher complexity [44].

1.2.3 Performance Criteria and Problem Formulation

The most popular approach for spectrum sensing in the literature is to use the *Neyman-Pearson* (NP) formulation, where the goal is to maximize the probability of correctly detecting the primary signal when it is indeed present, subject to a constraint on the false alarm probability, i.e., the probability of incorrectly declaring the primary to be present when it is actually absent. It is long established that the *Likelihood Ratio* (LR) is the optimal test statistic for any detection problem in the NP setup [36].

Alternatively, in a *Bayesian* approach, the effect of the prior probabilities are taken into

account and the detection threshold is chosen to minimize a convex combination of the false-alarm and signal detection probabilities.

When no knowledge about the primary and/or channel statistics is available, the class of *Goodness-of-Fit Tests* (GoFT) are the ideal choice for SS, where the goal is to either accept or reject the noise-only hypothesis, based on a test statistic constructed based only on the knowledge of the noise statistics.

1.2.4 Multi-Sensor Detection

Typically, a CR network consists of multiple CR nodes. These nodes can collaboratively detect the presence or absence of the primary, leading to greater detection accuracy or a lower time-to-detect at a given performance target. Multi-sensor detection offers the additional benefits of resilience to fading, the hidden node problem, etc [30]. In this scenario, multiple nodes record observations and share either a decision statistic, or their local decision with a central node, also known as the Fusion Center (FC) where an overall decision on the presence or absence of the primary signal is made. In a centralized scheme, each sensor shares its observations (or a sufficient statistic, if known) with the FC. In a decentralized scheme [45], the sensors make individual one-bit decisions on the presence or absence of the primary, and share their local decisions with the FC over a low-rate, dedicated channel. The FC combines all the individual decisions to arrive at the overall decision. For a given number of sensors, the centralized scheme outperforms decentralized scheme, but requires a high-rate communication overhead. In most cases, the decentralized scheme is preferred, given its simplicity and ease of implementation.

1.3 Challenges in Spectrum Sensing

1.3.1 Effect of Fading

One of the key aspects of wireless communication is the phenomenon of fading. The CR communication should consider the fading of the channel between the primary transmitter and the CR receiver for spectrum sensing. In the literature, the performance of ED under an NP framework for various fading models has been characterized ([46], [43]). On the other hand, Bayesian SS under fading has caught very little attention. The U.S. Federal Communications Commission (FCC), in its landmark study in 2002, showed that the licensed spectrum remains mostly unoccupied across space and time [34]. Specifically, across time, the probability of the spectrum being unoccupied was found to be as high as 70%. Bayesian SS accounts for this available prior information about the primary usage statistics to improve the average detection performance.

To illustrate the tradeoffs involved in considering the effect of fading on the detection performance, consider the following example. The IEEE 802.22 standard allows for opportunistic access in the Digital TV frequency band. In this case, the primary uses a wideband signal, occupying a bandwidth of 6MHz. There are two options for detecting the presence of such a primary signal. First, one could use a narrowband filter to capture the strong pilot tone present at 2.69MHz in the primary signal, and detect based on the pilot energy. This has the advantage of filtering out the wideband noise; but the detector has to contend with a narrowband signal undergoing small scale fading (for e.g., Rayleigh fading). Alternatively, one could use the energy in the entire wideband signal for detection, which averages out the small scale fading [47], but the detector has

to work against the slowly-varying large scale fading (modeled by a lognormal distribution). Since the statistical behavior of the Rayleigh fading and lognormal shadowing are different, the detection performance under the two models can be quite different. An important question that one can ask is as follows. Under the Bayesian framework, should one employ wideband (WB) sensing or narrowband (NB) sensing? An analysis of the probability of error does not give closed form expressions for the detection threshold, and the actual probability of error, even for the simplest case of Rayleigh fading. Further, analysis of the information theoretic quantities such as the error exponents that capture the large sample behavior of the detectors ([48], [49]), show that the exponents achieved on probability of error is zero for any practical fading model. Therefore, this question needs to be addressed with a different performance metric, one that captures the statistics of the fading distribution, and is yet amenable to analytical characterization.

Another interesting aspect related to signal fading is as follows. Most of the existing literature focuses on Rayleigh distributed fading, partly because it is indeed a common fading distribution encountered in nature, but mainly for analytical tractability. In practice, however, the fading could follow a variety of well accepted, although mathematically more complex distributions. Clearly, detectors designed under the Rayleigh fading assumption can be very suboptimal under other fading distributions. Hence, it is pertinent to ask whether one can obtain optimal or near-optimal detection thresholds for a variety of practical fading models such as Rayleigh, lognormal, Nakagami-m, Weibull and Suzuki. While this question has been answered under an NP approach [43], answering it under a Bayesian setup is significantly more challenging, as

the both the optimal threshold and the resulting performance depend on the fading statistics. In the NP framework, the threshold depends on the noise statistics, and the fading distribution only affects the the probability of detection.

1.3.2 Frequency-Hopping Primary Signals

Apart from wideband primary signals, another class of signals where spectrum sensing is challenging is when the primary employs frequency-hopping communication. Given the short hop-duration of the primary, there exists a tradeoff between the sensing duration, and the achieved throughput, which is known as the sensing-throughput tradeoff [50]. Increasing the sensing duration increases the sensing accuracy, but decreases the time remaining within the hop duration for data transmission. Therefore, determining the detection threshold and the sensing duration is a two parameter optimization problem. Additionally, synchronization of the secondary system with the hopping epochs of primary is required for effective sensing and maximizing the secondary throughput.

1.3.3 Robustness to Noise Models

Since a CR is envisioned to operate in various fading and interference environments [1], the fading distribution and the primary signal structure can be fairly general. Moreover, the noise and interference distributions can be only partially known. It is important, therefore, to design detectors that are robust to these model uncertainties. In such cases, the class of Goodness-of-Fit Tests (GoFT) is a natural choice for SS [51]. The detection threshold for a GoFT depends on the signal-absent hypothesis, and hence one

requires at least partial knowledge about the noise distribution. Contrary to the well-used assumption in the GoFT for SS in CR literature ([52–54]), the noise process in most communication systems is not i.i.d. Gaussian [55]. Presence of both controlled and impulsive noise components, with possibly unknown parameters, makes the design of a GoFT a challenging problem. Moreover, uncertainty in the knowledge of the noise distribution (for e.g., uncertainty in whether a controllable noise component is present or not, or in its temporal correlation) makes the design of a robust GoFT even more difficult.

Extending to the scenario where multiple CR nodes with multiple antennas each carry out SS, no GoFTs have been considered in the literature so far. Therefore, the design of a Multi-dimensional GoFT (MDGoFT) is also an interesting challenge.

1.4 Contributions of the Thesis

As highlighted in Fig. 1.2, in this thesis, we design and analytically study spectrum sensing algorithms for cognitive radios under the following scenarios:

1. When the primary signal, channel and noise statistics are known, for e.g., in the DTV signal detection problem that arises in the IEEE 802.22 standard. In particular, we consider the detection of wideband primary signals with a strong pilot tone (Chapters 2 and 3).

2. When the primary signal follows frequency-hopping communication. In such scenarios, the key challenge is to reliably sense for the presence of the primary signal within a fraction of the hop duration (Chapter 4).
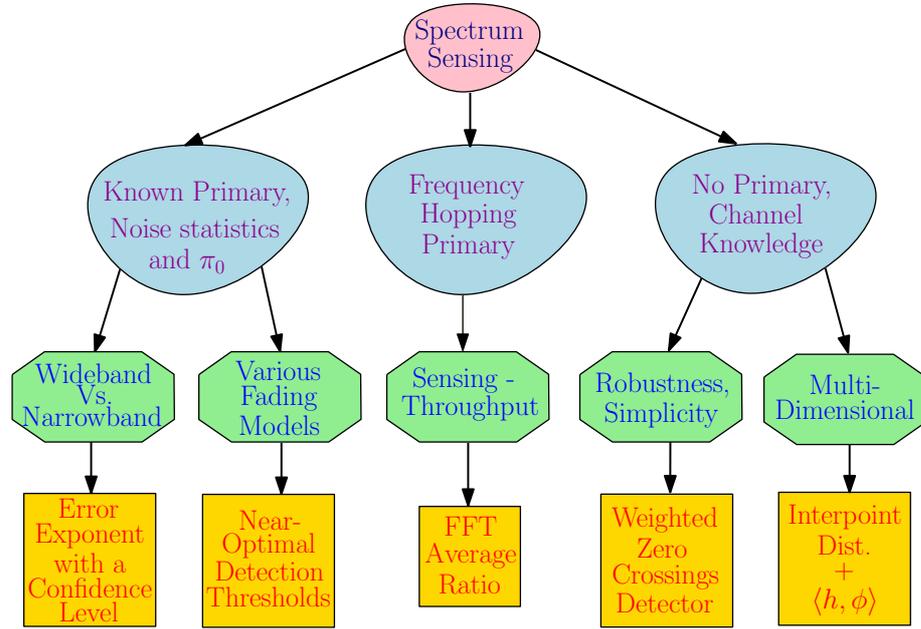
Figure 1.2: Contributions of the Thesis.

3. When no knowledge on the primary signal, and channel statistics is assumed. In such scenarios, the class of Goodness-of-Fit Tests (GoFT), which either accepts or rejects the noise-only hypothesis is an ideal choice (Chapters 5 and 6).

We now describe our specific contributions in each of these scenarios mentioned above in detail.

In Chap. 2, the impact of channel fading on the energy-based detection of signals is studied in detail. A novel concept of *Error Exponent with a Confidence Level* (EECL) is introduced, which captures the largest exponent on the probability of error that can be achieved when a small fraction $1 - q$ (with $0 < q \leq 1$) of the worst channel states are discounted. The EECL at an individual sensor is derived for a large class of fading distributions, and it is shown that as q approaches 1, the EECL approaches 0. The EECL for decentralized detection with N sensors and when the FC uses the OR (1 out of N) rule

is derived under the Rayleigh fading and lognormal shadowing channels. Closed-form lower bounds on the EECL are also derived, for both Rayleigh fading and lognormal shadowing channels. The bounds are easy to compute and become increasingly accurate as q approaches 1. The theoretical development is used to successfully address the question of NB versus WB sensing alluded to earlier (See Sec. 1.3.1), and a rigorous analysis is presented. Specifically, if the ratio of normalized NB and WB powers exceeds a threshold, NB sensing is better than WB sensing in terms of the EECL, and vice versa. The contents of this chapter has been published in part in [56].

Chapter 3 derives near-optimal thresholds for energy detection of signals under the commonly used fading models, namely Rayleigh, lognormal, Nakagami-m, Weibull and Suzuki distributions, for spectrum sensing under a Bayesian framework. For the Rayleigh fading case, the trade-off between the number of observations and the primary power for given error performance is found. Extending the analysis to the decentralized case, the error exponents at the Fusion Center (FC) as the number of sensors grows large is derived. For the decentralized detection with Rayleigh fading, the diversity gain on the overall probability of error is shown through simulations. The contents of this chapter have been published in part in [57] and [58].

In Chap. 4, we apply an existing technique called the FFT Average Ratio (FAR) algorithm for primary signal detection under a multiuser frequency-hopping primary scenario, and derive closed-form expressions for the probabilities of false alarm and detection as a function of the detection threshold, number of averaging frames, and the estimated SNRs of the primary signal in the occupied bands. We define a utility metric to quantify the throughput of the CR, and analytically obtain the CR sensing

duration that maximizes the throughput while satisfying a constraint on the maximum allowable interference to the PUs. We implement the FAR Algorithm on a Lyrtech Small Form Factor Software Defined Radio Development Platform (Lyrtech SFF SDR DP), and validate the implementation by comparing its performance with that obtained from the analysis and simulations. The contents of this chapter have been published in [59].

In Chap. 5, we formulate the problem of spectrum sensing as a Goodness-of-Fit test, and a detector based on the number of zero-crossings in the observations is proposed. Given a target false alarm probability, near-optimal detection thresholds are obtained for uniform and exponential weights. The proposed detector is shown to be robust to two types of noise uncertainties encountered in practice, namely, noise parameter uncertainty and the noise model uncertainty. In a detailed simulation study, the performance of the proposed detectors is compared with existing techniques under various primary signal models operating in different noise and fading environments. The contents of this chapter have been published in [60].

Finally, in Chap. 6, we propose two GoFTs in a multi-dimensional setup where multiple observations recorded in a multi-sensor, multi-antenna environment are used by the test. The proposed GoFTs are based on the properties of stochastic distances. The advantages of the proposed detectors are highlighted, and the performance benefits relative to existing techniques are illustrated through simulations. The contents of this chapter have been published in [61].

List of Publications From This Thesis

Journal Papers

1. S. Gurugopinath, C. R. Murthy and V. Sharma, "Error exponent analysis of energy-based bayesian decentralized spectrum sensing under fading," submitted to IEEE Transactions on Vehicular Technology.

2. S. Gurugopinath, R. Akula, C. R. Murthy, R. Prasanna and B. Amruthur, "Design and Implementation of Spectrum Sensing for Cognitive Radios with a Frequency-Hopping Primary System," submitted to IEEE Transactions on Instrumentation and Measurement.

3. S. Gurugopinath, C. R. Murthy and C. S. Seelamantula, "Zero-crossings based spectrum sensing under noise uncertainties," journal version under preparation.

Conference Papers

1. Sanjeev G., K. V. K. Chaythanya, and C. R. Murthy, "Bayesian decentralized spectrum sensing in cognitive radio networks," Proc. International Conference on Signal Processing and Communications (SPCOM), Bangalore, India, Jul. 2010.

2. S. Gurugopinath, C. R. Murthy, and V. Sharma, "Error exponent analysis of energy-based bayesian spectrum sensing under fading channels," Proc. IEEE Global Telecommunications Conference (GLOBECOM), Houston, USA, Dec. 2011.

3. S. Gurugopinath, Raghavendra Akula, C. R. Murthy, R. Prasanna and B. Amruthur, "Spectrum sensing with a frequency-hopping primary: from theory to practice," Proc. IEEE International Conference on Communications (ICC), Jun. 2014, Sydney, Australia.

4. S. Gurugopinath, C. R. Murthy and C. S. Seelamantula, "Zero-crossings based spectrum sensing under noise uncertainties," Proc. National Conference on Communications (NCC), Kanpur, India, Feb-Mar. 2014.

5. S. Gurugopinath, "Near-optimal detection thresholds for bayesian spectrum sensing under fading," Proc. International Conference on Signal Processing and Communications (SPCOM), Bangalore, India, Jul. 2014.

6. S. Gurugopinath, "Multi-dimensional goodness-of-fit tests for spectrum sensing based on stochastic distances," Proc. International Conference on Signal Processing and Communications (SPCOM), Bangalore, India, Jul. 2014.

Chapter 2

Error Exponent Analysis of Energy-Based Bayesian Decentralized Spectrum Sensing Under Fading

2.1 Introduction

Spectrum sensing, or the detection of the presence or absence of a primary signal in a given frequency band of interest, is a well-studied topic in recent literature on Cognitive Radios (CR) [1, 4]. Multi-sensor detection, or decentralized detection, is the preferred approach for spectrum sensing, because of its resilience to signal fading, the hidden node problem, etc. [10, 13, 31, 62–65]. In fixed sample-size decentralized detection, individual CR nodes make one-bit decisions about the availability of the spectrum using a given number of samples, and the individual decisions are combined at a Fusion Center (FC) to detect the presence or absence of the primary signal. Energy-based detection, popularly referred to as Energy Detection (ED), is a well known technique for spectrum sensing, wherein the signal energy in the band of interest is measured and compared with a threshold [43, 46, 66, 67]. The primary signal is declared to be present

if the measured energy exceeds the threshold.

The detection probability performance of ED when the channel between the primary transmitter and the secondary node undergoes narrowband Rayleigh fading has been analyzed under the Neyman-Pearson (NP) framework [43, 66, 68]. Although closed-form expressions for the probability of detection have been derived, due to the form of the integrals involved, it is cumbersome to obtain the detection threshold that meets a given minimum detection probability requirement. One way around this is to use an alternative performance metric such as the error exponent [48, 49], which essentially captures the asymptotic behavior of the probability of error performance of a detector as the number of samples used for making decisions gets large.[1] Mathematically, the error exponent is defined as $\lim_{M \to \infty} -\log(P_e)/M$, where M is the number of samples used for detection, and P_e is the corresponding probability of error. One of the early studies on the error exponent performance of decentralized detection was the seminal work of Tsitsiklis [45]. In the Bayesian framework, the exponent on the probability of error of decentralized detection has been analyzed in [69]. The Bayesian error exponent of mismatched likelihood ratio detectors was derived in [70]. The analysis uses the fact that the best achievable exponent in the Bayesian probability of error is the Chernoff information between the probability distribution functions under the two hypotheses. In turn, this implies that the optimal exponents associated with the probability of false alarm and the probability of missed detection must equal each other [48, Chap. 11], [71]. When the primary signal power or the noise variance at the secondary receiver are unknown, a robust and blind detection scheme based on the maximum eigenvalue of

---

[1]The number of samples can be considered to be large, for example, in Digital Television (DTV) signal detection, where the primary network changes its occupancy infrequently.

the sample covariance matrix has been proposed and studied through simulations [72]. In [73] and [74], multi-antenna assisted spectrum sensing is considered under the NP framework.

Decentralized detection for spectrum sensing under the Bayesian framework is considered in [57, 75, 76]. Here, the channel between the primary transmitter and the secondary sensors is assumed to undergo fading, while the channel between the sensors and the FC is assumed to be lossless but finite-rate. However, to the best of our knowledge, prior to this study, error exponents for energy-based decentralized spectrum sensing have not been derived in the literature. There are several advantages in using the error exponent as a performance metric under a Bayesian set-up. First, the optimal error exponent is independent of the specific values of the prior probabilities, provided they are nonzero [48]. Due to this, the optimal error exponent, and detection schemes based on maximizing the error exponent, are naturally robust to uncertainties in estimating the prior probabilities, unlike detectors designed with the goal of minimizing the probability of error. Further, error exponents allow one to contrast the performance of competing detectors over a range of target performance requirements, rather than at a single missed detection probability target. This is useful when choosing between detectors at the design phase of a hardware implementation.

Yet another reason for considering an error exponent analysis of spectrum sensing is related to the statistical properties of the fading experienced by the primary signal. For Narrow-Band (NB) signals, the multipath (Rayleigh) fading effect is dominant, in a non line-of-sight environment. On the other hand, Wide-Band (WB) signals span multiple coherence bandwidths, due to which, the Rayleigh fading component averages out

when the signal energy is accumulated across the wideband, resulting in the lognormal

shadowing as the dominant fading component [47, 77]. As a concrete example, in the

IEEE 802.22 (WRAN) standard, the primary (Digital Television (DTV)) signal is a wide-

band signal, with a strong pilot tone at 2.69 MHz (see Figure 2.1).[2] There are therefore

two options for detection. First, one could use an NB filter to capture just the pilot tone,

and detect based on the pilot energy. This has the advantage of filtering out the WB

noise; but the detector has to contend with a Rayleigh-faded NB signal. Alternatively,

one could use the energy in the entire WB signal for detection, which averages out the

Rayleigh fading [47, 77], but the detector has to work against the lognormal shadowing

and the added impairment due to the AWGN over the WB. Again, due to the complex

form of the integrals involved, direct comparison of the two options using conventional

performance metrics such as the probability of error is difficult. Hence, in this chapter,

we contrast these two options by analyzing the Bayesian error exponent performance

of energy-based detection.

The main contributions of this work are as follows:

- The concept of *Error Exponent with a Confidence Level* (EECL) is introduced, which

  captures the largest exponent on the probability of error that can be achieved if a

  fraction $1 - q$ (with $0 < q \leq 1$) of the worst channel states are discounted. The

  EECL at an individual sensor is derived for a large class of fading distributions,

  and it is shown that as q approaches 1, the EECL approaches 0.

- The EECL for decentralized detection with N sensors and when the FC uses the

---

[2]Note that, at the time of writing this chapter, in the U.S., spectrum sensing is made optional in the IEEE 802.22 standard. However, in many countries other than the U.S. and European countries, reliable databases may not be available [78]. In these cases, spectrum sensing is essential.
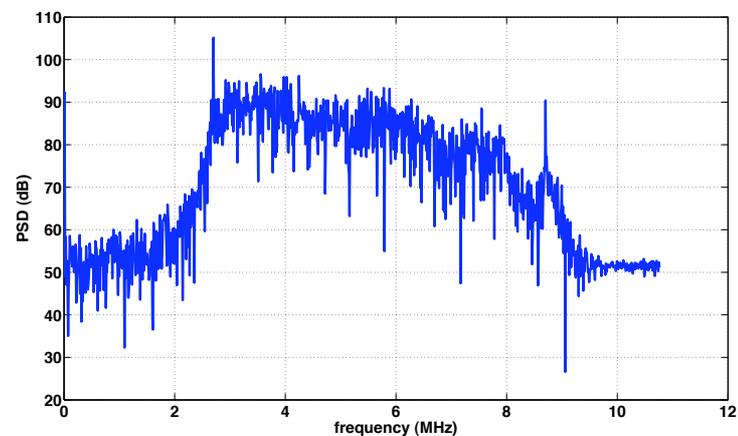
Figure 2.1: One sided PSD of IEEE 802.22 DTV wideband signal.

OR (1 out of N) rule is derived under the Rayleigh fading and lognormal shadowing channels.

- Closed-form lower bounds on the EECL are also derived, for both Rayleigh fading and lognormal shadowing channels. The bounds are easy to compute and become increasingly accurate as q approaches 1.

- The theoretical development is used to successfully address the question of NB versus WB sensing, and a rigorous analysis is presented. Specifically, if the ratio of normalized NB and WB powers exceeds a threshold, then NB sensing is better than WB sensing in terms of the EECL, and vice versa.

We show, through Monte Carlo simulations, that our proposed detector outperforms existing detectors in terms of the probability of error, when a small fraction of the worst channel states are discounted. The improved sensing performance can lead to better CR throughput and/or better primary user protection in CR implementations. Note that, joint design of the sensing scheme and the medium access protocol to maximize

the secondary throughput [50, 79], while an important topic of study, requires one to assume a specific model for the temporal behavior of the primary occupancy. Such a study is beyond the scope of this chapter.

The rest of this chapter is organized as follows. The problem set-up and the basics of error exponents are presented in Sec. 2.2. The EECL at a single node is introduced and analyzed in Sec. 2.3. Distributed detection is considered in Sec. 2.4, where the EECL at the FC with the OR rule is derived. The comparison between WB and NB spectrum sensing in terms of the EECL is discussed in Sec. 2.5. Simulation results are presented in Sec. 2.6, and Sec. 2.7 concludes the chapter. Proofs of the various theorems and corollaries are presented in the Appendix.

2.2 System Model

We consider a decentralized detection set-up where N sensors use the average energy measured from M independent observations each as the test statistic for making their individual decisions between the signal absent (denoted $\mathcal{H}_0$) and signal present (denoted $\mathcal{H}_1$) hypotheses [10, 13, 31, 63, 64, 73, 80]. Such an energy-based test is known to be optimal when no knowledge about the structure of the primary signal is available at the CR nodes [46]. When M is large, using the Central Limit Theorem (CLT), the test statistic can be well-approximated as being Gaussian distributed, resulting in the following hypothesis test at each sensor [38, 66, 81]:

$$
\begin{aligned}
\mathcal{H}_0 &: \quad V_y \sim \mathcal{N}\left(0, \frac{1}{M}\right) \\
\mathcal{H}_1 &: \quad V_y \sim \mathbb{E}_h \mathcal{N}\left(|h|^2 P, \frac{1}{M}\right),
\end{aligned}
\tag{2.1}
$$

where $V_y \triangleq \frac{1}{M} \sum_{k=1}^{M} |Y_k|^2 - 1$ is the test statistic, and Y_k is the kth observation at the sensor. Also, $\mathcal{N}(\mu, \sigma^2)$ represents a normal distribution with mean μ and variance σ^2. In writing the above, without loss of generality, we normalize the receiver noise variance to unity. The *average* received power of the primary signal, P, is also assumed to be known at the nodes. The noise variance and average received signal power can be estimated, for example, using a calibration phase, when the primary signal is known to be absent and present, respectively. Furthermore, for simplicity, we assume that the CR nodes are sufficiently close to each other that P is the same at all nodes [76]. This assumption is valid when the CR nodes involved in cooperative spectrum sensing are located in proximity with each other, and are relatively far from the primary transmitter. In such a situation, one can assume that the path loss from the primary transmitter to the CR nodes, which is the main contributor to the average received power, is essentially the same for all CR nodes.[3] The expectation $\mathbb{E}_h$ in the above is taken over the distribution over the channel gain, h, which is assumed to be random, unknown, and constant for the M observations. In (2.1), we have omitted the sensor index from V_y for notational convenience, since the observations are assumed to be independent and identically distributed (i.i.d.) conditioned on the true hypothesis.

In the literature, various statistical models have been proposed for the channel h, depending on the signal bandwidth and propagation environment. As mentioned earlier, when the primary signal is NB, the Rayleigh fading component typically dominates the

---

[3]In practice, the average received power may not be the same at the sensors. However, one could *design* the detectors assuming a certain minimum value of the average power at all sensors. If a particular sensor sees an average power larger than P, its detection probability will only be better than the designed value. Hence, this represents a conservative design approach in terms of protecting the primary users.

lognormal shadowing components, and hence $|h|^2$ can be modeled as exponentially distributed [82, 83]. When the primary signal is a WB signal, it spans multiple coherence bandwidths, due to which, the Rayleigh fading components average out, resulting in h being a lognormal shadowing random variable [47, 77]. Other models include the Nakagami-m distribution, the Weibull distribution, and the Suzuki distribution [77]. In this work, we focus on the two most commonly used models, namely, the Rayleigh and the lognormal shadowing distributions, for the NB and WB fading cases, respectively. However, our main results can be extended to handle any of the fading models mentioned above.

We assume that the sensors transmit their binary decisions to an FC through a finite rate, noiseless, delay-free CR control channel, as in [75, 76]. This simplifies the analysis, and the corresponding EECL represents an upper bound on the error exponent achievable in the general case. It is valid when the CRs use a low-rate dedicated control channel to forward their decisions to the FC. The FC combines the individual decisions using the K out of N fusion rule to detect the presence or absence of the primary signal. It is known that, when the individual sensor decisions are i.i.d. conditioned on the true hypothesis, the K out of N fusion rule is optimal in terms of probability of error [71, 84]. In particular, we will focus on the 1 out of N fusion rule, i.e., the OR fusion rule, in the sequel. We will show that the OR fusion rule has a certain optimality property in terms of the error exponents. In the next section, we present the main results on the EECL at an individual sensor. We extend it to multiple-node decentralized detection in Sec. 2.4.

2.3 Detection at the Sensors

We start by considering the single-sensor hypothesis testing problem in (2.1). The conventional error exponent is defined as $\lim_{M \to \infty} \frac{-\log p_e}{M}$, where p_e denotes the probability of error at the sensor, and is given by $\pi_0 p_f + (1 - \pi_0)p_m$, with π_0, p_f and p_m denoting the prior probability of hypothesis $\mathcal{H}_0$, the false alarm probability, and the missed detection probability, respectively. Below, we show that the exponent on the probability of missed detection is zero, provided the pdf of the channel gain is continuous and satisfies $\mathcal{P}(|h|^2 \leq |h_0|^2) > 0$ for arbitrarily small $|h_0| > 0$, which is satisfied by all of the distributions mentioned above. Therefore, the conventional error exponent analysis is not useful for answering the question of NB vs. WB spectrum sensing. Essentially, this happens because the deep fade instantiations, where the hypotheses are indistinguishable, dominate the average detection performance; and all detection techniques perform equally poorly in this scenario. Hence, in this chapter, we propose the following novel performance metric to evaluate and compare the performance of NB and WB spectrum sensing approaches. The EECL at a single sensor is defined as given below. We extend the definition to the N sensor case in the next section.

Definition 1. *Let S_q denote a set of channel instantiations such that $\mathcal{P}(|h|^2 \in S_q) = q$. The error exponent with a confidence level q, denoted EECL(q), is the maximum error exponent achievable conditioned on $|h|^2 \in S_q$, where the maximization is over all possible choices of S_q.*

The above definition of the error exponent, discounting the deep fade instantiations,

has practical relevance. For example, in the IEEE 802.22 standard, the primary sig-
nal detection is required to achieve a probability of miss ≤ 0.1 with a sensing dura-
tion ≤ 2 seconds, whenever the primary signal power at the secondary node exceeds
-116 dBm [35]. Thus, typically, the primary network would require the CR to guaran-
tee a given probability of missed detection target whenever the signal power level at
the CR exceeds a given threshold. Now, in the single sensor case, it is immediate to see
that, among all possible choices for S_q, the highest error exponent is achieved by letting
$S_q = \{|h|^2 : |h|^2 \geq |h_0|^2\}$, where the threshold $|h_0|^2$ depends on the minimum power
level at which the primary signal detection performance needs to be guaranteed by the
CR.

An alternative interpretation of the operational significance of the EECL is as follows.
Consider a given missed detection probability constraint, β, imposed by the primary
network. Pick $0 < \alpha < \beta$. For a fraction α of the channel states, the missed detection
probability can be upper bounded by unity. For the remaining fraction $1 - \alpha$ of the
channel states, we set the detection threshold such that the missed detection probability
is at most $\beta - \alpha$. Then, the overall missed detection probability is upper bounded by
β. As will be shown in the sequel, discounting a fraction α of the channel states allows
one to achieve a positive exponent on the probability of error. Hence, if one detection
scheme has a larger EECL than another, the detector with the larger EECL will have a
significantly smaller false alarm rate, and, consequently, better secondary throughput,
for the given missed detection probability constraint of β, as the number of observations
gets large. Hence, the EECL can be used as a metric for the design, and performance
comparison, of different detection schemes (e.g., NB vs. WB spectrum sensing).

The main result of this section is stated as the theorem below. It is valid as long as the distribution of the gain of the channel from the primary transmitter to the sensors is continuous and nonzero for infinitesimally small arguments. To obtain the result, we use the fact that, under a Bayesian setup, the optimal exponent on the probability of error is achieved when the exponents on the probability of false alarm and the probability of missed detection are equal [48, Chap. 11], [71].

Theorem 1. *Let $\alpha \triangleq |h|^2$. The Bayesian hypothesis test defined in (2.1) achieves an EECL(q) of $\frac{(\alpha_0 P)^2}{8}$, where α_0 satisfies $\mathcal{P}(\alpha \geq \alpha_0) = q$. Further, the optimal detection threshold on V_y asymptotically approaches $\frac{\alpha_0 P}{2}$ as M gets large.*

Proof. See Appendix A.1. □

Under Rayleigh fading, $f_\alpha(\alpha) = e^{-\alpha}, \alpha \geq 0$, and hence, $\mathcal{P}(\alpha \geq \alpha_0) = q$ leads to $e^{-\alpha_0} = q$, or $\alpha_0 = -\log(q)$. Under lognormal shadowing, the cumulative distribution function (cdf) of α is given by $F_\alpha(\alpha) = 1 - Q((\log(\alpha) - \mu_s)/\sigma_s)$, where $Q(\cdot)$ is the standard Gaussian tail function, and μ_s and σ_s are the shape and scale parameters of the shadowing distribution, respectively. Hence, $\alpha_0 = \exp(\sigma_s Q^{-1}(q) + \mu_s)$. Also, note that the both the error exponent and the detection threshold are independent of π_0, the prior probability of hypothesis $\mathcal{H}_0$, as expected ([48], Pg. 389). Therefore, the error exponent is unaffected by uncertainties in the knowledge of π_0. Conditioned on $\alpha > \alpha_0$, each individual sensor achieves an error exponent of $(\alpha_0 P)^2/8$ on p_f and p_m. An important corollary to the above theorem is that when $q = 1$, under the commonly used fading models such as the Rayleigh, Rician, lognormal, Nakagami distributions, etc., only $\alpha_0 = 0$ solves $\mathcal{P}(\alpha \geq \alpha_0) = 1$, and hence, the error exponent with $q = 1$ is zero. Thus, the

conventional error exponent cannot be used to choose between NB and WB sensing, as both result in a zero error exponent.

Another useful aspect of the above theorem is the determination of the asymptotically optimal detection threshold at an individual sensor, $\frac{\alpha_0 P}{2}$. Due to its asymptotic optimality, the local decision rule of comparing V_y to $\frac{\alpha_0 P}{2}$ will be assumed at all sensors in the next section, where the EECL(q) performance at the FC with the OR fusion rule is analyzed.

2.4   Detection at the Fusion Center

In this section, we consider an energy-based local decision rule at the individual sensors with the threshold set as $\frac{\alpha_0 P}{2}$, where α_0 is a parameter to be optimized. We consider the OR rule for combining the individual decisions at the FC. We use the OR fusion rule because it can detect the presence of the primary signal even if just one of the sensors is not in a deep fade, and also because it possesses an optimality property that we will show later in this section. The main result of this section is stated as the following theorem.

Theorem 2. *When the individual sensors employ energy detection and the FC employs the OR fusion rule, the EECL(q), denoted $\epsilon_E^{(N)}$, is given by $\epsilon_E^{(N)} = \frac{(\alpha_0 P)^2}{8}$, where α_0 satisfies*

$$\mathcal{P}\left\{\sum_{j=1}^{N}\left(\frac{2\alpha_j}{\alpha_0}-1\right)^2 \mathbb{I}_{\left\{\frac{2\alpha_j}{\alpha_0}>1\right\}} \leq 1\right\} = 1 - q. \tag{2.2}$$

In the above, α_j is the random channel power gain from the primary transmitter to the jth sensor, and $\mathbb{I}_A$ is the indicator function, taking value 1 when the event A is true and 0 otherwise.

Proof. See Appendix A.2. $\qquad\qquad\square$

Note that the condition in (2.2) to determine α_0 does not require the fading coefficients from the primary transmitter to the individual sensors to be independent or identically distributed. The joint distribution of the fading coefficients has to be used to evaluate the probability in (2.2) to find the value of α_0, and the solution cannot be obtained in closed-form in the general case. When the fading coefficients are i.i.d., simpler equations that determine α_0 for the cases of Rayleigh fading and lognormal shadowing are stated as the corollary below.

Corollary 1. *When the individual sensors employ energy detection with threshold $\frac{\alpha_0 P}{2}$ and the FC employs the OR fusion rule, with i.i.d. Rayleigh fading channels between the primary transmitter and the sensors, the EECL(q) is given by $\frac{(\alpha_0 P)^2}{8}$, where α_0 satisfies*

$$\left[1 - \exp\left(-\frac{\alpha_0}{2}\right)\right]^N + \sum_{l=1}^{N} \binom{N}{l} \left[1 - \exp\left(-\frac{\alpha_0}{2}\right)\right]^{N-l} \exp\left(-\frac{\alpha_0}{2}l\right) \mathcal{P}\left\{\sum_{k=1}^{l} a_k^2 \leq 1\right\} = 1 - q. \tag{2.3}$$

In (2.3), a_k is exponentially distributed with parameter $\frac{2}{\alpha_0}$. The same detector, under i.i.d. lognormal shadowing, with a threshold of $\frac{\ell_0 P}{2}$ at the individual sensors, achieves an EECL(q) of $\frac{(\ell_0 P)^2}{8}$, where ℓ_0 satisfies

$$P_{\mathcal{A}}^N + \sum_{l=1}^{N} \binom{N}{l} P_{\mathcal{A}}^{N-l} P_{\mathcal{A}^c}^l \mathcal{P}\left\{\sum_{k=1}^{l} (e^{y_k} - 1)^2 \leq 1\right\} = 1 - q. \tag{2.4}$$

In (2.4), y_k has a truncated Gaussian distribution with mean $\mu_s + \log\left(\frac{2}{\ell_0}\right)$ and variance σ_s^2, truncated to $[0, \infty)$. Also, $P_{\mathcal{A}^c} \triangleq Q\left(-\frac{\mu_s + \log\left(\frac{2}{\ell_0}\right)}{\sigma_s}\right)$, and $P_{\mathcal{A}} \triangleq 1 - P_{\mathcal{A}^c}$.

Proof. See Appendix A.3. $\qquad\square$

Note that both (2.3) and (2.4) need to be numerically solved to obtain α_0 and ℓ_0, respectively. This, in turn, requires the probability terms in the expressions to be evaluated.
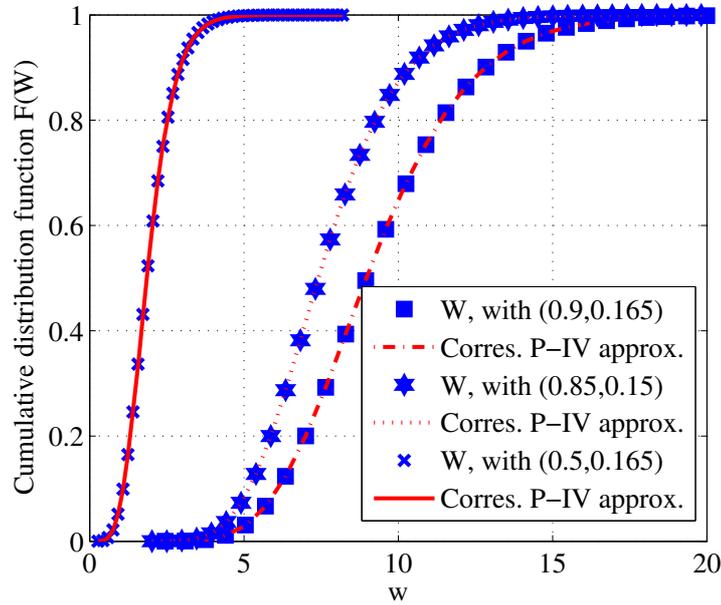
Figure 2.2: CDF of $W \triangleq \sum_{k=1}^{10} \left(e^{y_k} - 1\right)^2$, where y_k are i.i.d. and truncated Gaussian distributed with (mean, variance) = $(0.9, 0.165), (0.85, 0.15)$ and $(0.5, 0.165)$, and $y_k > 0$ with probability 1.

To this end, we use the simple and tight approximation to the cdf of the sum of Weibull random variates (with parameter $c = 2$) derived in [85] to evaluate the probability term in (2.3) in closed-form. Also, we use the Pearson type IV distribution approximation in [86] (see Fig. 2.2) to evaluate the probability term in (2.4), in closed-form. However, we do not present the expressions here for the sake of brevity.

2.4.1 Extension to Unequal Average Received Signal Powers

Recall that, in the above, we had assumed that the CR nodes are sufficiently close to each other that P is the same at all nodes. We now present an extension of our results to handle unequal average received powers at the secondary nodes. First, we consider the case of detection at individual sensors. If the average received power at the ith sensor is

P_i, the EECL(q) of that sensor is $(\alpha_0 P_i)^2/8$, where α_0 is as defined in Theorem 1. Next, for the detection at the fusion center, it is easy to see from the proof of Theorem 2 that the exponent on the false alarm rate is $(\alpha_0 P_{\min})^2/8$, where $P_{\min} \triangleq \min_{1 \le i \le N} P_i$. The rest of the proof of Theorem 2 also follows through, with P replaced by P_j in (A.7) and (A.8) in Appendix A.2. Correspondingly, the condition on α_0 in (2.2) gets modified as:

$$\mathcal{P}\left\{\sum_{j=1}^{N}\left(\frac{P_j}{P_{\min}}\right)^2\left(\frac{2\alpha_j}{\alpha_0}-1\right)^2 \mathbb{I}_{\left\{\frac{2\alpha_j}{\alpha_0}>1\right\}} \le 1\right\} = 1-q. \tag{2.5}$$

Note that, compared to (2.2), we have an extra $(P_j/P_{\min})^2$ factor in the summation, since the average received powers are unequal. Hence, with the OR fusion rule, an error exponent of $\epsilon_E^{(N)} = (\alpha_0 P_{\min})^2/8$ is achievable, where α_0 satisfies the condition in (2.5).

2.4.2 Lower Bounds on the EECL

In this subsection, we derive lower bounds on $\epsilon_E^{(N)}$ for NB and WB sensing. These lower bounds are easy to calculate and become tight as $q \to 1$. The values obtained from these lower bounds can also be used as a good initialization for solving (2.3) and (2.4).

Corollary 2. *In the set-up of Corollary 1, under i.i.d. Rayleigh fading channels, with a threshold of $\frac{\alpha_0^{LB}P}{2}$ at the individual sensors, a lower bound on the EECL(q) is given by $\frac{(\alpha_0^{LB}P)^2}{8}$, where α_0^{LB} satisfies*

$$\alpha_0^{LB} = 2\left(\frac{1-q}{C_N}\right)^{\frac{1}{N}}, \text{ with } C_N \triangleq \sum_{k=0}^{N}\binom{N}{k}\frac{\pi^{\frac{k}{2}}}{\Gamma\left(1+\frac{k}{2}\right)}\frac{1}{2^k}. \tag{2.6}$$

Under i.i.d. lognormal shadowing, with a threshold of $\frac{\ell_0^{LB}P}{2}$ at the individual sensors, a lower

bound on the EECL(q) is given by $\frac{(\ell_0^{LB} P)^2}{8}$, where ℓ_0^{LB} satisfies

$$\ell_0^{LB} = 2 \exp\left(-\sqrt{2\sigma_s^2 \log\left(\frac{1}{\sqrt{2\pi}}\left(\frac{C_N'}{1-q}\right)^{\frac{1}{N}}\right)}\right), \textit{with } C_N' \triangleq \sum_{k=0}^{N} \binom{N}{k} \frac{\pi^{\frac{k}{2}}}{(2\sigma_s)^k \Gamma\left(1+\frac{k}{2}\right)}.$$

$$(2.7)$$

Proof. See Appendix A.4.   □

2.4.3   Optimality of the OR rule

In this subsection, we show that the OR fusion rule satisfies a local optimality property. We show that, when the detection threshold $\frac{\alpha_0 P}{2}$ at the individual sensors is chosen to satisfy (2.2), the OR fusion rule minimizes the probability of error at the FC.

Theorem 3. *For sufficiently large M, the OR fusion rule is probability of error optimal for decentralized detection, when the individual sensors employ energy detection with threshold $\frac{\alpha_0 P}{2}$.*

Proof. See Appendix A.5.   □

2.5   Wideband Vs. Narrowband Spectrum Sensing

As discussed earlier, when the primary signal is a wideband signal containing a strong pilot tone, spectrum sensing can either be carried out by collecting the signal energy over its entire WB or over a small bandwidth around the pilot [47]. In this section, we characterize the relative performance of these two schemes in terms of the EECL. Let P_{NB} and P_{WB} denote the ratios of the energies of the NB and WB signals to their bandwidths, respectively. Typically, P_{NB} is significantly larger than P_{WB}.

2.5.1   NB vs. WB Sensing at Individual Sensors

Let ϵ_{NB} and ϵ_{WB} represent the EECL(q) achieved under NB and WB spectrum sensing at a single sensor, respectively. From Sec. 2.3, given q, setting $\alpha_0 = -\log q$ and $\ell_0 = \exp(\sigma_s Q^{-1}(q) + \mu_s)$ ensures $\mathcal{P}\{\alpha > \alpha_0\} = \mathcal{P}\{\ell > \ell_0\} = q$. Now, NB sensing outperforms WB sensing in terms of EECL, i.e., $\epsilon_{\text{NB}} > \epsilon_{\text{WB}}$, whenever

$$\left(\frac{P_{\text{NB}}}{P_{\text{WB}}}\right)^2 > \left(\frac{\exp(\sigma_s Q^{-1}(q) + \mu_s)}{-\log q}\right)^2. \tag{2.8}$$

2.5.2   NB vs. WB Sensing at the Fusion Center

Similar to the above, let $\epsilon_{\text{NB}}^{(N)}$ and $\epsilon_{\text{WB}}^{(N)}$ represent the EECL(q) achieved by the FC under NB and WB spectrum sensing, respectively. For a given q, $\epsilon_{\text{NB}}^{(N)} > \epsilon_{\text{WB}}^{(N)}$ if $\frac{(\alpha_0 P_{\text{NB}})^2}{8} > \frac{(\ell_0 P_{\text{WB}})^2}{8}$, i.e., when

$$\left(\frac{P_{\text{NB}}}{P_{\text{WB}}}\right)^2 > \left(\frac{\ell_0}{\alpha_0}\right)^2, \tag{2.9}$$

where α_0 and ℓ_0 satisfy (2.3) and (2.4), respectively.

Note that we have used the Rayleigh fading and the lognormal shadowing assumptions only in evaluating the numerical values of α_0 and ℓ_0 above. That is, the above procedure immediately extends to analyzing the EECL(q) of other fading distributions such as Rician, Nakagami-m, Weibull, Suzuki, etc., and the framework can be used to compare NB and WB sensing under various fading conditions.

Also note that, due to the difference in their bandwidths, the sampling rates under NB and WB fading can be different. In the above, we considered the behavior of the sensing performance with respect to M, the number of observations at each sensor. However, the analysis can be easily extended to study the behavior with respect to the

sensing duration, as follows. Let $f_{s,\text{NB}}$ and $f_{s,\text{WB}}$ denote the sampling rates of the NB and WB signals, respectively. Then, a given spectrum sensing duration of T_{ss} leads to a probability of error approximately given by $P_{E,NB} \triangleq \exp(-T_{\text{ss}} f_{s,\text{NB}} \epsilon_{\text{NB}})$ and $P_{E,\text{WB}} \triangleq \exp(-T_{\text{ss}} f_{s,\text{WB}} \epsilon_{\text{WB}})$ in the two cases. Suppose $f_{s,\text{WB}} = B f_{s,\text{NB}}$, where B is the ratio of bandwidths of the WB and NB signals. Thus, NB detection outperforms WB detection in terms of the EECL with the same confidence q and when both detectors sense for the same duration, if

$$\frac{(\alpha_0 P_{\text{NB}})^2}{8} > B \frac{(\ell_0 P_{\text{WB}})^2}{8}. \tag{2.10}$$

For a given signal bandwidth, as B is increased (i.e., as the bandwidth of the NB signal is decreased), P_{NB} also increases relative to P_{WB}, since the NB filter captures the energy in the pilot tone more accurately. If the NB signal consists of a pure pilot tone, the ratio $\frac{P_{\text{NB}}}{P_{\text{WB}}}$ increases linearly with B. Thus, by using a large enough B, NB sensing can be made to outperform WB sensing for a given sensing duration, since the factor B appears quadratically in the error exponent term, while it occurs only linearly in the detection delay term. However, increasing B comes at the cost of an increasing accuracy in the CR's knowledge of the frequency of the pilot tone in the primary signal.

2.6 Numerical Results and Simulations

In this section, we present simulation results to validate the analytical development in the preceding sections, and to illustrate the relative performance of NB and WB sensing schemes. For the NB and WB cases, we denote the signal powers by P_{NB} and P_{WB},

and we let the channel gains be Rayleigh distributed and lognormal distributed, respectively. The prior probability was chosen to be $\pi_0 = 0.5$ for all the simulations. For comparison with existing results, we extend the analysis in [66] to derive the probability of error with a confidence level, and then calculate the EECL(q) from it. We also compare the performance of our detector with the detector designed under the NP criterion [47], for both NB and WB cases, as well as for single sensor detection and decentralized detection with multiple sensors.

2.6.1 Detection at the Sensors

In Fig. 2.3, we plot EECL(q) as a function of the confidence level q, for the NB and WB fading models, with $\frac{P_{\text{NB}}}{P_{\text{WB}}} = 3$. In the WB fading case, we show the curves for three typical values of the shadowing parameter σ_s^2. First, note that all the curves approach an EECL of 0 as q approaches 1, i.e., the conventional error exponent is zero under both NB and WB fading, as expected. As the confidence level is decreased, the NB sensing outperforms the WB sensing. Also, in the single sensor case, the design in [66] corresponds to using an NB detector. The excellent match between our results and those derived from [66] is clear from the plot.

In order to show that it is possible to achieve a positive error exponent with a confidence level under fading, we simulated the probability of error with confidence $q = 0.9$ at very low error probabilities, using importance sampling [87]. Figure 2.4 shows the performance as a function of the average primary SNR, for various values of q. The waterfall-type behavior of the curve indicates a positive error exponent. As mentioned earlier, an advantage of the error exponent approach is that the threshold, $\tau = \frac{\alpha_0 P}{2}$,
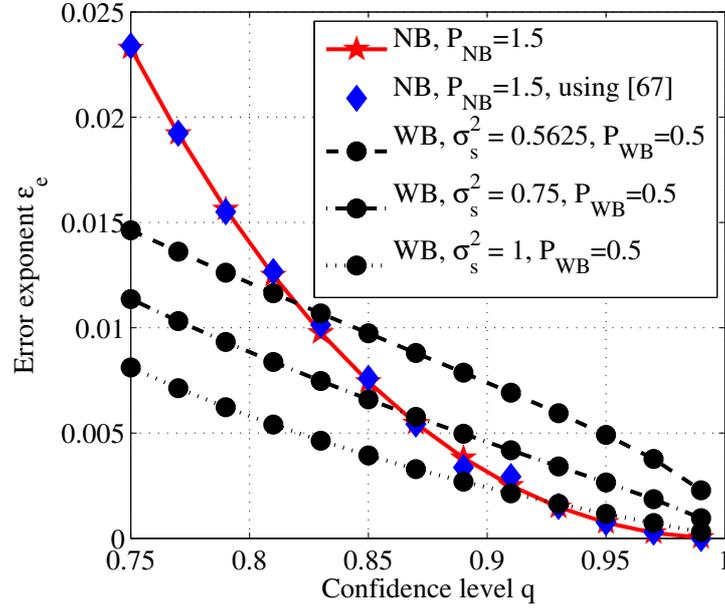
Figure 2.3: Trade-off between NB and WB sensing at a single sensor, with $\mu_s = 0$ in the WB case.

is independent of the prior probability π_0. In the figure, we see that the performance with $\tau = \frac{\alpha_0 P}{2}$ matches well with that obtained by using the near-optimal threshold derived in [57]. We also illustrate the effect of mismatched π_0 in Fig. 2.4. The performance loss due to lack of knowledge of π_0 is over 3 dB at a probability of error of 10^{-2}, when $M = 10^6$. For lower values of M, the performance loss would be much higher, because of the inverse square-root relationship between the number of samples and the SNR required to achieve a given performance [57].

2.6.2 Detection at the Fusion Center

We now consider the decentralized set-up with the OR fusion rule for combining the individual decisions from N sensors. In Fig. 2.5, we show the variation of the lower bound on $\epsilon_E^{(N)}$ with confidence $q = 0.99$. The detection threshold parameters α_0^{LB} and
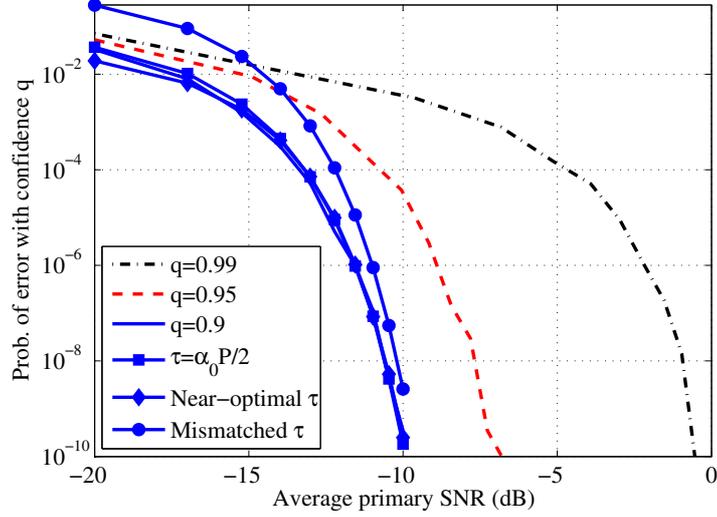
Figure 2.4: Variation of p_e with a confidence q as a function of SNR, under narrowband Rayleigh fading. Here, $N = 1$, $\pi_0 = 0.5$, $M = 10^6$. The curve labeled 'Mismatched τ' corresponds to using $\pi_0 = 0.5$ to design the detector, when the actual $\pi_0 = 0.01$.

ℓ_0^{LB} are obtained from (2.6) and (2.7). We see that the lower bound closely approximates the cross-over behavior of the NB and WB sensing schemes, shown in Fig. 2.6. For obtaining the latter curve, the detection thresholds are found by numerically solving (2.3) and (2.4) for the NB and WB cases, respectively.

We plot $\epsilon_E^{(N)}$ as a function of the power ratio $\frac{P_{\text{NB}}}{P_{\text{WB}}}$ in Fig. 2.7, for different values of q, and with $N = 4$. Both Figs. 2.6 and 2.7 show the cross-over between NB and WB sensing: as $\frac{P_{\text{NB}}}{P_{\text{WB}}}$ is increased, NB sensing outperforms WB sensing. Next, the variation of $\epsilon_E^{(N)}$ with the number of sensors N is shown in Fig. 2.8, with the power ratio $\frac{P_{\text{NB}}}{P_{\text{WB}}} = 1$. The plot shows an approximately linear improvement in the EECL(q) as the number of sensors is increased.

Next, we present simulation results of the probability of error at the FC, P_E, with the signal modeled as the sum of a sinusoidal component and an AWGN component,
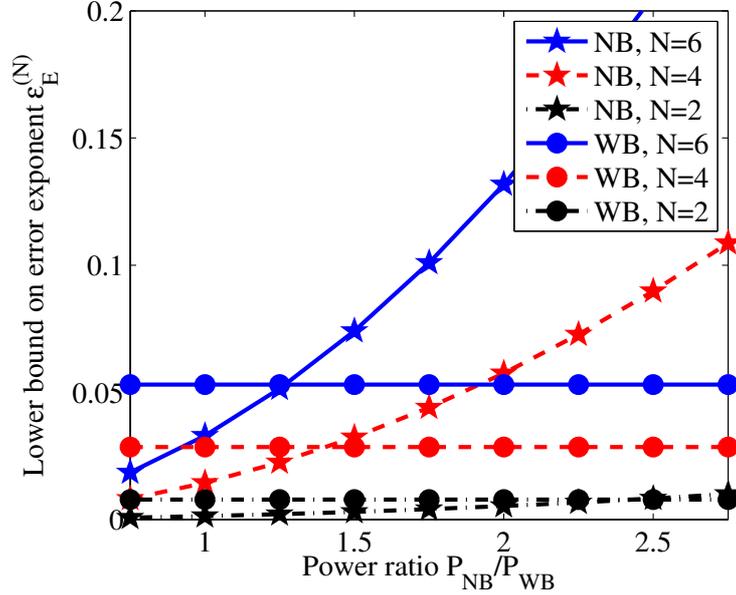
Figure 2.5: Variation of the lower bound on $\epsilon_E^{(N)}$ as a function of $\frac{P_{NB}}{P_{WB}}$, with $q = 0.99$, $\mu_s = 0, \sigma_s = 1$.

varying ratio of their powers according to $\frac{P_{NB}}{P_{WB}}$. The bandwidths of the NB and WB signals are fixed as 1 kHZ and 20 kHz, respectively. The sensing duration is chosen as 20 ms. We compute the probability of error with confidence q by computing the probability of error for 1000 i.i.d. channel states, and discounting a fraction $1 - q$ of the channel states that yield the highest probability of miss when averaged over $10,000$ noise instantiations. Under this set-up, we plot the probability of error with $N = 2, 4, 6$ and confidence level $q = 0.99$ in Figs. 2.9 and 2.10. From Fig. 2.9, we see that the power ratio at which the cross-over between NB and WB sensing occurs is roughly the same as the cross-over points in the EECL plot of Fig. 2.6, i.e., the EECL does capture the probability of error behavior of the detectors. In Fig. 2.10, we compare the performance of our design with that of the NP-based design adopted in ([10, 47]), for both single-sensor detection and decentralized detection, and for both the NB and WB cases. The
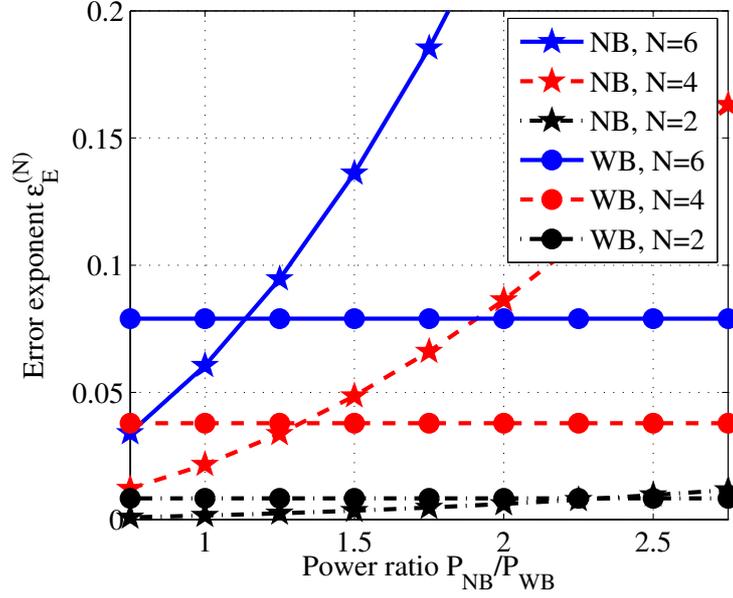
Figure 2.6: Variation of $\epsilon_E^{(N)}$ as a function of $\frac{P_{\text{NB}}}{P_{\text{WB}}}$, with $q = 0.99$, $\mu_s = 0$, $\sigma_s = 1$.

NP test is designed to meet a false alarm probability target of 0.01. We see that, while the probability of error with confidence $q = 0.99$ of the NP test saturates as the average primary SNR increases, the performance of our EECL-based design exhibits a waterfall-type drop with increasing SNR. Note that, for the settings in this simulation, the EECL of NB sensing is higher than that of WB sensing. Since having a larger average primary SNR is akin to having a larger number of observations at the individual sensors, one would expect that NB sensing should outperform WB sensing as the average primary SNR increases; this is also corroborated by Fig. 2.10.

Finally, Table 2.1 shows the values of α_0 and ℓ_0 for different q and N. It can be seen that both α_0 and ℓ_0 increase with N and decrease with q. Using importance sampling, the theoretical and experimental values of the error exponents obtained for different values of P, q and N are listed in Tables 2.2 and 2.3. We note the good agreement between the
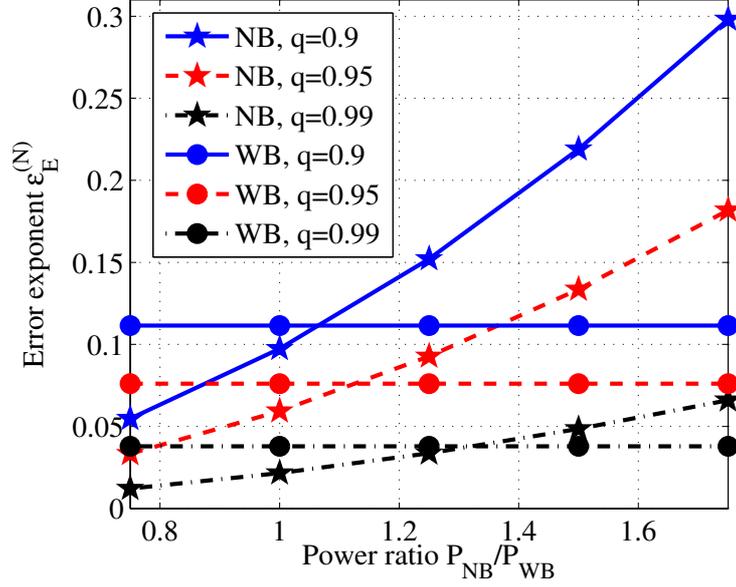
Figure 2.7: Variation of $\epsilon_E^{(N)}$ as a function of q, with $N = 4$, $\mu_s = 0$, $\sigma_s = 1$.

theoretical and simulated error exponents, even at very low exponent values.

2.7 Conclusions

In this chapter, we analyzed the performance of energy-based Bayesian decentralized detection for spectrum sensing in cognitive radios, with the exponent on the probability of error as the performance metric. We showed that, for various fading models, with the OR rule for decision fusion, the error exponent is equal to zero. We introduced a novel performance metric called the Error Exponent with a Confidence Level (EECL), and showed that the EECL at a given confidence level $q < 1$ is strictly positive. We used the EECL to answer the question of whether it is better to sense for the pilot tone in a narrow band, or to sense the entire wide-band signal. We also derived simplified
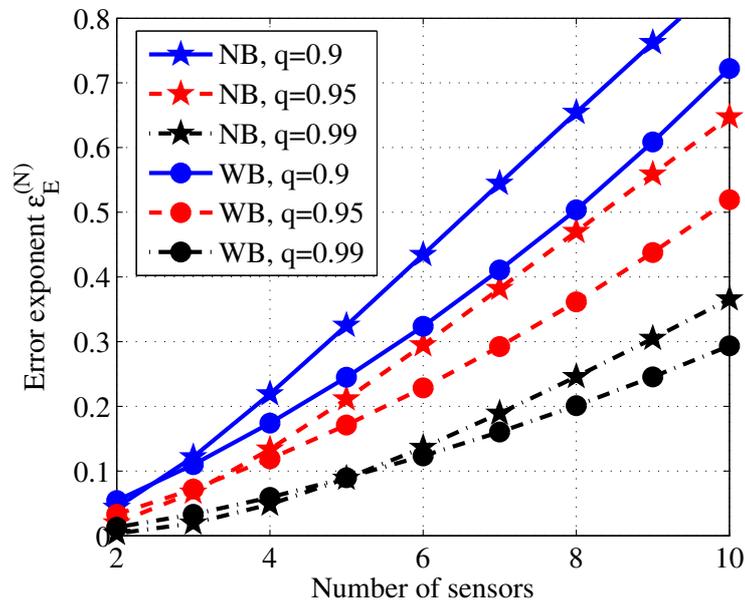
Figure 2.8: Variation of $\epsilon_E^{(N)}$ as a function of N, with $\frac{P_{\text{NB}}}{P_{\text{WB}}} = 1, \mu_s = 0, \sigma_s = 1$.

expressions for finding the detection threshold and the EECL for the i.i.d. Rayleigh fading and lognormal shadowing cases. We validated the theoretical expressions through simulations. Future work could include incorporating correlation in the signal or noise, extending the results to allow for time-varying channels, and optimally combining NB and WB spectrum sensing.
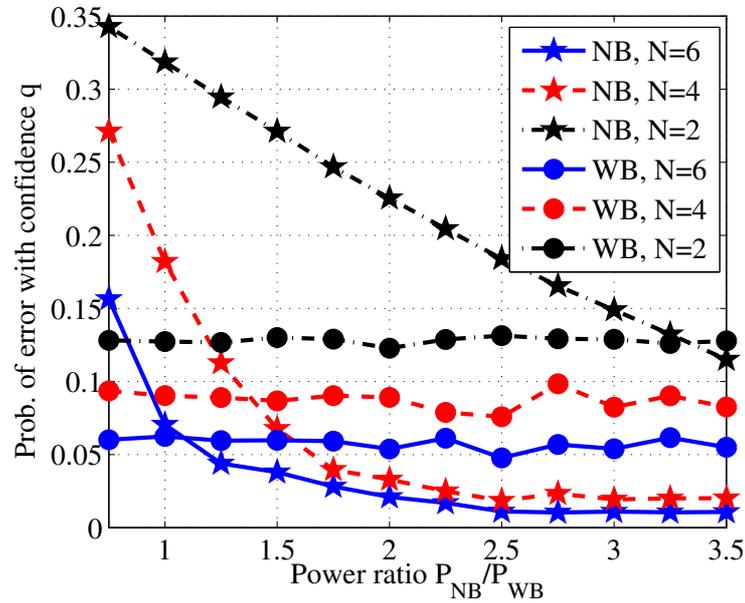
Figure 2.9: Variation of P_E with a confidence level as a function of $\frac{P_{NB}}{P_{WB}}$ with $q = 0.99$, $\mu_s = 0, \sigma_s = 1$ and $\pi_0 = 0.5$.
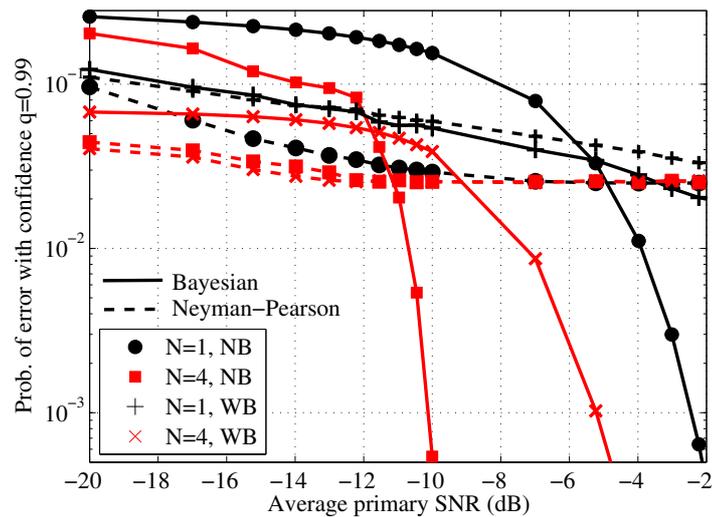
Figure 2.10: Comparison of the Bayesian and Neyman-Pearson approaches in terms of the P_E with a confidence $q = 0.99$, as a function of the average primary SNR, with $\mu_s = 0, \sigma_s = 1$ and $\pi_0 = 0.5$.

Table 2.1: Values of α_0 and ℓ_0 for different q and N.

N	$q = 0.9$		N	$q = 0.95$		N	$q = 0.99$	
	α_0	ℓ_0		α_0	ℓ_0		α_0	ℓ_0
2	0.39	0.53	2	0.26	0.41	2	0.11	0.26
3	0.66	0.75	3	0.49	0.61	3	0.26	0.41
4	0.88	0.94	4	0.69	0.78	4	0.42	0.55
5	1.07	1.12	5	0.87	0.94	5	0.56	0.68
6	1.24	1.29	6	1.02	1.08	6	0.70	0.79
7	1.39	1.45	7	1.16	1.22	7	0.82	0.91
8	1.52	1.60	8	1.29	1.36	8	0.93	1.01
9	1.65	1.77	9	1.41	1.50	9	1.04	1.12
10	1.76	1.92	10	1.52	1.63	10	1.14	1.23

Table 2.2: EECL(q) at a single sensor, with $N = 1$ and Rayleigh fading. All values have to be multiplied by 10^{-5}.

P (dB)	$q = 0.9$		P (dB)	$q = 0.95$		P (dB)	$q = 0.99$	
	Th	Sims		Th	Sims		Th	Sims
-10	1.39	2.20	-2.25	11.84	14.28	0	1.26	2.32
-7	5.55	6.57	-1.5	16.11	17.64	2	2.84	4.39
-5	12.49	13.63	-1	21.05	22.30	3	5.05	7.34
-4	22.02	23.56	-0.5	26.66	28.23	4	7.89	9.36
-3	34.69	35.91	0	32.89	34.19	5	11.36	14.49

Table 2.3: EECL(q) at the FC, with $P = -10$ dB and Rayleigh fading. All values have to be multiplied by 10^{-4}.

N	$q = 0.9$		N	$q = 0.95$	
	Th	Sims		Th	Sims
2	1.94	2.11	2	0.87	0.88
3	5.41	6.12	3	2.97	3.52
4	9.73	10.28	4	5.93	6.34

Chapter 3

Near-Optimal Detection Thresholds for Bayesian Spectrum Sensing

3.1 Introduction

As explained in Chapter 2, the under-utilization of the licensed radio spectrum, and the consequent scarcity of bandwidth for newer wireless applications, can be alleviated using cognitive radios. These are devices that can sense the radio spectrum and transmit only when a *primary* user is not using a given band. Spectrum Sensing (SS), or the problem of detecting the presence or absence of the primary transmissions, is typically modeled as a problem of signal detection, where the goal is to test the hypothesis H_1 that the primary is on, against the hypothesis H_0 that the primary is off. The performance of the SS algorithm in CR systems at low primary SNR is of particular interest, since one has to ensure that the CR transmissions do not cause harmful interference even at a primary receiver that receives a weak primary signal ([30,35,42]). This needs to be ensured under the different possible fading environments that the CR

may encounter. A design based on a single fading model such as Rayleigh is necessarily too restrictive. Hence, there is a need to develop detectors that offer guaranteed performance under a variety of fading conditions.

Among the techniques proposed for SS, Energy Detection (ED) is one of the simplest, from the point of view of implementability. The importance of the exploiting the fading model between the primary transmitter and CR using ED is well studied under the Neyman-Pearson (NP) formulation. One of the earliest works on ED for NP for deterministic signals was by Urkowitz [46], which was later extended in the context of SS under various fading models by Digham et al. [43] and Zeng et al. [72]. Under a mixture Nakagami distributed fading model (which has many well-known fading models as special cases), the optimum detector, a locally optimum detector under low SNR and a generalized likelihood ratio based detector have been studied in an NP context by Astaneh and Gazor [88]. Conditions under which all the above mentioned detectors perform better than ED are established. In all these works, the expressions of the PU detection probability (p_d), averaged over the channel distribution, for a given tolerance level on the false-alarm (p_f) were derived. A comprehensive survey of algorithms for ED under the NP formulation can be found in [30, 89].

Although the NP framework is commonly used in SS applications to design the test, it suffers from the significant disadvantage of incorrectly assigning the null and alternate hypothesis [90]. That is, it is more important, in the CR context, to find the threshold that minimizes the p_f for a given lower bound on p_d, but most of the literature focuses on maximizing p_d under and upper bound constraint on p_f. To overcome this issue, in this chapter, we formulate the SS problem under a Bayesian setup, where the goal is to

minimize the probability of error. To the best of authors' knowledge, a detailed study of the performance of ED under various fading channel models in a Bayesian setup has not been considered in the literature so far; the current chapter tries to fill this gap.

The primary challenge in Bayesian SS is to determine the detection threshold that minimizes the probability of error. The optimal threshold depends on the fading model between the primary transmitter and the CR node. In contrast, under the NP framework, the detection threshold depends only on the signal statistics under the noise-only hypothesis, not on the fading model. As mentioned in the previous chapter, the exact analysis of the detection threshold, even for the simplest case of Rayleigh fading, is difficult. In this chapter, we are interested in deriving the near-optimal detection thresholds for the Bayesian SS under various practical fading models such as Rayleigh [46], lognormal [47], Nakagami-m [43], Weibull [91]. We are also interested in uncovering the relationship between the number of samples used for detection, the detection SNR and the average probability of error achieved. In this context, our main contributions in this chapter are:

- We derive near-optimal detection thresholds for the commonly encountered fading models, viz., Rayleigh, lognormal, Nakagami-m, Weibull distributions, for spectrum sensing under a Bayesian framework. The conditions under which each approximation (near-optimality of the threshold) holds are highlighted.

- For the Rayleigh fading case, a fundamental relationship between the number of observations and the primary SNR for achieving a given probability of error is found.

- Extending the analysis to decentralized detection using ED at N sensors, the exponent on the probability of error at the Fusion Center (FC) is derived, as the number of sensors grows large. This is different from the previous chapter, where the error exponent was derived as the number of samples grows large, keeping the number of sensors fixed.

- In the case of Rayleigh flat-fading primary signals, the diversity gain due to employing decentralized detection using multiple sensors is demonstrated using Monte Carlo simulations. Also, an approximation for the Suzuki distribution [77] is used to obtain an easily computable expression for the near-optimal detection threshold under Suzuki fading.

The optimal thresholds derived in this chapter thus provide a comprehensive, near-optimal solution set for Bayesian ED, under a variety of signal fading models that are commonly encountered in CR applications.

The rest of the chapter is organized as follows. Section 3.2 explains the system model considered in this work. Near-optimal detection threholds under various fading models are discussed in Sec. 3.3. In particular, the near-optimal detection thresholds for SS under Rayleigh fading, lognormal shadowing, Nakagami-m fading, and Weibull fading are discussed in Secs. 3.3.1, 3.3.2, 3.3.3, and 3.3.4, respectively. The error exponent at the fusion center under a general fading model between the primary transmitter and individual sensors, and a lossless link between the sensors and the fusion center is derived in Sec. B.4. Simulation results are discussed in Sec. 3.4, and the conclusions are given in Sec. 3.5.

3.2 System Model

Consider a CR user which uses energy detection from M observations to carry out spectrum sensing. At low SNR, and when M is reasonably large, using the Central Limit Theorem (CLT) ([38, 66, 81]), the statistic under both hypotheses can be modeled as Gaussian distributed, leading to the hypothesis test

$$
\begin{aligned}
\mathcal{H}_0 &: V_y \sim \mathcal{N}\left(0, \frac{1}{M}\right) \\
\mathcal{H}_1 &: V_y \sim \mathbb{E}_h \mathcal{N}\left(|h|^2 P, \frac{1}{M}\right),
\end{aligned}
\tag{3.1}
$$

where $V_y \triangleq \frac{1}{M} \sum_{k=1}^{M} |Y_k|^2 - 1$, with Y_i representing the observations recorded by the CR node. Without loss of generality, in writing the above, we have assumed that the noise variance is unity. The notation $\mathbb{E}_h$ indicates the expectation operator over the random channel h, which is assumed to be constant throughout the M observations. For the problem in (3.1), the likelihood ratio (LR) is given by

$$
LR(V_y) = \frac{\int_0^\infty \exp\left[-\frac{(V_y - \alpha P)^2}{2/M}\right] f_\alpha(\alpha) d\alpha}{\exp\left[-\frac{V_y^2}{2/M}\right]},
\tag{3.2}
$$

where $\alpha \triangleq |h|^2$ represents the random channel gain. For the commonly encountered channel fading models for $|h|$ such as Rayleigh, lognormal, Nakagami-m, and Weibull, the channel gain $|h|^2$ is distributed as exponential, lognormal, gamma and Weibull, respectively. Additionally, $|h|^2$ can also be modeled as the Suzuki distribution, to consider the effect of both small scale and large scale fading (Rayleigh and lognormal, respectively). From elementary detection theory, it is known that the optimal test (that minimizes the probability of error) for the hypothesis test in (3.1) is a test on $LR(V_y)$,

which is of the form

$$LR(V_y) \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \frac{\pi_0}{1-\pi_0}, \tag{3.3}$$

where π_0 is the prior probability of $\mathcal{H}_0$, assumed to be known. In the next section, we show that, under fairly general conditions on the fading distribution $f_\alpha(\alpha)$ the above test reduces to a test on V_y, i.e., to a test of the following form

$$V_y \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \tau(\pi_0, M, P, \{p_h\}), \tag{3.4}$$

where $\tau(\pi_0, M, P, \{p_h\})$ is the optimal detection threshold, and is a function of parameters π_0, M, P and on the set of parameters of the fading distribution considered (which is denoted by $\{p_h\}$). We also show the conditions under which the test on V_y in (3.4) is optimal for a particular fading model. By a simple substitution, it can be easily shown that when the test is of the form (3.4), the probabilities of false-alarm, missed detection and overall error are given by

$$p_f \triangleq \mathcal{P}\{V_y > \tau(\pi_0, M, P, \{p_h\})|\mathcal{H}_0\} = Q\left(\tau(\pi_0, M, P, \{p_h\})\sqrt{M}\right), \tag{3.5}$$

$$p_m \triangleq \mathcal{P}\{V_y \leq \tau(\pi_0, M, P, \{p_h\})|\mathcal{H}_1\} = \int_0^\infty Q\left(-\sqrt{M}(\tau(\pi_0, M, P, \{p_h\}) - \alpha P)\right) f_\alpha(\alpha) d\alpha, \tag{3.6}$$

$$p_e = \pi_0 \, p_f + (1 - \pi_0) \, p_m, \tag{3.7}$$

where $Q(\cdot)$ represents the Gaussian Q-function.

3.3 Detection Under Various Fading Models

3.3.1 Detection Under Rayleigh Fading

When the channel $|h|$ is Rayleigh distributed, $|h|^2$ follows the standard exponential distribution. The following theorem gives an approximate solution to the threshold that minimizes the probability of error for the hypothesis testing problem in (3.1).

Theorem 4. *When the channel between primary transmitter and the CR node is Rayleigh distributed, the detector for the problem considered in* (3.1) *has an approximate critical region given by*

$$V_y \geq x_{CLT}^{(R)} \triangleq \frac{1}{MP} + \frac{1}{\sqrt{M}} \sqrt{2 \log \left(1 + \left(\frac{\pi_0}{1-\pi_0} \right) \frac{P}{\sqrt{\frac{2\pi}{M}}} \right)}. \tag{3.8}$$

Proof. See Appendix B.1 □

The above result allows one to determine how M needs to scale with the received average primary power P to obtain a given p_e, which is stated as the corollary below.

Corollary 3. *The probability of error for the hypothesis testing problem in* (3.1) *under the Rayleigh fading case (given in* (B.1)*) depends on P and M only through the product $P\sqrt{M}$.*

The above corollary follows directly by substituting the optimum threshold $x_{CLT}^{(R)}$ in (3.8) into (3.7), and verifying that p_e depends only upon the product $P\sqrt{M}$. The utility of the corollary is that it shows, for example, that if average received primary signal power reduces by 3 dB, the probability of error performance can be retained by increasing the number of observations by a factor of four.

3.3.2 Detection Under Lognormal Shadowing

In wideband SS (for instance, in the first DTV 802.22 proposed standard), the primary signal spans multiple coherence bandwidths, due to which, the small scale (for e.g., Rayleigh) fading components average out ([47], [77]). Under this model, the channel gain $\ell \triangleq |h|^2$ is modeled as lognormal distributed with parameters μ_s and σ_s^2. That is, $\ell = e^{\mu_s + \sigma_s S}$, where S is a zero mean, unit variance Gaussian distributed random variable. The following theorem gives an approximate solution to the near-optimal threshold, under lognormal shadowing. Before stating the theorem, we present a heuristic observation based on extensive simulations, which gives the conditions under which a lognormal distribution can be well approximated by a Gaussian distribution.

Result 1. *When the parameter σ_s^2 is small and around zero, then, for the purpose of designing detection thresholds, the distribution of a lognormal random variable with parameters μ_s and σ_s can be approximated by a Gaussian distribution with mean e^{μ_s} and variance $e^{2\mu_s}\sigma_s^2$, i.e.,*

$$\mathcal{LN}(\mu_s, \sigma_s) \stackrel{d.}{\approx} \mathcal{N}(e^{\mu_s}, e^{2\mu_s}\sigma_s^2). \tag{3.9}$$

The intuition behind the above observation is as follows. Let $X \sim \mathcal{LN}(\mu_s, \sigma_s)$. When the value of σ_s is low, defining $x = e^{\mu_s}(1 + \epsilon)$, we have

$$\mathcal{P}(X \leq x) = 1 - Q\left(\frac{\log|x| - \mu_s}{\sigma_s}\right) \approx 1 - Q\left(\frac{\mu_s + \log(1 + \epsilon) - \mu_s}{\sigma_s}\right) \approx 1 - Q\left(\frac{\epsilon}{\sigma_s}\right) \tag{3.10}$$

when ϵ is a small positive real number. Approximating the PDF of X as $X \sim \mathcal{N}(e^{\mu_s}, e^{2\mu_s}\sigma_s^2)$ gives:

$$\mathcal{P}(X \leq x) = 1 - Q\left(\frac{x - e^{\mu_s}}{e^{\mu_s}\sigma_s}\right) \approx 1 - Q\left(\frac{e^{\mu_s}(1 + \epsilon) - e^{\mu_s}}{e^{\mu_s}\sigma_s}\right) = 1 - Q\left(\frac{\epsilon}{\sigma_s}\right). \tag{3.11}$$

Therefore, for small values of ϵ (which will be the case of interest for the SS problem when σ_s is small), (3.11) and (3.10) are equal. The accuracy of the approximation is highlighted in Figs. 3.1 and 3.2.
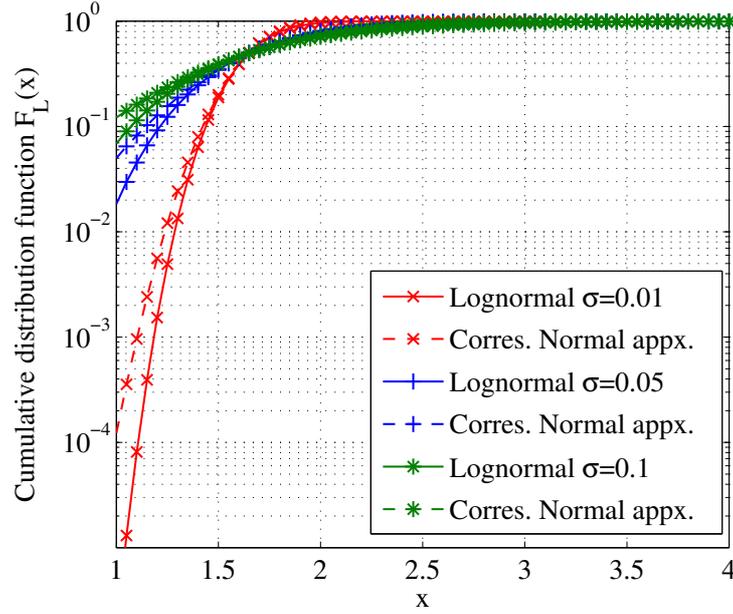
Figure 3.1: CDF of Lognormal distribution for various values of its log-shape parameter, and the corresponding Normal CDF approximation.

Theorem 5. *When the channel between primary transmitter and the CR node follows a lognormal distribution, the Bayesian detection problem considered in (3.1) has an approximate critical region given by*

$$\left\{ V_y : V_y \geq x_{CLT}^{(WB)} \triangleq \frac{1}{M}\left[\sqrt{2\left(M + \frac{1}{P^2\sigma_s^2}\right)\left(\log\left(\frac{\pi_0}{1-\pi_0} + K_c\right) - \log K_c\right)} - \frac{1}{P\sigma_s^2}\right]\right\}$$
(3.12)

where,

$$K_c \triangleq \frac{\exp\left(-\frac{1}{2\sigma_s^2}\right)}{\sqrt{1 + MP^2\sigma_s^2}}.$$
(3.13)
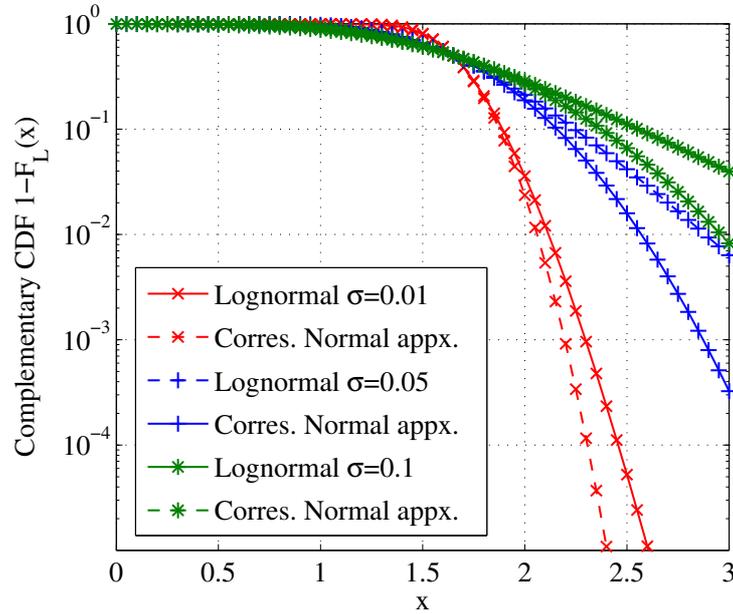
Figure 3.2: CCDF of Lognormal distribution for various values of its log-shape parameter, and the corresponding Normal CCDF approximation.

Proof. See Appendix B.2. □

3.3.3 Detection Under Nakagami-m Fading

Note that, in the Bayesian hypothesis testing problem in (3.1), if $|h|$ follows a Nakagami-m distribution [77] with a positive integer parameter K, then $g \triangleq |h|^2$ follows a gamma distribution, with PDF $f_G(g) = \frac{g^{K-1}e^{-g}}{\Gamma(K)}$. The following theorem gives the near-optimal detection threshold under Nakagami-m fading.

Theorem 6. *Let the channel between primary transmitter and the CR node be Nakagami-m distributed. For the Bayesian hypothesis testing problem considered in* (3.1), *the optimal test is of the form* (3.4), *i.e., ED is optimal. Further, the solution to the following equation yields the*

near-optimal detection threshold $x_{\text{CLT}}^{(Nm)}$:

$$\frac{K^K}{P^K \Gamma(K)} \exp\left(\frac{\left(x_{\text{CLT}}^{(Nm)} - \frac{K}{MP}\right)^2}{2/M}\right) \left(\frac{1}{M}\right)^{\frac{K-1}{2}} \left[\left[i\sqrt{2}\,sign\left(x_{\text{CLT}}^{(Nm)} - \frac{K}{MP}\right)\right]^{K-1} \sqrt{\frac{2\pi}{M}}\right.$$

$$\left\{\frac{\Gamma(1/2)}{\Gamma\left(1 - \frac{K}{2}\right)} + \frac{(K-1)\Gamma(1/2)}{\Gamma\left(1 - \frac{K}{2}\right)} \times \frac{\left(x_{\text{CLT}}^{(Nm)} - \frac{K}{MP}\right)^2}{2/M} - i\frac{\Gamma(-1/2)}{\Gamma\left(-\frac{K-1}{2}\right)} \times \frac{\left(x_{\text{CLT}}^{(Nm)} - \frac{K}{MP}\right)}{\sqrt{2/M}}\right\}$$

$$\left.- \left[\left(x_{\text{CLT}}^{(Nm)} - \frac{K}{MP}\right)^{K-1} Q\left(-\frac{x_{\text{CLT}}^{(Nm)} - \frac{K}{MP}}{\sqrt{1/M}}\right)\right]\right] - \frac{\pi_0}{1 - \pi_0} = 0, \qquad (3.14)$$

provided the solution is close to the value $\frac{K}{MP}$. Here, $i \triangleq \sqrt{-1}$. Otherwise, ED may not be optimal.

Proof. See Appendix B.3. $\qquad\qquad\square$

Although (3.14) cannot be solved in closed-form, numerically solving it to obtain the threshold $x_{\text{CLT}}^{(Nm)}$ is not difficult. Theorem 6 shows that ED is optimal when $x_{\text{CLT}}^{(Nm)}$ is close to $\frac{K}{MP}$. Following this, a simple, suboptimal detector can be proposed as follows (this can be used in the Rayleigh fading case also, by setting $K = 1$), which compares V_y to a threshold $x_{\text{CLT}}^{\text{sub}}$:

$$V_y \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} x_{\text{CLT}}^{\text{sub}} \triangleq \frac{K}{MP}. \qquad (3.15)$$

We will illustrate the performance of the above suboptimal detector via Monte Carlo simulations in Sec. 3.4.

3.3.4   Detection Under Weibull Fading

Weibull fading model is typically used to model the attenuation due to fading in wireless applications operating in 800-900 MHz range ([92], [91], [77]). In this case, the

channel $|h|$ is Weibull distributed, and $W \triangleq |h|^2$ is also a Weibull distributed random variable. If $\mathcal{W}(a_w, b_w)$ denotes a Weibull distributed variable with shape and scale parameters a_w and b_w respectively, it is known that when $|h| \sim \mathcal{W}(a_w, b_w)$, then $|h|^n \sim \mathcal{W}\left(\frac{a_w}{n}, b_w^n\right)$, for any positive integer n. Therefore, the PDF of W is given by

$$f_W(w; a_w; b_w) = \frac{a_w}{2b_w^2} \left(\frac{w}{b_w^2}\right)^{\frac{a_w}{2}-1} \exp\left(-\left(\frac{w}{b_w^2}\right)^{\frac{a_w}{2}}\right), \quad w \geq 0. \tag{3.16}$$

The resulting integral in the numerator of (3.2) is hard to obtain in closed form. Hence, we use a normal approximation on Weibull distribution ([93], [94]), which, as seen in Figs. 3.3 and 3.4, gives a good fit for large values of the shape parameter a_w. Restriction of the analysis for larger values of a_w is in agreement with a result that for fading in wireless communications around $900\mathrm{MHz}$, Weibull fit with larger values of the shape parameter a_w is sufficiently accurate [95]. The mean and the variance of the approximated normal random variable are given by

$$\begin{aligned}
\mu_w &\triangleq b_w \Gamma\left(1 + \frac{1}{a_w}\right) \\
\sigma_w^2 &\triangleq b_w^2 \left[\Gamma\left(1 + \frac{2}{a_w}\right) - \Gamma^2\left(1 + \frac{1}{a_w}\right)\right].
\end{aligned} \tag{3.17}$$

Theorem 7. *When the channel between primary transmitter and the CR node follows a Weibull distribution, the Bayesian detection problem in* (3.1) *has an approximate critical region*

$$\left\{V_y : V_y \geq x_{CLT}^{(W)} \triangleq \frac{1}{M}\left[\sqrt{2\left(M + \frac{1}{P^2\sigma_w^2}\right)\left(\log\left(\frac{\pi_0}{1-\pi_0} + K_w\right) - \log K_w\right)} - \frac{\mu_w}{P\sigma_w^2}\right]\right\} \tag{3.18}$$
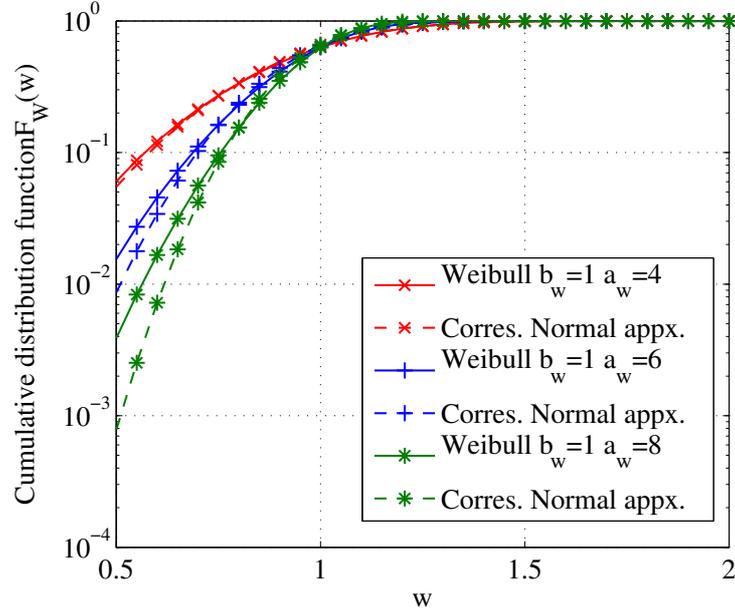
Figure 3.3: CDF of the Weibull distribution for various values of its scale and shape parameters, and the corresponding Normal CDF approximation.

where

$$K_w \triangleq \frac{\exp\left(-\frac{\mu_w^2}{2\sigma_w^2}\right)}{\sqrt{1 + MP^2\sigma_w^2}}. \tag{3.19}$$

Proof. The proof is similar to the proof of Theorem 5, and follows by using the approximation given in (3.17). □

In Sec. B.5, we discuss detection under Suzuki distribution [77]. Unfortunately, the analysis of the optimal detection threshold is very hard for the case of Suzuki distribution, because of its complicated integral form. In order to simplify the probability of detection expressions, we approximate the Suzuki distributed random variable with a generalized Gamma distribution [96]. However, the evaluation of $LR(V_y)$ is not
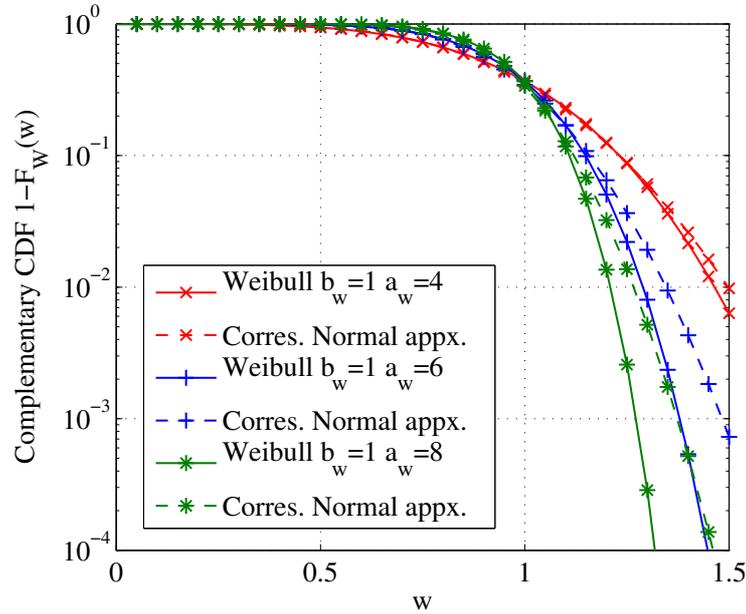
Figure 3.4: CCDF of the Weibull distribution for various values of its scale and shape parameters, and the corresponding Normal CCDF approximation.

straightforward even after this approximation, and hence, the optimal detection threshold has to be evaluated through numerical techniques.

In Appendix B.4, we consider a decentralized setup and derive the error exponent on the probability of error at a Fusion Center (FC), as a function of p_f and p_m at the individual sensors. With some algebra, it can be shown that the right hand side of (B.22) is strictly positive. A diversity-like improvement in P_E is obtained with increasing N, and further insights can be drawn from the simulation results, discussed in Sec. 3.4.

3.4 Simulation Results

Our simulation setup consists of N sensors, using M observations each, with a average primary SNR observed at each sensor denoted by P. The range of P and a target P_E is typically specified by the primary network. This, in turn, dictates the value of M
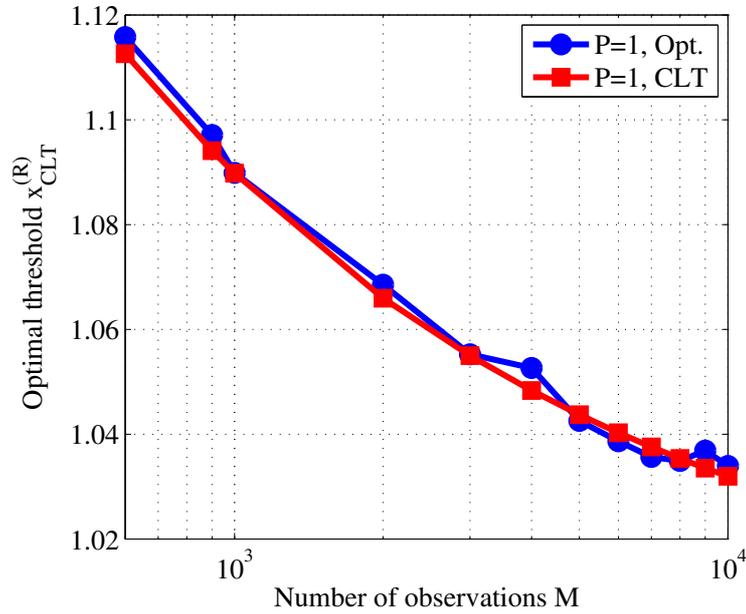
Figure 3.5: Variation of $x_{\text{CLT}}^{(R)}$ with M for the optimum and CLT schemes.

and the detection threshold, such that the resulting P_E is within the target specified by the primary network. For all plots, π_0 was fixed to be 0.5. First, consider the case with $N = 1$. Figure 3.5 shows the comparison of the optimum threshold (curve labeled `Opt.`, obtained through simulations by sweeping over a range of thresholds) with that of the near-optimal threshold for the Rayleigh fading case (curve labeled `CLT`), indicating that the $x_{\text{CLT}}^{(R)}$ given in (3.8) is a good approximation. As expected, when M increases, $x_{\text{CLT}}^{(R)}$ goes close to zero.

Figure 3.6 plots the probability of error p_e versus average primary SNR P to show that keeping $P\sqrt{M}$ constant as P varies maintains a fixed probability of error level at the detector. In this example, M was chosen as 500 for $P = 1$ and (P, M) are varied such that $P\sqrt{M}$ is constant. The two additional curves show the performance for a constant M (chosen as 222, and 2000, corresponding to the value of M needed to obtain the p_e at
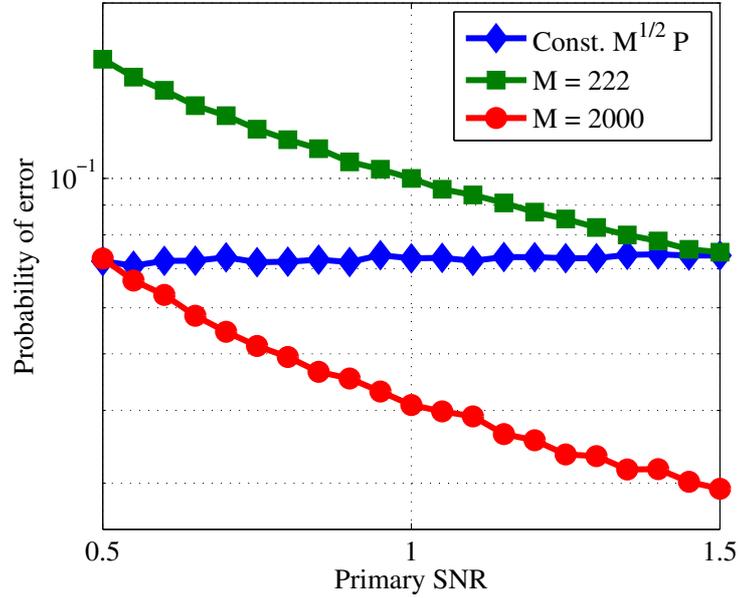
Figure 3.6: A constant p_e can be maintained if $P\sqrt{M}$ remains fixed.

the two extreme values of P).

Figures 3.7 and 3.8 show the variation of the optimal threshold for the shadowing case ($\sigma_{dB} = 0.5$) and weibull fading case ($a_w = 5, b_w = 1$), respectively, as a function of M, for various values of P. As expected, the approximations become tighter for high M and low P. A similar trend is seen in Fig. 3.9, where the simulated value of optimal threshold for the Nakagami-m fading case with $K = 10$ is compared with the corresponding theoretical near-optimal threshold derived in (3.14), as a function of M, for various values of P. Also, the value of suboptimal threshold indicated in (3.15) is plotted. It is observed that for the Nakagami-m case, the regime of number of observations where our derived results hold is seen to be about ten-fold higher ($\sim 10^5$) as compared with the corresponding regimes in the case of lognormal or Weibull fading ($\sim 10^4$), to achieve the same probability of error.

Now, consider the decentralized detection case with $N > 1$. In Fig. 3.10, the quantities $\frac{-\log P_E}{N}$ and the theoretical ϵ_E is plotted as a function of the number of sensors N for $M = 5, 80$, and $P = 0.1dB, -5.9dB$, respectively. As N becomes large, the quantity $\frac{-\log P_E}{N}$ approaches the theoretical exponent given in (B.22). For both values of P, the values of M are chosen such that $P\sqrt{M}$ is a constant. Therefore, the simulated curves corresponding to $\frac{-\log P_E}{N}$ are nearly equal, as expected from Corollary 3. Figure 3.11 plots the P_E at the FC *vs.* the average primary SNR P, using $N = 1, 3$ sensors, which illustrates the diversity advantage that can be obtained by multi-sensor decentralized detection. The optimum detection threshold and K_{opt} was obtained from simulations.
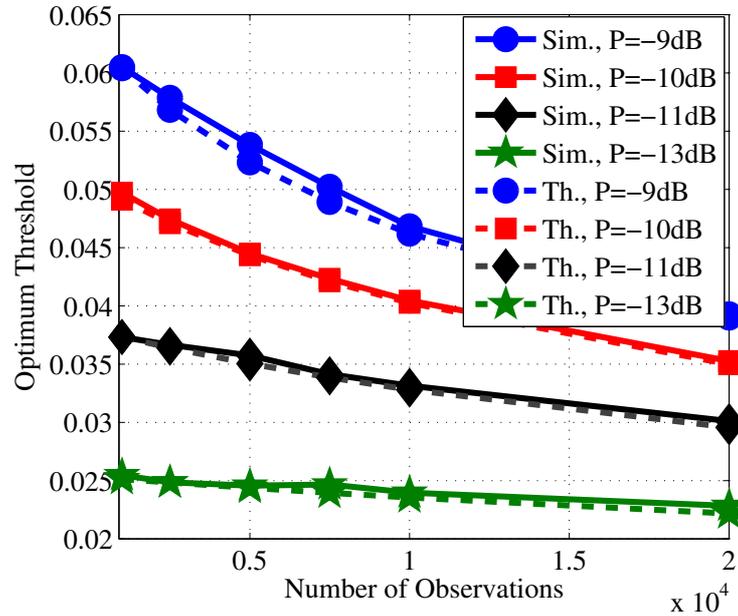
Figure 3.7: Simulated optimal thresholds and near-optimal theoretical thresholds for the shadowing fading case, with its log-scale parameter $\sigma_{dB} = 0.5$.
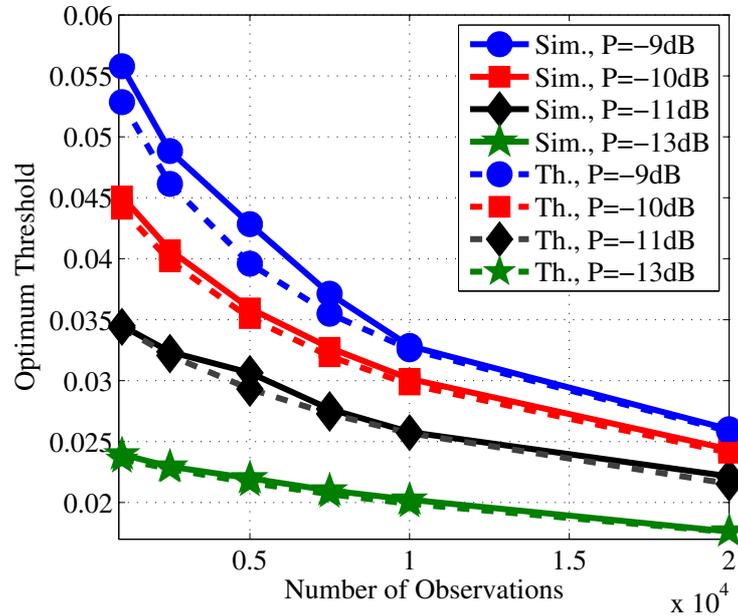
Figure 3.8: Simulated optimal thresholds and near-optimal theoretical thresholds for the Weibull fading case, with $a_w = 5, b_w = 1$.

3.5 Conclusions

This chapter considered the problem of Bayesian decentralized SS in CR networks under various fading environments. A CLT-based approximation to the problem was explored, which lead to analytically tractable expressions for near-optimal detection thresholds that minimize the probability of error under Rayleigh, Lognormal, Nakagami-m, Weibull fading cases. For the Suzuki fading case, a generalized gamma approximation was provided, which saves on the computation of an integral. Also, in the Rayleigh fading case, it was shown that the probability of error at the individual sensors, in the low SNR regime, is a function of $P\sqrt{M}$, where P is the received primary power and M is the number of observations needed for detection. Extending to the decentralized
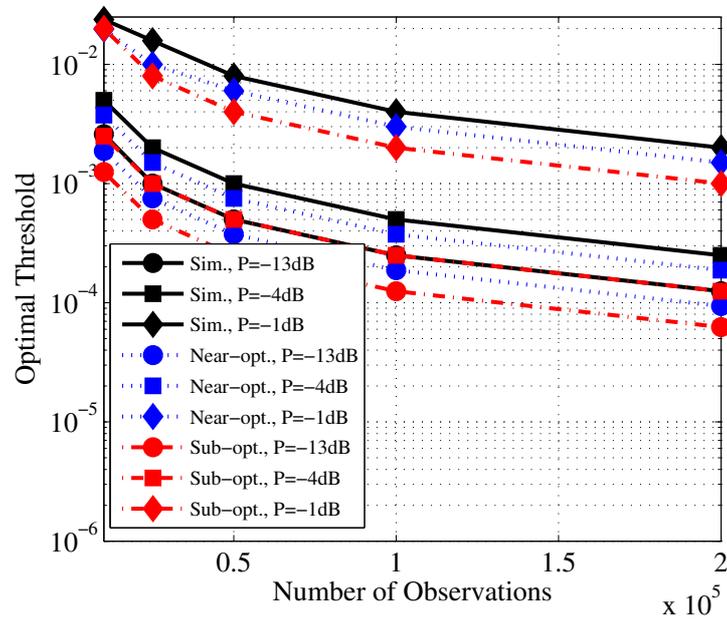
Figure 3.9: Simulated optimal thresholds and near-optimal theoretical thresholds for the Nakagami-m fading case, with $K = 10$.

case with N sensors, the optimal exponent on the probability of error at the FC was derived in closed form. The accuracy of the theoretical expressions and the diversity gain obtainable through the use of multiple sensors were illustrated through simulations.
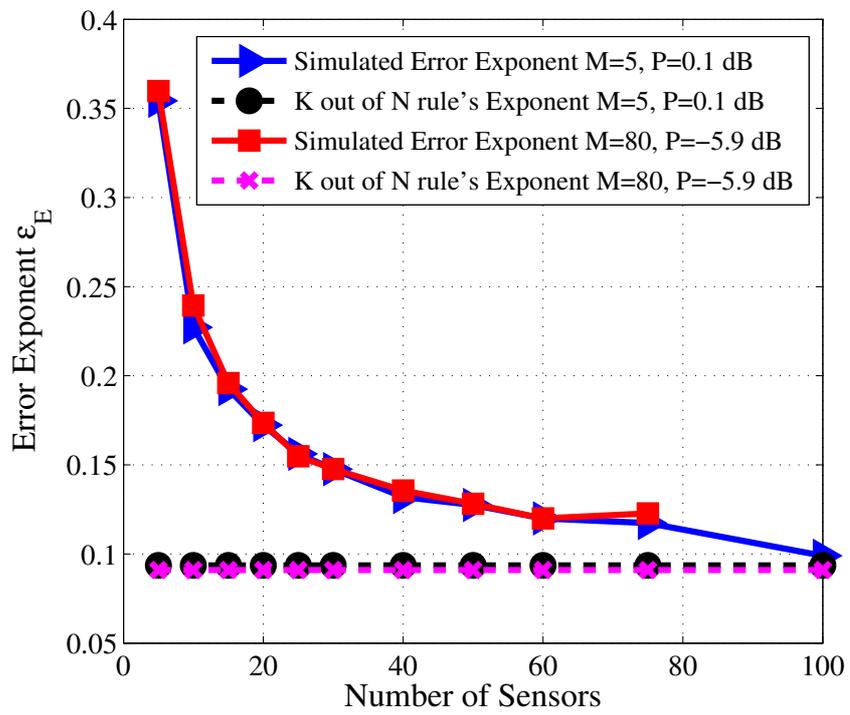
Figure 3.10: Variation of ϵ_E with N ($\pi_0 = 0.5$).
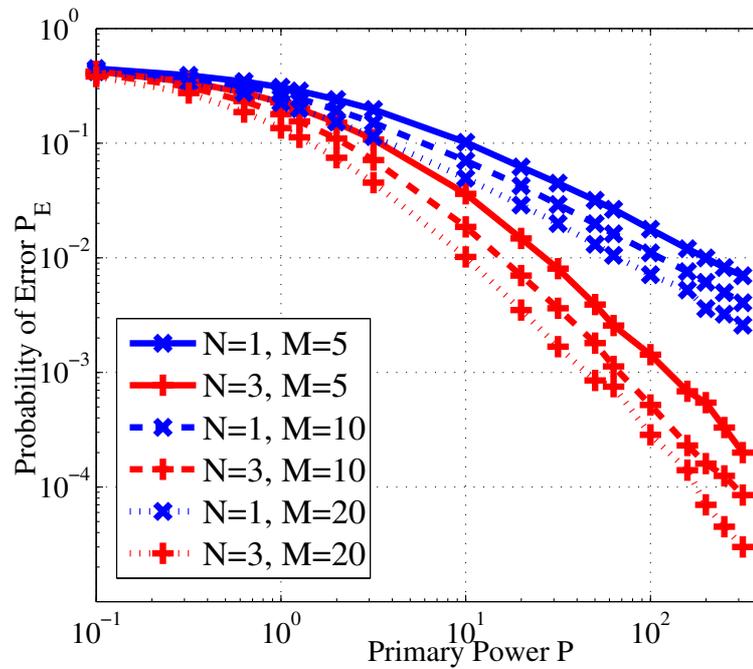
Figure 3.11: Variation of P_E with P ($\pi_0 = 0.5$).

Chapter 4

Design and Implementation of Spectrum Sensing with a Frequency-Hopping Primary System

4.1 Introduction

A *Cognitive Radio* (CR) [1] operates by *sensing* the spectrum for *Primary User* (PU) activity, and transmitting only in those frequency bands where the PU signal is observed to be absent. The detection of PU activity is accomplished via Spectrum Sensing (SS), which is the binary hypothesis testing problem of detecting the presence or absence of a PU in the frequency-band of interest. In Chapter 2, we presented a detailed survey of recent literature on SS.

It is clear that, in order to maximize the CR's throughput while offering adequate protection to the primary, it is important to tune the various sensing and CR transmission related parameters such as the sensing duration, threshold, transmit power, etc., considering the primary signal characteristics. The problem of SS is particularly challenging when the primary users employ Frequency Hopping (FH) signaling, since

the CR needs to detect and exploit the available frequency bins within the short hop duration of the primary.

The focus of this chapter is on the detection of FH primary signals. FH signals are typically used in secure/military applications and in standards such as IEEE 802.15.1/Bluetooth. SS of FH primary signals is challenging due to the frequently changing nature of the primary frequency band [97], and has received relatively less attention in the CR literature. Some early work on the detection of frequency-hopped signals include [98–100]. In these studies, the received signal is passed through a bank of Band Pass Filters (BPFs), and a decision on the signal presence in each frequency bin is made from the energy computed in each BPF, in time domain. Also, implementations of spectrum sensing in a non-FH primary environment have been demonstrated using off-the-shelf hardware platforms in [101–103].

To the best of our knowledge, however, past work on detecting FH signals has focussed primarily on detecting the presence or absence of an FH primary signal itself, not on detecting unoccupied bands within each hop duration. Detecting and exploiting unoccupied bands within the wider bandwidth over which primary users execute their hopping patterns can lead to the harvesting of significant RF spectrum for CR communications. To this end, in this chapter, we consider energy-based detection of unoccupied bands in the presence of FH primary signals, and evaluate its efficacy through rigorous theoretical analysis, Monte Carlo simulations, and implementation on a hardware platform.

In this chapter, we study performance of the Fast Fourier Transform (FFT) Averaging Ratio (FAR) algorithm [101, 104], an energy-based detection scheme, for the problem of

sensing FH signals in CR applications. We consider the FAR algorithm for detection be-
cause of its computational simplicity and ease of implementation in a software defined
radio environment. Our novel contributions are as follows:

- We extend the FAR algorithm to the multiuser FH-PU scenario, and derive closed-
  form expressions for the probabilities of false alarm and detection, as a function
  of the detection threshold, number of averaging frames, and the estimated SNRs
  of the primary signal in the occupied bands (Sec. 4.3.1).

- We define a utility metric to quantify the throughput of the CR, and obtain the CR
  sensing duration that maximizes the throughput while satisfying a constraint on
  the maximum allowable interference to the PUs (Sec. 4.3.2).

- We implement the FAR Algorithm on a Lyrtech Small Form Factor Software De-
  fined Radio Development Platform (Lyrtech SFF SDR DP), and validate the im-
  plementation by comparing its performance with that obtained from the analysis
  and simulations (Appendix C.3 and Sec. 4.4).

Allowing for implementation losses, we show that the results obtained from the hard-
ware corroborate well with those obtained through theory as well as Monte Carlo sim-
ulations. We conclude, therefore, that the FAR algorithm is an easy-to-implement and
effective solution to the SS problem with an FH-PU network.

The rest of the chapter is organized as follows. The system model, and the FAR algo-
rithm are given in Sec. 4.2. The associated probabilities of false-alarm and detection of
the FAR algorithm as applied to SS with FH primary signals are derived in Sec. 4.3.1,
and the optimum sensing duration that maximizes the CR throughput is derived in
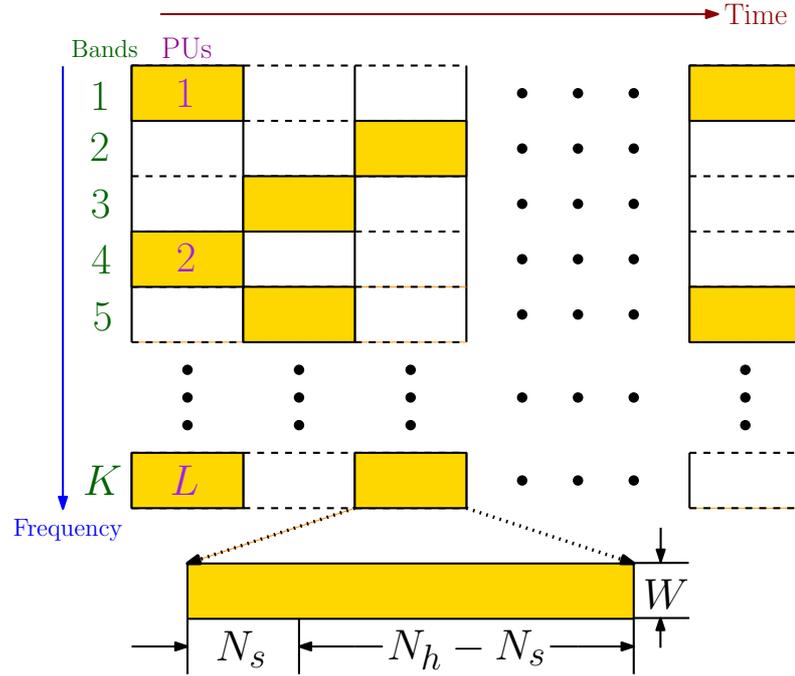
Figure 4.1: Typical frequency-band occupancies in a multiple FH-PU network.

Sec. 4.3.2. The implementation of the FAR algorithm on the Lyrtech SFF SDR DP is discussed in Appendix C.3. Finally, Sec. 4.4 discusses the Monte-Carlo simulations and the experimental results, and Sec. 4.5 concludes the chapter.

4.2   System Model and FAR Algorithm

4.2.1   System Model

In an FH network, each PU occupies a frequency bin for a period of time, known as the hopping period, or the hop duration (N_h s). In successive hop periods, the PUs synchronously switch to new frequency bins chosen according a pseudo-random sequence. The hop sequence followed by the primary users is not available at the CR nodes. Hence, in each hop duration, the task of the CR is to identify unoccupied frequency bins as quickly as possible and transmit its data over them, before sensing again

during the next hop period of the primary network.

Suppose the CR wishes to find a spectrum hole of bandwidth W Hz within a wider band of interest of bandwidth B Hz. Let $B = KW$ i.e., the bandwidth B consists of K contiguous, non-overlapping bins. We assume that, at any given hop duration, each active primary user occupies one of the K bins, as shown in Fig. 4.1. The shaded bins in Fig. 4.1 represent the bins occupied by the PUs at some given hop interval; there are L active PUs in the figure. The primary signal hopping pattern is modeled as an i.i.d. random sequence both across users and across time, and uniformly distributed over the K bins. The CR down-converts the received signal, band-limits it to B Hz, and samples it at a rate of $f_s \geq 2B$ samples/s. In this chapter, for simplicity, we assume that the number of active PUs, denoted by L, is known. In practice, since the number of active PUs typically varies very slowly compared to the bin occupancy patterns, it can be estimated and tracked based on the sensing outcomes. Let $\mathbf{u} \in \{0, 1\}^K$ represent the primary occupancy pattern, where $\mathbf{u}(k)$ takes the value 0 or 1, depending on whether the kth frequency band is free or occupied, respectively.

The CR collects $N_s = NM$ data samples, groups them into M frames of N samples each, and applies an N-point FFT on each frame (possibly, with windowing, to control the side-lobes). Thus, the sensing duration is N_s/f_s s. Also, N is chosen to be a positive integer multiple of K. In each frame, multiple FFT bins are grouped to represent the samples from each of the K PU bands. This is done to reduce the spectral leakage due to the FFT. For every N_s, there are many combinations of M and N possible. In our analysis, we fix N and vary M, since, in practical implementations, the FFT size N is fixed based on the hardware capability. The received baseband samples in m^{th} frame

are represented as

$$\bar{y}_m = \bar{x}_m + \bar{z}_m, \quad m = 0, 1, \ldots, M - 1,$$

where $\bar{y}_m \triangleq \left[y_m[0], y_m[1], \ldots, y_m[N-1] \right]^T$, $\bar{x}_m \triangleq \left[x_m[0], x_m[1], \ldots, x_m[N-1] \right]^T$ and $\bar{z}_m$ $\triangleq \left[z_m[0], z_m[1], \ldots, z_m[N-1] \right]^T$ represent the received samples, the received PU signal component and the AWGN component at the receiver, respectively. We assume that $z_m[n] \sim \mathcal{CN}(0, \sigma^2)$, and i.i.d. across all m, n, where $\mathcal{CN}(\mu, \nu)$ represents a circularly symmetric complex Gaussian distribution with mean μ and variance ν.

Let Q represent the $N \times N$ FFT matrix, with $(p, q)^{\text{th}}$ entry equal to $\frac{1}{\sqrt{N}} \exp\left\{ -j \frac{2\pi pq}{N} \right\}$, $0 \leq p, q \leq N - 1$. Let $\bar{Y}_m \triangleq Q\bar{y}_m$, $\bar{X}_m \triangleq Q\bar{x}_m$ and $\bar{Z}_m \triangleq Q\bar{z}_m$. When the kth band is vacant, the samples in the (N/K) frequency bins corresponding to the kth band are modeled as

$$\bar{Y}_m(\ell) = \bar{Z}_m(\ell), \ \ell = \frac{kN}{K}, \frac{kN}{K} + 1, \ldots, \frac{(k+1)N}{K} - 1, \tag{4.1}$$

with $m = 0, \ldots, M - 1$. Here, m denotes the frame index within the sensing duration. On the other hand, when the kth band is occupied, the corresponding received samples at the CR are modeled as

$$\bar{Y}_m(\ell) = \bar{X}_m(\ell) + \bar{Z}_m(\ell), \ \ell = \frac{kN}{K}, \frac{kN}{K} + 1, \ldots, \frac{(k+1)N}{K} - 1, \tag{4.2}$$

with $m = 0, \ldots, M - 1$, where $\bar{X}_m(\ell)$ represents the received PU signal in the ℓth FFT bin and the Mth frame, including the effect of path loss, shadowing and multipath fading. Let SNR(p), $p = 0, \ldots, L - 1$ denote the SNR values of the L PUs, which are assumed to be known. The unknown SNR case can be handled by considering a conservative design that assumes a certain minimum SNR on occupied bins. The goal of the SS

module in the CR is to determine the presence (denoted $\mathcal{H}_1$) or absence (denoted $\mathcal{H}_0$) of the primary signal using the observations $\bar{Y}_m(k)$ described above.

Let $P(k)$ be the average energy over M consecutive frames in the kth band, and P_{tot} be the average energy over all K bands and M frames. Therefore, for $k = 0, 1, \ldots, K - 1$,

$$P(k) \triangleq \frac{1}{M} \sum_{m=0}^{M-1} \sum_{q=0}^{\frac{N}{K}-1} \left| Y_m \left(\frac{kN}{K} + q \right) \right|^2 , \ P_{\text{tot}} = \sum_{k=0}^{K-1} P(k).$$

Note that, the above model assumes that the CR is aware of the hop instant of the PU network, i.e., it is aware of the time instants when the PUs possibly change their frequency bands. The hop instants can be determined using the FAR algorithm itself during an initial sensing/calibration phase, where the CR node detects changes in the hop pattern over time. This technique, described in detail in Sec C.3.1, is used in the hardware implementation of FAR in this chapter.

4.2.2 The FAR algorithm

In this subsection, we describe the detection technique investigated in this chapter, namely, the FAR algorithm. The FAR algorithm has been previously proposed for detecting the presence or absence of an FH signal, as it offers a constant false alarm rate irrespective of the noise variance at the receiver [104]. The FAR algorithm is also computationally simple, and, therefore, easily implementable on a hardware platform with limited resources. The FAR decision statistic for the kth band is given by

$$T_M(k) \triangleq \frac{P(k)}{P_{\text{tot}}}, \ \ k = 0, \ldots, K - 1. \tag{4.3}$$

$$P_D(k, \gamma, M, N, K, SNR_{tot}) = 1 - \frac{\left(\frac{E_{1k}}{G_{1k}}\gamma\right)^{D_{1k}/2}}{\frac{D_{1k}}{2}\mathcal{B}\left(\frac{D_{1k}}{2}, \frac{F_{1k}}{2}\right)}$$
$$\times {}_2\mathcal{F}_1\left(\frac{D_{1k}}{2}, \frac{D_{1k} + F_{1k}}{2}; 1 + \frac{D_{1k}}{2}, -\frac{E_{1k}}{G_{1k}}\gamma\right), \quad (4.5)$$

$$P_{FA}(\gamma, M, N, K, SNR_{tot}) = 1 - \frac{(G_0\gamma)^{\frac{MN}{K}}}{\frac{MN}{K}\mathcal{B}\left(\frac{MN}{K}, \frac{D_0}{2}\right)}$$
$$\times {}_2\mathcal{F}_1\left(\frac{MN}{K}, \frac{MN}{K} + \frac{D_0}{2}; 1 + \frac{MN}{K}, -G_0\gamma\right), \quad (4.6)$$

The presence of a PU on the kth band is detected by comparing $T_M(k)$ with a threshold τ, as follows:

$$T_M(k) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau. \tag{4.4}$$

In the following section, we derive the per-band probabilities of false-alarm (P_{FA}) and signal detection (P_D) of the FAR algorithm, as a function of τ and M, for a given N.

4.3 Performance Analysis and Optimization

4.3.1 Probabilities of False Alarm and Detection

The following lemma presents the expressions for the false alarm and detection probabilities of the FAR algorithm, denoted by P_{FA} and P_D, respectively.

Lemma 1. *For the FAR algorithm-based detection scheme in (4.4), the signal detection and false-alarm probabilities are given by (4.5) and (4.6), where the parameters are as defined in (4.7), with $\gamma \triangleq \frac{\tau}{1-\tau} \in [0, \infty)$, $\mathcal{B}(\cdot, \cdot)$ denoting the beta function, and ${}_2\mathcal{F}_1(\cdot, \cdot; \cdot; \cdot)$ representing the Gauss' hypergeometric function.*

Proof. See Appendix C.1. □

$$G_0 \triangleq 2 - \frac{K-1}{K-1+\frac{1}{N}SNR_{tot}}, \quad D_0 \triangleq \frac{M}{K}SNR_{tot} + MN\left\{\frac{3}{2} - \frac{3}{2K} + \frac{K-2+1/K}{2K-2+\frac{4}{N}SNR_{tot}}\right\},$$

$$G_{1k} \triangleq 2 - \frac{1}{1+\frac{1}{N}SNR(k)}, \quad D_{1k} \triangleq \frac{MN}{K}\left\{\frac{3}{2} + \frac{1}{N}SNR(k) + \frac{1}{2+\frac{1}{N}SNR(k)}\right\},$$

$$E_{1k} \triangleq 2 - \frac{K-1}{K-1+\frac{1}{N}SNR_{tot}^{(k)}}, \quad F_{1k} \triangleq \frac{M}{K}SNR_{tot}^{(k)} + MN\left\{\frac{3}{2} - \frac{3}{2K} + \frac{K-2+1/K}{2K-2+\frac{4}{N}SNR_{tot}^{(k)}}\right\},$$

$$SNR_{tot} \triangleq \sum_{p=0}^{L-1} SNR(p), \quad SNR_{tot}^{(k)} \triangleq \sum_{p=0,\neq k}^{L-1} SNR(p), \quad and \; SNR(p) \triangleq \frac{\frac{1}{M}\sum_{m=0}^{M-1}|X_m(p)|^2}{\frac{\sigma^2}{K}}, \quad (4.7)$$

With the expressions for the false alarm and detection probabilities in hand, the per-band detection threshold to satisfy a target P_{FA} (or P_D) constraint can be fixed by using numerical techniques. Note that, $_2\mathcal{F}_1(a, b; c; d)$ converges if the real part of $c - a + b$ is > 0. Hence, K, N and M should satisfy $\{1 - \frac{2MN}{K} - \frac{D_0}{2}\} > 0$. Substituting for D_0, it is easy to verify that the function converges for all M and N, provided $K > 1$ and $L \geq 1$.

4.3.2 Optimum Sensing Duration

Clearly, a longer sensing duration results in more accurate sensing, but leaves less time within each hop duration for data transmission; we wish to find the right trade-off between the two effects. To this end, we now derive the sensing duration that maximizes the CR throughput, subject to a constraint on PU protection. When there are L active PUs, at any time, $K - L$ bands are available, and L bands are busy. On average, the CR correctly declares $(K - L)(1 - P_{FA}(\gamma, M, N, K, \text{SNR}_{tot}))$ bands as available, and it incorrectly declares $\sum_{k:\mathbf{u}(k)=1}(1 - P_D(k, \gamma, M, N, K, \text{SNR}_{tot}))$ bands as available. Let $0 \leq \alpha(k) < 1, k = 0, \ldots, K - 1$, represents the fractional data rate obtained by the CR when it transmits on the bands occupied by the PUs, after incorrectly declaring them to

be free. The value $\alpha(k) = 0, \forall k$ denotes the case where the CR node obtains no usable throughput when it transmits in bands that are actually occupied by PUs. In practice, the data rate achieved on bands occupied by PUs would depend on the relative locations of the PU transmitter, CR transmitter, CR receiver, and the CR and PU transmit powers. Assuming zero rate on such bands is a conservative approach, and serves as an additional protection to the PUs, along with the interference constraint, which will be elaborated on below. In this chapter, we consider the product of the time available for data transmission and the average bandwidth harvested by the CR, denoted by Π, as the performance metric of interest:

$$
\begin{aligned}
\Pi &\triangleq \mathbb{E}\{K^{(r)} + \alpha K^{(w)}\}W \times (N_h - N_s) \\
&= \Big[(K - L)(1 - P_{FA}(\gamma, M, N, K, \text{SNR}_{\text{tot}})) \\
&\quad + \sum_{k:\mathbf{u}(k)=1} \alpha(k)\,(1 - P_D(k, \gamma, M, N, K, \text{SNR}_{\text{tot}}))\Big]W(N_h - N_s),
\end{aligned}
\tag{4.8}
$$

Analytically optimizing the sensing duration and detection threshold to maximize the above cost function for a general $\alpha(k)$ is difficult. Therefore, we consider the special case of $\alpha(k) = 0, \ \forall k$. Then, the objective function reduces to

$$
\Pi = (K - L)(1 - P_{FA}(\gamma, M, N, K, \text{SNR}_{\text{tot}}))W(N_h - N_s), \tag{4.9}
$$

Observe that, in the above, as N_s increases, $1 - P_{FA}(\gamma, M, N, K, \text{SNR}_{\text{tot}})$ increases, while $N_h - N_s$ decreases; and hence there exists an optimal sensing duration that maximizes Π. Thus, we state the optimization problem as follows:

$$
\max_{N_s,\gamma}\{\Pi\} \text{ subject to } \min_{k:\mathbf{u}(k)=1} P_D(k, \gamma, M, N, K, \text{SNR}_{\text{tot}}) \geq P_{\min} \tag{4.10}
$$

$$\frac{G_0\gamma_{\min}}{{}_2\mathcal{F}_1\left(1, 1 - BM; 1 + AM; -G_0\gamma_{\min}\right)} \times$$

$$\left\{ \frac{B}{1 + AM} {}_2\Theta^{(1)}\left(\begin{matrix} 1, & 1 \middle| 1 - BM, & 2 - BM, & 2 \\ & 2 - BM \middle| 2, & 2 + AM \end{matrix} \middle| ; G_0\gamma_{\min}, \; G_0\gamma_{\min} \right) \right.$$

$$\left. + \frac{A(1 - BM)}{(1 + AM)^2} {}_2\Theta^{(1)}\left(\begin{matrix} 1, & 1 \middle| 1 + AM, & 2, & 2 - BM \\ & 2 + AM \middle| 2, & 2 + AM \end{matrix} \middle| ; G_0\gamma_{\min}, \; G_0\gamma_{\min} \right) \right\}$$

$$+ (A + B)\log\left(1 + G_0\gamma_{\min}\right) - A\log\left(G_0\gamma_{\min}\right)$$

$$+ (A + B)\psi^{(0)}(AM + BM) - A\psi^{(0)}(AM) - B\psi^{(0)}(BM) - \frac{N_h}{M(N_h - NM)} = 0. \quad (4.11)$$

where $P_{\min}$ is the minimum detection probability performance that the CR detector is required to satisfy. Since $N_s = NM$, finding the optimum N_s reduces to finding the optimum M, for a given FFT size N. The value of N can be considered to be fixed, as it is generally taken to be the largest value supported by the SS hardware. Now, for a given γ, it can be shown that Π is concave in $0 \leq M \leq \frac{N_h}{N}$. Also, for a given M, both $P_D(k, \gamma, M, N, K, \mathrm{SNR_{tot}})$ and $P_{FA}(\gamma, M, N, K, \mathrm{SNR_{tot}})$ decrease with γ. Hence, Π is maximized when γ is such that the constraint in (4.10) is satisfied with equality. The following lemma gives the equation which needs to be numerically solved to find the optimum value of M.

Lemma 2. *Let $\gamma_{\min}$ denote the value of γ that satisfies $\min\limits_{k:\mathbf{u}(k)=1} P_D(k, \gamma, M, N, K, SNR_{tot}) \geq P_{\min}$ with equality. Then, the value of M which maximizes the cost function in (4.9) is the solution to (4.11), with*

$$A \triangleq \frac{N}{K}, \quad B \triangleq \frac{\left[\left(N - \frac{N}{K}\right) + \frac{1}{K}SNR_{tot}\right]^2}{\left(N - \frac{N}{K}\right) + \frac{2}{K}SNR_{tot}}, \quad (4.12)$$

and where $\psi^{(0)}$ is the digamma function, and ${}_2\Theta^{(1)}(\cdot)$ is a Kampé de Fériet-like function [105],

$$
{}_2\Theta^{(1)}\left(\begin{matrix} a_1, \; a_2 \big| b_1, \; b_2, \; b_3 \\ c_1 \big| d_1, \; d_2 \end{matrix} \; \middle| \; ; x_1, \; x_2 \right) \triangleq \sum_{m=0}^{\infty} \frac{(a_1)_m (b_1)_m (b_2)_m (b_3)_m}{(c_1)_m (d_1)_m (d_2)_m} \; \frac{x_1^m}{m!}
$$
$$
\times \, {}_3\mathcal{F}_2(a_2, b_2 + m, b_3 + m; d_1 + m, d_2 + m; x_2) \quad (4.13)
$$

defined in (4.13). Also, ${}_3\mathcal{F}_2(\cdot, \cdot, \cdot; \cdot, \cdot; \cdot)$ is a hypergeometric function, and $(a)_m \triangleq \frac{\Gamma(a+m)}{\Gamma(a)}$ is the Pochhammer symbol.

Proof. See Appendix C.2. □

Note that the infinite series of the function ${}_2\Theta^{(1)}(\cdot)$ as given by (4.13) converges very fast. In our experiments, the result obtained from a truncated series with 30 terms was found to be accurate to four decimal places.

4.4 Results

4.4.1 Monte Carlo Simulations

Our simulation setup is chosen to match the hardware setup explained in the previous section, with $N = 64, M = 16, K = 8$ and $L = 2$ PUs. We consider the performance of the detector for each of the different bands. For ease of presentation, suppose that the two PUs are active in bands C_0 and C_7, at a given point in time. For evaluating the algorithms, it is sufficient to condition on this particular occupying pattern, by the symmetry of the problem. That is, we get the same CR performance conditioning on any pair of occupied bins. In Fig. 4.2, the accuracy of theoretical expressions derived for P_{FA} and P_D are compared with simulations. The results are presented for the primary SNR values of -10 and -5 dB at C_0 and C_7, respectively. The P_{FA} curve is shown for

the empty band C_1. The accuracy of expressions in (4.5) and (4.6) is clear from the plot.

In Fig. 4.3, we compare the detection performance of the FAR algorithm with ED [106], with and without uncertainty in the noise variance. The noise uncertainty model assumed is the same as in past work [38], namely, that the noise variance is unknown, but lies in a range of $[\sigma_n^2 - x \text{ dB}, \sigma_n^2 + x \text{ dB}]$, where x is the noise variance uncertainty, and σ_n^2 is the nominal noise variance. Then, the detector is designed to meet a false alarm probability target of 0.01 at a noise variance of $\sigma_n^2 + x$ dB, and the probability of detection performance is evaluated at a noise variance of $\sigma_n^2 - x$ dB. The plot shows that the FAR algorithm outperforms ED, and offers about 0.5 to 1 dB improvement in the primary SNR required to achieve a given probability of detection. Thus, the FAR is a better decision statistic compared to the energy in the band, for detection of FH primary signals. In Fig. 4.4, we plot the effective CR throughput as a function of the sensing duration. For larger primary SNR, the highest CR throughput is obtained at a shorter sensing duration, as expected. Also, in terms of the effective throughput, the FAR and ED perform almost equally well. This is because the throughput is a relatively insensitive function of the detector performance, and, hence, detectors with similar performance would yield average throughputs that are only marginally different from each other.

In Fig. 4.5, we plot the simulated optimal throughput (i.e., simulated value of the cost function in (4.9)) and its corresponding theoretical throughput calculated using the expressions in (4.5) and (4.6), for various SNR values. It is seen that in the low SNR regime, the accuracy of theoretical calculations become looser. This happens because of the inaccuracies in the approximation used in Lemma 1, as highlighted by Patnaik [107].

As SNR increases, the approximation becomes more and more tight.

Figure 4.6 shows the variation of the optimal value of M as a function of the interference limit $P_{\min}$. The hopping duration N_h is set to 1024, $L = 2$ primary users, and the SNR values are fixed to be $[-5, -5]$ dB at $[C_0, C_7]$, respectively. The theoretical curves are obtained by numerically solving (4.11) to obtain a real-valued M. We then evaluate the throughput for the two nearby integer values of M, and pick the optimal M as the value that offers the better throughput. For obtaining the simulated curves, we sweep over a range of detection thresholds and different values of M, and pick the combination that offers the best CR throughput. The good match between theoretical and simulated curves validates the optimization of the CR throughput presented in Sec. 4.3.2. Also, we notice that as N varies, for each given $P_{\min}$, the optimal M is such that NM is roughly constant. For example, at $P_{\min} = 0.9$, the optimal M is 5, 10 and 21 for $N = 16, 32$ and 64, respectively. This is because the detection performance, and, consequently, the effective throughput, is primarily determined by the sensing duration, which equals NM.

In Fig. 4.7, the variation of theoretical throughput, calculated using (4.8) is plotted as a function of threshold τ, for $N_h = 1024$, $N = 64$ and SNR values $[-5, -5]$ dB at $[C_0, C_7]$, respectively. For illustration purposes, the value of α is fixed to be 0.5 in both $[C_0, C_7]$. The region of τ over which the objective function is concave varies as M increases. As mentioned earlier, we need to resort to numerical techniques to find the region of (γ, M) over which the optimization is concave. For any positive $\alpha(\cdot)$, the throughput achieved through the FAR algorithm is better than the case of $\alpha(k) = 0, \forall k$.

4.4.2   Experimental Results from the Lyrtech SFF SDR DP

For the experimental results, we generated a pure sinusoidal FH primary signal using the National Instruments PXI signal generator, and evaluated the performance at the band corresponding to C_0, with a center frequency of 393.5 MHz.

Figure 4.8 shows the variation of probability of detection obtained through simulations and experiments, at various values of SNR. It can be seen that the trend observed in our experiments match well with the trend seen through simulations, allowing for an implementation loss of about 1 dB. As expected, the probability of detection decreases as the threshold increases in both cases, with nearly the same trend.

In Fig. 4.9, we plot the the Receiver Operating Characteristic (ROC) curves for different values of M and primary SNR. As expected, the detection performance improves with M and SNR. We observe that the experimental curves follow the same trends as the theoretical curves, allowing for an implementation loss of about 1 dB in the primary SNR.

Finally, in Fig. 4.10, we show the normalized optimal throughput of the CR, normalized to its maximum attainable value at the given $P_{\min}$, as a function of the sensing duration N_s, comparing the throughput observed from the DP with that observed via simulations. The experimental results were generated by using a CR transmitter that sends data at a rate of 20.833 Msps, a primary transmit power of -107.5 dBm, a hopping duration of $N_h = 6.5$ ms, and about 5 m distance between the primary transmitter and CR spectrum sensing node. The simulation results were generated using the setup described in the previous subsection, at a primary SNR of -10 dB at the CR node. The
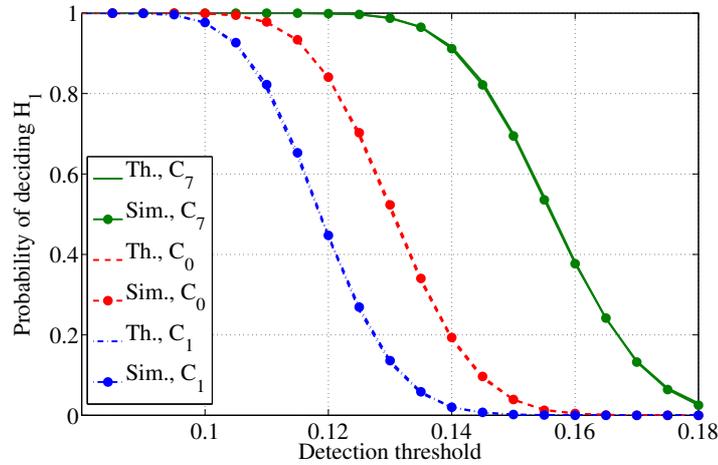
Figure 4.2: Comparison of theoretical and simulation results for the probability of deciding $\mathcal{H}_1$, for C_0, C_1 and C_7, as a function of the detection threshold. The curve marked C_1 corresponds to the false alarm probability curve, as the PU is not present on bin C_1.

good match between the two sets of plots is clear from the graph, validating our implementation. Also, the optimal sensing duration is larger for larger minimum detection probability performance $P_{\min}$.

4.5 Conclusions

In this chapter, we considered the problem of spectrum sensing in the presence of a multiuser frequency-hopping primary network. We theoretically analyzed the performance of the FAR algorithm, and validated the results through simulations. The sensing duration that maximizes the throughput of the CR system, under a constraint on the interference to the primary network was derived. A technique to synchronize the CR system with the primary hopping instants was presented. The FAR algorithm was implemented on Lyrtech SFF SDR DP and its performance was benchmarked by the ROCs obtained from Monte Carlo simulations. An implementation loss of about 1dB was observed in the hardware implementation.
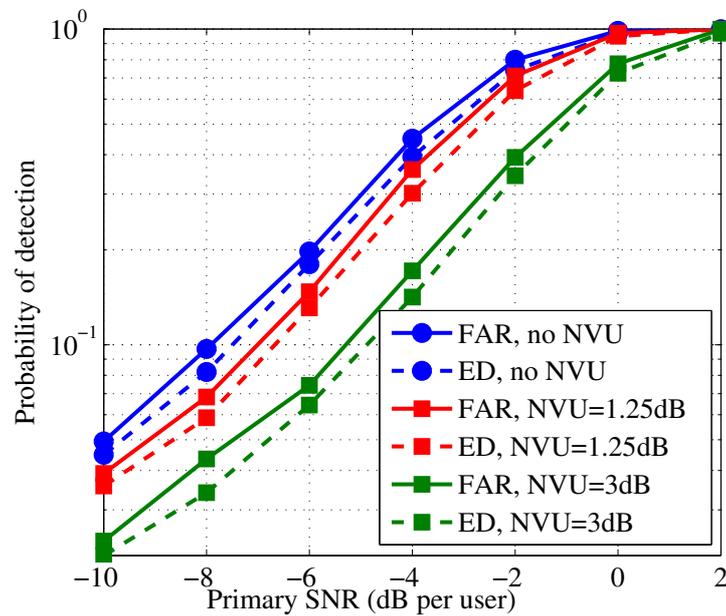
Figure 4.3: Comparison of FAR with the conventional ED, with and without noise variance uncertainty. Here, $N = 64$, $M = 128$, $L = 3$, and the detectors are designed with a target false alarm probability of 0.01.
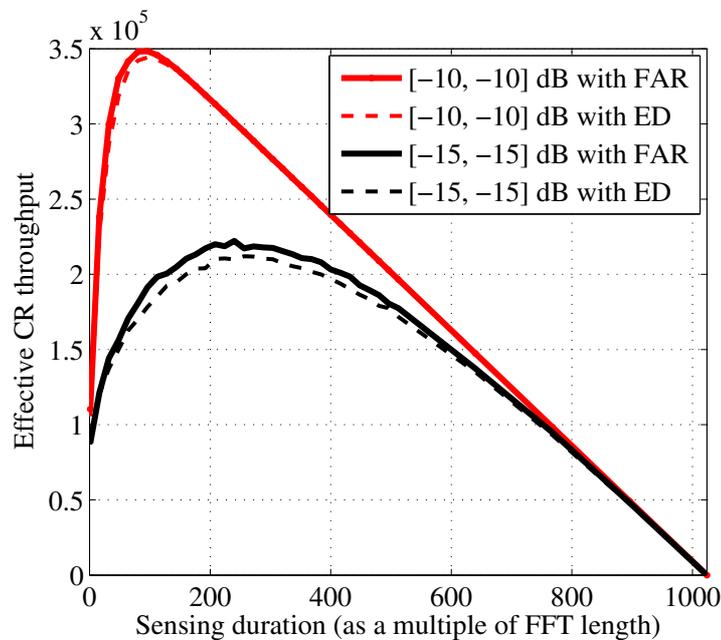
Figure 4.4: Comparison of CR throughput obtained by FAR algorithm with that of ED, obtained through hardware implementation.
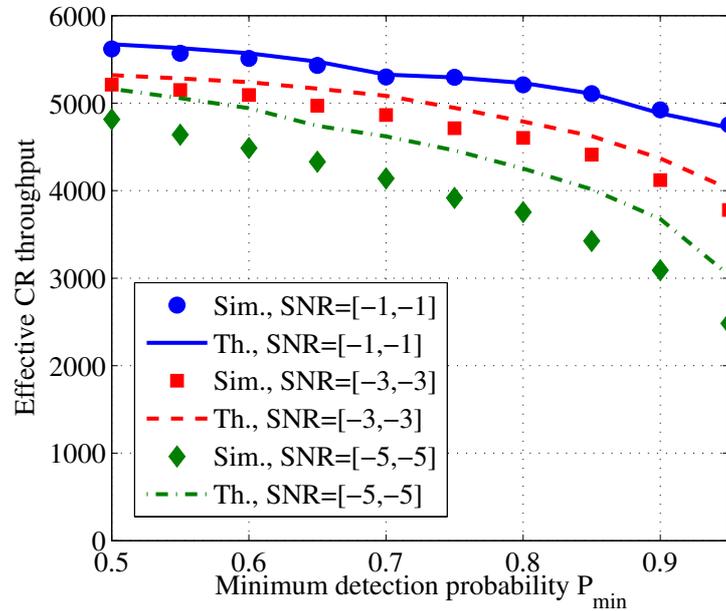
Figure 4.5: Optimal throughput for $N = 64$, $N_h = 1024$. For the simulation result, the optimal throughput was obtained by sweeping a range of M and threshold, and choosing the pair that offered the best throughput.
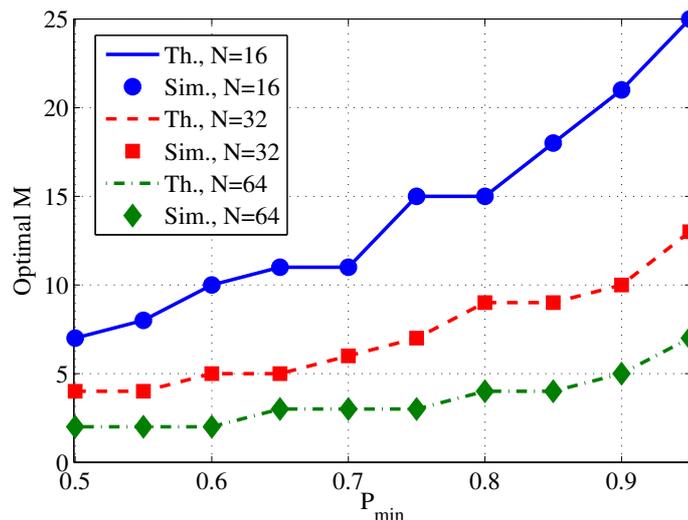
Figure 4.6: Comparison of optimal number of frames M for different values of the FFT size N, for $L = 2$, and $N_h = 1024$ samples. Notice that as N varies, the optimal M varies such that NM is roughly the same, for each given $P_{\min}$.
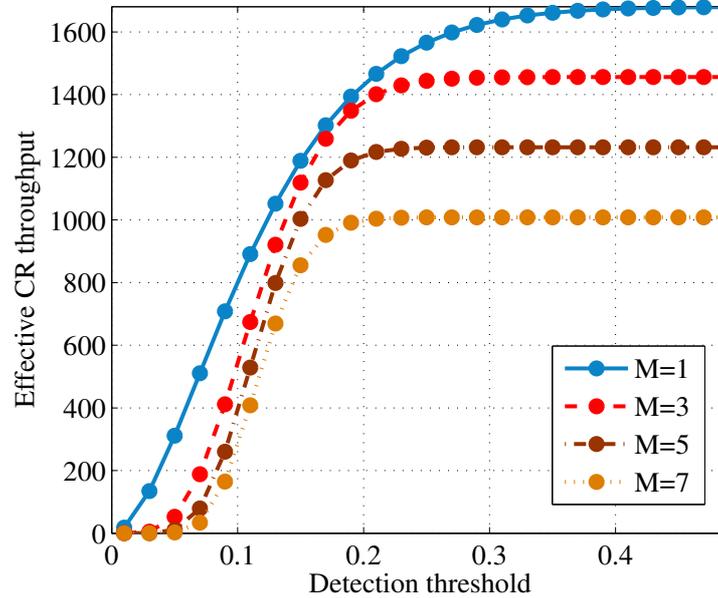
Figure 4.7: Variation of theoretical throughput Vs. τ, for $N_h = 1024$, $N = 64$, with SNR=$[-5, -5]$ dB, and $\alpha = [0.5, 0.5]$ for $[C_0, C_7]$.
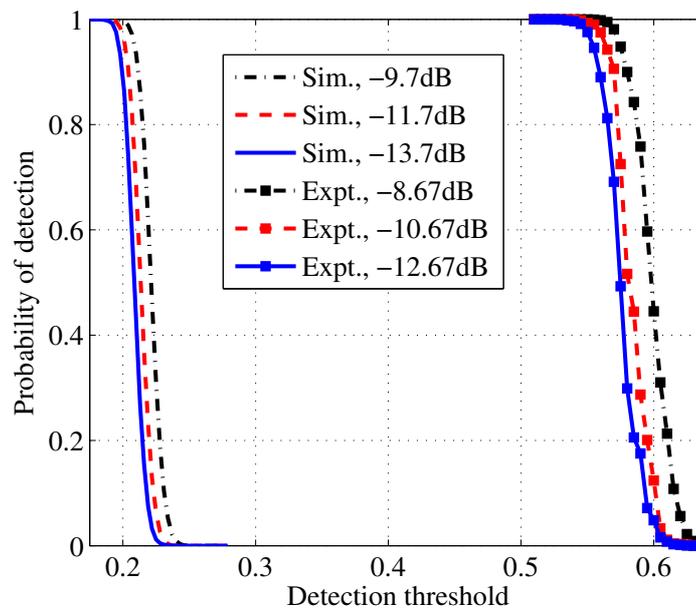
Figure 4.8: Comparison of P_{FA} and P_D from simulations and experiments, for $M = 128$ at different SNRs. The implementation loss is about 1 dB.
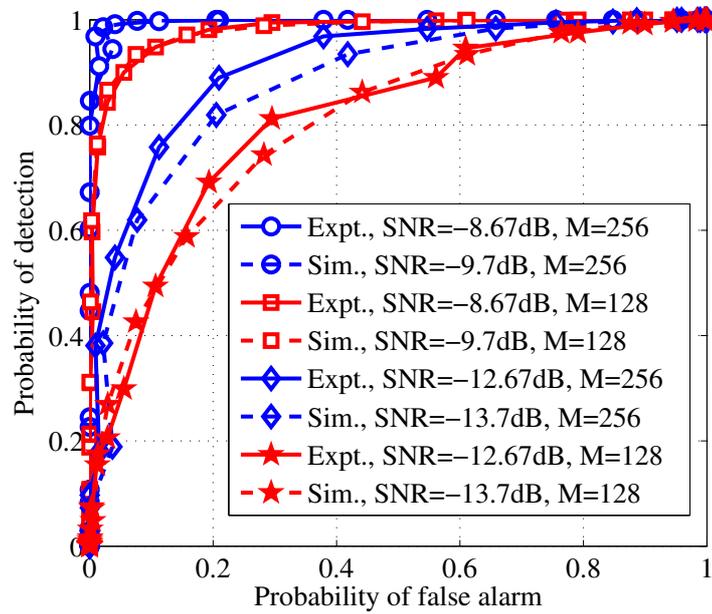
Figure 4.9: Comparison of ROCs from simulations and experiments, at different M and SNRs. The implementation loss is about 1 dB.
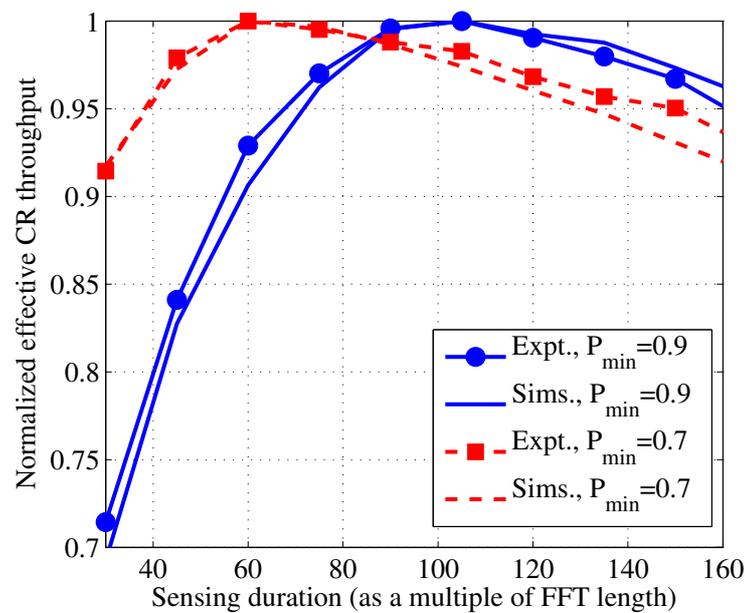
Figure 4.10: Optimum CR throughput Vs. N_s, comparing the hardware implementation with simulated curves.

Chapter 5

Zero-Crossings Based Spectrum Sensing Under Noise Uncertainties

5.1 Introduction

One of the key challenges in Cognitive Radios (CR) [1] is Spectrum Sensing (SS), which is the well-studied binary hypothesis testing problem of determining the presence or absence of a primary signal in a given frequency band of interest [30, 89]. In the future, CRs are envisioned to operate in various wireless environments, and in the presence of interference, changing noise statistics, etc ([62], [30], and references therein). Therefore, techniques used for SS need to be capable of handling various fading environments, primary signal models and different types of noise distributions. Hence, the class of Goodness-of-Fit Tests (GoFT) [51] is a natural choice for SS, where the problem reduces to accepting or rejecting the noise-only hypothesis, under a constraint on the false alarm probability. This chapter explores the benefits and drawbacks of GoFTs for CR-SS applications, in the presence of different types of model uncertainties.

Construction of a GoFT-based detector requires knowledge of the noise statistics. The

noise process in most communications-related applications consists of a Gaussian component (also known as the *background noise* or *thermal noise*), a controllable interference component (the so-called *class A* noise), and an uncontrollable impulsive component (the so-called *class B* noise) [55]. Depending on the application and communication scenario, either class A, class B, or both exist in the noise model. Much of the existing works in the CR literature consider SS in the Gaussian component alone, i.e., they assume the noise distribution to be i.i.d. Gaussian. Further, for the most part, the literature assumes the noise variance to be known [89].[1] An exception to this approach is [108], which studies SS under only non-Gaussian noise model, which ignores the Gaussian component. However, even if the non-Gaussian component is weak compared to the Gaussian component, the effect of the latter cannot be ignored in practice [55]. Another important aspect for the SS design is the need for knowledge of the parameters of the noise model, such as its variance [38,55,109–111]. Therefore, the performance of SS is affected by the two kinds of uncertainties in the knowledge of the noise: imperfect knowledge of the parameters of the distribution, which we refer to as *Noise Parameter Uncertainty (NPU)*, and imperfect knowledge of the distribution itself, which we refer to as *Noise Model Uncertainty (NMU)*. Additionally, in many scenarios, the noise process may be correlated either spatially or temporally [112], or due to digital filtering at the receiver [113]. To the best of authors' knowledge, SS under the above stated noise uncertainties, and GoFT under colored noise is not addressed in the GoFT based SS literature so far.

---

[1]The GoFT literature for testing against Gaussianity spans over a century now, and is an active area of research even today. A brief survey of the well-known and widely used techniques is presented in Appendix E.

Energy Detection (ED) is a simple technique for SS, where the signal energy in the frequency of interest is measured over a sensing duration and compared to a threshold. However, in the presence of the background noise alone, it performs poorly in the low SNR regime under the NPU [38]. In the presence of the non-Gaussian components, ED fails to satisfy the false-alarm probability constraint because of the underlying heavy-tailed, non-Gaussian distribution with infinite variance. Recently, an Anderson-Darling statistic [114] based test (ADD) [52] and an Ordered Statistics based Detector (OSD) [54] were proposed, which outperform ED (in the Neyman-Pearson sense) when the primary signal is a constant with Rayleigh fading and i.i.d. Gaussian noise, with known variance. However, because of the underlying assumptions and construction, ADD and OSD are susceptible to both NPU and NMU. The Blind Detector (BD), proposed in [53], is robust to NPU, but only handles i.i.d. Gaussian noise, i.e., it is not robust to NMU.

In this chapter, a detector called *weighted Zero-Crossings Detector* (WZCD), based on the Zero-Crossings (ZC) and Higher Order Crossings (HOC) in the received observations is proposed, which is a generalization of a detector proposed by Kedem and Slud [115]. The main contributions of this chapter are

- A weighted ZC based detector is proposed. Given a target false alarm probability, near-optimal detection thresholds are obtained for uniform and exponential weights (Sec. 5.4).

- The proposed detector is shown to be robust to both NMU and NPU (Sec. 5.5).

- For the specific case where the noise has two components namely the background component and class A component and with both components being modeled by Gaussian distribution with different variances, it is analytically shown that the ED

and ADD do not satisfy the given false-alarm constraint. The actual applicability and utility/efficacy of the BD, which is much wider than that discussed in [53] is highlighted.

- Under colored noise, the expected number of ZCs and HOCs when the noise distribution is Gaussian, are derived through a generalized level-crossings lemma (Sec. 5.6). The robustness of the proposed detector to both NMU and NPU in the case of colored noise is discussed.

- In a detailed simulation study, the performance of the proposed detectors is compared with the BD under various primary signal models operating in different noise and fading environments (Sec. 5.7).

Thus, we conclude, in Sec. 5.9, that the proposed WZCD is a promising technique for detecting primary signals at low SNR, under both NPU and NMU.

We start by describing the system model.

5.2 System Model

Consider a CR node with M observations denoted by $Y_i, i \in \mathcal{M} \triangleq \{1, 2, \cdots, M\}$. It is assumed that each Y_i is real valued ([52, 53]). In the GoFT formulation, the problem is to either reject or accept the null hypothesis

$$\mathcal{H}_0 \quad : \quad Y_i \sim f_{\mathbb{N}}, \quad i \in \mathcal{M}$$

and the threshold is chosen such that a constraint $\alpha_f \in [0,1]$ on the probability of false alarm is satisfied, i.e.,

$$p_f \triangleq \mathcal{P}\{\text{reject } \mathcal{H}_0 | \mathcal{H}_0 \text{ is true}\} \leq \alpha_f. \tag{5.1}$$

Here, $f_{\mathbb{N}}$ represents the distribution of the noise process, i.e., the distribution of Y_i under $\mathcal{H}_0$. First, consider the case where the noise observations are i.i.d. Following Middleton [55], the observations under $\mathcal{H}_0$ are modeled as

$$Y_i = Y_i^{(\mathcal{G})} + Y_i^{(\mathcal{A})} + Y_i^{(\mathcal{B})}, \quad i \in \mathcal{M}, \tag{5.2}$$

where the Gaussian component $Y_i^{(\mathcal{G})} \sim f_{\mathcal{G}} \overset{d.}{=} \mathcal{N}(0, \sigma_{\text{G}}^2)$, and $\mathcal{N}(\mu, \sigma^2)$ represents a Gaussian distribution with mean μ and variance σ^2. The class A and B noise components are denoted by $Y_i^{(\mathcal{A})}$ and $Y_i^{(\mathcal{B})}$, respectively. Middleton has shown that the PDF of class B component alone can be well-approximated by a two-parameter symmetric α-stable ($\mathcal{S}\alpha\mathcal{S}$) distribution, i.e., the characteristic function of $Y_i^{(\mathcal{B})}$, denoted by Φ_B, is given by [55, eq. (88)]

$$\Phi_B(\omega, \gamma_0, \alpha) = \exp\left(-\gamma_0 |\omega|^\alpha\right), \quad \gamma_0 > 0, \ 0 < \alpha \leq 2. \tag{5.3}$$

Vastola [109] has shown that the PDF of $Y_i^{(\mathcal{G})} + Y_i^{(\mathcal{A})}$ can be well-approximated by the *ϵ-mixture model* [110]. Using this approximation, we can write

$$Y_i^{(\mathcal{G})} + Y_i^{(\mathcal{A})} \sim (1 - \epsilon)f_{\mathcal{G}} + \epsilon f_{\mathcal{I}}, \tag{5.4}$$

where $f_{\mathcal{I}}$ represents a distribution which has heavier tails as compared to $f_{\mathcal{G}}$, for example, a Laplace distribution [110], or another Gaussian distribution with a variance

larger than that of $f_\mathcal{G}$ [111]. Typically, σ_I^2, the variance of $f_\mathcal{I}$, satisfies $\sigma_\mathrm{I}^2 \gg \sigma_\mathrm{G}^2$. The mixing parameter $0 < \epsilon \ll 1$ depends on the parameters of the PDF of the class A model ([55], [109]).

In this chapter, our aim is to design a GoFT that is robust to the following two noise uncertainties:

1. The *noise model uncertainty* (NMU) is caused due to the imperfect knowledge of noise distribution $f_\mathbb{N}$. The presence of either class A, B or both in $f_\mathbb{N}$ depends on the physical environment [55, Tab. I]. Also, the distribution of $f_\mathcal{I}$ in (5.4) can be modeled to be Gaussian, Laplace, or Cauchy ([110], [111]). However, the background Gaussian noise is always present [55].

2. The *noise parameter uncertainty* (NPU) arises due to the inaccurate knowledge of the parameters in $f_\mathbb{N}$, i.e., σ_G^2 [38], σ_I^2 [110], ϵ [109], and α [55].

In the following section, we present a brief note on some of the existing GoFTs, proposed and studied in the context of SS in CR.

5.3 Existing GoFT for Spectrum Sensing

5.3.1 Energy Detector (ED)

The ED can be proposed as a GoFT, and has the following critical region

$$\left\{ \{Y_1, Y_2, \cdots, Y_M\} \in \mathbb{R}^M : E \triangleq \sum_{i=1}^{M} Y_i^2 > \tau_{ED} \right\}, \tag{5.5}$$

where τ_{ED} is the detection threshold, chosen such that the p_f is at a given desired level α. In particular, when the observations are i.i.d. and $f_\mathbb{N} = f_\mathcal{G} \overset{d.}{=} \mathcal{N}(0, \sigma_\mathrm{G}^2)$ with a known

σ_{G}^2, the statistic E is chi-square distributed with $2M$ degrees-of-freedom. In this case, it is easy to show that

$$\tau_{ED} = \gamma_{\mathrm{inc}}^{-1}\left(1 - \alpha, \frac{M-1}{2}, \frac{2\sigma_{\mathrm{G}}^2}{M}\right),\tag{5.6}$$

where $\gamma_{\mathrm{inc}}^{-1}(x, A, B)$ represents the normalized inverse gamma cumulative distribution function evaluated at x, with parameters A and B [116].

5.3.2 Anderson-Darling Statistic Based Detector (ADD)

The Anderson-Darling statistic [114] based GoFT was proposed in the CR context by Wang et al. [52]. When $f_{\mathbb{N}}$ is completely known (with i.i.d. observations, and known variance), and the observations are ordered such that $Y_1 \leq Y_2 \leq \cdots \leq Y_M$, the Anderson-Darling statistic is defined as

$$A_{c,M}^2 \triangleq -\frac{\sum_{i=1}^{M}(2i-1)(\ln Z_i + \ln(1 - Z_{M+1-i}))}{M} - M\tag{5.7}$$

with $Z_i \triangleq F_{\mathbb{N}}$, the CDF of noise observations. The ADD has the following critical region [52]

$$\left\{Y_i, i \in \mathcal{M} : A_{c,M}^2 \geq \tau_{ADD}\right\},\tag{5.8}$$

where τ_{ADD} is chosen such that p_f is set to a level α. For a given α, and for moderate values of M, τ_{ADD} satisfies [52]:

$$1 - \frac{\sqrt{2\pi}}{\tau_{ADD}}\sum_{\ell=0}^{\infty}\frac{(-1)^{\ell}\Gamma(0.5+\ell)}{\Gamma(0.5)\ell!}\exp\left(-\frac{\pi^2(4\ell+1)^2}{8\tau_{ADD}}\right)$$
$$\times (4\ell+1)\int_0^{\infty}\exp\left(\frac{\tau_{ADD}}{8(w^2+1)} - \frac{\pi^2 w^2(4\ell+1)^2}{8\tau_{ADD}}\right)dw = \alpha.\tag{5.9}$$

A table of thresholds for different values of p_f is given by Stephens [117].

5.3.3 Blind Detector (BD)

The BD was proposed by Shen et al. [53] as a robust detector under noise uncertainty. When $f_{\mathbb{N}} = f_{\mathcal{G}} \overset{d.}{=} \mathcal{N}(0, \sigma_{\mathrm{G}}^2)$, the construction of the BD is such that the test statistic is independent of σ_{G}^2. M observations are divided into n windows of m observations each, and the test statistic is constructed as follows. Define

$$X_l \triangleq \sum_{u=0}^{m-1} \frac{Y_{ml-u}}{m}, \quad S_l^2 \triangleq \sum_{u=0}^{m-1} \frac{(Y_{ml-u} - X_l)^2}{m-1}, \tag{5.10}$$

$$\text{and } B_l \triangleq \frac{X_l}{S_l/\sqrt{m}}, l = 1, \cdots, n. \tag{5.11}$$

Now, the Anderson-Darling statistic is formed based on $B_l, l = 1, \cdots, n$ using (5.7), which is defined as

$$A_{c,n}^2 \triangleq -\frac{\sum_{i=1}^{n}(2i-1)(\ln C_i + \ln(1 - C_{n+1-i}))}{n} - n, \tag{5.12}$$

where $C_l = F_s(B_l)$, where F_s represents the CDF of a student-t distributed random variable with parameter $m - 1$. Observe that F_s does not depend σ_{G}^2. Therefore, the BD is robust to noise variance uncertainty, and has the following critical region

$$\left\{ Y_i, i \in \mathcal{M} : A_{c,n}^2 \geq \tau_{BD} \right\}. \tag{5.13}$$

For large enough n, the optimal threshold τ_{BD} is calculated in a similar way to ADD, using (5.9).

Remark: In our simulations, we have found that BD can be applied even in much weaker scenarios that those noted by Shen et al [53]. In this chapter, along with a new zero-crossings based detector, we also present an analysis of BD that captures some of

its key strengths. The applicability of BD is discussed in Appendix D.4.

In the next section, we propose a detector based on the weighted Zero-Crossings (ZC) in the observations $Y_i, i \in \mathcal{M}$. We discuss the robustness of the proposed detector to both NMU and NPU in Sec. 5.5.

5.4   Weighted Zero-Crossings Based Detection

Zero-crossings based detection was first proposed by Kedem and Slud [115] as a simple non-parametric detection strategy for testing against Gaussian samples. Here, we generalize this approach by considering zero-crossings, and study its performance in the presence of NMU and NPU. In our simulations, for a large class of primary signal models, the first few values of $\Delta_{k,M}$ were found to be significantly larger than those for higher values of k. Therefore, if the weights are chosen such that the lower order ZCs are weighed larger than the higher order ZCs, the detection performance can be significantly improved. The corresponding test statistic is constructed as follows. Let ∇^k denote the kth order difference operator on Y_i, defined as

$$
\begin{aligned}
\nabla Y_i &\triangleq Y_i - Y_{i-1} \\
\nabla^2 Y_i &= \nabla(\nabla Y_i) = Y_i - 2Y_{i-1} + Y_{i-2} \\
&\vdots \\
\nabla^k Y_i &= \sum_{j=0}^{k} \binom{k}{j}(-1)^j Y_{i-j}, \ \ i \geq k+1.
\end{aligned}
\tag{5.14}
$$

A kth *order zero-crossing* in the observations $\{Y_i, i \in \mathcal{M}\}$ is said to occur if the sign of $\nabla^{k-1}Y_i$ is different from that of $\nabla^{k-1}Y_{i+1}$. Let $D_{k,M}$ denote the number of kth order zero-crossings across M samples. Note that, we define $\nabla^0 Y_i \triangleq Y_i$. Now, let $\Delta_{j,M}$, and $\mu_{j,M}$ be defined as

$$\Delta_{j,M} \triangleq \begin{cases} D_{1,M}, & j = 1, \\ D_{j,M} - D_{j-1,M}, & j = 2, \cdots, k-1 \\ (M-1) - D_{k-1,M}, & j = k, \end{cases} \tag{5.15}$$

$$\mu_{j,M} \triangleq \mathbb{E}\Delta_{j,M}, \qquad\qquad j = 1, \cdots, k, \tag{5.16}$$

where $\mathbb{E}(\cdot)$ denotes the expectation operator. Observe that $\sum_{j=1}^{k} \Delta_{j,M} = M - 1$. A goodness-of-fit measure Ψ_w^2 up to a given order k can be defined as

$$\Psi_w^2 \triangleq \sum_{j=1}^{k} w_j \frac{(\Delta_{j,M} - \mu_{j,M})^2}{\mu_{j,M}}. \tag{5.17}$$

For a given set of weights w_j, a Ψ_w^2 Statistic based Detector (Ψ_wSD) is given by

$$\Psi_w^2 \underset{\sim\mathcal{H}_0}{\overset{\nsim\mathcal{H}_0}{\gtrless}} \tau_w^\Psi, \tag{5.18}$$

and τ_w^Ψ is chosen such that $\mathcal{P}\{\Psi_w^2 > \tau_w^\Psi | \mathcal{H}_0\} \le \alpha_f$, for some target false alarm probability $\alpha_f \in [0, 1]$. We refer to the detector based on (5.17) as a *Weighted Zero-Crossings based Detector* (WZCD). Note that, when $w_j = 1, j = 1, 2, \ldots, k$, the detector reduces to the classical ZCD [115].

Now, as seen from (5.16) and (5.17), for a specific set of weights, the construction of the WZCD depends only on the knowledge of $\mu_{j,M}$ for $j = 1, 2, \cdots$. When $f_{\mathbb{N}} = f_{\mathcal{G}} \overset{d.}{=}$

$\mathcal{N}(0, \sigma_\mathrm{G}^2)$, for moderately large M, it is known that [118]

$$\mathbb{E}D_{k,M} = (M-1)\left\{\frac{1}{2} + \frac{1}{\pi}\sin^{-1}\left(\frac{k-1}{k}\right)\right\}, \tag{5.19}$$

Hence, $\mu_{k,M}$ can be easily calculated by substituting (5.19) in (5.15) and (5.16). Also, it has been observed that for most processes, calculating and using $D_{1,M}, \cdots, D_{8,M}$, i.e., up to $k = 8$ are enough, in the sense that for $k > 9$, the ZCs do not contribute much to the performance [118].

Next, we consider the following two cases.

A) Equal and unit weights: The past work [115] studied the statistic in (5.17) for $w_j = 1, j = 1, \cdots, k$, which will be denoted as Ψ_1^2. It was observed that the PDF of Ψ_1^2 is approximately the same for any discrete-time stationary ARMA process, and hence can be used to construct a GoFT against any such process. Additionally, it is known that for moderately large M, the Ψ_1^2 statistic can be approximated by a Pearson type III (chi-squared or gamma) distribution [118]. These PDF approximations have been studied in detail earlier, and are known to be highly accurate [119]. In particular,

$$\Psi_1^2 \sim \chi_3^2(11), \tag{5.20}$$

where $\chi_D^2(\lambda)$ is a non-central chi-square distribution with D degrees of freedom and non-centrality parameter λ. Hence, for a given target false alarm probability α_f, the detection threshold corresponding to using Ψ_1^2, denoted by τ_1^Ψ, satisfies

$$Q_{\frac{3}{2}}\left(\sqrt{11}, \sqrt{\tau_1^\Psi}\right) = \alpha_f, \tag{5.21}$$

where $Q_\kappa(\cdot, \cdot)$ represents the Marcum-Q function of order κ.

B) Exponential weights: As mentioned earlier, in our simulations, we found that weighing the lower order ZCs higher than the higher order ZCs can lead to significantly better performance. Motivated by this, we consider the exponential weighting case, i.e., $w_j \triangleq e^{-(j-1)}$. Let the corresponding WZCD test statistic be denoted by Ψ_e^2. Through simulations, it was observed that the tail of the distribution of Ψ_e^2 follows closely to that of an F-distribution with parameters 17.5 and 7 (denoted by $\mathcal{F}(17.5, 7)$), for moderately large M. The loss due to this approximation is negligible, and is quantified in the Sec. 5.7. Therefore, for a test based on Ψ_e^2, the near-optimal detection threshold τ_e^{Ψ} satisfies

$$1 - \mathcal{I}\left(\frac{17.5\tau_{m\Psi SD}}{17.5\tau_{m\Psi SD} + 7}, 8.75, 3.5\right) = \alpha_f, \tag{5.22}$$

where $\mathcal{I}(\kappa, a, b)$ represents the regularized incomplete beta function with parameters κ, a and b [116].

We note that the optimal threshold calculation for ADD [52] and BD [53] require the evaluation of an integral of an infinite series, which is computationally intensive, as opposed to the single integral calculation in (5.21) and (5.22).

To summarize, the detection procedure using the WZCD is as follows:

1. Fix $k = 9$. Collect $M = M' + k$ observations, $\{Y_i\}$.

2. Calculate the first-k ZCs and HOCs of $\{Y_i\}$ using (5.14). Using (5.19), calculate the expected zero-crossings. Denote this by $\mu^{(g)}$.

3. Construct the statistic Ψ_w^2 for appropriately chosen weights.

4. Compare Ψ_w^2 to τ_w^{Ψ}. For unit and exponential weights, use τ_1^{Ψ}, and τ_e^{Ψ} from (5.21) and (5.22), respectively. Declare $\mathcal{H}_0$ if $\Psi_w^2 < \tau_w^{\Psi}$, and *not* $\mathcal{H}_0$ otherwise.

In the following section, we will discuss the advantages offered by the WZCD, more specifically, its robustness to NMU and NPU.

5.5 Robustness to Noise Uncertainties

5.5.1 Noise Model Uncertainty

As mentioned earlier, NMU arises because of imperfect knowledge of $f_{\mathbb{N}}$. Along with the Gaussian noise, either class A, class B, or both can be present in $f_{\mathbb{N}}$. It is easy to see that, with class A noise, when Y_i is distributed as given in (5.4), with $f_{\mathcal{I}} \sim \mathcal{N}(0, \sigma_{\mathrm{I}}^2)$, the proposed detector is robust to both NMU and NPU, as the test statistic Ψ_w^2 is independent of both σ_{G}^2 and σ_{I}^2 (see (5.19)). To get some insight into the general $f_{\mathcal{I}}$ case, we present the following heuristic argument. Note that, if X represents a random variable from the $\mathcal{S}\alpha\mathcal{S}$ family or exponential family, then its PDF $p(X)$ can be written as [120]

$$p(X) = \int_0^\infty \left(\frac{1}{\sigma}\right) g\left(\frac{X}{\sigma}\right) h(\sigma)\,d\sigma, \tag{5.23}$$

where $g(\cdot)$ is the standard Gaussian PDF, and $h : \mathbb{R}^+ \to \mathbb{R}^+$ is a function that determines the distribution of X. For example, the Cauchy distribution ($\alpha = 1$) can be generated using (5.23), by choosing $h(\cdot)$ to be the Lévy distribution function with parameter 0.5. Similarly, the Laplace distribution can be generated using (5.23), by choosing $h(\cdot)$ to be the exponential distribution function. Therefore, the distribution of Ψ_w^2 can be expressed as scale-mixture of the Gaussian PDF. Since the proposed detector is independent of the variance of a Gaussian PDF, and can even handle infinite variance distributions, it is robust to NMU.

5.5.2 Noise Parameter Uncertainty

Since every distribution in the $\mathcal{S}\alpha\mathcal{S}$ and exponential families can be generated as scale-mixture of Gaussian distributions, when M is sufficiently large, the WZCD can be made robust to uncertainty in the parameter set (α, ϵ, $f_{\mathcal{I}}$, σ_{I}^2). Intuitively, the number of ZCs in the observations for a unimodal, symmetric distribution is approximately $M/2$, irrespective of its variance. That is, the probability that the distribution would take positive values will be approximately 0.5. Therefore, for large M, the distribution of the test statistic defined by (5.17) is approximately the same for all such symmetric distributions, thus making the statistic relatively independent of the parameters of the noise such as its variance.

In the following, we will show that ED and ADD do not satisfy their respective false-alarm constraints, even in the case where class B model is not present in $f_{\mathbb{N}}$. It is straightforward to see that when $f_{\mathcal{G}}$ and $f_{\mathcal{I}}$ are $\mathcal{N}(0, \sigma_{\mathrm{G}}^2)$ and $\mathcal{N}(0, \sigma_{\mathrm{I}}^2)$ respectively, $f_{\mathbb{N}} \sim \mathcal{N}(0, \epsilon\sigma_{\mathrm{G}}^2 + (1-\epsilon)\sigma_{\mathrm{I}}^2)$. As already seen, since the the statistics under BD and WZCD are not dependent on the Gaussian noise variance, they still meet the required false-alarm constraint, and their probability of detection remains unchanged. However, ED and ADD do not satisfy false-alarm constraints, as highlighed below.

Result 2. *Under the presence of only class A noise in $f_{\mathbb{N}}$, when $f_{\mathcal{I}}$ is $\mathcal{N}(0, \sigma_I^2)$, and when ED and ADD are designed oblivious to the presence of the impulsive noise i.e., their thresholds are fixed for $f_{\mathcal{G}}$, while the observations have the distribution $f_{\mathbb{N}}$, they satisfy their respective false-alarm constraints if and only if $\epsilon = 0$.*

In Appendix D.3, we will provide a sketch of the proof of the above result.

5.6 Expected HOCs for Correlated Gaussian Noise

In this section, we discuss the applicability of the WZCD for the case when the noise-only observations are correlated, and follows a Gaussian distribution. From the discussions in earlier section, for Ψ_eSD and Ψ_1SD, $\mu_{j,M}, j = 1, \cdots, k$ and the distribution of Ψ^2 statistic needs to be known. In the following, we derive a lemma which gives analytical expressions for the expected level-crossings in a Gaussian, correlated process of which, the expected zero-crossings is a special case. Later, some comments on the case of general $f_{\mathbb{N}}$, and the effect of NMU and NPU are studied.

Consider the case where the noise process is a (p, q) ordered Auto-Regressive Moving-Average (ARMA) Gaussian, i.e., $f_{\mathbb{N}} \overset{d.}{=} f_{\mathcal{G}}$, with a correlation matrix $\mathbf{R}_{\mathbb{G}}$. When $\mathbf{R}_{\mathbb{G}}$ is known, $\mu_{j,M}$ can be obtained in closed form, as shown through the following general Level-Crossings (LC) lemma, which highlights the connection between the expected number of LC (at any level ℓ) of a Gaussian process of known correlation structure and its normalized autocorrelation function. The expected number of ZC is a special case of this result, for $\ell = 0$.

Lemma 3. *For the observations $Y_1, Y_2, \cdots, Y_M$ from a known, stationary, Gaussian process $f_{\mathcal{G}}(\underline{y}; \mathbf{R}_{\mathbb{G}})$, the expected number of level crossings for any given level ℓ is given by*

$$\mathbb{E}D_{1,M} = 2(M - 1)\mathcal{Q}(\ell)(1 - \mathcal{Q}(\ell))(1 - \rho_X), \tag{5.24}$$

where $\mathcal{Q}(\cdot)$ is the Gaussian Q-function, and ρ_X is the first order normalized autocorrelation

value of a process $\{X_i, i \in \mathcal{M}\}$ which is defined as

$$X_i \triangleq \begin{cases} 1, & Y_i > \ell, \\ 0, & Y_i \leq \ell, \quad i \in \mathcal{M}, \quad and \end{cases} \tag{5.25}$$

$$\rho_X \triangleq \frac{cov(X_i, X_{i+1})}{var(X_i)} = \frac{\mathbb{E}X_i X_{i+1} - \mathbb{E}X_i \mathbb{E}X_{i+1}}{var(X_i)} \tag{5.26}$$

where $cov(\cdot, \cdot)$ and $var(\cdot)$ denote the cross- and auto-covariance functions, respectively.

Proof. See Appendix D.1. □

Corollary 4. *For $\ell = 0$, the result in (5.24) reduces to a known result on the expected number of ZCs [118]*

$$\mathbb{E}D_{1,M} = (M-1)\left(\frac{1}{2} - \frac{1}{\pi}\sin^{-1}\rho_{1,\mathbb{G}}\right), \tag{5.27}$$

where $\rho_{1,\mathbb{G}}$ is the first order normalized autocorrelation of $f_{\mathcal{G}}(\underline{y}; \mathbf{R}_{\mathbb{G}})$, given by

$$\rho_{1,\mathbb{G}} \triangleq \frac{cov(Y_i, Y_{i+1})}{var(Y_i)}, \quad i \in \mathcal{M} \tag{5.28}$$

Therefore, the expected number of first order ZCs depends on $\rho_{1,\mathbb{G}}$.

Proof. See Appendix D.2. □

Since $\{Y_i, i \in \mathcal{M}\}$ are Gaussian, $\nabla^k Y_i$ are also Gaussian for all k. Using this property and above result for $\ell = 0$, Kedem has shown that the HOC depends on $\rho_{2,\mathbb{G}}, \cdots, \rho_{M-1,\mathbb{G}}$ as [118]

$$\cos\left(\frac{\pi \mathbb{E}D_{k+1,M}}{M-1}\right) = \frac{\nabla^{2k}\rho_{k-1,\mathbb{G}}}{\nabla^{2k}\rho_{k,\mathbb{G}}}, \tag{5.29}$$

where $\nabla(\cdot)$ operates on the sequence $\rho_{k,\mathbb{G}}, k = 1, 2, \cdots$.

Now, consider the general case, where the noise-only observations are correlated and distributed as $f_{\mathbb{N}}$ along with the presence of class A and B models. As in the previous case, we can argue that the distribution of background noise and class A component combined follows a correlated Gaussian distribution. Also, let the implusive noise (class B) follow a correlated $\mathcal{S}\alpha\mathcal{S}$ distribution. Similar to the case of i.i.d. observations discussed in Sec. 5.5, even in this case, approximating the non-Gaussian noise models as a mixture Gaussian process is found to be sufficiently accurate [121]. Hence, the WZCD is robust to both NMU and NPU, even for correlated observations.

5.7 Simulation Results

5.7.1 Performance Under I.I.D. Noise

The suitability of a GoFT in the context of SS in a CR can only be validated through extensive simulations, i.e., by studying its performance against various primary signal models, and different channel conditions. We consider a Rayleigh fading channel from the primary transmitter to the CR node; the number of observations $M = 300$, and target false alarm probability $p_f = 0.05$. The channel gain is assumed to remain constant throughout the M observations. For the primary signal, we consider the following models, with SNR = -5 dB in all experiments:

1. Model 1 - constant primary: The primary signal is a known constant. This model was considered previously in the GoFT for SS [52], [53].

2. Model 2 - sinusoidal primary: This simulates the scenario where the primary signal contains a strong pilot tone signal at a known frequency, similar to pilot-based

detection in Digital TV (IEEE 802.22) signals [56, 122]. The frequency of the signal is set to be 4kHz.

We have fixed $\sigma_{\mathrm{G}}^2 = 1$. The uncertainty in σ_{G}^2, i.e., the noise variance uncertainty is assumed to be 3 dB [38], [123]. For the class B noise model, we let $\gamma_0 = 1/\sqrt{2}$. Note that the value of γ_0 matters only when $\alpha = 2$, i.e., when the distribution is Gaussian, in which case its variance is $2\gamma_0^2 = 1$. For the BD, we set the number of windows as $m = 30$ [53].

Consider SS under the class B ($\mathcal{S}\alpha\mathcal{S}$) and Gaussian noise under hypothesis $\mathcal{H}_0$. In Fig. 5.1, the performance of all the detectors for constant primary is shown, with varying α. It is seen that around $1 \leq \alpha \leq 2$, BD outperforms the proposed detectors. However, as α reduces, the performance of BD deteriorates and approaches the $p_f = p_d$ line. Therefore, when α is low, BD does not satisfy the false alarm constraint. However, under model 2, both the proposed detectors outperform BD and satisfy the required false alarm constraint. This is shown in Fig. 5.2.

Now, consider detection under class A model with the Gaussian noise, i.e., the ϵ-mixture model. In Fig. 5.3, the performance of the detectors is shown as a function of σ_{I}^2, with $f_{\mathcal{I}}$ modeled as i.i.d. Gaussian, and for $\epsilon = 0.05$. It is seen that the performance of all the detectors remains constant for different values of σ_{I}^2 and ϵ. In all the cases, the performance of $\Psi_1 \mathrm{SD}$ and $\Psi_e \mathrm{SD}$ are comparable to that of BD under primary signal model 1, and improves significantly relative to that of BD under model 2. A similar trend is observed under the same setup with the $f_{\mathcal{I}}$ being modeled as Laplacian, as seen in Fig. 5.4. This shows that the performance of all the detectors designed for Gaussian noise are also valid under Laplacian noise, and with similar performance trends.

Figure 5.5 shows the performance of BD, Ψ_1SD and Ψ_eSD under i.i.d. Gaussian noise alone, and when the primary signal follows Models 1 and 2. Under Model 1, it is seen that BD performs better than Ψ_1SD and Ψ_eSD, especially in the low SNR regime. This is expected, as the underlying Anderson-Darling statistic of BD is powerful for testing against mean change in Gaussian signals [114]. However, it is seen that Ψ_1SD and Ψ_eSD outperform the BD when the primary signal follows Model 2. Additionally, Ψ_eSD performs better than Ψ_1SD. In Fig. 5.6, the performance of the detectors are plotted with correlated Rayleigh fading with the correlation modeled as a first order Auto-Regressive (AR) model [124] with $\rho = 0.5$, under i.i.d. Gaussian noise alone. It is seen that Ψ_eSD outperforms all the detectors. These simulations confirm that the proposed detectors are well suited for testing pilot based signals; for example, under a setup similar to the primary signal detection in IEEE 802.22 DTV standard.

Finally, in Figs. 5.7 and 5.8, the performance of the detectors under all the noise models combined is plotted, for both primary models. The results seen earlier, for the class B noise model with Gaussian noise, hold in these cases as well. There is a performance degradation due to presence of the heavy-tailed class A noise component, but the performance of Ψ_1SD and Ψ_eSD are better than the chance line. For the same setup, Fig. 5.9 shows the agreement between theoretical and simulated threshold values for all the detectors, thereby validating our analysis in (5.21) and (5.22).

As a final remark, we note that, due to the CFAR property of BD, Ψ_1SD, and Ψ_eSD (and any detector from the WZCD family), they fail to distinguish between the hypotheses when the primary signal has the same distribution as the noise process, with a different variance. Then, all the detectors would operate on the chance line ($p_f = p_d$ line), for all

SNR values. Therefore, for this particular case, WZCD is not a viable choice.

5.7.2   Performance Under Colored Noise

In the case of colored noise, ZCs and HOCs can be obtained by using the generalized level crossings lemma. However, the distribution of the statistics Ψ_M^2 for both uniform and equal, and exponential weights are not known in closed form and the optimal thresholds need to be obtained through numerical techniques, provided the correlation structure is correctly known. The ADD and BD cannot be applied here, because their design is valid only for the i.i.d. case [114]. In this section, the performance of ED, Ψ_1SD and Ψ_eSD are considered in the following scenarios, where the noise observations are correlated. For the following results, we assume an uncertainty in the knowledge of $\mathbb{E}Y_i^2, 1 \leq i \leq M$, of 3dB.

Correlation Model C_1

In this section, we study the performance of WZCD when the noise observations are correlated with a constant correlation coefficient. Such a correlation model is observed in, for e.g., an antenna array, where the correlation structure depends on the impedances of the antennas ([125], [112]). Also, with this model, fast fading is considered in under the signal-present hypothesis, where the fading gains are different and i.i.d. across each observation. The performance of the detectors under such correlated noise is as shown in Fig. 5.10. It is observed that both Ψ_1SD and Ψ_eSD outperform ED. Under constant primary, the performances of both Ψ_1SD and Ψ_eSD decrease as the noise correlation increases. However, under model 2, the performance of both Ψ_1SD and Ψ_eSD first decrease with the correlation value, and later increases.
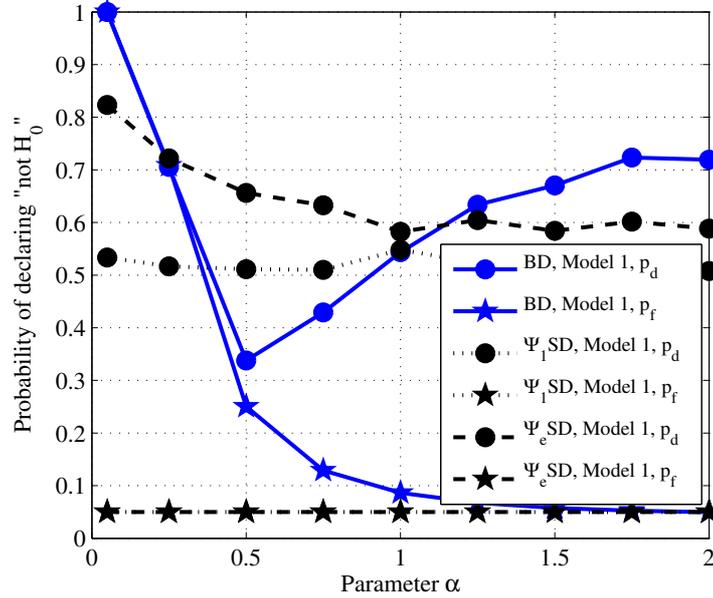
Figure 5.1: Detection of constant primary under Rayleigh fading, with Gaussian + $\mathcal{S}\alpha\mathcal{S}$ model.

Correlation Model C_2

Here, the correlation structure considered is the geometric model as described in Aalo and Viswanathan [126]. Following this model, the correlation decreases as a polynomial of the order or correlation. As in the previous case, fast fading is considered in the signal present hypothesis. The performance curves recorded in Fig. 5.11 show a similar trend to the equal correlation case. In all cases, Ψ_eSD outperforms the other detectors.

5.8 Probability of Detection for Constant Primary

In general, the distribution of the Ψ_M^2 statistic under $\mathcal{H}_1$ can be difficult to calculate. In the following, we show that even for the simple case of the constant signal model under $\mathcal{H}_1$ (under constant primary), a closed form solution for probability of detection might be hard to obtain. Following Bartlett's procedure [115], the first two moments of
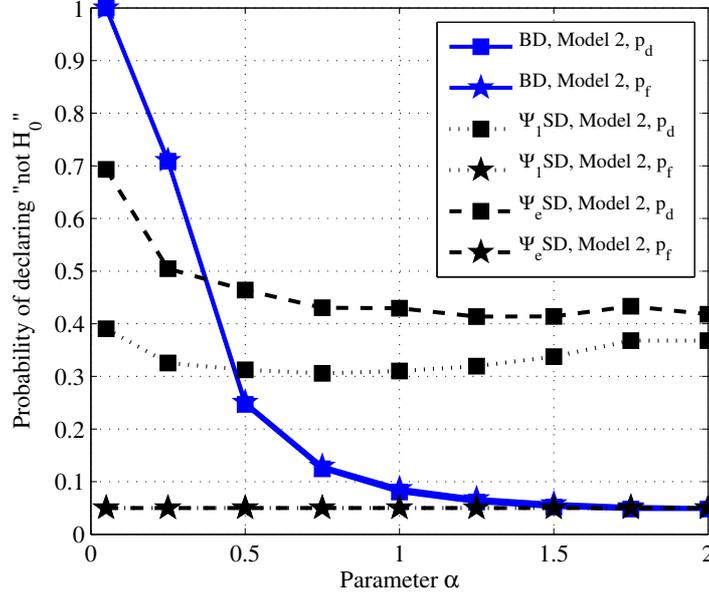
Figure 5.2: Detection of sinusoidal primary under Rayleigh fading, with Gaussian + $\mathcal{S}\alpha\mathcal{S}$ model.

the statistic can be used to fit a chi-squared statistic, by the moment matching method.

Let $m_M \triangleq \mathbb{E}(\Psi_M^2)$ and $v_M \triangleq \frac{1}{2}var(\Psi_M^2)$ denote the sample mean and sample variances

of the the Ψ_M^2 statistic under $\mathcal{H}_1$ for model 1. Then, the statistic [115]

$$\frac{m_M}{v_M}\Psi_M^2 \sim \chi_{\frac{m_M^2}{v_M}}^2(0). \tag{5.30}$$

For analytical simplicity, we assume that k (the number of HOCs) is large enough that

we can invoke the central limit theorem ($k = 8$ suffices). With this assumption, $\Psi^2|\mathcal{H}_1 \sim$

$\mathcal{N}(m_M, v_M)$. Now, the average p_d, for the case of Ψ_1SD, can be calculated as

$$
\begin{aligned}
p_d &= \int_0^\infty \int_{\tau_1^\Psi}^\infty \frac{1}{\sqrt{2\pi v_M}} \exp\left(-\frac{(\psi_M - m_M)^2}{2v_M}\right) d\psi_M f_h(h) dh \\
&= \int_0^\infty \left(\mathcal{Q}\left(\frac{\tau_1^\Psi - m_M}{\sqrt{v_M}}\right)\right) f_h(h) dh
\end{aligned}
\tag{5.31}
$$
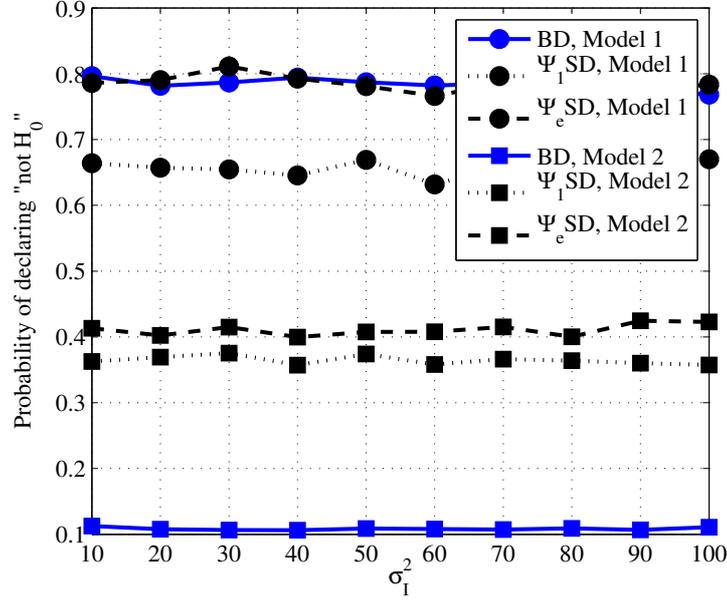
Figure 5.3: Detection of primary models 1 and 2 under Rayleigh fading, with ϵ-mixture model, $\epsilon = 0.05$, and $f_{\mathcal{I}} \sim \mathcal{N}(0, \sigma_{\mathrm{I}}^2)$.

where $f_h(\cdot)$ is the PDF of the channel. Note that m_M and v_M are functions of the channel gain h. A similar analysis can be carried out for the Ψ_eSD, in which case the probability of detection $p_d^{(m)}$ is given by

$$p_d^{(m)} = \int_0^\infty \left(\mathcal{Q}\left(\frac{\tau_e^\Psi - m_M^{(m)}}{\sqrt{v_M^{(m)}}} \right) \right) f_h(h)dh, \tag{5.32}$$

where $m_M^{(m)} \triangleq \mathbb{E}(_m\Psi^2)$, and $v_M^{(m)} \triangleq \frac{1}{2}var(_m\Psi^2)$. The agreement between theoretical and simulated values of p_m is shown in Fig. 5.12, where we have numerically evaluated the integral. It is seen that the agreement becomes tighter as SNR increases. Alternatively, the Q function can be approximated by an exponential term, to evalute the integral in closed form, as suggested by Lopez-Benitez and Casadevall [66].
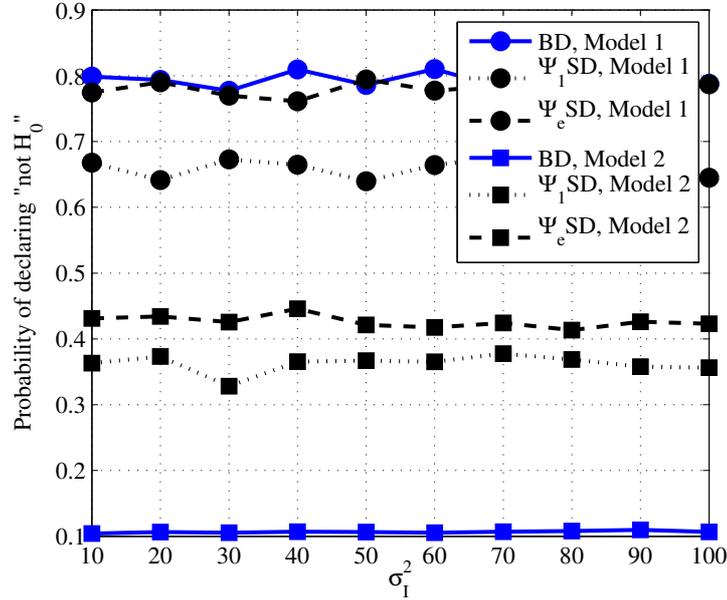
Figure 5.4: Detection of primary models 1 and 2 under Rayleigh fading, with ϵ-mixture model, $\epsilon = 0.05$, and $f_{\mathcal{I}} \sim \mathcal{L}(\sigma_I^2)$.

5.9 Conclusion

In this chapter, a weighted zero-crossings based goodness-of-fit test for spectrum sensing was proposed. A near-optimal detection threshold was derived for the specific choices of uniform and exponential weights. It was shown that this detector is robust to the noise model, and parameter uncertainties. Through simulations, it was shown that the proposed detectors outperform the existing tests in the CR literature in a variety of noise and primary signal conditions of practical interest. Also, the computational simplicity of the proposed test was highlighted. Therefore, the proposed detector is a promising choice for spectrum sensing in CR, and can be used in a wide range of communication scenarios.
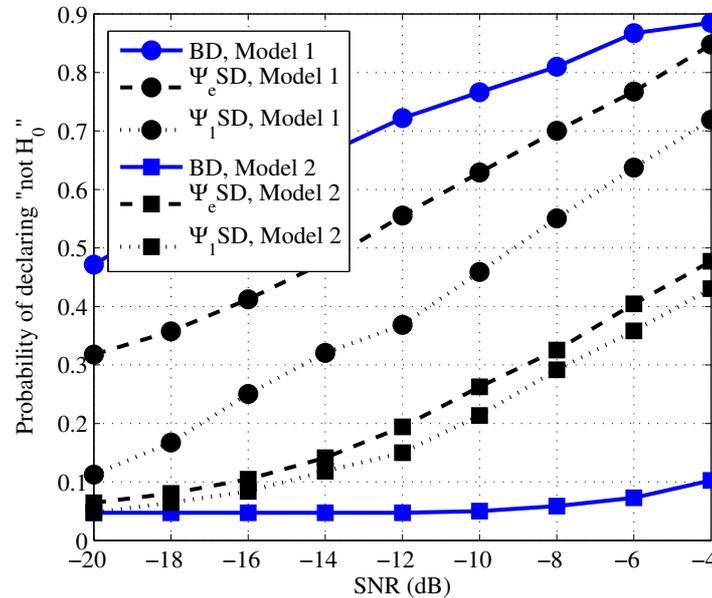
Figure 5.5: Detection of primary models 1 and 2 under pure Gaussian noise, with noise variance uncertainty= 3dB, $M = 300, \alpha_f = 0.05$. Average p_f obtained through simulations for BD, Ψ_1SD and Ψ_eSD are $0.0498, 0.05$, and 0.0501, respectively.
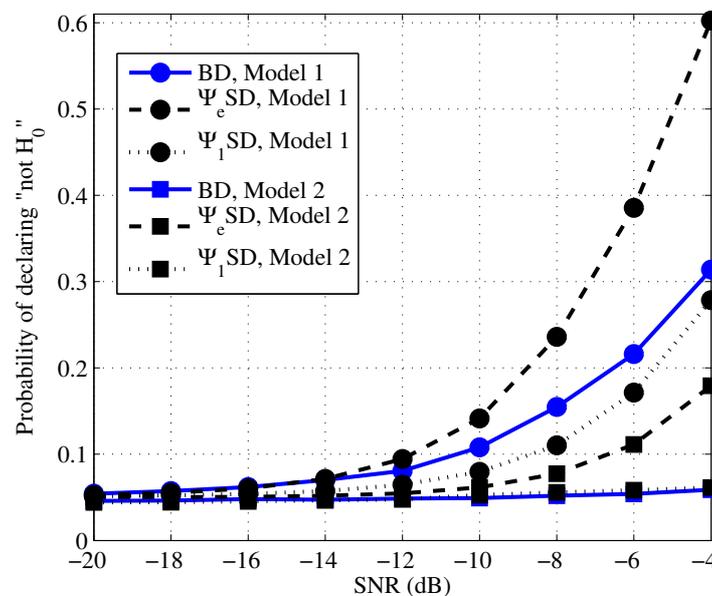
Figure 5.6: Detection of primary models 1 and 2 under first order AR correlated fading (with $\rho = 0.5$) and pure Gaussian Noise, with noise variance uncertainty= 3dB, $M = 300, \alpha_f = 0.05$.
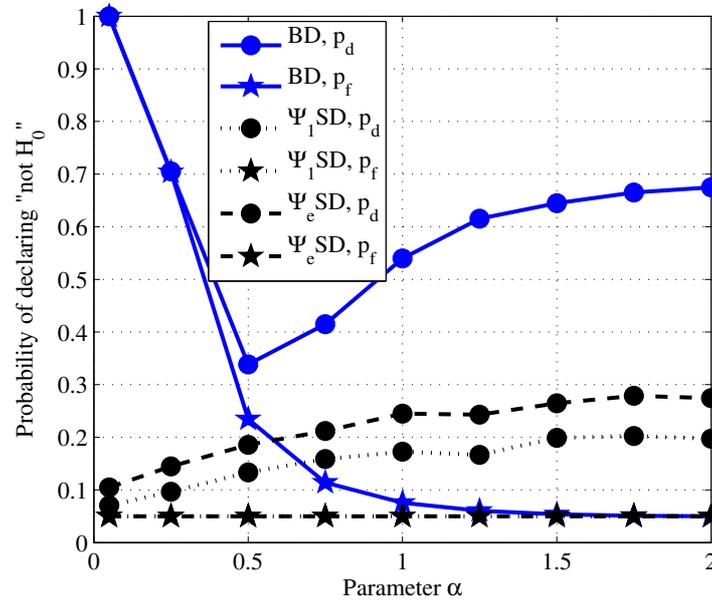
Figure 5.7: Detection of constant primary under Gaussian + class A + class B noises, with noise variance uncertainty= 3dB, $M = 300, \alpha_f = 0.05, \epsilon = 0.05, f_{\mathcal{I}} \sim \mathcal{N}(0, 100\sigma_{\mathrm{G}}^2)$.
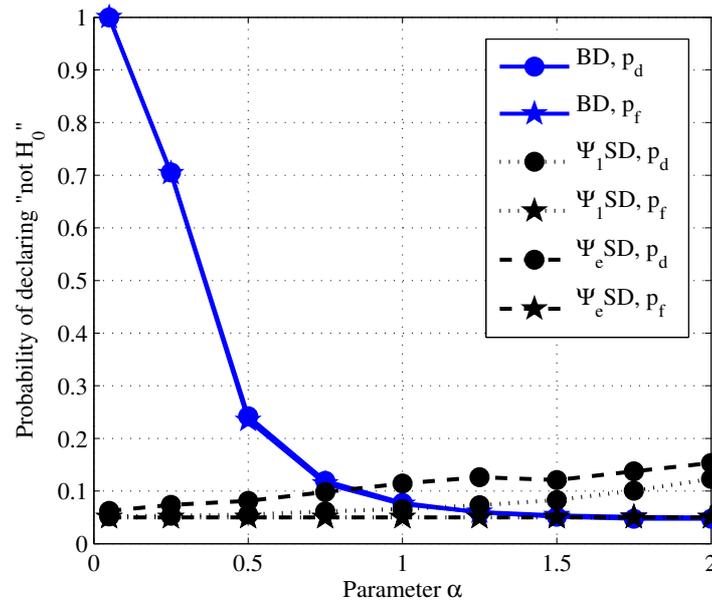
Figure 5.8: Detection of sinusoidal primary under Gaussian + class A + class B noises, with noise variance uncertainty= 3dB, $M = 300, \alpha_f = 0.05, \epsilon = 0.05, f_{\mathcal{I}} \sim \mathcal{N}(0, 100\sigma_{\mathrm{G}}^2)$.
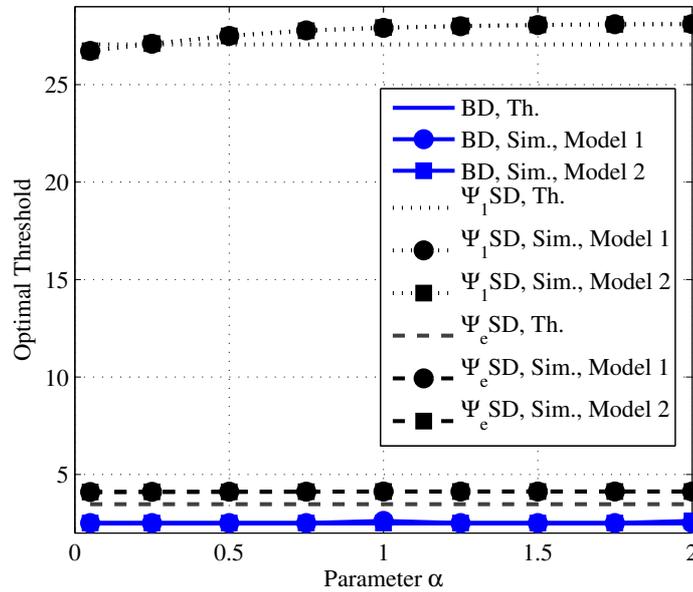
Figure 5.9: Optimal threshold calculation under Gaussian + class A + class B noises, with noise variance uncertainty= 3dB, $M = 300, \alpha_f = 0.05, \epsilon = 0.05, f_{\mathcal{I}} \sim \mathcal{N}(0, 100\sigma_{\mathrm{G}}^2)$.
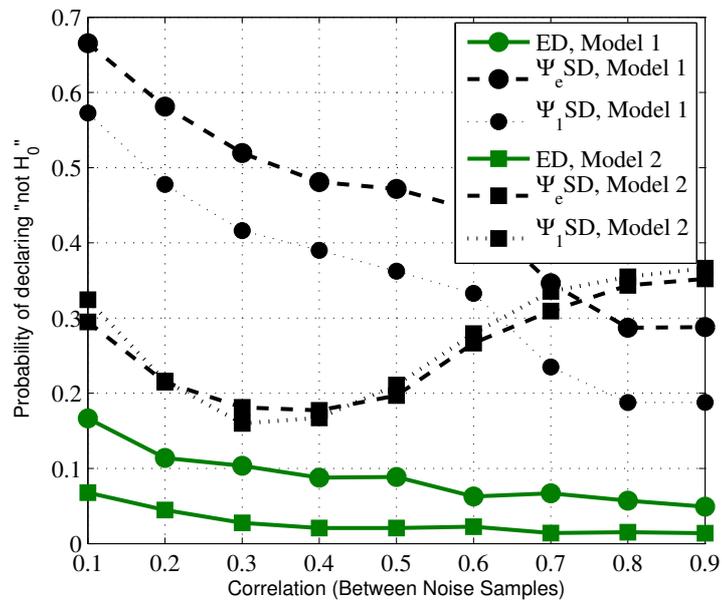
Figure 5.10: Detection of primary models 1 and 2 under equal correlated noise as a function of correlation co-efficient. Average p_f obtained through simulations for ED, Ψ_1SD and Ψ_eSD are $0.05, 0.05$, and 0.0501, respectively.
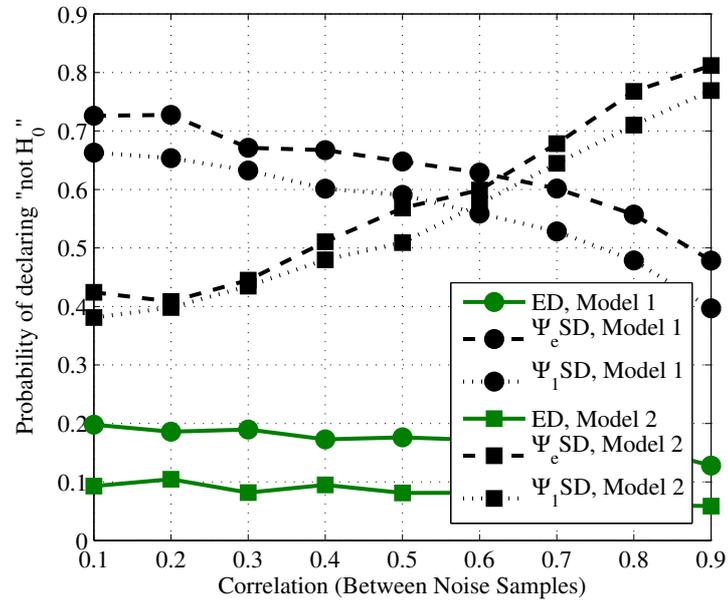
Figure 5.11: Detection of primary models 1 and 2 under geometric correlated noise as a function of correlation co-efficient. Average p_f obtained through simulations for ED, Ψ_1SD and Ψ_eSD are 0.05, 0.05, and 0.0501, respectively.
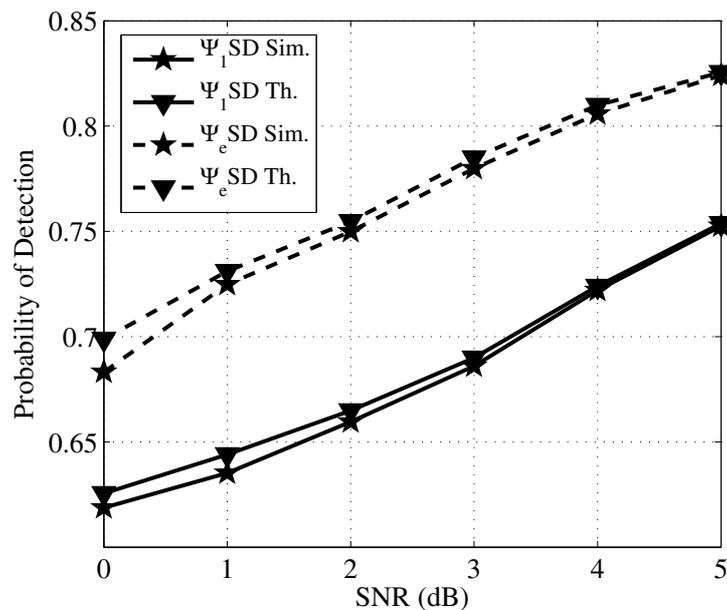
Figure 5.12: Comparison of theoretical and simulated p_d values for constant primary under Rayleigh fading, with $M = 300, \alpha = 0.05$. The agreement becomes stronger at high SNR.

Chapter 6

Multi-dimensional Goodness-of-Fit Tests Based on Stochastic Distances For Spectrum Sensing

6.1 Introduction

As explained earlier in Chap. 5, the class of Goodness-of-Fit Tests (GoFT) form the natural choice for spectrum sensing when very little or no knowledge is available about the distribution of the test statistic under the signal-present hypothesis.

The existing GoFT techniques in literature such as the Anderson-Darling based Detector (ADD) [52], Blind Detector (BD) [53], and Weighted Zero-Crossings based Detector (WZCD) have all been studied for spectrum sensing on a single CR node. However, it is known that using measurements from multiple CR nodes offers better performance, due to the diversity advantage ([8, 10, 57]). Also, when multiple antennas are available at each CR node, one can obtain a centralized-like system at each node, and additional diversity gains ([73, 74, 127]). In such a setup, several eigenvalue based tests have been proposed in the literature ([73, 128, 129]). Other widely used eigenvalue based

detection include John's test [130], and an Eigenvalue Ratio test (ER) ([131], [132]).
However, all the above techniques perform poorly in the presence of multiple primary
users, operating in the same set of bands that the secondary users intend to use [133].
The algorithm presented in [133], based on the *Sphericity Test* (ST) (originally proposed
in [134]), considers the effect of multiple users; but its analysis is restricted to the case
where primary signal is i.i.d. Gaussian distributed, and assuming that the channel re-
mains constant throughout the observations. To the best of our knowledge, GoFT for
SS in a multi-dimensional setup (multiple sensors, multiple antennas, multiple obser-
vations and multiple primary users) has not been developed in the literature so far. In
this chapter, we design and develop two Multi-Dimensional GoFT (MDGoFT) based on
stochastic distances for SS in CR, namely, a *Interpoint Distance based test* (ID) [135], and
a $\langle h, \phi \rangle$ distance based test [136]. The ID is useful in the scenario of multiple antennas
making multiple observations with multiple primary users. The $\langle h, \phi \rangle$ test considers
an additional dimension of multiple sensors. Through extensive simulations, we show
that both the tests perform better than the existing techniques, in several scenarios. To
summarize, the main contributions of this chapter are:

- For the case where a single CR node with multiple antennas records multiple
  observations for detection of multiple primary users, we propose and study the
  interpoint distance test [135]. We also mention a possible extension to handle
  multiple CR nodes.

- For a setup similar to above with multiple CR nodes, we propose the $\langle h, \phi \rangle$ dis-
  tance based test [136]. We analytically obtain the detection threshold for achieving

a given false alarm probability, and discuss a noise-robustness feature of a particular case of the $\langle h, \phi \rangle$ distance, namely, the Kullback-Leibler distance based test. This is related to the robustness to the presence of the class A noise component studied in the previous chapter.

- The performance of both the MDGoFTs, which are based on the statistical properties of stochastic distances, are studied through Monte Carlo simulations. The tests are shown to outperform the existing techniques, viz., the John's test [130], eigenvalue ratio based test [131], and the sphericity test [133].

6.2   System Model

Consider the cooperative model for Primary User (PU) detection [133], with a setup as shown in Fig. 6.1. Let L represent the number of CR nodes, with N antennas each, and P represent the number of Primary Users (PU). The CR nodes collect their observations and send a statistic based on the collected information, over a lossless channel to a Fusion Center (FC). The FC combines these statistics, and comes up with a decision on whether the spectrum is vacant or not. Under the signal-present hypothesis, the received vector $\mathbf{x}_l \in \mathbb{C}^{N \times 1}$ at each sensor is given by [133]

$$\mathbf{x}_l = \mathbf{H}_l \mathbf{s}_l + \sigma_n \mathbf{n}_l, \quad l = 1, \cdots, L \tag{6.1}$$

where $\mathbf{s}_l \in \mathbb{C}^{P \times 1}$ represents the transmitted signal from the P PUs. Also, $\mathbf{H}_l \in \mathbb{C}^{N \times P}$ represents the matrix of channel gains between N antennas and P PUs. Both $\mathbf{H}_l$ and $\mathbf{s}_l$ are unknown at CRs, and can be arbitrarily correlated across time, space and CR nodes. Finally, $\sigma_n \mathbf{n}_l$ represents the $N \times 1$ length complex Gaussian noise vector, with zero mean

and a known covariance matrix $\sigma_n^2 \mathbf{I}_N$, where $\mathbf{I}_N$ represents the identity matrix of size N.

Each CR node collects M observations following the model given in (6.1), into a $N \times M$

matrix $\mathbf{X}_l \triangleq [\mathbf{x}_l^{(1)}, \mathbf{x}_l^{(2)}, \cdots, \mathbf{x}_l^{(M)}]$.
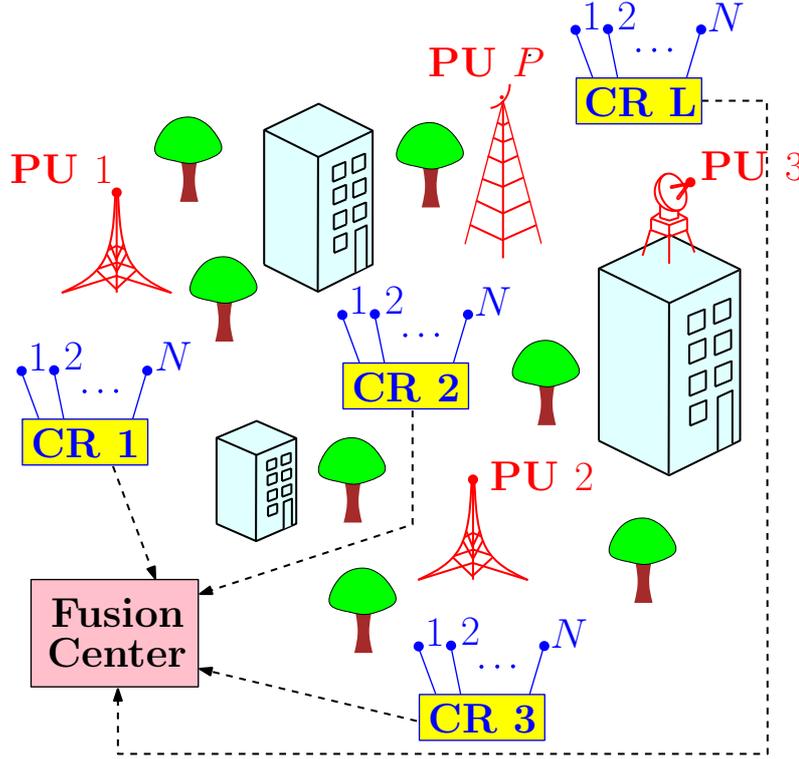
Figure 6.1: System Model

At each CR node, define $\mathbf{R}_l \triangleq \mathbf{X}_l \mathbf{X}_l^H$, and $\mathbf{\Sigma} \triangleq \frac{1}{M}\mathbb{E}\mathbf{X}_l \mathbf{X}_l^H$. When the primary signal is

absent, i.e., when $\mathbf{x}_l = \sigma_n \mathbf{n}_l$, it is known that the random matrix $\mathbf{R}_l$ is complex Wishart

distributed [133] with parameters M and $\mathbf{\Sigma}$, which is denoted by $\mathcal{W}_N(M, \mathbf{\Sigma})$. When

$L = 1$, the hypothesis testing problem for the above setup is formulated as the following

Goodness-of-Fit Test (GoFT):

$$
\begin{aligned}
\mathcal{H}_0 \;&:\; \mathbf{\Sigma} = \sigma_n^2 \mathbf{I}_N \\
\mathcal{H}_1 \;&:\; \mathbf{\Sigma} \neq \sigma_n^2 \mathbf{I}_N.
\end{aligned}
\tag{6.2}
$$

A variation on the above test is well studied in the CR literature [133], with the following assumptions:

1. The channel $\mathbf{H}_l$ is constant throughout the M observations.

2. The vector $\mathbf{s}_l$ follows a Gaussian distribution, i.i.d. across all dimensions and is independent of the receiver noise.

With the above assumptions, the condition under $\mathcal{H}_1$ can be shown to be equal to $\boldsymbol{\Sigma} > \sigma_{\mathrm{G}}^2 \mathbf{I}_N$, where the term ">" is in a positive definite sense. The above assumptions may not be true in practice. Therefore, in our work, we do not restrict to the above, or any such assumptions on the signal model under $\mathcal{H}_1$.

The test given in (6.2), is referred to as *Sphericity Test* (ST), or *deviation against sphericity*, in the literature. Let $\lambda_1^{(l)} \geq \lambda_2^{(l)} \geq \cdots \geq \lambda_N^{(l)}$ represent the eigenvalues of the matrix $\mathbf{R}_l$. Then, the following detectors have been proposed to solve the sphericity test in the CR SS context viz., ST [133], Eigenvalue Ratio (ER) test [131], and John's Test (JT) [130], whose test statistics are defined as

$$T_{\mathrm{ST}} \triangleq \frac{\prod_{i=1}^N \lambda_i}{\left(\frac{1}{N}\sum_{i=1}^N \lambda_i\right)^N}, \quad T_{\mathrm{J}} \triangleq \frac{\sum_{i=1}^N \lambda_i^2}{\left(\sum_{i=1}^N \lambda_i\right)^2}, \quad T_{\mathrm{ER}} \triangleq \frac{\lambda_1}{\lambda_N}. \tag{6.3}$$

With the knowledge of the statistics of $\mathbf{X}$ under the noise-only hypothesis alone, all the above tests can be used as non-parametric GoFTs. In the following sections, we propose two detectors viz., *Interpoint Distance* (ID) based Test, and $\langle h, \phi \rangle$-*distance based test*, which are based on stochastic distances. Later, these tests are shown to perform better than all the tests in (6.3) with various assumptions on the signal and channel characteristics under $\mathcal{H}_1$, for a given false-alarm level.

6.3   Interpoint Distance Based GoFT

Let $L = 1$. We will drop the sensor index "l" from all the related notations, for ease of presentation. Let $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M$ represent N dimensional observation vectors recorded by the CR node. Under $\mathcal{H}_0$, $\mathbf{x}_i, 1 \leq i \leq M$, follows a PDF $f_\mathbb{N}$, which is assumed to be a Gaussian distribution with a known mean vector, and a known covariance matrix. Assuming i.i.d observations under $\mathcal{H}_0$, the goal of the GoFT is to accept $\mathcal{H}_0$ if $\mathbf{x}_i, 1 \leq i \leq M$ follows $f_\mathbb{N}$, and reject otherwise. Mathematically,

$$
\begin{aligned}
\mathcal{H}_0 &: \quad \mathbf{x}_i \sim f_\mathbb{N} \\
\mathcal{H}_1 &: \quad \mathbf{x}_i \nsim f_\mathbb{N}.
\end{aligned}
\tag{6.4}
$$

In this section, an interpoint distance based GoFT is proposed, which is based on a test proposed by Bartoszynski et al. [135]. Let us define a distance function $\delta(\cdot, \cdot)$, on the probability space of $\mathbf{x}_i$, which satisfies the non-negativity, symmetry and triangle inequalities. In this context, we recall the following theorem, due to Maa et al. [137]:

Theorem 8 (Maa-Pearl-Bartoszynski). *Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be arbitrary countable sets, and let $\mathbf{X}$ and $\mathbf{Y}$ be N-dimensional random vectors with values in $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively. If $\delta(\mathbf{X}, \mathbf{Y})$ is any real valued, non-negative function on $\mathcal{S}_1 \times \mathcal{S}_2$, such that $\delta(\mathbf{X}, \mathbf{Y}) = 0$, if and only if $\mathbf{X} = \mathbf{Y}$. Also, let $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ and $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$ be random vectors chosen from the distributions $\mathcal{F}$ and $\mathcal{G}$, respectively. Then,*

$$
\delta(\mathbf{x_1}, \mathbf{x_2}) \stackrel{d.}{=} \delta(\mathbf{y_1}, \mathbf{y_2}) \stackrel{d.}{=} \delta(\mathbf{x_3}, \mathbf{y_3}), \; \textit{iff } \mathcal{F} = \mathcal{G}.
\tag{6.5}
$$

One implication of the above theorem is the following

$$\mathcal{P}\{\delta(\mathbf{x_1}, \mathbf{x_2}) \le \tau\} = \mathcal{P}\{\delta(\mathbf{x_1}, \mathbf{y_1}) \le \tau\}, \quad \tau \in \mathbb{R} \tag{6.6}$$

Another implication is that the data points $\mathbf{x_1}$, $\mathbf{x_2}$ and $\mathbf{y_1}$ come from the same distribution if and only if the lengths of the sides of a triangle formed by them (as measured by $\delta(\cdot, \cdot)$) have the same distribution.
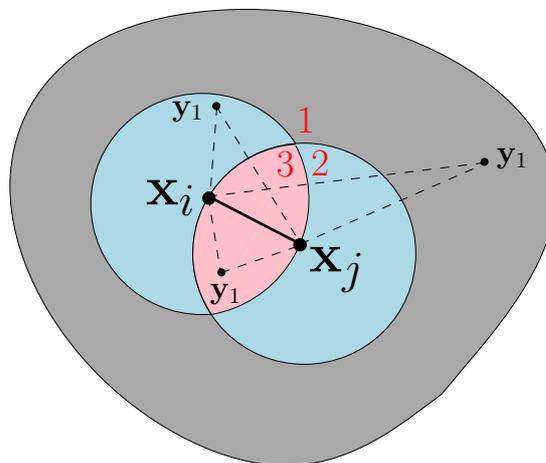
Figure 6.2: The regions defining p_1, p_2 and p_3.

The ID test is devised as follows [135]. Let $\mathbf{x_i}$ and $\mathbf{x_j}$ be two different N-dimensional samples, with $i, j \in \{1, \cdots, M\}$, and $i \ne j$. Let $p_1(\mathbf{x_i}, \mathbf{x_j})$, $p_2(\mathbf{x_i}, \mathbf{x_j})$ and $p_3(\mathbf{x_i}, \mathbf{x_j})$ denote the probabilities that in a triangle formed by points $\mathbf{x_i}$, $\mathbf{x_j}$ (sampled from the $f_{\mathbb{N}}$) and a given $\mathbf{y_1}$, the side joining $\mathbf{x_i}$ and $\mathbf{x_j}$ is the smallest, intermediate and longest, respectively. In other words, these probabilities correspond to the point $\mathbf{y_1}$ falling in the

regions $1, 2$ and 3, respectively, as shown in Fig. 6.2. Mathematically,

$$p_1(\mathbf{x_i}, \mathbf{x_j}) \triangleq \mathcal{P}\{\delta(\mathbf{x_i}, \mathbf{x_j}) < \min(\delta(\mathbf{x_i}, \mathbf{y_1}), \delta(\mathbf{x_j}, \mathbf{y_1}))\}$$

$$+ \frac{1}{2}\mathcal{P}\{\delta(\mathbf{x_i}, \mathbf{x_j}) = \delta(\mathbf{x_i}, \mathbf{y_1}) < \delta(\mathbf{x_j}, \mathbf{y_1})\}$$

$$+ \frac{1}{2}\mathcal{P}\{\delta(\mathbf{x_i}, \mathbf{x_j}) = \delta(\mathbf{x_j}, \mathbf{y_1}) < \delta(\mathbf{x_i}, \mathbf{y_1})\}$$

$$+ \frac{1}{2}\mathcal{P}\{\delta(\mathbf{x_i}, \mathbf{x_j}) = \delta(\mathbf{x_i}, \mathbf{y_1}) = \delta(\mathbf{x_j}, \mathbf{y_1})\} \tag{6.7}$$

$$p_2(\mathbf{x_i}, \mathbf{x_j}) \triangleq \mathcal{P}\{\delta(\mathbf{x_i}, \mathbf{x_j}) > \max(\delta(\mathbf{x_i}, \mathbf{y_1}), \delta(\mathbf{x_j}, \mathbf{y_1}))\}$$

$$+ \frac{1}{2}\mathcal{P}\{\delta(\mathbf{x_i}, \mathbf{x_j}) = \delta(\mathbf{x_i}, \mathbf{y_1}) > \delta(\mathbf{x_j}, \mathbf{y_1})\}$$

$$+ \frac{1}{2}\mathcal{P}\{\delta(\mathbf{x_i}, \mathbf{x_j}) = \delta(\mathbf{x_j}, \mathbf{y_1}) > \delta(\mathbf{x_i}, \mathbf{y_1})\}$$

$$+ \frac{1}{2}\mathcal{P}\{\delta(\mathbf{x_i}, \mathbf{x_j}) = \delta(\mathbf{x_i}, \mathbf{y_1}) = \delta(\mathbf{x_j}, \mathbf{y_1})\} \tag{6.8}$$

$$p_3(\mathbf{x_i}, \mathbf{x_j}) \triangleq 1 - p_1(\mathbf{x_i}, \mathbf{x_j}) - p_2(\mathbf{x_i}, \mathbf{x_j}) \tag{6.9}$$

For a given distance measure $\delta(\cdot, \cdot)$, and underlying probability distributions $\mathcal{F}$ and $\mathcal{G}$ respectively, deriving these probabilities in closed form might be difficult. In practice, they can be evaluated using Monte Carlo simulations. Now, define

$$U_k \triangleq \frac{1}{\binom{M}{2}} \sum_{i,j} p_k(\mathbf{x_i}, \mathbf{x_j}); \ \ i, j = 1, \cdots, M, \ k = 1, 2, 3, \tag{6.10}$$

Under the noise-only hypothesis (i.e., when $\mathcal{F} = \mathcal{G}$), asymptotic properties of $U_k, k = 1, 2, 3$ (as M grows large) are known [135]. For each U_k, define a corresponding Z_k as

$$Z_k \triangleq \frac{U_k - \frac{1}{3}}{\sqrt{\mathrm{var}(U_k|\mathcal{H}_0)}}; \ \ k = 1, 2, 3. \tag{6.11}$$

For large enough N, any Z_r and Z_s, $r, s = 1, 2, 3, r \neq s$, and for $\rho \triangleq \mathrm{cov}(Z_r, Z_s)$, it is

known that

$$Q_{r,s} \triangleq \frac{Z_r^2 + Z_s^2 - 2\rho Z_r Z_s}{1 - \rho^2}, \quad r, s = \{1, 2, 3\}; r \neq s, \tag{6.12}$$

closely follows a central chi-squared distribution with 2 degrees of freedom, i.e., $Q_{r,s} \sim \chi_2^2$, under $\mathcal{H}_0$. This statistic can be used for testing against a given $f_{\mathbb{N}}$. Therefore, the statistic $Q \triangleq Q_{1,3} + Q_{2,3} + Q_{3,1}$ follows a central chi-square distribution with 6 degrees of freedom. Using this result, the hypothesis testing problem (6.4), reduces to

$$Q \underset{\mathcal{H}_0}{\overset{\sim \mathcal{H}_0}{\gtrless}} \tau_{ID}. \tag{6.13}$$

Since the above test statistic is constructed depending on the distance measure $\delta(\cdot, \cdot)$ between the points $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M$, it is called as the Interpoint Distance (ID) based test. The threshold τ_{ID} is chosen such that the following constraint is satisfied.

$$p_f \triangleq \mathcal{P}\{\text{declaring } \sim \mathcal{H}_0 | \mathcal{H}_0\} = \alpha_f, \tag{6.14}$$

where $\alpha_f \in (0, 1)$ is given. Under $\mathcal{H}_0$, since $Q \sim \chi_6^2$, it is easy to see that the above condition is satisfied when τ_{ID} is chosen such that

$$1 - \frac{\gamma\left(\frac{\tau_{ID}}{2}, 3\right)}{\Gamma(3)} = \alpha_f, \tag{6.15}$$

where $\gamma(\cdot, \cdot)$, and $\Gamma(\cdot)$ are lower incomplete, and complete gamma functions, respectively [116].

The above test can be extended to handle the presence of class A and class B noise components [55], as long as the distribution and their parameters are known. If the parameters are unknown, they can be directly estimated from the observations. However,

this test is known to perform poorly, if there is an uncertainty in the knowledge of the noise-only distribution parameters [135].

6.3.1 Choice of $\delta(\cdot, \cdot)$

A distance function is said to be *invariant* under a given transformation $\mathcal{T}$ [135], if

$$\delta(\mathbf{x}_1, \mathbf{x}_2) \geq \delta(\mathbf{x}_3, \mathbf{x}_4) \Rightarrow \delta(\mathcal{T}(\mathbf{x}_1), \mathcal{T}(\mathbf{x}_2)) \geq \delta(\mathcal{T}(\mathbf{x}_3), \mathcal{T}(\mathbf{x}_4)). \tag{6.16}$$

If a distance function $\delta(\cdot, \cdot)$ is invariant to a transformation $\mathcal{T}(\cdot)$, then the values of the parameters U_k and ρ are not affected by $\mathcal{T}$ [135]. Since the relative ℓ_p-norm distances are invariant to linear transformations such as scaling, rotation, etc., they are considered in this work. As will be elaborated in Sec. 6.5, the performance of the ID detector can be improved by appropriately choosing the parameter "p".

6.3.2 Extension to Multiple Sensors

One way to extend the above analysis to the multi-sensor case (with L sensors), is to use the sum $\mathbf{X} \triangleq \sum_{l=1}^{L} X_l$, as the matrix from which the test statistic is computed. Given that every vector in each $\mathbf{X}_l$ is a Gaussian vector, every vector in $\mathbf{X}$ also follows a Gaussian distribution. In other words, this scenario is statistically equivalent to a centralized detector in which the fusion center has access to $\mathbf{X}_1, \cdots, \mathbf{X}_L$, which uses $\mathbf{X}$ as the test statistic. Therefore, the above procedure can now be applied on each vector of $\mathbf{X}$. The test, however, would remain the same as given in (6.13).

6.4 $\langle h, \phi \rangle$ Distance Based GoFT

Consider the scenario where L sensors transmit $\mathbf{R}_l = \mathbf{X}_l \mathbf{X}_l^H$ to a Fusion Center (FC), through a dedicated, lossless channel. The FC sums all the matrices $\mathbf{R}_l$ into a single matrix $\mathbf{R} \triangleq \sum_{l=1}^{L} \mathbf{R}_l$. Under the noise-only hypothesis, given that each $\mathbf{R}_l \sim \mathcal{W}_N(M, \boldsymbol{\Sigma})$, it is easy to see that $\mathbf{R} \sim \mathcal{W}_N(LM, \boldsymbol{\Sigma})$ [138].

In this section, we study the properties of $\langle h, \phi \rangle$ distance metric for probability distributions proposed by Salicru et al. [136], and propose a GoFT based on this metric for SS. Let $\mathbf{Y}$ and $\mathbf{Z}$ be two positive definite, Hermitian random matrices of size $N \times N$, with their distributions characterized by the densities $f_{\mathbf{Y}}(\cdot; \theta_1)$, and $f_{\mathbf{Z}}(\cdot; \theta_2)$, parametrized by θ_1 and θ_2, respectively. Then, the $\langle h, \phi \rangle$ divergence (not necessarily a distance metric, as explained later) between $f_{\mathbf{Y}}$ and $f_{\mathbf{Z}}$ is defined as [136]

$$D_\phi^h(\mathbf{Y}, \mathbf{Z}) \triangleq h\left(\int_{\mathbb{H}} \phi\left(\frac{f_{\mathbf{Y}}(\mathbf{Y}'; \theta_1)}{f_{\mathbf{Z}}(\mathbf{Y}'; \theta_2)} \right) f_{\mathbf{Z}}(\mathbf{Y}'; \theta_2) d\mathbf{Y}' \right), \tag{6.17}$$

where $h : [0, \infty) \to [0, \infty)$ is a strictly increasing function with $h(0) = 0$, and $\phi : [0, \infty) \to [0, \infty)$ is a convex function, with $\phi\left(\frac{0}{0}\right) \triangleq 0$. The space of all positive definite Hermitian matrices of size $N \times N$ is denoted by $\mathbb{H}$. The differential element $d\mathbf{Y}'$ is defined as

$$d\mathbf{Y}' = dY_{11}' dY_{22}' \cdots dY_{MM}' \prod_{i,j=1; i<j}^{M} d\mathrm{Re}(Y_{i,j}') d\mathrm{Im}(Y_{i,j}'), \tag{6.18}$$

where $Y_{i,j}'$ is the (i, j)th entry of the matrix $\mathbf{Y}$. Also, $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and imaginary parts of a complex number.

Some of the well-known information-theoretic divergences are special cases of the $\langle h, \phi \rangle$ divergence, with appropriate choices of $h(\cdot)$ and $\phi(\cdot)$. These divergence measures

Table 6.1: Various information-theoretic divergences as special cases of $\langle h, \phi \rangle$ distance, and their related functions $h(\cdot)$ and $\phi(\cdot)$.

Divergence	$h(x)$	$\phi(y)$
χ^2	$\frac{x}{4}$	$(y-1)^2 \frac{y+1}{y}$
Kullback-Leibler	$\frac{x}{2}$	$(y-1)\log y$
Rényi (order α)	$\frac{\log(\alpha x - x + 1)}{\alpha - 1}, 0 \leq x \leq \frac{1}{\alpha-1}$	$\frac{y^{1-\alpha} + y^\alpha - \alpha(y-1) - 2}{2(\alpha-1)}, 0 < \alpha < 1.$
Bhattacharyya	$-\log(1-x), \ 0 \leq x < 1$	$-\sqrt{y} + \frac{y+1}{2}.$
Hellinger	$\frac{x}{2}, 0 \leq x < 2$	$(\sqrt{y}-1)^2$

need not be distance measures, as they do not satisfy necessarily the triangle inequality. A symmetric, $\langle h, \phi \rangle$ distance metric based on the D_ϕ^h divergence is defined as follows

$$d_\phi^h(\mathbf{Y}, \mathbf{Z}) \triangleq \frac{D_\phi^h(\mathbf{Y}, \mathbf{Z}) + D_\phi^h(\mathbf{Z}, \mathbf{Y})}{2} \tag{6.19}$$

The above $\langle h, \phi \rangle$ metric, viz., $d_\phi^h(\cdot, \cdot)$ is a distance metric, i.e., it satisfies non-negativity property, symmetric property and triangle inequalities, for all possible choices of $h(\cdot)$ and $\phi(\cdot)$, subjected to the conditioned mentioned earlier [138]. Some of the commonly used information-theoretic divergence measures, and the corresponding $h(\cdot)$ and $\phi(\cdot)$ functions for them to be a valid D_ϕ^h divergence, are listed in Table 6.1.

Next, we recall the following theorem ([136], [139]), which establishes the distribution of the statistic d_ϕ^h under $\mathcal{H}_0$, for any $h(\cdot)$ and $\phi(\cdot)$ for our problem formulation.

Theorem 9 (Salicru et al.). *Let $\mathbf{R} \sim \mathcal{W}_N(LM, \boldsymbol{\Sigma})$, and $\mathbf{R}'$ be another random matrix of size $N \times N$. Then, under $\mathcal{H}_0$ (i.e., when $\mathbf{R} \stackrel{d.}{=} \mathbf{R}'$), and under the regularity conditions given by Salicru et al. [136, Pg. 375],*

$$\mathcal{S}(\mathbf{R}, \mathbf{R}') \triangleq LM \frac{d_\phi^h(\mathbf{R}, \mathbf{R}')}{h'(0)\phi''(1)} \xrightarrow[LM \to \infty]{d.} \chi^2_{\frac{N^2+N}{2}}. \tag{6.20}$$

Using the above theorem, an L sensor extension of the hypothesis testing problem in (6.4) reduces to a test on the statistic $\mathcal{S}(\mathbf{R}, \mathbf{R}')$, which is of the form

$$\mathcal{S}(\mathbf{R}, \mathbf{R}') \underset{\sim\mathcal{H}_0}{\overset{\sim\mathcal{H}_0}{\gtrless}} \tau_h^{(\phi)}, \tag{6.21}$$

where $\tau_h^{(\phi)}$ is chosen such that the constraint $p_f = \alpha_f \in (0, 1)$ is satisfied. Since $\mathcal{S}(\cdot, \cdot) \sim \chi^2_{\frac{N^2+N}{2}}$, it is easy to see that the above condition is satisfied when $\tau_h^{(\phi)}$ is chosen such that

$$1 - \frac{\gamma\left(\frac{\tau_h^{(\phi)}}{2}, \frac{N^2+N}{4}\right)}{\Gamma\left(\frac{N^2+N}{4}\right)} = \alpha_f. \tag{6.22}$$

Note that the above result holds for every $h(\cdot)$ and $\phi(\cdot)$, such that $d_\phi^h(\cdot, \cdot)$ is a valid distance metric. In the next section, closed from expressions for the metric d_ϕ^h for some of the divergence measures listed in Table 6.1 are provided.

6.4.1 Expressions for Various $d_\phi^h(\cdot, \cdot)$ Distances

For further analysis and simulation study, we will consider the following distance metrics. Since the metric $\mathcal{S}(\cdot, \cdot)$ for any $d_\phi^h(\cdot, \cdot)$ follows the same distribution for large enough L or M, it is sufficient to consider any one of the metrics. Later, through simulations, we confirm that each of the following metrics give the same performance. The expressions given in this section are special cases of the results given by Frery et al [139, Sec. 3.1].

1. *The Kullback-Leibler distance* (KL)

$$d_{KL}(\mathbf{R}, \mathbf{R}') = LM\left(\text{tr}\left\{\frac{\mathbf{R}'^{-1}\mathbf{R} + \mathbf{R}^{-1}\mathbf{R}'}{2}\right\} - N\right), \tag{6.23}$$

where $\text{tr}(\mathbf{R})$ represents the trace of the matrix $\mathbf{R}$.

2. *The Bhattacharyya distance* (B)

$$d_B(\mathbf{R}, \mathbf{R}') = \frac{LM}{2}\left[\log\det(\mathbf{R}) + \log\det(\mathbf{R}')\right] - LMN\log(LM)$$

$$-LM\log\det\left[\left(\frac{LM\mathbf{R}^{-1} + LM\mathbf{R}'^{-1}}{2}\right)^{-1}\right], \qquad (6.24)$$

where $\det(\mathbf{R})$ represents the determinant of a matrix $\mathbf{R}$.

3. *The Hellinger distance* (H)

$$d_H(\mathbf{R}, \mathbf{R}') = 1 - \left\{\frac{\det\left[\left(\frac{LM\mathbf{R}^{-1} + LM\mathbf{R}'^{-1}}{2}\right)^{-1}\right]}{\mathbf{R}^{(LM/2)}\mathbf{R}'^{(LM/2)}}\right\} \times LM^{LMN}. \qquad (6.25)$$

6.4.2 Robustness of the KL Distance Metric $d_{KL}(\cdot, \cdot)$

In this section, we discuss a robustness feature of the Kullback-Leibler distance metric $d_{KL}(\cdot, \cdot)$. As observed by Frery et al. [139, Sec. 4.3], the test statistic $\mathcal{S}(\mathbf{R}, \cdot)$ is robust to small "contaminations" in the observations under $\mathcal{H}_0$, i.e., it satisfies the false-alarm constraint under the following condition. When the matrix $\mathbf{R}$ under $\mathcal{H}_0$ is of the following form

$$\mathbf{R} \stackrel{d.}{=} \epsilon \mathcal{W}_N(LM, \kappa\boldsymbol{\Sigma}) + (1 - \epsilon)\mathcal{W}_N(LM, \boldsymbol{\Sigma}), \qquad (6.26)$$

where κ is a large number (~ 1000), and $0 \leq \epsilon \ll 1$. In other words, under $\mathcal{H}_0$, with probability $1 - \epsilon$, the noise observations come from the regular Wishart distribution, and with probability ϵ, the observations follow a Wishart distribution, whose underlying Gaussian distribution has a much larger variance. The model in (6.26) is closely related to the ϵ-mixture model, which characterizes the PDF of the background noise

and Middleton's Class A model ([55], [109], [111]), which was studied in Chap. 5. The main motivation for introducing the Weighted Zero-Crossings Detector (WZCD) in Chap. 5 was the requirement of robustness of the GoFT to the presence of Class A noise component. Therefore, the KL distance based $\langle h, \phi \rangle$ detector can be seen as a solution for robustness to the presence of Class A noise component in the multi-dimensional SS problem setup.

6.5   Simulation Results

As seen ealier, the performance of a GoFT for SS in a CR network needs to be studied through extensive simulations. We consider a multiple sensor setup, with multiple primary users with a Rayleigh fading channel from each primary transmitter to each CR node, and is i.i.d. across space and time. For both ID and $\langle h, \phi \rangle$ detectors, and in each case, the thresholds are chosen such that the target false alarm probability is fixed to be $p_f = 0.01$. Unless mentioned, the distributions of primary-only observations are modeled as Gaussian, i.i.d. across sensors. The legend entries ID, $\langle h, \phi \rangle$ KL, $\langle h, \phi \rangle$ B, $\langle h, \phi \rangle$ H, John, ER and ST represent Interpoint distance, KL distance, Bhattacharyya distance, Hellinger distance, John's, eigenvalue ratio based, and sphericity test based detectors, respectively.

6.5.1   ID Test

Figure 6.3 shows the performance of ID with John's, ER and ST for varying values of the total primary average SNR, for $L = 1$, $M = 100$, $N = 5$, $P = 2$. The SNR received due to each primary user is assumed to be equal. The distance metric was chosen to be an

l_p norm, with $p = 3$. It is seen that the ID performs better than all the other techniques.
Under the same conditions, the performance of all detectors for $p = 2$ is shown in Fig.
6.4. In the low SNR regime, ID performs better than all the other techniques, and as
the SNR increases, ER performs slightly better. In Fig. 6.5, detection of a single PU is
considered with $L = 1$, $M = 100$, and $N = 5$. In this scenario, ID performs slightly
better that ER in low SNR regime, and ER performs better as SNR increases. Detection
of a single PU with $L = 1$, $M = 80$, and $N = 5$, as considered in Fig. 6.6 shows a
different trend where ER performs better that ID. However, ID performs better that
both John and ST detectors.
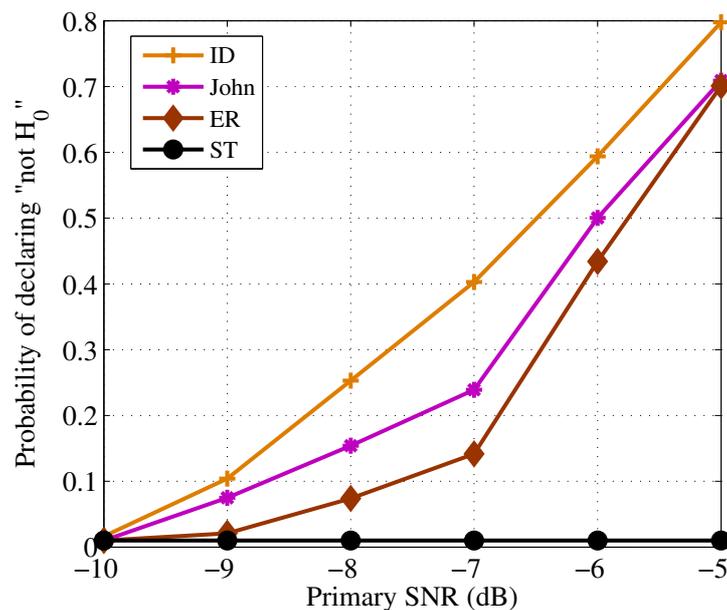
Figure 6.3: Performance comparison of detection of primary under Rayleigh fading,
with $L = 1$, $M = 100$, $N = 5$, $P = 2$, and $p = 3$.
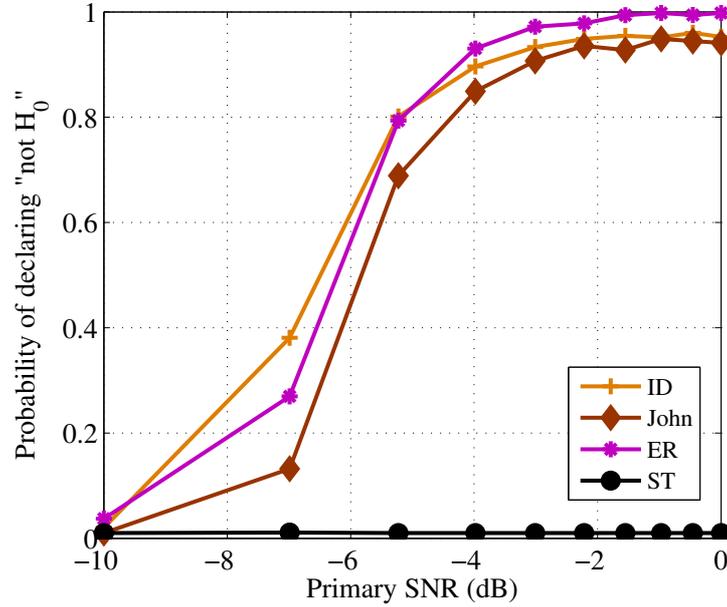
Figure 6.4: Performance comparison of detection of primary under Rayleigh fading, with $L = 1$, $M = 100$, $N = 5$, $P = 2$, and $p = 2$.

6.5.2 $\langle h, \phi \rangle$ Test

Figure 6.7 shows the performance comparison of the Kullback Liebler (KL) and Bhattacharyya (B) based $\langle h, \phi \rangle$ detectors with the ER, John and ST detectors, with respect to the total primary SNR, for $L = 10$, $M = 200$, $N = 4$ and $P = 3$. As expected, the performances of KL and B detectors are nearly equal. As the SNR increases, a huge performance improvement is observed in using the $\langle h, \phi \rangle$ detectors, as compared to other techniques. Similar trends are observed in Fig. 6.8, where the parameters are chosen to be $L = 10$, $M = 200$, $N = 4$ and $P = 5$. Since the presence of an extra PU increases the detection SNR, performance improvements are seen in Fig. 6.8, as compared to Fig. 6.7. In Fig. 6.9, where the parameters are $L = 10$, $M = 50$, $N = 5$ and $P = 4$, it is seen that as the SNR increases, improvements in the performances of all the detectors are seen. However, even in this case, the $\langle h, \phi \rangle$ based detectors perform better than the
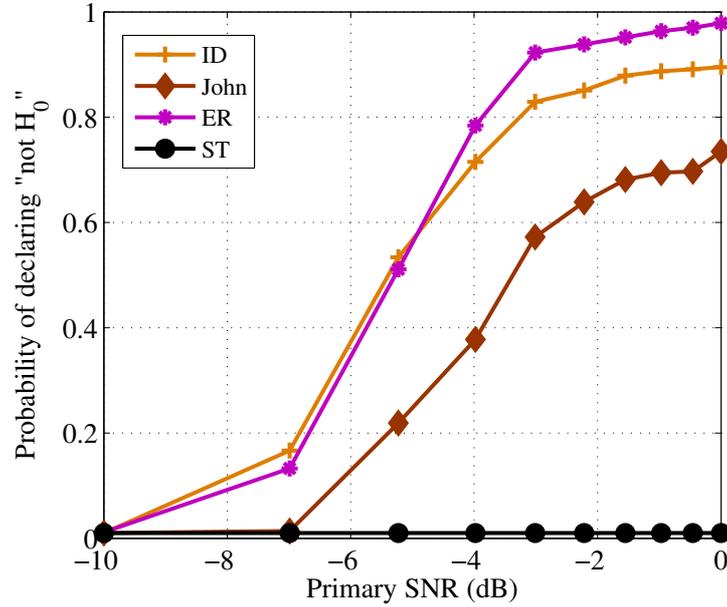
Figure 6.5: Performance comparison of detection of primary under Rayleigh fading, with $L = 1$, $M = 100$, $N = 5$, $P = 1$, and $p = 2$.

others. Similar trends are carried over in Fig. 6.10, with $L = 10$, $M = 50$, $N = 5$ and $P = 5$. As seen earlier, a performance improvement in all the detectors are seen due to the presence of extra PUs. Finally, Fig. 6.11 shows the performance curves of the all the $\langle h, \phi \rangle$ detectors studied in Sec. 6.4.1. As expected, the performance of all the detectors are nearly the same, across different values of P.

6.6 Conclusions

In this chapter, we studied two multi-dimensional Goodness-of-Fit tests for spectrum sensing in cognitive radios. Both the tests, viz., the Interpoint Distance (ID) based test and the $\langle h, \phi \rangle$ distance based tests were constructed based on the properties of stochastic distances. The construction of the ID test was studied for a single CR node case with multiple antenna, multiple observations from multiple primary users. The $\langle h, \phi \rangle$ test
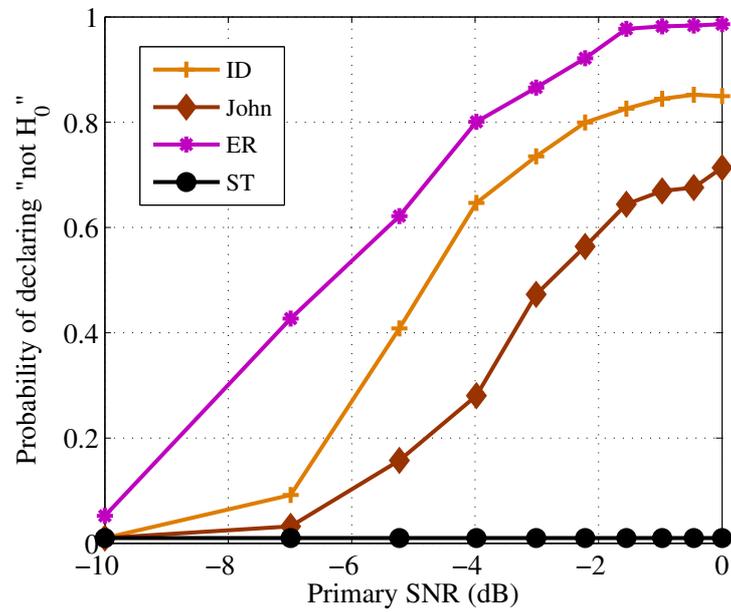
Figure 6.6: Performance comparison of detection of primary under Rayleigh fading, with $L = 1$, $M = 80$, $N = 5$, $P = 1$, and $p = 2$.

was studied for the multiple CR nodes. Also, a robustness feature of the KL distance

based test was studied, which has connections with Middleton's Class A noise model.

The proposed tests were shown to perform better that the existing techniques such as

the eigenvalue ratio based test, John's test, and the sphericity test, in several scenarios.
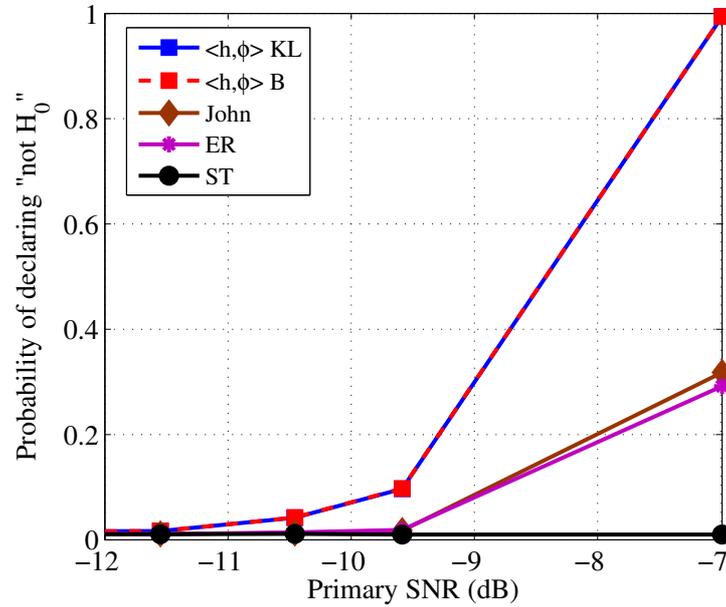
Figure 6.7: Performance comparison of detection of primary under Rayleigh fading, with $L = 10$, $M = 200$, $N = 4$, $P = 3$.
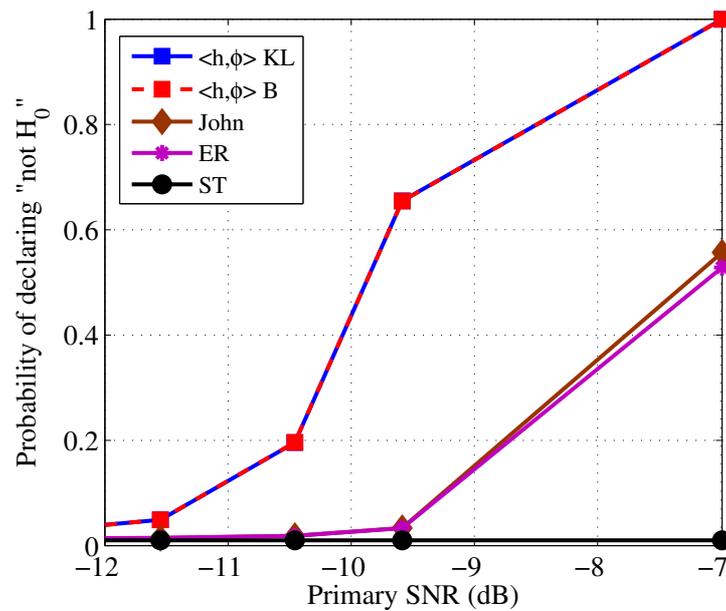
Figure 6.8: Performance comparison of detection of primary under Rayleigh fading, with $L = 10$, $M = 200$, $N = 4$, $P = 5$.
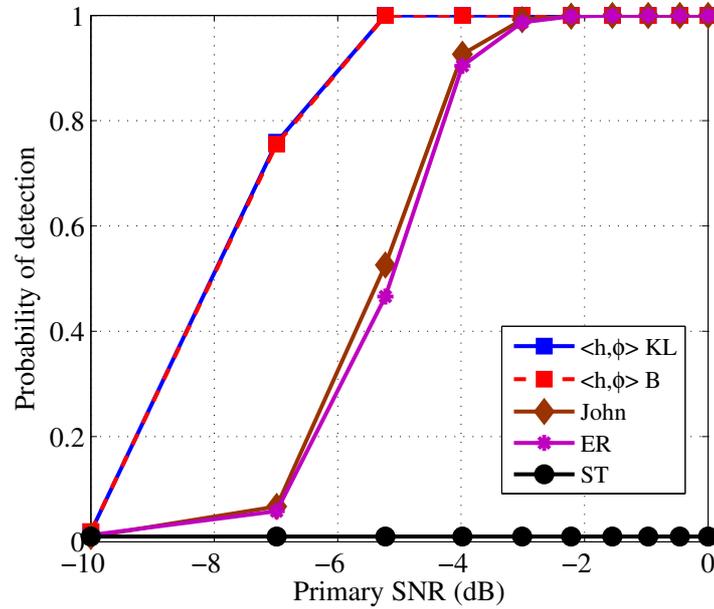
Figure 6.9: Performance comparison of detection of primary under Rayleigh fading, with $L = 10$, $M = 50$, $N = 5$, $P = 4$.
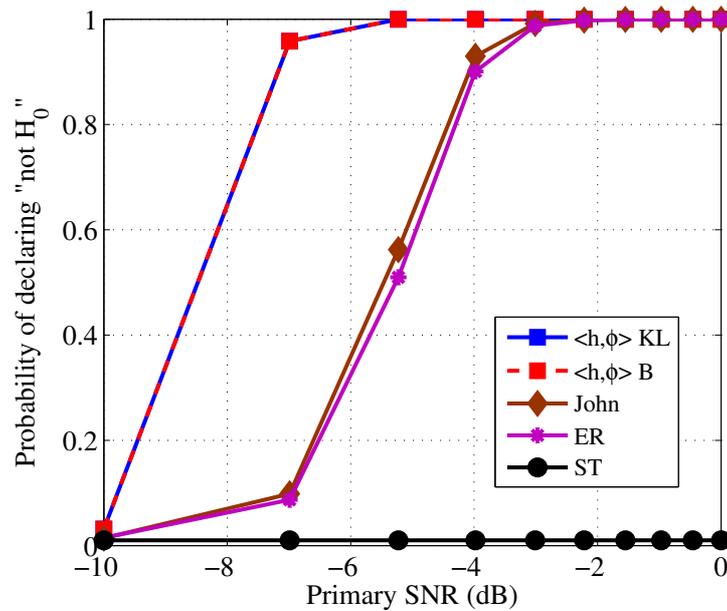
Figure 6.10: Performance comparison of detection of primary under Rayleigh fading, with $L = 10$, $M = 50$, $N = 5$, $P = 5$.
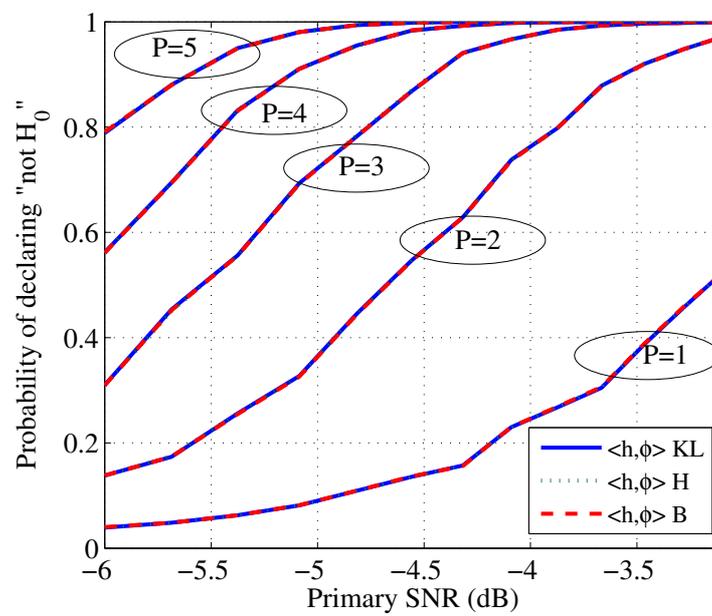
Figure 6.11: Performance comparison of detection of primary under Rayleigh fading, with $L = 20$, $M = 15$, $N = 1$.

Chapter 7

Conclusions and Future Work

In this thesis, we investigated the problem of spectrum sensing in cognitive radios. The main contributions of this thesis are summarized below.

7.1   Contributions

In chapter 2, we analyzed the performance of energy-based Bayesian decentralized detection for spectrum sensing in cognitive radios. We showed that, for various fading models, with the OR rule for decision fusion, the conventional error exponent is equal to zero. We introduced a novel performance metric called the Error Exponent with a Confidence Level (EECL), and showed that the EECL at a given confidence level $q < 1$ is strictly positive. We used the EECL to answer the question of whether it is better to sense for the pilot tone in a narrow band, or to sense the entire wide-band signal. We also derived simplified expressions for finding the detection threshold and the EECL for the i.i.d. Rayleigh fading and lognormal shadowing cases. We validated the theoretical expressions through Monte Carlo simulations.

Chapter 3 considered the problem of Bayesian decentralized SS in CR networks under

various fading environments. A CLT-based approximation was explored, which led to analytical expressions for near-optimal detection thresholds for Rayleigh, Lognormal, Nakagami-m, Weibull fading cases. For the Suzuki fading case, a generalized gamma approximation was provided, which saves on the computation of an integral. Also, in the Rayleigh fading case, a structural property of the detector, viz., the trade-off between M and P to maintain a given p_e at the individual sensors, in the low SNR regime, was discussed. Extending to the decentralized case with N sensors, the optimal exponent on P_E was derived in closed form. The accuracy of the theoretical expressions and the diversity gain obtainable through the use of N sensors were illustrated through simulations.

In chapter 4, we considered the problem of spectrum sensing in the presence of a multiuser frequency-hopping primary network. We theoretically analyzed the performance of the FAR algorithm, and validated the results through simulations. We derived the sensing duration that maximizes the throughput of the CR system, under a constraint on the interference to the primary network. We also presented a technique to synchronize the CR system with the primary hopping instants. The FAR algorithm was implemented on Lyrtech SFF SDR DP and its performance corroborated well with the ROCs obtained from Monte Carlo simulations.

In chapter 5, a weighted zero-crossings based goodness-of-fit test for spectrum sensing was proposed. A near-optimal detection threshold was derived for the specific choices of uniform and exponential weights. It was shown that this detector is robust to uncertainties in the noise model and parameters. Through simulations, it was shown

that the proposed detectors outperform the existing tests in the CR literature in a variety of noise and primary signal conditions of practical interest. Also, the computational simplicity of the proposed test was highlighted. Therefore, the proposed detector is a promising choice for spectrum sensing in CR, and can be used in a wide range of communication scenarios.

Finally, in chapter 6, we studied two multi-dimensional Goodness-of-Fit tests for spectrum sensing in cognitive radios. Both the tests, viz., the Interpoint Distance (ID) based test and the $\langle h, \phi \rangle$ distance based tests were constructed based on the properties of stochastic distances. The construction of the ID test was studied for a single CR node case with multiple antenna, multiple observations from multiple primary users. The $\langle h, \phi \rangle$ test was studied for the multiple CR nodes. Also, a robustness feature of the KL distance based test was studied, which has connections with Middleton's Class A noise model. The proposed tests were shown to perform better that the existing techniques such as the eigenvalue ratio based test, John's test, and the sphericity test, in several scenarios.

7.2 Future Work

Future work for spectrum sensing in cognitive radios could include the following issues. Some of them are already being addressed by the author.

- For the Bayesian decentralized detection, incorporating correlation in the signal or noise, extending the results to allow for time-varying channels, and optimally combining the outcomes from NB and WB spectrum sensing, could be interesting extensions.

- In Bayesian SS, the low-rate channel between the individual sensors and the FC was assumed to be lossless. Design and analysis of decentralized SS accounting for the effect of fading and noise in this channel would be useful under practical scenarios.

- For the detection of frequency-hopping primary signals, there exists a tradeoff: depending on the total number of available bands, active primary and secondary users, a CR can either continue to sense for a channel throughout the sensing duration, or to jump into another possibly vacant channel and perform fresh sensing. This exploration-exploitation tradeoff is challenging, and could lead to interesting results.

- The WZCD, or in general, any CFAR detector was found to fail when the primary signal also follows another Gaussian distribution with a different variance. Given the advantages of the WZCD, modifying the Ψ_M^2 statistic or proposing a new detector based on the statistics of zero-crossings to incorporate detection of the Gaussian signals could make the GoFT stronger.

- In general, spectrum sensing under an energy efficiency constraint have not been dealt in greater detail in the literature so far. Given the importance of green communications, such a study could lead to important designs that consume less energy and yet offer good performance.

Appendix A

Appendix for Chapter 2

A.1 Proof of Theorem 1

It is straightforward to show that, under Rayleigh fading, the likelihood ratio test corresponding to (2.1) is monotonically increasing in V_y, and hence, the optimum test reduces to a threshold test on V_y itself. That is, it declares H_1 when $V_y \geq x$, where x is the detection threshold. Let $f_\alpha(\alpha)$ denote the pdf of α. Conditioned on $\mathcal{A} \triangleq \{\alpha \geq \alpha_0\}$, the pdf of α is $f_{\alpha|\mathcal{A}}(\alpha|\mathcal{A}) = f_\alpha(\alpha)/\mathcal{P}(\mathcal{A}), \alpha \geq \alpha_0$. By construction, $\mathcal{P}(\mathcal{A}) = q$. The probability of error is given by

$$p_e = \pi_0 Q(x\sqrt{M}) + \pi_1 \int\limits_{-\infty}^{x} \int\limits_{\alpha_0}^{\infty} f_{\mathcal{N}}\left(v - \alpha P, \frac{1}{\sqrt{M}}\right) \frac{f_\alpha(\alpha)}{q} \mathrm{d}\alpha \mathrm{d}v, \tag{A.1}$$

where $f_{\mathcal{N}}(x, \sigma)$ is the Gaussian pdf with mean zero and variance σ^2 evaluated at x, $\pi_0 \triangleq \mathcal{P}(H_0)$, and $\pi_1 \triangleq \mathcal{P}(H_1) = 1 - \pi_0$. To find the optimum threshold, we differentiate the above w.r.t. x and equate to 0. After some simplification, we get

$$\frac{q\pi_0}{\pi_1} = \int\limits_{\alpha_0}^{\infty} \exp\left(M\left(x\alpha P - \frac{\alpha^2 P^2}{2}\right)\right) f_\alpha(\alpha) \mathrm{d}\alpha. \tag{A.2}$$

Let x_M denote the solution to the above equation for a given value of M.[1] First, we show that x_M converges to $\alpha_0 P/2$. To do this, we show that neither $x_M < \alpha_0 P/2$ nor $x_M > \alpha_0 P/2$ are possible for large M, as they lead to a contradiction. Define $g(x, \alpha) \triangleq x\alpha P - \alpha^2 P^2/2$. Note that $g(\alpha_0 P/2, \alpha) \leq 0$ for $\alpha \geq \alpha_0$. If $x_M < \alpha_0 P/2$, since $g(x, \alpha)$ is monotonic in x, we have $g(x_M, \alpha) < 0$ for $\alpha \geq \alpha_0$. Let $g_{\max} \triangleq \max_{\alpha \geq \alpha_0} g(x_M, \alpha)$, and note that $g_{\max} < 0$. Then, using $g_{\max} \geq g(x_M, \alpha)$ in (A.2) results in the following upper bound on the right hand side (RHS): RHS $\leq \exp(Mg_{\max}) \int_{\alpha_0}^{\infty} f_\alpha(\alpha) \mathrm{d}\alpha$. Since $g_{\max} < 0$, the upper bound can be made as small as desired by choosing M sufficiently large. Thus, if $x_M < \alpha_0 P/2$, the right hand side goes to zero as M gets large, and hence, attaining equality in (A.2) is not possible. Hence, x_M must satisfy $x_M \geq \alpha_0 P/2$.

Next, we show that $x_M \geq x_0 > \alpha_0 P/2$ also leads to a contradiction. Consider α such that $g(x_0, \alpha) > 0$. This corresponds to $\alpha < 2x_0/P$. By the assumption, we have $\alpha_0 < 2x_0/P$, so that, $g(x_0, \alpha) > 0$ for $\alpha_0 \leq \alpha < 2x_0/P$. Further, if $g(x_0, \alpha) > 0$ and $x_M \geq x_0$, we have $g(x_M, \alpha) > 0$. Therefore, there exists an $\epsilon > 0$ such that $g(x_M, \alpha) > 0$ for $\alpha_0 \leq \alpha \leq 2x_0/P - \epsilon$. Let $g_{\min} \triangleq \min_{\alpha \in [\alpha_0, 2x_0/P - \epsilon]} g(x_M, \alpha)$, and note that $g_{\min} > 0$. Then, the right hand side in (A.2) can be lower bounded as

$$\text{RHS} \geq \int_{\alpha_0}^{\frac{2x_0}{P} - \epsilon} \exp\left(Mg(x_M, \alpha)\right) f_\alpha(\alpha) \mathrm{d}\alpha \tag{A.3}$$

$$\geq \exp\left(Mg_{\min}\right) \int_{\alpha_0}^{\frac{2x_0}{P} - \epsilon} f_\alpha(\alpha) \mathrm{d}\alpha. \tag{A.4}$$

Since $g_{\min} > 0$, the above lower bound can be made as large as desired by choosing M sufficiently large, since the integral term is a strictly positive constant. This implies that if $x_M \geq x_0 > \alpha_0 P/2$, the right hand side grows unbounded as M gets large, and hence,

---

[1]That a unique solution exists can be seen from simple monotonicity arguments.

attaining equality in (A.2) is not possible. Hence, x_M converges to $\alpha_0 P/2$ as M goes to infinity.

Now, consider the exponent due to the false alarm term. This is simply given by

$$\epsilon_f \triangleq \lim_{M \to \infty} \frac{-\log\left(Q(x_M \sqrt{M})\right)}{M} = \frac{\alpha_0^2 P^2}{8}. \tag{A.5}$$

In the above, $Q(y)$ is the standard Gaussian tail probability evaluated at y. The second equality above is obtained by upper and lower bounding $Q(y)$ for large y and showing that both limits equal as $M \to \infty$. Since the exponents due to the false alarm and the missed detection are equal in a Bayesian set-up [48, Chap. 11], [71], it follows that the EECL(q) on the probability of error is $\frac{\alpha_0^2 P^2}{8}$, where α_0 is chosen to satisfy $\mathcal{P}(\alpha > \alpha_0) = q$.

A.2 Proof of Theorem 2

Suppose that the hypothesis $\mathcal{H}_0$ is true. With the OR fusion rule, a false alarm at any of the sensors results in a false alarm at the FC. Since, conditioned on $\mathcal{H}_0$, the sensor decisions are independent, the false alarm probability at the FC, denoted by P_F, is simply $1 - (1 - p_f)^N$, where p_f is the false alarm probability at an individual sensor. Now, given the detection threshold $\frac{\alpha_0 P}{2}$ at the sensors, the exponent ϵ_F at the FC is determined by the p_f term in the expansion of $1 - (1 - p_f)^N$. Thus, the error exponent at the FC is the same as that at the individual sensors, i.e., $\epsilon_F = \frac{(\alpha_0 P)^2}{8}$.

Suppose that the hypothesis $\mathcal{H}_1$ is true. Conditioned on α_j, the channel power gain from the primary transmitter to sensor j, the decision statistic V_y at the jth sensor is

distributed as $\mathcal{N}\left(\alpha_j P, \frac{1}{M}\right)$. Since the jth sensor uses a threshold of $\frac{\alpha_0 P}{2}$ for detection, using well-known bounds on the Q-function,[2] it is easy to show that the missed detection probability at the jth sensor conditioned on α_j, denoted $p_{m_j|\alpha_j}$, is given by

$$p_{m_j|\alpha_j} = \mathcal{P}\left\{V_y < \frac{\alpha_0 P}{2}\,\middle|\,\alpha_j\right\}$$

$$= Q\left(\sqrt{M}\left(\frac{\alpha_0 P}{2} - \alpha_j P\right)\right) \doteq \exp\left(-\frac{M}{2}\left(\alpha_j P - \frac{\alpha_0 P}{2}\right)^2 \mathbb{I}_{\left\{\alpha_j > \frac{\alpha_0}{2}\right\}}\right) \quad \text{(A.6)}$$

where the notation $f(M) \doteq \exp(-M\beta)$ is used to mean $\lim_{M\to\infty} \frac{-\log f(M)}{M} = \beta$. That is, the jth sensor achieves an error exponent of $\frac{1}{2}\left(\alpha_j P - \frac{\alpha_0 P}{2}\right)^2$ if $\alpha_j > \frac{\alpha_0 P}{2}$, and zero otherwise.

With the OR fusion rule, when hypothesis $\mathcal{H}_1$ is true, the FC makes an error and declares $\mathcal{H}_0$ only if all the sensors make an error. Hence, given $\alpha_1, \cdots, \alpha_N$, the missed detection probability at the FC $P_{M|\alpha_1,\cdots,\alpha_N}$ is given by

$$P_{M|\alpha_1,\cdots,\alpha_N} = \prod_{j=1}^{N} p_{m_j} \doteq \exp\left(-\frac{M}{2}\sum_{j=1}^{N}\left(\alpha_j P - \frac{\alpha_0 P}{2}\right)^2 \mathbb{I}_{\left\{\alpha_j > \frac{\alpha_0}{2}\right\}}\right). \quad \text{(A.7)}$$

Now consider the case where $\alpha_1, \alpha_2, \ldots, \alpha_N$ are random. The FC attains an EECL(q) of ϵ_M, provided

$$\mathcal{P}\left\{\frac{1}{2}\sum_{j=1}^{N}\left(\alpha_j P - \frac{\alpha_0 P}{2}\right)^2 \mathbb{I}_{\left\{\alpha_j > \frac{\alpha_0}{2}\right\}} \le \epsilon_M\right\} \le 1 - q, \quad \text{(A.8)}$$

where the probability is taken over the distribution of $\alpha_1, \alpha_2, \ldots, \alpha_N$. The best error exponent is obtained, i.e., ϵ_M is maximized, when the left hand side above equals $1 - q$, since, otherwise, ϵ_M (and α_0) can be increased to improve the error exponent.

---

[2]For example, $\frac{y}{1+y^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}} \le Q(y) \le \frac{1}{y\sqrt{2\pi}}e^{-\frac{y^2}{2}}$.

Finally, for optimal Bayesian detection, the exponent associated with the false alarm and missed detection must be equal, i.e., $\epsilon_F = \epsilon_M$ [48, Chap. 11], [71]. Hence, substituting $\epsilon_M = \frac{(\alpha_0 P)^2}{8}$ in (A.8) and simplifying, we get (2.2), which completes the proof.

A.3 Proof of Corollary 1

Let α_0 and ℓ_0 denote the solution to (2.2) under the Rayleigh fading and lognormal shadowing cases, respectively. Let $\texttt{expn}(\lambda)$ and $\texttt{LN}(\mu, \sigma)$ denote the exponential distribution with parameter λ and the lognormal distribution with parameters μ and σ, respectively. Now, under Rayleigh fading, $t_j \triangleq \frac{2\alpha_j}{\alpha_0} \sim \texttt{expn}\left(\frac{2}{\alpha_0}\right)$, while under lognormal shadowing, with a slight abuse of notation, $t_j \triangleq \frac{2\alpha_j}{\ell_0} \sim \texttt{LN}\left(\mu_s + \log\left(\frac{2}{\ell_0}\right), \sigma_s\right)$. For notational convenience, let $Z \triangleq \sum_{j=1}^{N} (t_j - 1)^2 \, \mathbb{I}_{\{(t_j - 1) \geq 0\}}$. From Theorem 2, note that we need to find α_0 such that $F_Z(1) = 1 - q$, where $F_Z(\cdot)$ is the CDF of Z.

$$
\begin{aligned}
\mathcal{P}\{Z \leq 1\} &= \sum_{l=0}^{N} \mathcal{P}\{l \text{ out of } N \ t_j\text{'s are } \geq 1\} \mathcal{P}\{Z \leq 1 | l \text{ out of } N \ t_j\text{'s are } \geq 1\} \quad \text{(A.9)} \\
&= \sum_{l=1}^{N} \binom{N}{l} (\mathcal{P}\{t_j \leq 1\})^{N-l} (\mathcal{P}\{t_j > 1\})^l \\
&\quad \mathcal{P}\left\{\sum_{k=1}^{l} (t_k - 1)^2 \leq 1 \,\middle|\, t_k > 1, k = 1, \ldots, l\right\} + (\mathcal{P}\{t_k \leq 1\})^N,
\end{aligned}
$$

which should equal $1 - q$ by requirement. For Rayleigh and shadowing caess, $t_k \sim \texttt{expn}\left(\frac{2}{\alpha_0}\right)$ and $t_k \sim \texttt{LN}\left(\mu_s + \log\left(\frac{2}{\ell_0}\right), \sigma_s\right)$, respectively. In the Rayleigh fading case, by the memoryless property of exponential random variables, it is easy to show that

$$
\mathcal{P}\left\{\sum_{k=1}^{l} (t_k - 1)^2 \leq 1 \,\middle|\, t_k > 1, k = 1, \ldots, l\right\} = \mathcal{P}\left\{\sum_{k=1}^{l} a_k^2 \leq 1\right\}, \quad \text{(A.10)}
$$

where $a_k \sim \mathtt{expn}\left(\frac{2}{\alpha_0}\right)$ are independent and exponentially distributed. Since $\mathcal{P}\left\{t_k > 1\right\}$ $= e^{-\frac{2}{\alpha_0}}$, (A.10) reduces to the expression in (2.3).

The proof for the lognormal shadowing case is similar to the Rayleigh fading case, and follows by noting that $\mathcal{P}\left\{t_k \leq 1\right\} = \mathcal{P}\left\{\log t_k \leq 0\right\} = Q\left(\frac{\mu_s + \log\left(\frac{2}{\ell_0}\right)}{\sigma_s}\right)$. Further,

$$\mathcal{P}\left\{\sum_{k=1}^{l}(t_k - 1)^2 \leq 1 \,\middle|\, t_k > 1, k = 1, \ldots, l\right\} = \mathcal{P}\left\{\sum_{k=1}^{l}(e^{y_k} - 1)^2 \leq 1\right\}. \qquad \text{(A.11)}$$

In the above, $y_k \triangleq \log t_k$, and, due to the conditioning on $t_k > 1$, we have that y_k has a truncated Gaussian distribution, with pdf $\dfrac{\mathcal{N}\left(\mu_s + \log\left(\frac{2}{\ell_0}\right), \sigma_s^2\right)}{Q\left(-\frac{\mu_s + \log\left(\frac{2}{\ell_0}\right)}{\sigma_s}\right)}$ for $y_k > 0$ and zero otherwise.

A.4   Proof of Corollary 2

Consider the left hand side of (2.3). Upper bounding the terms in the expression would lead to a lower bound on α_0, and, consequently, on the EECL(q). First, note that $1 - \exp\left(\frac{\alpha_0}{2}\right) \leq \frac{\alpha_0}{2}$. Also, a_k in (2.3) is distributed as $f_{a_k}(a_k) = \frac{\alpha_0}{2}\exp\left(-\frac{\alpha_0 a_k}{2}\right)$, $a_k \geq 0$, and hence, $f_{a_k}(a_k) \leq \frac{\alpha_0}{2}$. Thus, by replacing the pdf of a_k with its upper bound, we get

$$\mathcal{P}\left(\sum_{k=1}^{l} a_k^2 \leq 1\right) = \int_{\sum_{k=1}^{l} a_k^2 \leq 1, a_k \geq 0} f_{a_k}(a_1) f_{a_k}(a_2) \cdots f_{a_k}(a_l) \mathrm{d}a_1 \mathrm{d}a_2 \cdots \mathrm{d}a_l$$

$$\leq \left(\frac{\alpha_0}{2}\right)^l \int_{\sum_{k=1}^{l} a_k^2 \leq 1, a_k \geq 0} \mathrm{d}a_1 \mathrm{d}a_2 \cdots \mathrm{d}a_l = \left(\frac{\alpha_0}{2}\right)^l \frac{\mathcal{V}_l}{2^l}, \qquad \text{(A.12)}$$

where $\mathcal{V}_l = \frac{\pi^{\frac{k}{2}}}{\Gamma\left(1 + \frac{k}{2}\right)}$ is the volume of the l-dimensional unit sphere, with $\Gamma(\cdot)$ being the Gamma function. The 2^l factor in the denominator arises because only the volume of the first orthant is relevant here, since $a_k \geq 0$. Substituting in (2.3), we get a lower

bound on α_0 by solving

$$\left(\frac{\alpha_0^{LB}}{2}\right)^N + \sum_{l=1}^{N} \binom{N}{l} \left(\frac{\alpha_0^{LB}}{2}\right)^{N-l} \left(\frac{\alpha_0^{LB}}{2}\right)^l \frac{\mathcal{V}_l}{2^l} = 1 - q. \tag{A.13}$$

The result in (2.6) follows from rearranging the above equation.

The proof for the lognormal shadowing case is similar. Starting from (2.4), using a well-known bound on the Q-function, we upper bound $P_{\mathcal{A}}$ as

$$P_{\mathcal{A}} \leq \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\left(\log\left(\frac{2}{\ell_0}\right)\right)^2}{2\sigma_s^2}\right). \tag{A.14}$$

Next, conditioned on $y_k > 0$, it is easy to show that $z_k \triangleq e^{y_k} - 1$ is distributed as

$$f_{z_k}(z_k) = \frac{1}{(z_k+1)\sigma_s\sqrt{2\pi}} \frac{\exp\left(-\frac{(\log(z_k+1)-\log(2/\ell_0))^2}{2\sigma_s^2}\right)}{Q\left(\frac{-\log(2/\ell_0)}{\sigma_s}\right)}, \, z_k \geq 0. \tag{A.15}$$

Further, since $\ell_0 \leq 1$, setting $z_k = 0$ in the right hand side above leads to an upper bound on $f_{z_k}(z_k)$. Hence, we have

$$\mathcal{P}\left(\sum_{k=1}^{l} z_k^2 \leq 1\right) \leq \left(\frac{1}{\sigma_s\sqrt{2\pi}} \frac{\exp\left(-\frac{(\log(2/\ell_0))^2}{2\sigma_s^2}\right)}{Q\left(-\frac{\log(2/\ell_0)}{\sigma_s}\right)}\right)^l \frac{\mathcal{V}_l}{2^l}. \tag{A.16}$$

Substituting the upper bounds in (A.14) and (A.16) into (2.4), and using the fact that $P_{\mathcal{A}^c} = Q\left(-\frac{\log(2/\ell_0)}{\sigma_s}\right)$, and simplifying, we get the result in (2.7).

A.5   Proof of Theorem 3

It is known that, with conditionally i.i.d. observations at the sensors, the probability of error at the FC is minimized by the K out of N rule, and the optimum K is given by [84]

$$K_{opt} = \min\left(N, \left\lceil \frac{\log\left(\frac{\pi_0}{1-\pi_0}\right) + N\log\left(\frac{1-p_f}{p_m}\right)}{\log\left\{\left(\frac{1-p_m}{p_f}\right)\left(\frac{1-p_f}{p_m}\right)\right\}} \right\rceil\right), \tag{A.17}$$

where p_f and p_m are the false alarm and missed detection probabilities, respectively, at the individual nodes. Now, given the detection threshold $\frac{\alpha_0 P}{2} > 0$ at the individual sensors, p_f clearly decreases with an exponent $\frac{(\alpha_0 P)^2}{8}$. On the other hand, whenever $\alpha < \alpha_0$, the missed detection probability of the hypothesis test in (2.1) is lower bounded by $\frac{1}{2}$. Since the event $\alpha < \alpha_0$ occurs with a nonzero probability, the exponent on p_m is 0. Thus,

$$\left\lceil \frac{\log\left(\frac{\pi_0}{1-\pi_0}\right) + N\log\left(\frac{1-p_f}{p_m}\right)}{\log\left\{\left(\frac{1-p_m}{p_f}\right)\left(\frac{1-p_f}{p_m}\right)\right\}} \right\rceil \to 1, \tag{A.18}$$

since the numerator approaches a constant, while the denominator is linearly increasing with M. Thus, for sufficiently large M, $K_{opt} = 1$, i.e., the OR fusion rule is optimal.

A.6   Expressions for Approximations in Sec. 2.4, Cor. 1

A.6.1   Weibull Sum Approximation in Rayleigh Fading Case

Consider the probability term in (2.3), viz., $\mathcal{P}\left\{\sum_{k=1}^{l} a_k^2\right\} \leq 1$, where a_k is a exponentially distributed random variable with parameter $\frac{2}{\alpha_0}$. Following the genesis [116], a random variable X is said to be Weibull distributed with shape and scale parameters a_w and b_w, respectively, if $\left(\frac{X}{b_w}\right)^{a_w}$ is a standard exponential random variable. Using this result, it is easy to see that a_k^2 follows a Weibull distribution with parameters $a_w = 0.5$,

and $b_w = \alpha_0/2$.

Now, following [85], the PDF and CDF of the term $A_l \triangleq \sum_{k=1}^{l} a_k^2$ can be tightly approximated by a $\alpha - \mu$ distribution, whose expressions are given by

$$f_{A_l}(a) = \frac{\alpha \mu^\mu a^{\alpha\mu-1}}{\Omega^\mu \Gamma(\mu)} \exp\left(-\frac{\mu a^\alpha}{\Omega}\right) \text{; and } F_{A_l}(a) = \frac{\gamma_{inc}\left(\mu, \frac{\mu a^\alpha}{\Omega}\right)}{\Gamma(\mu)}, \tag{A.19}$$

where $\gamma_{inc}(\cdot, \cdot)$ is the lower incomplete gamma function.

Using the following equations give moment based estimators for α, μ and Ω.

$$\frac{\Gamma^2\left(\mu + \frac{1}{\alpha}\right)}{\Gamma(\mu)\Gamma\left(\mu + \frac{2}{\alpha}\right) - \Gamma^2\left(\mu + \frac{1}{\alpha}\right)} = \frac{\mathbb{E}^2 A_l}{\mathbb{E}A_l^2 - \mathbb{E}^2 A_l}$$

$$\frac{\Gamma^2\left(\mu + \frac{2}{\alpha}\right)}{\Gamma(\mu)\Gamma\left(\mu + \frac{4}{\alpha}\right) - \Gamma^2\left(\mu + \frac{2}{\alpha}\right)} = \frac{\mathbb{E}^2 A_l^2}{\mathbb{E}A_l^4 - \mathbb{E}^2 A_l^2}$$

$$\Omega = \left[\frac{\mu^{1/\alpha}\Gamma(\mu)\mathbb{E}A_l}{\Gamma\left(\mu + \frac{1}{\alpha}\right)}\right]^\alpha. \tag{A.20}$$

The expectation terms can be found out by using multinomial expansion as

$$\mathbb{E}A_l^n = \sum_{n_1=0}^{n} \sum_{n_2=0}^{n_1} \cdots \sum_{n_{l-1}=0}^{n_{l-2}} \binom{n}{n_1}\binom{n_1}{n_2} \cdots \binom{n_{l-2}}{n_{l-1}} \mathbb{E}a_1^{n-n_1} \mathbb{E}a_2^{n_1-n_2} \cdots \mathbb{E}a_l^{n_{l-1}}, \text{with}$$

$$\mathbb{E}a_k^n = b_w^{n/a_w} \Gamma\left(1 + \frac{n}{a_w}\right), \tag{A.21}$$

for any positive integer n.

The above approximation has been found to be very tight [85], and is applicable for all values of $a_w > 0$, unlike other exact expressions or approximations available in the literature. For e.g., Yilmaz and Alouini [140] have derived the exact PDF of Weibull sums, but are applicable only when $a_w > 1$.

A.6.2 Pearson Type IV Approximation in Lognormal Shadowing Case

Consider the probability term in (2.4), viz., $\mathcal{P}\left\{\sum_{k=1}^{l}\left(e^{y_k}-1\right)^2\right\} \leq 1$, where y_k has a truncated Gaussian distribution with mean $\mu_s + \log\left(\frac{2}{\ell_0}\right)$ and variance σ_s^2, truncated to $[0, \infty)$. Following [86], the PDF of the above sum $\mathcal{E}_l \triangleq \sum_{k=1}^{l}\left(e^{y_k}-1\right)^2$ can be approximated by a Pearson Type IV distribution, whose PDF is given by

$$f_{\mathcal{E}_l}(x; u, m, d, \nu) = \frac{\Gamma(m)}{\sqrt{\pi}d\Gamma(m-0.5)}{}_2\mathcal{F}_1(-i\nu/2, i\nu/2; m; 1)\left[1 + \frac{(x+u)^2}{d^2}\right]^{-m}$$

$$\times \exp\left[-\nu\tan^{-1}\left(\frac{x+u}{d}\right)\right], \tag{A.22}$$

where $i = \sqrt{-1}$ and ${}_2\mathcal{F}_1(\cdot, \cdot; \cdot; \cdot)$ is the Gauss' hypergeometric function. Estimating the parameters u, m, d and ν using method of moments gives:

$$u = \frac{a}{2b_2} - m_1; \quad m = \frac{1}{2b_2}; \quad d = \sqrt{\frac{(4b_0b_2 - a)^2}{4b_2^2}}; \quad \text{and} \quad \nu = \frac{a(1+2b_2)}{2db_2^2}, \tag{A.23}$$

with the values of a, b_0, b_1, b_2, m_1 are chosen through the moments $\{\mu_n\}_{n=2}^4$ as

$$a = b_1 = -A\mu_3(\mu_4 + 3\mu_2^2); \quad b_0 = -A\mu_2(4\mu_2\mu_4 - 3\mu_3^2);$$

$$b_2 = -A(2\mu_2\mu_4 - 3\mu_3^2 - 6\mu_2^3); \quad A \triangleq 10\mu_2\mu_4 - 18\mu_2^3 - 12\mu_3^2, \tag{A.24}$$

The moments $\{\mu_n\}_{n=2}^4$ can be calculated through Moment Generating Function formulae, given in Sec. III in [86].

Appendix B

Appendix for Chapter 3

B.1   Proof of Theorem 4

Since the test is of the form (3.4), the optimal threshold can be found by differentiating the probability of error with respect to the threshold, and equating it to zero. Following (3.3) and (3.7), a straightforward substitution and simplification of the LR test results in the expression for p_e under Rayleigh fading for the problem in (3.1) as

$$p_e = \pi_0 \mathcal{P}(V_y > x | \mathcal{H}_0) + (1 - \pi_0) \mathcal{P}(V_y \leq x | \mathcal{H}_1)$$

$$= \pi_0 Q\left(\sqrt{M}x\right) + (1 - \pi_0) \int_0^\infty Q\left(-\frac{x - \alpha P}{\frac{1}{\sqrt{M}}}\right) e^{-\alpha} d\alpha \tag{B.1}$$

In the above equation, $\alpha \triangleq |h|^2$ has the exponential distribution, and x is the detection threshold. The $x_{\text{CLT}}^{(R)}$ which minimizes (B.1) can be obtained by equating $\frac{\partial p_e}{\partial x}$ to zero, which gives

$$\frac{\partial p_e}{\partial x} = -\pi_0 \sqrt{\frac{M}{2\pi}} \exp\left(-\frac{Mx^2}{2}\right) + \frac{(1 - \pi_0)}{P} \exp\left(-\frac{Mx^2}{2}\right) \exp\left(\frac{\left(x - \frac{1}{MP}\right)^2}{2/M}\right)$$

$$\times Q\left(-\frac{\left(x - \frac{1}{MP}\right)}{1/\sqrt{M}}\right) = 0. \tag{B.2}$$

Since M is assumed to be sufficiently high, using the approximation $Q(y) \approx \exp(-\frac{y^2}{2})$, and further simplification gives the required result.

B.2 Proof of Theorem 5

For small values of P, using Result 1, the lognormal distribution can be approximated by $\mathcal{N}(P, P^2\sigma_s^2)$. Let $s \triangleq |h|^2$. Using this, the PDF under $\mathcal{H}_1$, denoted by $P(V_y|\mathcal{H}_1)$ is shown to be

$$P(V_y|\mathcal{H}_1) = \int_0^\infty Q\left(-\frac{V_y - sP}{\frac{1}{\sqrt{M}}}\right) \frac{1}{s\sqrt{2\pi\sigma_s^2}} \exp\left(-\frac{(\log s - \log P)^2}{2\sigma_s^2}\right) ds \qquad \text{(B.3)}$$

$$\approx \int_0^\infty Q\left(-\frac{V_y - sP}{\frac{1}{\sqrt{M}}}\right) \frac{1}{\sqrt{2\pi P^2\sigma_s^2}} \exp\left(-\frac{(s - P)^2}{2P^2\sigma_s^2}\right) ds \qquad \text{(B.4)}$$

$$= \frac{1}{2\pi\sqrt{\frac{P^2\sigma_s^2}{M}}} \exp\left(-\frac{V_y^2}{2/M} - \frac{1}{2\sigma_s^2}\right)$$

$$\times \int_0^\infty \exp\left(-\left[\frac{M}{2} + \frac{1}{2P^2\sigma_s^2}\right]s^2 + \left[\frac{V_y}{1/M} + \frac{1}{P\sigma_s^2}\right]\right) ds. \qquad \text{(B.5)}$$

The above integral has a closed form solution [116, Pg. 336, eq. 3.322(1)]. Following this result,

$$\int_0^\infty \exp\left(-\left[\frac{M}{2} + \frac{1}{2P^2\sigma_s^2}\right]s^2 + \left[\frac{V_y}{1/M} + \frac{1}{P\sigma_s^2}\right]s\right) ds = \sqrt{\frac{\pi}{4\left(\frac{M}{2} + \frac{1}{2P^2\sigma_s^2}\right)}}$$

$$\times \exp\left(\frac{\left[\frac{V_y}{1/M} + \frac{1}{P\sigma_s^2}\right]^2}{4\left(\frac{M}{2} + \frac{1}{2P^2\sigma_s^2}\right)}\right) \left\{1 - erf\left(\frac{\left[\frac{V_y}{1/M} + \frac{1}{P\sigma_s^2}\right]}{2\sqrt{\left(\frac{M}{2} + \frac{1}{2P^2\sigma_s^2}\right)}}\right)\right\} \qquad \text{(B.6)}$$

Using the above result and (3.3), the likelihood ratio can be written as

$$LR(V_y) = \frac{2}{\sqrt{2\pi \frac{P^2 \sigma_s^2}{M}}} \exp\left(-\frac{1}{2\sigma_s^2}\right) \sqrt{\frac{\pi}{4\left(\frac{M}{2} + \frac{1}{2P^2 \sigma_s^2}\right)}}$$

$$\times \exp\left(\frac{\left[\frac{V_y}{1/M} + \frac{1}{P\sigma_s^2}\right]^2}{4\left(\frac{M}{2} + \frac{1}{2P^2 \sigma_s^2}\right)}\right) Q\left\{-\frac{\left(\frac{V_y}{1/M} + \frac{1}{P\sigma_s^2}\right)}{\sqrt{2\left(\frac{M}{2} + \frac{1}{2P^2 \sigma_s^2}\right)}}\right\}$$

where, we have used the result $1 - erf(y) = 2Q(\sqrt{2}y)$.

The $LR(V_y)$ can be shown to be monotone in V_y, by examining the derivative to be positive $\forall\ V_y$. The corresponding mathematical expressions are lengthy and therefore omitted. Hence, the test can be written in terms of V_y, as in (3.4). Writing out the derivative of the missed detection probability p_m, and by using (B.6), it can be shown that

$$\frac{\partial p_m}{\partial x} = \frac{1}{\pi \sqrt{\frac{P^2 \sigma_s^2}{M}}} \exp\left(-\frac{x^2}{2/M} - \frac{1}{2\sigma_s^2}\right) \sqrt{\frac{\pi}{4\left(\frac{M}{2} + \frac{1}{2P^2 \sigma_s^2}\right)}}$$

$$\times \exp\left(\frac{\left[\frac{x}{1/M} + \frac{1}{P\sigma_s^2}\right]^2}{4\left(\frac{M}{2} + \frac{1}{2P^2 \sigma_s^2}\right)}\right) Q\left\{-\frac{\frac{x}{1/M} + \frac{1}{P\sigma_s^2}}{\sqrt{2\left(\frac{M}{2} + \frac{1}{2P^2 \sigma_s^2}\right)}}\right\}.$$

Upon equating $\frac{\partial p_e}{\partial x}\left\{= \pi_0 \frac{\partial p_f}{\partial x} + (1 - \pi_0)\frac{\partial p_m}{\partial x}\right\} = 0$, and further simplification gives

$$K_c \exp\left(\frac{\rho^2}{2}\right) Q(-\rho) = \frac{\pi_0}{1 - \pi_0}, \tag{B.7}$$

where K_c is defined as in (3.13) and $\rho \triangleq \frac{Mx + \frac{1}{P\sigma_s^2}}{\sqrt{\left(M + \frac{1}{P^2 \sigma_s^2}\right)}}$. For high M and low σ_s^2, we use the approximation $\exp\left(\frac{\rho^2}{2}\right) Q(-\rho) \approx \exp\left(\frac{\rho^2}{2}\right) - 1$. Rearranging the terms gives the required result.

B.3   Proof of Theorem 6

Let $\zeta \triangleq |h|^2$. Following (3.2), the likelihood ratio $LR(V_y)$ (upon expanding and completing the squares on the exponential term) can be written as

$$LR(V_y) = \frac{K^K}{P^K \Gamma(K)} \exp\left\{ \left(\frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M} \right) \right\} \int_0^\infty \zeta^{K-1} e^{\left(-\frac{M}{2}\left(\zeta - V_y + \frac{K}{MP}\right)^2\right)} \, d\zeta. \qquad \text{(B.8)}$$

The integral term $\int_0^\infty (\cdot)$ in the above equation can be written as a difference between two integrals $\int_{-\infty}^\infty (\cdot)$ and $\int_{-\infty}^0 (\cdot)$. The first integral is known to be equal to [141, Sec. III, eq. (11) and (12)]

$$\int_{-\infty}^\infty \zeta^{K-1} \exp\left(-\frac{M}{2}\left(\zeta - V_y + \frac{1}{MP} \right)^2 \right) d\zeta$$

$$= \left(\frac{1}{M} \right)^{\frac{K-1}{2}} \sqrt{\frac{2\pi}{M}} \left[i\sqrt{2}\, sign\left(V_y - \frac{K}{MP} \right) \right]^{K-1} \mathcal{U}\left(-\frac{K-1}{2}; \frac{1}{2}; -\frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M} \right) \qquad \text{(B.9)}$$

$$= \begin{cases} \left(\frac{2}{M} \right)^{\frac{K-1}{2}} \frac{\Gamma\left(\frac{K}{2}\right)}{\sqrt{\pi}} \mathcal{M}\left(-\frac{K-1}{2}; \frac{1}{2}; -\frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M} \right), & K \text{ odd.} \\[2ex] 2^{\frac{K}{2}} \left(\frac{1}{M} \right)^{\frac{K-2}{2}} \left(V_y - \frac{K}{MP} \right) \frac{\Gamma\left(\frac{K+1}{2}\right)}{\sqrt{\pi}} \mathcal{M}\left(-\frac{2-K}{2}; \frac{3}{2}; -\frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M} \right), & K \text{ even.} \end{cases} \qquad \text{(B.10)}$$

where, $i = \sqrt{-1}$, $\mathcal{U}(\cdot; \cdot; \cdot)$ and $\mathcal{M}(\cdot; \cdot; \cdot)$ are the Tricomi's, and Kummer's confluent hypergeometric functions; also called as the hypergeometric function of the second, and first kind, respectively [116]. We can use either (B.9) or (B.10) to solve the above integral. For convenience, we will use (B.9).

Now, the second integral represents the $(K-1)^{th}$ partial moment of a Gaussian random variable, and is derived in Winkler et al. [142]. Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then, let the nth-ordered partial moment be defined as

$$\mathbb{E}_{-\infty}^z(X^n) \triangleq \int_{-\infty}^z x^n \sqrt{\frac{M}{2\pi}} \exp\left(-\frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M} \right), \qquad \text{(B.11)}$$

and let X_* represent the standard Gaussian corresponding to X. In our problem, we have $X \sim \mathcal{N}\left(V_y - \frac{K}{MP}, \frac{1}{M}\right)$, and we can write [142, Sec. 2]

$$\int_{-\infty}^{0} \zeta^{K-1} \exp\left(-\frac{\left(\zeta - V_y + \frac{1}{MP}\right)^2}{2/M}\right) d\zeta = \sum_{k=0}^{K-1} \binom{K-1}{k}$$
$$\left(V_y - \frac{1}{MP}\right)^k \left(\frac{1}{M}\right)^{\frac{K-1-k}{2}} \mathbb{E}_{-\infty}^{-M\left(V_y - \frac{K}{MP}\right)}\left(X_*^{K-1-k}\right) \tag{B.12}$$

where,

$$\mathbb{E}_{-\infty}^{-M\left(V_y - \frac{K}{MP}\right)}\left(X_*^{K-1-k}\right) \triangleq \frac{A_{K-1-k}^{-M\left(V_y - \frac{K}{MP}\right)}}{\sqrt{2\pi}} \exp\left(-\frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M}\right)$$
$$+ B_{K-1-k}^{-M\left(V_y - \frac{K}{MP}\right)} Q\left(-\frac{\left(v_y - \frac{K}{MP}\right)}{\sqrt{1/M}}\right), \tag{B.13}$$

with $Q(\cdot)$ being the Gaussian-Q function, and

$$A_K^{-M\left(V_y - \frac{K}{MP}\right)} \triangleq \begin{cases} -\left(-\frac{\left(V_y - \frac{K}{MP}\right)}{\sqrt{1/M}}\right)^{K-1} - \sum_{q=1}^{\frac{K-1}{2}}\left[\prod_{r=1}^{q}(K-2r+1)\right]\left(-\frac{\left(V_y - \frac{K}{MP}\right)}{\sqrt{1/M}}\right)^{K-1-2q}, K \text{ odd.} \\ -\left(-\frac{\left(V_y - \frac{K}{MP}\right)}{\sqrt{1/M}}\right)^{K-1} - \sum_{q=1}^{\frac{K-2}{2}}\left[\prod_{r=1}^{q}(K-2r+1)\right]\left(-\frac{\left(V_y - \frac{K}{MP}\right)}{\sqrt{1/M}}\right)^{K-1-2q}, K \text{ even.} \end{cases}$$
$$\tag{B.14}$$

$$B_K^{-M\left(V_y - \frac{K}{MP}\right)} \triangleq \begin{cases} 0, & K \text{ odd.} \\ \prod_{r=1}^{\frac{K}{2}}(K-2j+1), & K \text{ even.} \end{cases} \tag{B.15}$$

Substituting the above expression in (B.8), we get

$$LR(V_y) = \frac{K^K}{P^K \Gamma(K)} \exp\left(\frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M}\right)\left\{\left(\frac{1}{M}\right)^{\frac{K-1}{2}}\sqrt{\frac{2\pi}{M}}\right.$$
$$\left[i\sqrt{2}\,sign\left(V_y - \frac{K}{MP}\right)\right]^{K-1}\mathcal{U}\left(-\frac{K-1}{2}; \frac{1}{2}; -\frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M}\right)$$
$$\left. -\sum_{k=0}^{K-1}\binom{K-1}{k}\left(V_y - \frac{K}{MP}\right)^k\left(\frac{1}{M}\right)^{\frac{K-1-k}{2}}\mathbb{E}_{-\infty}^{-M\left(V_y - \frac{K}{MP}\right)}\left(X_*^{K-1-k}\right)\right\} \tag{B.16}$$

The ED is an optimal test when $LR(V_y)$ is monotone in V_y, so that it can be written as a test on V_y. The function $\mathcal{U}(\cdot;\cdot;\cdot)$ can be represented as a $(K-1)^{\text{th}}$ order polynomial in V_y. Therefore, the above function $LR(V_y)$ need not be monotone in V_y. Hence, in general, no conclusion on the optimality of a test of the form (3.4), under Nakagami-m fading can be drawn. Note that the parameter of the Nakagami-m distribution viz., K, signifies the "number of paths" in a multipath fading environment. In an earlier result [143], the authors have shown that the ED for detection using an equal gain combining receiver when the channel undergoes Rayleigh fading, is only locally optimal. The structure of a K-fold equal gain combining receiver under Rayleigh fading is statistically equivalent to a single receiver with Nakagami-K fading. Therefore, the ED is not optimal for detection under Nakagami-m fading channel.

In the following, we show that the ED, or a test of the form (3.4) is a *Locally Most Powerful* (LMP) test [44] around $x = \frac{K}{MP}$. In the analysis of the Rayleigh fading case (which is a special case of this problem with $K = 1$), it was seen that the optimal threshold was around $\frac{1}{MP}$, as shown in (3.8). Intuitively, for the Nakagami-m case with parameter K, we expect that $x_{\text{CLT}}^{(Nm)}$ will be in and around $\frac{K}{MP}$. Writing out p_f and p_m from (3.5) and (3.6), respectively, following a similar approach used in deriving $LR(V_y)$, and equating $\frac{\partial p_e}{\partial x} = 0$, gives

$$\pi_0 \sqrt{\frac{M}{2\pi}} \exp\left(-\frac{MV_y^2}{2}\right) + (1-\pi_0)\left\{ \sqrt{\frac{M}{2\pi}} \frac{K^K}{P^K \Gamma(K)} \exp\left(-\frac{MV_y^2}{2} + \frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M}\right)\right.$$

$$\left\{\left(\frac{1}{M}\right)^{\frac{K-1}{2}} \sqrt{\frac{2\pi}{M}} \left[i\sqrt{2}\, sign\left(V_y - \frac{K}{MP}\right)\right]^{K-1} \mathcal{U}\left(-\frac{K-1}{2}; \frac{1}{2}; -\frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M}\right)\right.$$

$$\left.\left. -\sum_{k=0}^{K-1} \binom{K-1}{k}\left(V_y - \frac{K}{MP}\right)^k \left(\frac{1}{M}\right)^{\frac{K-1-k}{2}} \mathbb{E}_{-\infty}^{-M\left(V_y - \frac{K}{MP}\right)}\left(X_*^{K-1-k}\right)\right\}\right\} = 0. \qquad \text{(B.17)}$$

Using the Taylor series expansion on the Tricomi confluent hypergeometric function $\mathcal{U}(a; b; z)$ around $z = 0$ gives [116]

$$\mathcal{U}\left(\frac{K-1}{2}; \frac{1}{2}; -\frac{\left(x - \frac{K}{MP}\right)^2}{2/M}\right) \approx \frac{\Gamma(1/2)}{\Gamma\left(1 - \frac{K}{2}\right)} + \frac{\left(x - \frac{K}{MP}\right)^2}{2/M} \times \frac{(K-1)\Gamma(1/2)}{\Gamma\left(1 - \frac{K}{2}\right)}$$
$$- i\frac{\left(x - \frac{K}{MP}\right)}{\sqrt{2/M}} \times \frac{\Gamma(-1/2)}{\Gamma\left(-\frac{K-1}{2}\right)}, \tag{B.18}$$

where $i = \sqrt{-1}$. Considering the last term of the integral in (B.12), we can approximate,

$$\mathbb{E}_{-\infty}^{-M\left(x - \frac{K}{MP}\right)}\left(X_*^{K-1-k}\right) \approx \left(V_y - \frac{K}{MP}\right)^{K-1} Q\left(-\sqrt{M}\left(V_y - \frac{K}{MP}\right)\right), \forall k \tag{B.19}$$

By using (B.19) and (B.18), $LR(V_y)$ is given by,

$$LR(V_y) \approx \frac{K^K}{P^K \Gamma(K)} \exp\left(\frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M}\right) \left(\frac{1}{M}\right)^{\frac{K-1}{2}} \left[\left[i\sqrt{2}\,\text{sign}\left(V_y - \frac{K}{MP}\right)\right]^{K-1}\sqrt{\frac{2\pi}{M}}\right.$$
$$\times \left\{\frac{\Gamma(1/2)}{\Gamma\left(1 - \frac{K}{2}\right)} + \frac{(K-1)\Gamma(1/2)}{\Gamma\left(1 - \frac{K}{2}\right)} \times \frac{\left(V_y - \frac{K}{MP}\right)^2}{2/M} - i\frac{\Gamma(-1/2)}{\Gamma\left(-\frac{K-1}{2}\right)} \times \frac{\left(V_y - \frac{K}{MP}\right)}{\sqrt{2/M}}\right\}$$
$$\left. - \left[\left(V_y - \frac{K}{MP}\right)^{K-1} Q\left(-\frac{V_y - \frac{K}{MP}}{1/\sqrt{M}}\right)\right]\right] \tag{B.20}$$

To show that a test on the above $LR(V_y)$ reduces to an ED, we need to prove that the LHS in (B.20) is monotone in V_y. To this end, it can be verified that

$$\frac{\Gamma(1/2)}{\Gamma\left(1 - \frac{K}{2}\right)} = \begin{cases} \text{positive,} & K = 1, 5, 9, \cdots, \\ 0, & K \text{ even,} \\ \text{negative,} & K = 3, 7, 11, \cdots, \end{cases}$$

$$\text{and } \frac{\Gamma(-1/2)}{\Gamma\left(-\frac{K-1}{2}\right)} = \begin{cases} \text{positive,} & K = 2, 6, 10, \cdots, \\ 0, & K \text{ odd,} \\ \text{negative,} & K = 0, 4, 8, \cdots. \end{cases}$$

To show that $LR(V_y)$ is monotone in V_y, it is enough to check if $\frac{\partial LR}{\partial V_y} \gtrless 0$, for all values of V_y. We will only brief the rest of the proof. Consider the case where K is even. When $V_y > \frac{K}{MP}$, the "sign($\cdot$)" term equals 1. For $K = 2, 6, \cdots$, $i^{K-1} = i$, and $i \times -i = 1$. Hence the terms with "sign" and "i" are positive real for all values of V_y, for all K. The rest of the terms are positive real for every V_y, for all K. A similar argument holds when $K = 4, 8, \cdots$. Hence $LR(V_y)$ is monotonically increasing with V_y, for even K and $V_y > \frac{K}{MP}$. Now, when $V_y < \frac{K}{MP}$, the "sign($\cdot$)" term equals -1. The rest of the terms are positive for every V_y, following similar arguments as above, for all K. Hence $LR(V_y)$ is monotonically decreasing with V_y, for even K and $V_y < \frac{K}{MP}$.

Similar arguments hold when K is odd, where "i" in the "sign($\cdot$)" term will be raised to an even number, which results in a real power. Therefore, for each K, the RHS of (B.20) is real valued. Since $LR(V_y)$ is monotone in V_y, the test can be written as a test on V_y. The equation for the solution of $x_{\text{CLT}}^{(Nm)}$, i.e., (3.14) is obtained by differentiating p_e w.r.t. x, and equating it to zero, which is a simplified form of (B.17)

B.4 Error Exponent at the FC using the K-out-of-N Rule

Now, let us consider the case where N sensors make local decisions based on conditionally independent observations, and send their binary decisions to an FC through an error-free channel. The FC combines the decisions using the K out of N rule, which is optimum in the setup considered here. The optimum value of K is given by [144]:

$$K_{opt} = \min\left(N, \left\lceil \frac{\log\left(\frac{\pi_0}{1-\pi_0}\right) + N\log\left(\frac{1-p_f}{p_m}\right)}{\log\left\{\left(\frac{1-p_m}{p_f}\right)\left(\frac{1-p_f}{p_m}\right)\right\}} \right\rceil\right)$$

(B.21)

The probabilities of false alarm (P_F) and missed detection (P_M) at the FC can be obtained using a simple summation of binomial terms with parameters p_f and p_m at the single sensor. The following theorem gives the error exponent of the K out of N fusion rule.

Theorem 10. *The error exponent achieved by using K out of N rule at the FC when $N \to \infty$ is given by*

$$\lim_{N \to \infty} -\frac{1}{N} \log P_E \triangleq \epsilon_E = \min(\epsilon_F, \epsilon_M) = \epsilon_F = \epsilon_M, \quad where, \tag{B.22}$$

$$\epsilon_F \triangleq -(1 - k_1) \log(1 - p_f) - k_1 \log p_f - \mathbb{H}(k_1), \tag{B.23}$$

$$\epsilon_M \triangleq -(1 - k_1) \log p_m - k_1 \log(1 - p_m) - \mathbb{H}(k_1), \tag{B.24}$$

$$with \; \mathbb{H}(k_1) \triangleq -k_1 \log(k_1) - (1 - k_1) \log(1 - k_1), \tag{B.25}$$

$$and \; k_1 \triangleq \frac{\log\left(\frac{1 - p_f}{p_m}\right)}{\log\left\{\left(\frac{1 - p_m}{p_f}\right)\left(\frac{1 - p_f}{p_m}\right)\right\}}. \tag{B.26}$$

Proof. Since the channel between the individual sensors and the FC is assumed to be lossless, and the individual sensor decisions are i.i.d., the transmission of bits from all N sensors to the FC can be modeled as a N channel use on a Binary Asymmetric Channel (BAC), with parameters p_f and p_m, as shown in the Fig. B.1. For the BAC model, let p_0 and p_1 represent the probability distributions on the bits b_0 and b_1, respectively. Following this notation, $p_0(0)$, for example, represents the probability of decoding the bit $b_0(= 0)$ at the FC, when the transmitted bit was also $b_0(= 0)$ (representing a decision favoring $\mathcal{H}_0$ at the sensors).

The best exponent on the probability of error at the FC, P_E, is known to be the Chernoff
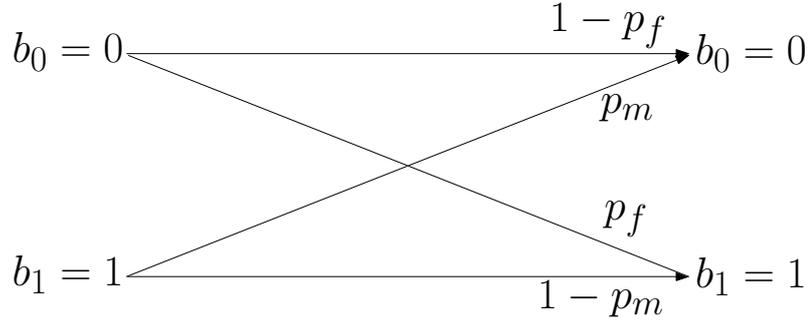
Figure B.1: The binary asymmetric channel model for communication between sensors and the fusion center.

information between p_0 and p_1 [44, Pg. 387], and is defined as

$$
\begin{aligned}
C(p_0, p_1) &\triangleq -\min_{0 \leq s \leq 1} \log \left(\sum_{x \in (0,1)} p_0^s(x) p_1^{(1-s)}(x) \right), \\
&= -\min_{0 \leq s \leq 1} \log \left[p_0^s(0) p_1^{(1-s)}(0) + p_0^s(1) p_1^{(1-s)}(1) \right], \\
&= -\min_{0 \leq s \leq 1} \log \left[(1 - p_f)^s p_m^{(1-s)} + p_f^s (1 - p_m)^{(1-s)} \right],
\end{aligned}
\tag{B.27}
$$

parametrized by s. The best error exponent is found by differentiating (B.27) w.r.t. the

parameter s and equating it to zero. Upon differentiation,

$$
(1 - p_f)^s p_m^{(1-s)} \log p_m(-1) + p_m^{(1-s)}(1 - p_f)^s \log(1 - p_f)
$$

$$
+ p_f^s (1 - p_m)^{(1-s)} \log(1 - p_m)(-1) + (1 - p_m)^{(1-s)} p_f^s \log p_f = 0.
\tag{B.28}
$$

On further simplification and rearranging the terms gives the optimum value of s as,

$$
s_{opt} = \frac{\log \left(\frac{1 - p_m}{p_m} \right) - \log \left(-\frac{\log \left(\frac{1 - p_f}{p_m} \right)}{\log \left(\frac{p_f}{1 - p_m} \right)} \right)}{\log \left(\frac{1 - p_f}{p_f} \times \frac{1 - p_m}{p_m} \right)}
\tag{B.29}
$$

Therefore, the optimal exponent on P_E is given by $C(p_0, p_1)|_{s=s_{opt}}$. Next, we show that

the optimum exponent at the FC can be achieved by using the K-out-of-N rule. This

result should not surprising, given that the K-out-of-N rule is optimal in the P_E sense [71]. For the K-out-of-N rule, the expressions for P_F and P_M are given by

$$P_F = \sum_{k=K_{opt}}^{N} \binom{N}{k} p_f^k (1-p_f)^{N-k} \tag{B.30}$$

$$P_M = \sum_{k=0}^{K_{opt}-1} \binom{N}{k} (1-p_m)^k p_m^{N-k}, \tag{B.31}$$

where the formula for K_{opt} is given in (A.17). Using a result in [145], for large N,

$$\binom{N}{k_1 N} \geq \left(\frac{1}{N+1}\right)^2 \exp\left(NH(k_1)\right) \tag{B.32}$$

where k_1 is as defined in (B.26), and

$$H(k_1) \triangleq -k_1 \log(k_1) - (1-k_1)\log(1-k_1) \tag{B.33}$$

The above lower bound given is becomes increasingly tight, as N becomes large. Using the above bound, we can write (B.30) and (B.31) as

$$P_M \geq \left(\frac{1}{N+1}\right)^2 \exp\left(NH(k_1)\right) \left(\frac{1-p_m}{p_m}\right)^{k_1 N - 1} p_m^N \tag{B.34}$$

$$P_F \geq \left(\frac{1}{N+1}\right)^2 \exp\left(NH(k_1)\right) \left(\frac{p_f}{1-p_f}\right)^{k_1 N} (1-p_f)^N \tag{B.35}$$

Simplifying further, and examining the exponential terms in P_F and P_M gives the required result. As $N \to \infty$, the bounds become tight, and does not affect the exponential term. The accuracy of the lower bound to the actual value of the exponent is highlighted in Fig. 3.10. It can be easily verified that the exponents on P_F and P_M viz., ϵ_F and ϵ_M respectively, are equal.

$\square$

B.5   Detection Under Suzuki Fading

Suzuki fading model has been widely considered as one of the best fits for fading channels [77], as it accounts for both small and large scale fading. The pdf of a Suzuki distributed random variable Z is available only in an integral form ([146], [96])

$$f_Z(z; \nu_z; \mu_z; \sigma_z^2) = \int_0^\infty \frac{z}{\sigma_z^2} \exp\left(-\frac{z^2}{2\sigma_z^2}\right) \frac{1}{\sqrt{2\pi}\sigma_z\nu_z} \exp\left(-\frac{(\log\sigma_z - \mu_z)^2}{2\nu_z^2}\right) d\sigma_z. \quad \text{(B.36)}$$

Unavailability of a closed form for $f_Z(z)$ makes it intractable to obtain a closed form expression for the optimal threshold. It has been shown that a Suzuki distribution can be approximated by numerous distributions, for example, by a K-distribution [147], or by a generalized gamma distribution [96]. It is observed that the generalized gamma distribution fits Suzuki model to a satisfactory degree of accuracy, as shown in Figs. B.2 and B.3. A generalized gamma distribution f_{GG} is characterized by three parameters viz. a_z, b_z, and c_z, as shown below [96]

$$f_{GG}(g; a_g; b_g; c_g) = \frac{c_g g^{c_g a_g - 1}}{b_g^{c_g a_g} \Gamma(a_g)} \exp\left(-\left[\frac{g}{b_g}\right]^{c_g}\right) \quad \text{(B.37)}$$

where the product $c_z a_z$ controls the lower tail of Z and accounts for fast fading, while c_z controls the upper tail and accounts for shadowing. Even with the approximated distribution, obtaining an analytical expression for the near-optimal threshold is hard, because of the need to take an expectation of the likelihood ratio over the PDF in (B.37). Therefore, relying on numerical techniques becomes essential. However, with the approximation given in (B.37), the two dimensional integral reduces to a single dimensional integral, which can be solved numerically, without difficulty.
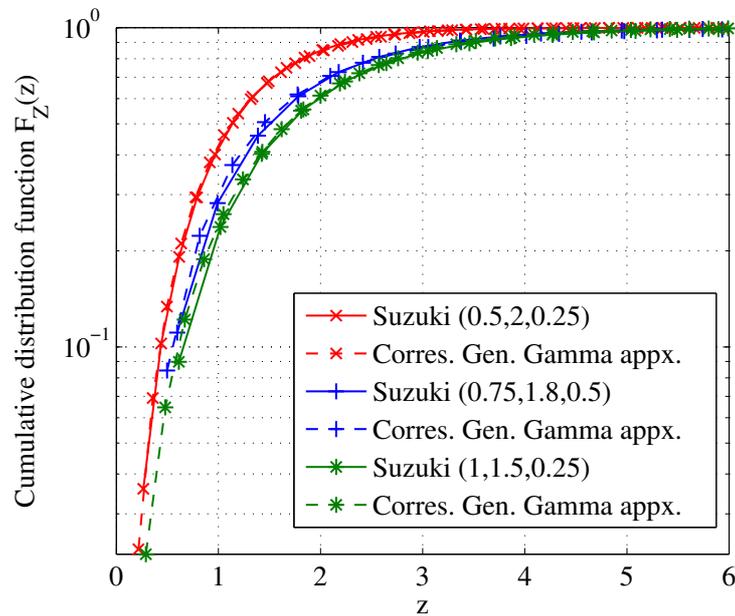
Figure B.2: CDF of the Suzuki distribution for values $(\nu_z, \mu_z, \sigma_z^2) = (0.5,2,0.25)$, $(0.75,1.8,0.5)$, $(1,1.5,0.25)$, $(1.25,1.5,0.5)$, and the corresponding Generalized Gamma approximations.
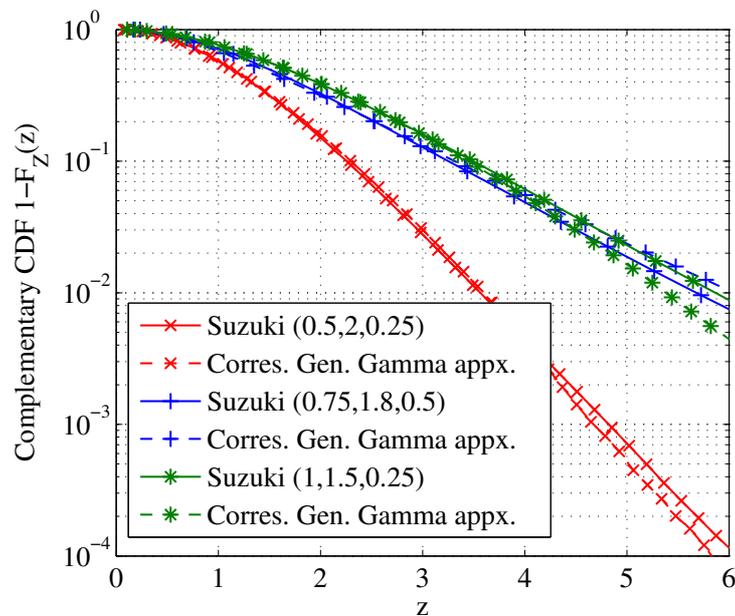
Figure B.3: CCDF of the Suzuki distribution for values $(\nu_z, \mu_z, \sigma_z^2) = (0.5,2,0.25)$, $(0.75,1.8,0.5)$, $(1,1.5,0.25)$, $(1.25,1.5,0.5)$, and the corresponding Generalized Gamma approximations.

Appendix C

Appendix for Chapter 4

C.1 Proof of Lemma 1

Under $\mathcal{H}_0$, $\bar{Y}_m = Q\bar{y}_m = Q\bar{z}_m$ is i.i.d. Gaussian with mean 0 and covariance $\sigma^2 \mathbf{I}_N$, since Q is a unitary transform. Similarly, under $\mathcal{H}_1$, $\bar{Y}_m$ is jointly Gaussian with mean $\bar{X}_m = Q\bar{x}_m$ and covariance $\sigma^2 \mathbf{I}_N$. Now, the statistic

$$T_M(k) = \frac{P(k)}{P_{tot}} = \frac{P(k)}{P(k) + \sum\limits_{y=0,\neq k}^{K-1} P(y)} = \frac{1}{1 + \frac{\sum\limits_{y=0,\neq k}^{K-1} P(y)}{P(k)}}; \text{ and therefore,}$$

$$T_M(k) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau, \quad \Rightarrow \frac{P(k)}{P_{tot} - P(k)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{\tau}{1-\tau}. \tag{C.1}$$

Following the above result, Let

$$T_k \triangleq \sum_{m=0}^{M-1} \sum_{q=0}^{\frac{N}{K}-1} \left| \frac{Y_m\left(\frac{N}{K} \times k + q\right)}{\sigma/\sqrt{2}} \right|^2, \quad S_k \triangleq \sum_{\ell=0,\ell\neq k}^{K-1} T_\ell. \tag{C.2}$$

Then, if $F_{T_M(k)}^{\mathcal{H}_0}(\tau)$ represents the CDF of $T_M(k)$ under $\mathcal{H}_0$,

$$F_{T_M(k)}^{\mathcal{H}_0}(\tau) = \Pr\left\{ \frac{T_k}{S_k} \leq \frac{\tau}{1-\tau} \mid \mathcal{H}_0 \right\}. \tag{C.3}$$

Now, let $\mathcal{K}\mathcal{X}_\Theta^2(\Psi)$ represent a chi-squared random variable (RV) with Θ Degrees of Freedom (DoF), non-centrality parameter Ψ and a scaling factor $\mathcal{K}$. It is easy to see that under $\mathcal{H}_0$, $T_k \sim \mathcal{X}_{\frac{2MN}{K}}^2(0)$, $k = 1, 2, \ldots, K - 1$. Also, $S_k \sim \mathcal{X}_{2M\left(N-\frac{N}{K}\right)}^2\left(\frac{2M}{K}\sum_{p=0}^{L-1}\text{SNR}(p)\right)$. Next, we use a result due to Patnaik ([107], [106]), which approximates a $\mathcal{X}_\Theta^2(\Psi)$ RV with a $G\mathcal{X}_\Omega^2(0)$ RV, where $\Omega \triangleq \frac{(\Theta+\psi)^2}{\Theta+2\psi}$ and $G \triangleq \frac{\Theta+\psi}{\Theta+2\psi}$. Using this, and the notations in (4.7), it follows that $S_k|\mathcal{H}_0 \sim G_0\mathcal{X}_{D_0}^2(0)$, and $T_k|\mathcal{H}_0 \sim \mathcal{X}_{\frac{2MN}{K}}^2(0)$.

Let $\gamma \triangleq \frac{\tau}{1-\tau}$. Since the statistic $T_M(k)$ is a ratio of scaled chi-squared RVs, it follows a four-parameter beta prime distribution, i.e., [148, chap. 25],

$$T_M(k) = \frac{T_k}{S_k} \sim \beta'\left(\frac{MN}{K}, \frac{D_0}{2}, 1, \frac{1}{G_0}\right), \quad \text{with PDF}$$

$$f_{T_M(k)}^{\mathcal{H}_0}(\nu) = \frac{(G_0\,\gamma)^{\frac{MN}{K}}\,(1+G_0\gamma)^{-\frac{MN}{K}-\frac{D_0}{2}}}{\frac{MN}{K}\mathcal{B}\left(\frac{MN}{K}, \frac{D_0}{2}\right)} \tag{C.4}$$

for $\gamma \in [0, \infty)$. The expression for P_{FA} follows from calculating the CDF. A similar analysis for the statistic under $\mathcal{H}_1$ gives the expression of P_D.

C.2   Proof of Lemma 2

First, note that Π is concave in $0 \le M \le \frac{N_h}{N}$. Hence, it suffices to pick the M that satisfies $\frac{\partial\Pi}{\partial M} = 0$. Define A and B, as in (4.12). Using (4.6), (4.9), and a transformation result for the Gauss' hypergeometric function [116, Sec. 9.131], Π can be rewritten as:

$$\Pi = K'\left[\frac{\Gamma(AM + BM)}{AM\,\Gamma(AM)\Gamma(BM)} \times (G_0\gamma_{\min})^{AM}\right.$$
$$\left. \times (1+G_0\gamma_{\min})^{1-AM-BM}{}_2\mathcal{F}_1(1, 1-BM; 1+AM; -G_0\gamma_{\min})\right],$$

where $K' \triangleq (K-L)(N_h-NM)$. Ancarani and Gasaneo [105] have derived the following

partial derivatives of the Gauss' hypergeometric function ${}_2\mathcal{F}_1(a,b;c;d)$ with respect to b

and c:

$$\frac{\partial_2 \mathcal{F}_1}{\partial b} = \frac{d}{b}\frac{ab}{c}{}_2\Theta^{(1)}\left(\begin{array}{c|ccc} 1, & 1 & b, & 1+b, & 1+a \\ & 1+b & 2, & 1+c \end{array}\middle|; d, \; d\right), \tag{C.5}$$

$$\frac{\partial_2 \mathcal{F}_1}{\partial c} = -\frac{d}{c}\frac{ab}{c}{}_2\Theta^{(1)}\left(\begin{array}{c|ccc} 1, & 1 & c, & 1+a, & 1+b \\ & 1+c & 2, & 1+c \end{array}\middle|; d, \; d\right). \tag{C.6}$$

Now, calculating $\frac{\partial \Pi}{\partial M}$ and equating it to zero gives,

$$\frac{\partial \Pi}{\partial M} = AM \frac{\Gamma(AM)\Gamma(BM)}{\Gamma^2(AM+BM)}\left\{(K-L)(N_h-NM)\frac{G_0\gamma}{(1+G_0\gamma)^{-1+AM+BM}}\left[\frac{B}{1+AM}(G_0\gamma)\right.\right.$$

$${}_2\Theta^{(1)}\left(\begin{array}{c|ccc} 1, & 1 & 1-BM, & 2-BM, & 2 \\ & 2-BM & 2, & 2+AM \end{array}\middle|; G_0\gamma_{\min}, \; G_0\gamma_{\min}\right) + \frac{A(1-BM)}{(1+A)^2}(G_0\gamma)$$

$$\left.{}_2\Theta^{(1)}\left(\begin{array}{c|ccc} 1, & 1 & 1+AM, & 2, & 2-BM \\ & 2+AM & 2, & 2+AM \end{array}\middle|; G_0\gamma_{\min}, \; G_0\gamma_{\min}\right)\right] + (K-L)(N_h-NM)$$

$$\frac{(G_0\gamma)^{AM}(-A-B)\log(1+G_0\gamma)}{(1+G_0\gamma)^{-1+AM+BM}}{}_2\mathcal{F}_1\left(1, 1-BM; 1+AM; -G_0\gamma_{\min}\right) + (K-L)$$

$$\frac{A(N_h-NM)(G_0\gamma)^{AM}\log(G_0\gamma)}{(1+G_0\gamma)^{-1+AM+BM}}{}_2\mathcal{F}_1\left(1, 1-BM; 1+AM; -G_0\gamma_{\min}\right) + (K-L)$$

$$\frac{-N(N_h-NM)(G_0\gamma)^{AM}}{(1+G_0\gamma)^{-1+AM+BM}}{}_2\mathcal{F}_1\left(1, 1-BM; 1+AM; -G_0\gamma_{\min}\right)\right\} - {}_{(K-L)(N_h-NM)}$$

$$\frac{(G_0\gamma)^{AM}}{(1+G_0\gamma)^{-1+AM+BM}}{}_2\mathcal{F}_1\left(1, 1-BM; 1+AM; -G_0\gamma_{\min}\right)\left\{\frac{\Gamma(AM+BM)}{\Gamma^2(AM+BM)}\right.$$

$$\left[AM\Gamma(AM)\Gamma(BM)\{B\psi^{(0)}(BM) + A\psi^{(0)}(AM)\} + \Gamma(AM)\Gamma(BM)A\right]$$

$$\left.-\frac{AM\Gamma(AM)\Gamma(BM)}{\Gamma^2(AM+BM)}\left[(A+B)\Gamma(AM+BM)\psi^{(0)}(AM+BM)\right]\right\} = 0. \tag{C.7}$$

The rest of the proof follows by simplifying further by taking out the common factors.

C.3 FAR Algorithm on Lyrtech SFF SDR DP

We now describe our implementation of the FAR algorithm on the Lyrtech SFF SDR DP (from here on, called DP for short). The block diagram of the DP is as shown in Fig. C.1, and the hardware circuitry is shown in Fig. C.2. The DP consists of the following three modules:

1. A *Digital Processing Module* (DPM), with a Xilinx Virtex-4 $SX35$ FPGA, TMS320 DM6446 system-on-chip DSP, MSP430 MCU for power management.

2. A *Data Conversion Module* (DCM), with a 14-bit, 125Msps input channel (ADC), and dual, 16-bit, 500Msps output channels (DAC). For synchronization, a 10 MHz onboard reference clock is provided, along with two external clock inputs for ADC and DAC.

3. An *RF module* (RFM) with a half-duplex (stackable for full-duplex) receiver operating at RF frequency range of 360-960MHz, selectable bandwidth of 5/20MHz, IF at 70MHz, with an RF input and output gains of up to 22dB.

Several software development tools are supported by the DP. In particular, we implemented the FAR algorithm on the DPM, using a Model Based Design Kit (MBDK) released by Lyrtech. Since the MBDK works in association with MATLAB® Simulink^{TM}, the implementation of all the modules are done in MATLAB Simulink.

The parameters chosen for the implementation are as follows. We chose $B = 5$ MHz with $K = 8$ bands, denoted (in the increasing order of their center frequencies) by C_4, C_5, C_6, C_7, C_0, C_1, C_2, C_3. The center frequencies of these bands are 395.6, 396.1, 396.65, 397.23, 393.5, 394.15, 394.8, 395.1 MHz, respectively, representing a total bandwidth

of 3.73 Mhz. We set $N = 64$, $M = 128$, 256; larger values were not feasible due to the limitations of the in-built DSP multipliers, in the DPM. We use the NI PXIe1062Q instrument (Fig. C.3) to generate sinusoids, that model the primary user signals.

The received signal is filtered to a passband of 5 MHz at with center frequency 395.4 MHz, and is down-converted to an Intermediate Frequency (IF) of 30MHz. The IF signal is sampled at a rate of 125 Msps. The sampled IF signal is digitally down-converted to baseband. The signal is then down-sampled by a factor 25 (since the Nyquist rate required is 5Msps), each for the in-phase and quadrature components, and is passed to the FFT block, which outputs the corresponding frequency domain signal in groups of N samples. These values are sent to the decision statistic block, where the power in the each band is computed by grouping $\frac{N}{K}$ bins for each band and averaging them over M frames. This calculation corresponds to the FAR statistic, i.e., calculating $T_M(k)$, $k = 1, \cdots, K$, following (4.4). Then, the detection is carried out by comparing the power in each bin with a user-defined threshold. The decisions made on each band can be made to be seen on a display (800×600) provided in the DPM (VPBE and VPFE).

C.3.1 Primary Hop-Instant Identification

It is important for the CR user to know, and synchronize its operation to the hopping instants of the frequency-hopping primary user. A method for estimating the FH boundary for bluetooth signal, with help of a Short Time Fourier Transform (STFT), has been discussed in [149]. Based on this approach, we propose a technique to identify the FH boundary by using the already implemented FAR algorithm. This method works under the assumption that all the primary users are hopping synchronously.
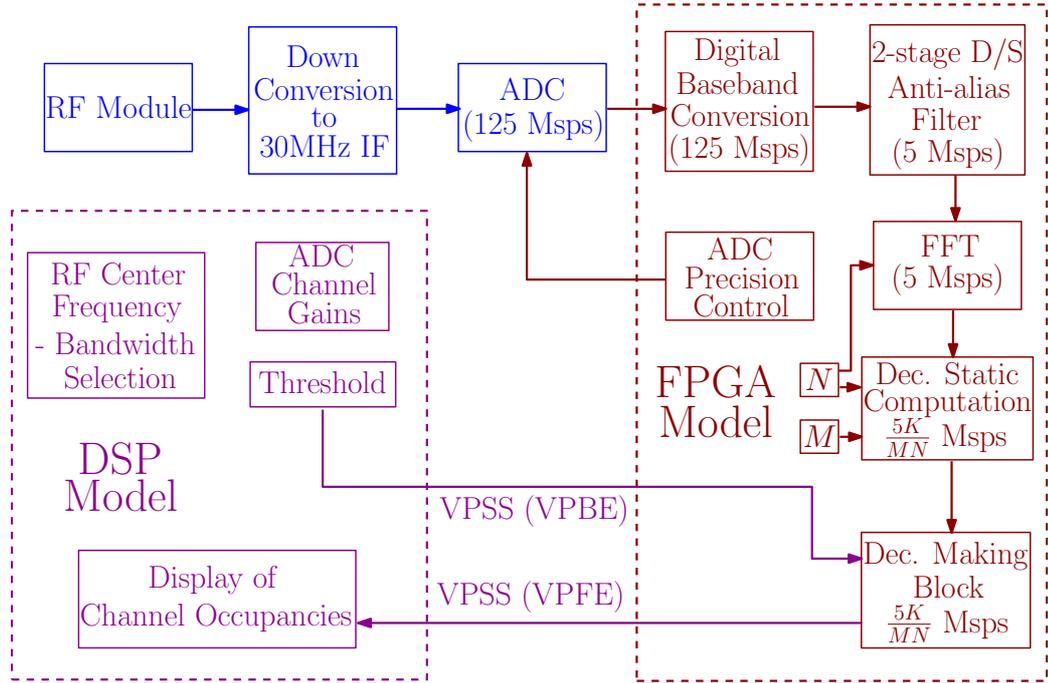
Figure C.1: Block Diagram for the Implementation on Lyrtech SFF SDR DP

The time instant at which the CR user powers up and starts sensing the spectrum will be the reference point in time (from here on, called as the "reference") for all the secondary operations. The operational time of the secondary user is also divided into successive durations of N_h samples starting from the reference, each of which is termed as *Virtual Hopping Period* (VHP) of the secondary user. The CR needs to estimate the difference in time between start of VHP and the start of the hopping period of the primary, within that particular VHP. Let this time lag be defined as N_{offset} samples. The idea is to identify the location in each VHP, where a change in the occupancies of the primary channels occurs. To this end, in each VHP, the spectrum has to be sensed repeatedly to know the occupancies. Let N_{acc} be the difference between the start of successive sensing operations by the CR (ideally, N_{acc} should be equal to 1, but due to hardware and processing limitations, one may have to use a larger N_{acc}). Since the

Figure C.2: Lyrtech SFF SDR DP circuit board.

sensing duration of the FAR algorithm is N_s, the samples in the successive sensing operations will be overlapped when $N_s > N_{\mathrm{acc}}$.

The procedure to estimate N_{offset} is as follows. Let an occupancy vector $\bar{U}$ of length K represent the presence or absence of the primary user on each channel, as declared by the CR, with an initial value set to the all zero vector $[0, \cdots, 0]$ for all the K bands. Suppose, N_{est} VHPs are used for estimating the hopping boundary. Starting from the reference, the spectrum is sensed repeatedly after each N_{acc} samples. The threshold for the FAR algorithm is set to satisfy a given, low value of P_{FA}. A vector $\bar{H}$ of length $\frac{N_h}{N_{\mathrm{acc}}}$ is defined with all elements as zeros. Let i_d, $0 \leq i_d \leq \frac{N_h}{N_{\mathrm{acc}}} - 1$ denote the index of the sensing operations performed in one VHP. The occupancy vector $(\bar{U})$ for $(i_d + 1)^{\mathrm{th}}$ sensing operation is logically XOR-ed with that of the $(i_d)^{\mathrm{th}}$ operation. If any one or more entries of the resultant vector is one, then $\bar{H}(i_d + 1)$ is incremented by one. This process is repeated for all values of i_d for N_{est} VHPs. Later, each value of $\bar{H}$ is compared
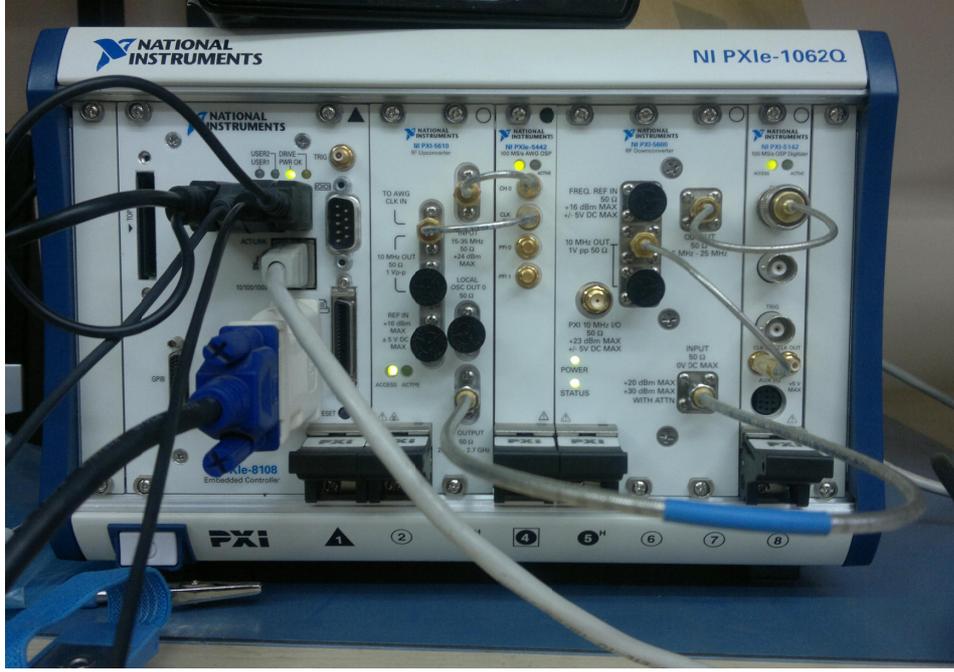
Figure C.3: NI PXIe1062Q, used for generating primary signals.

with a threshold h_{thr} and the estimated offset is given by

$$\hat{N}_{\text{offset}} = \{N_{\text{acc}} \times i_d : \ H(i_d) \geq h_{\text{thr}}\} \tag{C.8}$$

The value of h_{thr} is chosen through simulations.

As an example, consider a case where $N_{\text{offset}} = 1040$. It is assumed that $N_h = 2^{12}$, $N_{\text{est}} = 100$, and the SNR $= 6$ dB, for each active user. Since $N_{\text{acc}} = NM$, and the choice of N is constrained by the hardware, $N = 16$ is chosen as the FFT size. Therefore, the choosing N_{acc} depends on the choice of M. Note that as M increases, the detection accuracy increases, but the resolution of the boundary detection decreases. For this example, we will consider $M = 1$, and $M = 32$. A threshold value of 0.3 is chosen for FAR algorithm. The histogram $\bar{H}$ is computed, and the value of the threshold h_{thr} is set to be 85 (chosen through simulations). With these values, for the case of $M = 1$, it is

seen that the above mentioned boundary identification procedure gave the boundary change index as 65, which results in $65 \times 16 \times M = 1040 = N_{\text{offset}}$. Similar exercise with $M = 32$ gave the boundary change index as 2, which implies $2 \times 16 \times M = 1040 = N_{\text{offset}}$.

Thus, a secondary user can identify the hopping boundary of the FH primary signal where all the primary users are hopping synchronously. This method can be applied with slight modifications to the case where the primary users when they are not hopping synchronously, by suitably choosing the threshold for the histogram. More details can be found in [149].

Appendix D

Appendix for Chapter 5

D.1 Proof of Lemma 3

It is known that the LC rate for $\{Y_i, i \sim \mathcal{M}\}$ at any given level ℓ is given by [150]

$$\frac{\mathbb{E}D_{1,M}}{M-1} = 2[\mathcal{P}(Y_i > \ell) - \mathcal{P}(Y_i > \ell, Y_{i+1} > \ell)] \tag{D.1}$$

from (5.25), it is clear that

$$\mathcal{P}(Y_i > \ell) - \mathcal{P}(Y_i > \ell, Y_{i+1} > \ell) = \mathcal{P}(X_i = 1) - \mathcal{P}(X_i = 1, X_{i+1} = 1) \tag{D.2}$$

$$= \mathcal{Q}(\ell) + \mathcal{P}(X_i = 1, X_{i+1} = 1) \tag{D.3}$$

From (5.25), it can be readily shown that

$$\mathbb{E}X_i X_{i+1} = \mathcal{P}(X_i = 1, X_{i+1} = 1), \text{ and} \tag{D.4}$$

$$\mathbb{E}X_i \mathbb{E}X_{i+1} = \mathcal{Q}^2(\ell) \tag{D.5}$$

$$\text{var}(X_i) = \mathcal{Q}(\ell)(1 - \mathcal{Q}(\ell)) \tag{D.6}$$

Substituting the above into (5.26),

$$\mathcal{P}(X_i = 1, X_{i+1} = 1) = \mathcal{Q}^2(\ell) + \mathcal{Q}(\ell)\left(1 - \mathcal{Q}(\ell)\right)\rho_X \tag{D.7}$$

Equating (D.1) and (D.7), using (D.3) yields the result.

D.2   Proof of Corollary 4

When $\ell = 0$, $\mathcal{Q}(\ell) = \mathcal{Q}(0) = \frac{1}{2}$. This implies

$$\mathbb{E}D_{1,M} = \frac{M-1}{2}(1 - \rho_X), \tag{D.8}$$

Following a classical result by Rice [151], Kedem has shown that [118]

$$\frac{\mathbb{E}D_{1,M}}{M-1} = \frac{1}{\pi}\cos^{-1}\rho_X. \tag{D.9}$$

Substituting for ρ_X in (D.8), gives the required result.

D.3   Analysis on the non-applicability of ADD and ED

We will use contradiction. Consider any $\epsilon > 0$. Let $p_{f,\mathcal{G}}$, $p_{f,\mathcal{I}|\mathcal{G}}$ and $p_{f,\mathbb{N}|\mathcal{G}}$ denote the p_f values when the observations come from the distributions $f_\mathcal{G}$, $f_\mathcal{I}$ and $f_\mathbb{N}$ respectively, while the threshold is calculated and fixed based on $f_\mathcal{G}$. We need to show that $p_{f,\mathbb{N}|\mathcal{G}} > p_{f,\mathcal{G}}$, for any $\epsilon > 0$. Suppose $p_{f,\mathcal{I}|\mathcal{G}} > p_{f,\mathcal{G}}$, for any $\epsilon > 0$. Then, it is easy to see that

$$\begin{aligned} p_{f,\mathbb{N}|\mathcal{G}} &= (1 - \epsilon)p_{f,\mathcal{G}} + \epsilon p_{f,\mathcal{I}|\mathcal{G}} \\ &\geq (1 - \epsilon)p_{f,\mathcal{G}} + \epsilon p_{f,\mathcal{G}} = p_{f,\mathcal{G}}, \end{aligned} \tag{D.10}$$

with equality only if $\epsilon = 0$. Therefore, for both ED and ADD, it is enough to show that

$p_{f,\mathcal{I}|\mathcal{G}} > p_{f,\mathcal{G}}$.

First consider ED. It was seen earlier that τ_{ED} is given by (5.6). Therefore,

$$p_{f,\mathcal{G}} = 1 - \frac{\gamma\left(\frac{M-1}{2}, \frac{2\sigma_{\mathrm{G}}^2}{M}\right)}{\Gamma\left(\frac{M-1}{2}\right)}, \quad p_{f,\mathcal{I}|\mathcal{G}} = 1 - \frac{\gamma\left(\frac{M-1}{2}, \frac{2(\sigma_{\mathrm{G}}^2+\sigma_{\mathrm{I}}^2)}{M}\right)}{\Gamma\left(\frac{M-1}{2}\right)},$$

where $\gamma(\cdot, \cdot)$ and $\Gamma(\cdot)$ are the lower incomplete gamma and gamma functions, respectively. By using the well known properties of these functions, it can be easily shown that [116]

$$p_{f,\mathcal{G}} = \exp\left\{-\frac{2\sigma_{\mathrm{G}}^2}{M}\right\} \sum_{a=0}^{\frac{M-1}{2}-1} \frac{(2\sigma_{\mathrm{G}}^2)^a}{M^a a!}$$

$$p_{f,\mathcal{I}|\mathcal{G}} = \exp\left\{-\frac{2(\sigma_{\mathrm{G}}^2 + \sigma_{\mathrm{I}}^2)}{M}\right\} \sum_{a=0}^{\frac{M-1}{2}-1} \frac{(2(\sigma_{\mathrm{G}}^2 + \sigma_{\mathrm{I}}^2))^a}{M^a a!}.$$

By direct comparison, for same τ_{ED}, $p_{f,\mathcal{I}|\mathcal{G}} \geq p_{f,\mathcal{G}}$, with equality only if $\sigma_{\mathrm{I}}^2 = 0$.

Next, consider the ADD. Let A_M^2 represent the Anderson-Darling statistic and let $A_{M,\mathcal{G}}^2$ and $A_{M,\mathcal{I}|\mathcal{G}}^2$ represent the A_M^2 values when the observations come from $f_{\mathcal{G}}$ and $f_{\mathcal{I}}$ respectively, while the Z_i are calculated using $F_{\mathbb{G}}$. Using (5.9), it suffices to show that the tail probabilities $\mathcal{P}\left\{A_{M,\mathcal{I}|\mathcal{G}}^2 > \tau_{ADD}\right\} > \mathcal{P}\left\{A_{M,\mathcal{G}}^2 > \tau_{ADD}\right\}$, for large values of τ_{ADD} where $p_{f,\mathcal{I}|\mathcal{G}}$ and $p_{f,\mathcal{G}}$ are calculated.

The tail probabilities of $A_{M,\mathcal{G}}^2$ and $A_{M,\mathcal{I}|\mathcal{G}}^2$, for large τ_{ADD}, can be calculated by using Hoeffding's classical result [152]:

$$\mathcal{P}\left\{A_{M,\mathcal{G}}^2 > \tau_{ADD}\right\} = \prod_{v \geq 2} \left(1 - \frac{\lambda_{1,0}}{\lambda_{v,0}}\right)^{-1/2}$$

$$\mathcal{P}\left\{A_{M,\mathcal{I}|\mathcal{G}}^2 > \tau_{ADD}\right\} = \prod_{v \geq 2} \left(1 - \frac{\lambda_{1,1}}{\lambda_{v,1}}\right)^{-1/2}, \tag{D.11}$$

where $1/\lambda_{1,0} = \max_{v \geq 1}(1/\lambda_{v,0})$ and $1/\lambda_{1,1} = \max_{v \geq 1}(1/\lambda_{v,1})$.

From (D.11), it can be immediately seen that

$$\mathcal{P}\{A^2_{M,\mathcal{I}|\mathcal{G}} > \tau_{ADD}\} \geq \mathcal{P}\{A^2_{M,\mathcal{G}} > \tau_{ADD}\}, \text{ if } \frac{1}{\lambda_{v,0}} \geq \frac{1}{\lambda_{v,1}}, \forall v \geq 1$$

We will show that $\lambda_{v,0} \leq \lambda_{v,1} \forall v \geq 1$, by using the technique described by Stephens [153] to derive the A^2_M for the case of unknown parameters, the eigenvalues $\lambda_{v,0}$ and $\lambda_{v,1}$ can be obtained as a solution to the corresponding Fredholm determinants of $K_0(\cdot, \cdot)$ and $K_1(\cdot, \cdot)$. The eigenfunctions are chosen to be Ferrer associated Legendre functions, using which we can write [114]

$$\lambda_{v,0} = 2v(2v - 1), \forall v \geq 1. \tag{D.12}$$

However, the values of $\lambda_{v,1}$ have to be obtained numerically, and the first few values of $\lambda_{v,1}$ are listed by Stephens [153, Table 2, Column 8]. A direct comparison shows that $\lambda_{v,0}$ and $\lambda_{v,1}$ for any desired level of accuracy (for any finite, large v) show that $\lambda_{v,0} \leq \lambda_{v,1}$.

D.4   On the Wider-Applicability of the Blind Detector

As detailed in Sec. 5.3.3, the statistic used in BD is constructed using the sequence $\{B_l, l \in \{1, \cdots, n\}\}$, which is given by

$$B_l = \frac{\sum_{u=0}^{m-1} \frac{Y_{ml-u}}{m}}{\sqrt{\sum_{u=0}^{m-1} \frac{(Y_{ml-u} - X_l)^2}{m-1}}/m}, \quad l = 1, \cdots, n. \tag{D.13}$$

Following a classic result in statistics [51], when $Y_i \sim \mathcal{N}(\mu, \sigma^2), i \in \mathcal{M}$,

(i) The random variables, viz., the sample mean and the sample variance are uncorrelated. In fact, they are independent.

(ii) The ratio of sample mean to sample variance, for a small number of samples m, is student-t distributed with parameter $m - 1$.

Shen et al. [53] used the above results to propose the BD, by constructing an Anderson-Darling detector on the sequence $\{B_l\}$, assuming Gaussian noise, i.e., samples $\{Y_i\}$ to be Gaussian distributed. In this section, we discuss the cases when BD can be applied even when $\{Y_i\}$ are not Gaussian.

Kendall and Stuart [51, Chap. 30, Sec. 3] have shown that for $\{Y_i\}$ with an arbitrary PDF, the asymptotic correlation (as $n \rightarrow \infty$) between the sample mean and sample variance (denoted by $\rho_{m,v}$) is given by

$$\rho_{m,v} = \frac{\kappa_3}{\sqrt{\kappa_2(\kappa_4 + 2\kappa_2^2)}}, \qquad (D.14)$$

where κ_p is the pth-order cumulant of the PDF of $\{Y_i\}$. Therefore, as long as the parent distribution is symmetric and n is large enough, $\kappa_3 = 0$, and hence $\rho_{m,v} \rightarrow 0$. Therefore, provided that the PDF of $\{Y_i\}$ is symmetric, the samples $\{B_l\}$ would still follow the student-t distribution for large enough n. Additionally, several studies have shown that this convergence is rapid [51], and is even valid for heavy tailed distributions such as the $\mathcal{S}\alpha\mathcal{S}$ distribution.

The t test (i.e., the test on $\{B_l\}$) was originally introduced by Student as a GoFT test against mean change in the parent distribution [154]. By the property of the $\mathcal{S}\alpha\mathcal{S}$ distribution, the pth-ordered moment exists only when $\alpha \geq p$. When $\alpha < 1$, no moment exists and hence a test on the mean fails. Therefore, when $\alpha < 1$, $\{B_l\}$ no longer follows

a student-t distribution, and BD fails to satisfy the false alarm constraint.

All the above presented arguments are in agreement with our simulation results, where it was seen that BD performs well even in the presence of non-Gaussian components such as class A and B, except for the case when $\alpha < 1$.

Appendix E

Popular Goodness-of-Fit Tests for the Gaussian Distribution

In this section, we give mathematical details of some of the popular GoFTs for testing against samples which are Gaussian distributed with mean 0 and variance σ^2. Starting from the celebrated Pearson's χ^2 test [155], the history of research on designing GoFTs for testing against various distributions spans over a century, and remains an active area even today [156]. We choose some of the well-known GoFTs from the literature and classify them based on an empirical distribution function calculation, correlation between the samples, higher moments such as skewness and kurtosis, and L-moments. Unless mentioned otherwise, the main idea in all the tests is to calculate a statistic based on the observations and derive its distribution under $\mathcal{H}_0$ when the number of observations grows large. This distribution is used to determine the detection threshold that satisfies a given probability of false-alarm constraint.

In the context of the CR, the tests described below are useful in the scenario when no knowledge on the primary-only signal, and channel statistics are assumed, in addition to either of the following scenarios:

(i) When only the *background noise* is present ([53], [52]).

(ii) When the class A [55] component is present along with the background noise, where the PDF of the combined process can be well approximated by another Gaussian process ([109], [111]).

Let the observations be denoted by $Y_i \in \mathbb{R}$, $i \in \mathcal{M} \triangleq \{1, \cdots, M\}$. Also, let $Y_{(i)}, i \in \mathcal{M}$ represent the ordered samples from Y_i such that $Y_{(1)} < \cdots < Y_{(M)}$. Let the observation vector and the vector of the ordered observations be represented as $\underline{Y} = [Y_1, \cdots, Y_M]$, and $\underline{\widehat{Y}} = [Y_{(1)}, \cdots, Y_{(M)}]$. Given Y_i, $i \in \mathcal{M}$, every GoFT listed in this Appendix answers whether the hypothesis

$$\mathcal{H}_0 : Y_i \sim \mathcal{N}(0, \sigma^2) \tag{E.1}$$

is true or not, and the detection threshold is chosen such that a constraint on the probability of false-alarm, defined as

$$\mathcal{P}\{\text{declare "not } \mathcal{H}_0\text{"}|\mathcal{H}_0 \text{ true}\} = \alpha_f, \tag{E.2}$$

is satisfied, for a given $\alpha_f \in (0, 1)$.

E.1 Regression and Correlation Based Tests

The basic idea in regression based tests is as follows. Consider the regression between the ordered observations and the expected ordered statistics of the standardized version of the underlying hypothesized distribution. On a probability plot, this regression

tends to be linear if the hypothesis is true. The tests considered in this section are designed to evaluate the accuracy of the linear fit. The term "correlation" refers to the strength of the relationship between the sample statistics and their expected value.

E.1.1   The W Test (*Shapiro-Wilk* Test)

The Shapiro-Wilk test [157], or the W test, is one of the popular and widely used GoFT for testing against Gaussianity. The idea here is to obtain the best linear unbiased estimates of the mean and variance, for the linear regression. The test is devised as follows. For the ordered observations $Y_{(i)}, i \in \mathcal{M}$, let

$$m_i \triangleq \mathbb{E}Y_{(i)}, \quad v_{ij} \triangleq \mathbb{E}(Y_{(i)}Y_{(j)}), \quad \mathbf{m} \triangleq [m_1, \cdots, m_M], \quad \mathbf{V} \triangleq (v_{ij})$$

$$\widehat{Y} \triangleq \frac{1}{M}\sum_{i=1}^{M} Y_i, \quad \widehat{\sigma} \triangleq \frac{\mathbf{m}^T\mathbf{V}^{-1}\widehat{\underline{Y}}}{\mathbf{m}^T\mathbf{V}^{-1}\mathbf{m}}, \quad S^2 \triangleq \sum_{i=1}^{M}(Y_{(i)} - \widehat{Y})^2, \quad b \triangleq \frac{\mathbf{m}^T\mathbf{V}^{-1}\mathbf{m}\widehat{\sigma}}{(\mathbf{m}^T\mathbf{V}^{-1}\mathbf{V}^{-1}\mathbf{m}^T)^{1/2}} \quad \text{(E.3)}$$

where $(\cdot)^T$ represents the transpose of a matrix. The variable b represents the best linear unbiased estimate of the slope of the linear regression of the ordered observations $Y_{(i)}$ on the expected values m_i. The statistic W is defined as

$$W \triangleq \frac{b^2}{S^2} = \frac{\left(\sum_{i=1}^{M} a_i Y_{(i)}\right)^2}{\sum_{i=1}^{M}\left(Y_{(i)} - \widehat{Y}\right)^2}, \quad \text{with } \mathbf{a}^T = \frac{\mathbf{m}^T\mathbf{V}^{-1}}{(\mathbf{m}^T\mathbf{V}^{-1}\mathbf{V}^{-1}\mathbf{m}^T)^{1/2}}. \quad \text{(E.4)}$$

It is important to note that [157]

(1) W is scale and origin invariant.

(2) The distribution of W depends only on M.

(3) W is statistically independent of S^2 and $\widehat{Y}$, when Y_i are Gaussian.

(4) When $M = 2K$, $b = \sum_{i=1}^{M} a_{M-i+1}(Y_{M-i+1} - Y_i)$, and when $M = 2K + 1$, $b = \sum_{i=1}^{M} a_{M-i+1}(Y_{M-i+1} - Y_i)$, with $a_{K+1} = 0$. Observe the b is same for both even and odd M.

(5) If a_i represents the ith element of vector $\mathbf{a}$, then $-a_i = a_{M-i+1}$.

The distribution of the W statistic has to be obtained through tables [157]. Some other approximations to the W statistic are also discussed in [157]. The W statistic is found to be one of the best tests in literature for testing against Gaussianity. The disadvantages of this test are the need for a lookup table, and the fact that the tables are available only upto several tens of number of observations ($M \sim 50$).

E.1.2   The Y Test (*D'Agostino* Test)

The Y test ([158], [159]) focuses on the correlation between $Y_{(i)}, i \in \mathcal{M}$ and their expected value. Let

$$Y \triangleq \frac{\sum_{i=1}^{M} \left[i - \frac{1}{2}(M+1)\right] Y_{(i)}}{M^2 \sqrt{\frac{1}{M} \sum_{i=1}^{M} \left(Y_{(i)} - \widehat{Y}\right)}}, \tag{E.5}$$

where $\widehat{Y}$ is the sample mean of Y_i, $i \in \mathcal{M}$. Now, if $Y_i, i \in \mathcal{M}$ are from a Gaussian distribution, then [158]

$$\mathbb{E}Y = \frac{(M-1)\Gamma\left(\frac{M-1}{2}\right)}{2\sqrt{2\pi M}\Gamma\left(\frac{M}{2}\right)} \approx \frac{1}{2\sqrt{\pi}} = 0.28209479 \tag{E.6}$$

$$asd(Y) \triangleq \sqrt{\mathbb{E}Y^2 - \mathbb{E}^2 Y} = \left(\frac{12\sqrt{3} - 37 + 2\pi}{24\pi M}\right)^{\frac{1}{2}} \approx \frac{0.02998598}{\sqrt{M}} \tag{E.7}$$

where $asd(Y)$ is the asymptotic standard deviation of the statistic Y.

Therefore, for large enough M,

$$T(Y) \triangleq \frac{Y - \mathbb{E}Y}{asd(Y)} \sim \mathcal{N}(0, 1). \tag{E.8}$$

Hence, the Y-test on the observations Y_i has the critical region:

$$\left\{ Y_{(i)} : T(Y) > \tau_Y \right\}, \tag{E.9}$$

where the threshold τ_Y is chosen to meet the criterion $p_f = \alpha_f$. It is straightforward that for a given α_f, τ_Y is such that

$$\tau_Y = \mathcal{Q}^{-1}(\alpha_f), \tag{E.10}$$

where $\mathcal{Q}^{-1}(\cdot)$ is the inverse of the Gaussian Q-function.

E.1.3   The Z Test

The Z test [160] follows from the W test, and also exploits the slope of the line between the sample order statistics and their expected value under Gaussianity. As an approximation on the statistic W, the statistic Z is defined as

$$Z \triangleq \frac{\left(\sum\limits_{i=1}^{M} \tilde{c}_i Y_{(i)} \right)^2}{\sum\limits_{i=1}^{M} \left(Y_{(i)} - \widehat{Y} \right)^2}, \tag{E.11}$$

where $\tilde{c}_i \triangleq (\tilde{\mathbf{m}}^T \tilde{\mathbf{m}})^{1/2} \tilde{m}_i$, with $\tilde{m}_i = \phi \left[\frac{i - \frac{3}{8}}{M + \frac{1}{4}} \right]$, and $\phi(\cdot)$ is the CDF of a standard normal distribution. A three parameter lognormal distribution was found to fit the statistic Z, for various values of M, details of which are lengthy, and is provided in [160]. The advantages of this test as compared to W test includes less computational requirements, and applicability to larger values of M (up to several thousands).

E.1.4 The QH* Test

Consider the statistic [161]

$$QH \triangleq \frac{1}{(M-1)S_1} \sum_{j=1}^{M-1} \frac{Y_{(i+1)} - Y_{(i)}}{H_{i+1} - H_i}, \tag{E.12}$$

where

$$H_i \triangleq \phi^{-1} \left[\frac{\left(i - \frac{3}{8}\right)}{\left(M + \frac{1}{4}\right)} \right], \; S_1 \triangleq \left[\frac{1}{M-1} \sum_{i=1}^{M} \left(Y_{(i)} - \widehat{Y} \right)^2 \right]^{\frac{1}{2}}, \tag{E.13}$$

and $\phi^{-1}(\cdot)$ is the inverse of CDF of the standard normal distribution. It is seen that when $\mathcal{H}_0$ is true, the distribution of QH will have a mean close to unity. Now, consider another statistic

$$\text{QH*} \triangleq \sqrt{M}(1 - QH), \tag{E.14}$$

which can be used as the test statistic. The distribution of QH* and the detection thresholds are chosen using a lookup table [161, Tab. 2]. The main advantage of this test is simplicity in the construction of the test statistic, and the disadvantages include applicability to moderately large values of M, and the need of a lookup table.

E.1.5 The Q Test

The Q test [162] uses similar ideas that are used in the QH* test, and is applicable for moderate and large values of M. Let

$$q_1 = \sum_{i=1}^{M} a_i Y_{(i)}, \; q_2 = \sum_{i=1}^{M} b_i Y_{(i)}, \tag{E.15}$$

where

$$a_i = \frac{1}{(M-1)(u_i - u_1)}, i = 2, \cdots, M, \ a_1 = -\sum_{i=2}^{M} a_i \tag{E.16}$$

$$b_1 = -b_M = \frac{1}{(M-4)(u_1 - u_5)}, \quad b_2 = -b_{M-1} = \frac{1}{(M-4)(u_2 - u_6)} \tag{E.17}$$

$$b_3 = -b_{M-2} = \frac{1}{(M-4)(u_3 - u_7)}, \quad b_4 = -b_{M-3} = \frac{1}{(M-4)(u_4 - u_8)} \tag{E.18}$$

$$b_i = -b_{M-i+1} = \frac{1}{M-4} \left(\frac{1}{u_i - u_{i+4}} - \frac{1}{u_{i-1} - u_i} \right), \quad i = 5, \cdots, M-4, \tag{E.19}$$

$$\text{and } u_i \triangleq \phi^{-1} \left[\frac{\left(i - \frac{3}{8}\right)}{\left(M + \frac{1}{4}\right)} \right] \tag{E.20}$$

The statistic for the Q-test is given by

$$Q \triangleq \log \left(\frac{q_1}{q_2} \right) \tag{E.21}$$

It is also known that

$$\mathbb{E}Q = -0.00176 - \frac{1.06}{M} + \frac{6.03}{M^2} \tag{E.22}$$

$$\sigma_Q^2 = 0.011 + \frac{2.9}{M} - \frac{73.8}{M^2} + \frac{1150.1}{M^3} - \frac{6022.5}{M^4} \tag{E.23}$$

The Q-test on the observed samples have a critical region:

$$\left\{ Y_{(i)} : T(Q) \triangleq \frac{Q - \mathbb{E}Q}{\sigma_Q} > \tau_Q \right\}, \tag{E.24}$$

where τ_Q is chosen such that $p_f = \alpha_f$. The statistic $T(Q)$ follows $\mathcal{N}(0, 1)$ closely. The test is suitable for sample sizes between 10-2000 [162].

E.2 Empirical Distribution Function (EDF) Based Tests

The EDF based tests, as the name suggests, are based on comparison of the Cumulative Distribution Function (CDF) under the true hypothesis with the EDF estimated from the samples. These tests are applicable mainly for testing for i.i.d. random variables with a continuous, known distribution. In this case, it is a Gaussian distribution. When the parameters of the distribution are unknown and have to be estimated from the samples, these tests can be used with minor modifications as explained in detail by Stephens [117].

E.2.1 The D Test (*Kolmogorov-Smirnov* Test)

Let $\mathcal{F}(\cdot)$ represent the CDF of a normal distribution, which is $\mathcal{F}(x) = \phi(x)$. Given $Y_{(i)}$, $i \in \mathcal{M}$, let $Z_{(i)} \triangleq \mathcal{F}(Y_{(i)})$. The statistic for the D test is constructed as follows [117] [163, and ref. within]:

$$
\begin{aligned}
D^+ &\triangleq \max_{1 \leq i \leq M} \left[\left(\frac{i}{M} \right) - Z_{(i)} \right] \\
D^- &\triangleq \max_{1 \leq i \leq M} \left[Z_{(i)} - \left(\frac{i-1}{M} \right) \right] \\
D &\triangleq \max(D^+, D^-).
\end{aligned}
\tag{E.25}
$$

The GoFT based on the statistic D against Gaussianity has the following form [117]

$$
\left(\sqrt{M} + 0.12 + \frac{0.11}{\sqrt{M}} \right) D \underset{\sim \mathcal{H}_0}{\overset{\sim \mathcal{H}_0}{\gtrless}} \tau_D,
\tag{E.26}
$$

where the threshold τ_D is chosen such that $p_f = \alpha_f$. The values of τ_D for some values of α_f is due to Stephens [117].

E.2.2 The W^2 Test (*Cramér-Von Mises* Test)

Let $\mathcal{F}_M(\cdot)$ represent the EDF estimated from M samples, which is defined as

$$\mathcal{F}_M(t) = \frac{1}{M}\sum_{i=1}^{M}\mathbb{I}_{\{Y_i \leq t\}}, \tag{E.27}$$

where $t \in \mathbb{R}$, and $\mathbb{I}_{\{\cdot\}}$ is the indicator function. Then the Cramér-Von Mises statistic is defined as [117]

$$W^2 \triangleq M \int_{-\infty}^{\infty} [\mathcal{F}_M(x) - \mathcal{F}(x)]^2 \, \psi(\mathcal{F}(x))d\mathcal{F}, \tag{E.28}$$

where $\psi(\cdot) \geq 0$, is a non-negative weight function. With the transformation $u = \mathcal{F}(x)$, the statistic reduces to

$$W^2 = M \int_0^1 [\mathcal{G}_M(u) - u]^2 \, \psi(u)du, \tag{E.29}$$

where $\mathcal{G}_M(\cdot)$ represents the EDF calculated from $\{u_1, \cdots, u_M\}$. It is observed that in terms of the given observations $Z_{(i)} = \mathcal{F}(Y_{(i)})$ the statistic reduces to

$$W^2 = \sum_{i=1}^{M} \left[Z_{(i)} - \frac{2i-1}{2M} \right]^2 + \left(\frac{1}{12M} \right), \tag{E.30}$$

and the test is devised as below [117].

$$\left(W^2 + \frac{0.4}{M} + \frac{0.6}{M^2} \right) \times \left(1 + \frac{1}{M} \right) \underset{\sim \mathcal{H}_0}{\overset{\sim \mathcal{H}_0}{\gtrless}} \tau_W. \tag{E.31}$$

Even in this case, the threshold are to be chosen from the available lookup tables [117], such that $p_f = \alpha_f$.

E.2.3   The A^2 Test (*Anderson-Darling* Test)

The A^2 test, or the Anderson-Darling statistic based test [114] is a powerful test for testing against an alternative hypothesis $\mathcal{H}_1$, which is another Gaussian with a positive mean. The statistic A^2 used for this test is a special case of the W^2, with the function $\psi(t) = \frac{1}{\mathcal{F}(t)(1-\mathcal{F}(t))}$. See Chap. 5, Sec. 5.3 for details on this test.

E.3   Omnibus Tests

Omnibus tests are used for testing against Gaussianity based on the moments such as skewness, or kurtosis, calcuated from the observations. The tests mentioned below are noted to perform better compared to some of the other tests discussed in the previous sections.

E.3.1   The K^2 Test

The K^2-test [164] is an extension Pearson's χ^2 test [155], and detects deviations from Gaussianity by considering the test statistic K^2 to be a linear combination of skewness and kurtosis. Let $m_K \triangleq \frac{1}{M} \sum_{i=1}^{M}(Y_{(i)} - \widehat{Y})^K$, where $\widehat{Y}$ is the sample mean of $Y_i, i \in \mathcal{M}$. The test is devised as follows

  i) The skewness-based statistic $Z(b_1)$

    (a) Calculate $b_1 \triangleq \frac{m_3}{m_2^{3/2}}$.

    (b) Calculate $y = b_1 \left\{ \dfrac{(M+1)(M+3)}{6(M-2)} \right\}^{1/2}$,
$$\beta(b_1) = \frac{3(M^2 + 27M - 70)(M+1)(M+3)}{(M-2)(M+5)(M+7)(M+9)},$$
$w^2 \triangleq -1 + [2(\beta(b_1) - 1)]^{(1/2)}, \delta = \dfrac{1}{\sqrt{\log w}}$, and $\alpha = \left[\frac{2}{(w^2-1)} \right]^{1/2}$.

(c) Compute $Z(b_1) \triangleq \delta \log \left\{ \frac{y}{\alpha} + \left[1 + \left(\frac{y}{\alpha} \right)^{1/2} \right] \right\}$

ii) The kurtosis-based statistic $Z(b_2)$

    (a) Calculate $b_2 \triangleq \frac{m_4}{m_2^2}$.

    (b) Calculate $\mathbb{E}b_2 = \dfrac{3(M-1)}{(M+1)}$ and $var(b_2) = \dfrac{24M(M-2)(M-3)}{(M+1)^2(M+3)(M+5)}$

    (c) Compute $b \triangleq \dfrac{b_2 - \mathbb{E}b_2}{\sqrt{var(b_2)}}$.

    (d) Compute $\beta_1 = \dfrac{6(M^2-5M+2)}{(M+7)(M+9)} \sqrt{\dfrac{6(M+3)(M+5)}{M(M-2)(M-3)}}$.

    (e) Calculate $A = 6 + \dfrac{8}{\beta_1} \left[\dfrac{2}{\beta_1} + \sqrt{1 + \dfrac{4}{\beta_1^2}} \right]$.

    (f) $Z(b_2) \triangleq \sqrt{\dfrac{9A}{2}} \left[\left(1 - \dfrac{2}{9A} \right) - \left(\dfrac{1-2/A}{1+b\sqrt{\frac{2}{A-4}}} \right)^{1/3} \right]$

Finally, the K² statistic is computed as

$$K^2 \triangleq Z^2(b_1) + Z^2(b_2) \tag{E.32}$$

It is known that $Z(b_1)$ and $Z(b_2)$ closely follow the normal distribution and therefore, the K² statistic closely follows a χ^2 distribution with 2 degrees of freedom. Therefore, the K² test is of the form

$$K^2 \underset{\sim \mathcal{H}_0}{\overset{\sim \mathcal{H}_0}{\gtrless}} \tau_K, \tag{E.33}$$

with τ_K chosen such that $1 - \gamma \left(\frac{\tau_K}{2}, 1 \right) = \alpha_f$. Here, $\gamma(\cdot, \cdot)$ represents the lower incomplete Gamma function.

E.3.2 The G_w^2 Test

The G_w^2 statistic is constructed from the statistic $Z(b_1)$ defined in Sec. E.3.1, as follows [165]. Let

$$Z_w = \frac{(M+2)^{1/2}(\widehat{w}-3)}{3.54},$$ (E.34)

where $\widehat{w} = 13.29(\log \sigma - \log \tau)$, with

$$\sigma \triangleq \sqrt{\frac{1}{M}\sum_{i=1}^{M}(Y_{(i)} - \widehat{Y})^2}, \quad \tau \triangleq \frac{1}{M}\sum_{i=1}^{M}|Y_{(i)} - \widehat{Y}|,$$ (E.35)

where $\widehat{Y}$ is the sample mean of Y_i, $i \in \mathcal{M}$. Then,

$$G_w^2 \triangleq Z^2(b_1) + Z_w^2.$$ (E.36)

Under $\mathcal{H}_0$, G_w^2 closely follows a χ^2 distribution with 2 degrees of freedom. The test, and the corresponding threshold are calculated similar to that explained in Sec. E.3.1.

E.3.3 The G_w^{2*} Test

The G_w^{2*} test is a simple, easy to calculate approximation to the G_w^2 statistic, and is given by

$$G_w^{2*} \triangleq \left(\frac{M}{(M-2)\sqrt{\frac{6}{M+1}}}b_1\right)^2 + Z_w^2$$ (E.37)

The rest of the analysis remain similar to the G_w^2-test.

E.3.4 The L-Moment Skewness-Kurtosis (LSK) Based Test

The test discussed in this section [156] uses a combination of L-skewness and L-kurtosis to construct the test statistic. The L-moments, introduced by Hosking [166], have several advantages as compared to the regular moments. Let $Y_{(i):j}$ represent the ith ordered sample in a sample size of j. The first few L-moments λ_l, $l = 1, \cdots, 4$, the L-skewness κ_3, and the L-kurtosis κ_4 are defined as

$$\lambda_1 \triangleq \mathbb{E}Y_i, \quad \lambda_2 \triangleq \frac{1}{2}\mathbb{E}\left(Y_{(2):2} - Y_{(1):2}\right), \quad \lambda_3 \triangleq \frac{1}{3}\mathbb{E}\left(Y_{(3):3} - 2Y_{(2):3} + Y_{(1):3}\right)$$

$$\lambda_4 \triangleq \frac{1}{4}\mathbb{E}\left(Y_{(4):4} - 3Y_{(3):4} + 3Y_{(3):4} - Y_{(1):4}\right), \quad \kappa_3 \triangleq \frac{\lambda_3}{\lambda_2}, \quad \kappa_4 \triangleq \frac{\lambda_4}{\lambda_2}. \tag{E.38}$$

Based on the samples $Y_{(i)}$, the estimates for the L-moments are given as

$$\widehat{\lambda_1} = \frac{1}{M}\sum_{i=1}^{M}Y_{(i):M}, \quad \widehat{\lambda_2} = \frac{1}{2}\binom{M}{2}^{-1}\sum\sum_{i>j}\left(Y_{(i):M} - Y_{(j):M}\right),$$

$$\widehat{\lambda_3} = \frac{1}{3}\binom{M}{3}^{-1}\sum\sum\sum_{i>j>k}\left(Y_{(i):M} - 2Y_{(j):M} + Y_{(k):M}\right),$$

$$\widehat{\lambda_4} = \frac{1}{4}\binom{M}{4}^{-1}\sum\sum\sum\sum_{i>j>k>m}\left(Y_{(i):M} - 3Y_{(j):M} + 3Y_{(k):M} - Y_{(m):M}\right) \tag{E.39}$$

For a Gaussian distribution, it is known that $\kappa_3 = 0$, and $\kappa_4 = 0.1226$. Let $\widehat{\kappa_3}$ and $\widehat{\kappa_4}$ be the corresponding estimates from $\widehat{\lambda_3}$, $\widehat{\lambda_4}$ and $\widehat{\lambda_2}$, respectively. Now each of the following statistics [156]

$$Z_3 \triangleq \left(\frac{0.1886}{M} + \frac{0.8}{M^2}\right)^{-1/2}\widehat{\kappa_3},$$

$$Z_4 \triangleq \left(\frac{0.0883}{M} + \frac{0.68}{M^2} + \frac{4.9}{M^3}\right)^{-1/2}\left(\widehat{\kappa_4} - 0.1226\right), \tag{E.40}$$

closely follow a standard Gaussian distribution. Therefore, a test statistic

$$L_{3,4} \triangleq Z_3^2 + Z_4^2 \tag{E.41}$$

follows a central χ^2 distribution with two degrees of freedom. This test is simple in construction and is independent of the variance σ^2. In detection theory, such tests are referred to as the Constant False Alarm Rate (CFAR) tests. The construction of the test and selection of the optimal threshold is similar to that used in the K^2 test.

As an ending note, it is worth mentioning that when the sample size is small (< 50), the Shapiro-Wilk (or W) test, and the Kolmogorov-Smirnov (or D) test may be the best choices ([160] [117]). For larger sample sizes, the choice of an appropriate test depends on the class of distributions under the alternative hypothesis. For e.g., if the alternative hypothesis is another Gaussian with an unknown, but non-zero mean, then the Anderson-Darling (or A^2) test could be a suitable choice ([52] [53]). Some of the above mentioned tests can also be used for a few distributions other than the Gaussian, but the scenarios under which it is applicable needs to be carefully examined. For e.g., the EDF based tests can be used for testing against known, non-Gaussian distributions with finite mean and variance, provided that the observations are i.i.d.

Bibliography

[1] J. Mitola and G. Maguire, "Cognitive radio: Making the software radios more personal," *IEEE Pers. Commun.*, vol. 6, pp. 13–18, Aug. 1999.

[2] J. Mitola, "Cognitive radio for flexible mobile multimedia communications," in *IEEE International Workshop on Mobile Multimedia Communications (MoMuC)*, San Diego, USA, 1999, pp. 3–10.

[3] ——, "Software radio architecture: a mathematical perspective," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 514–538, 1999.

[4] J. Mitola III, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Royal Institute of Technology (KTH), Sweden, May 2000.

[5] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radio," *Proc. Asilomar Conf. on Signals, Syst., and Comput.*, vol. 1, pp. 772–776, Nov. 2004.

[6] D. Cabric and R. Brodersen, "Physical layer design issues unique to cognitive radio systems," in *Proc. PIMRC*, vol. 2, Berlin, DE, 2005, pp. 759–763.

[7] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio networks," in *Proc. DySPAN*, Baltimore, USA, 2005, pp. 137–143.

[8] ——, "Agility improvement through cooperative diversity in cognitive radio," in *Proc. Globecom*, vol. 5, St. Louis, USA, 2005, pp. 1–5.

[9] H. Tang, "Some physical layer issues of wide-band cognitive radio systems," in *Proc. DySPAN*, Baltimore, USA, 2005, pp. 151–159.

[10] A. Ghasemi and E. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. DySPAN*, Baltimore, USA, 2005, pp. 131–136.

[11] ——, "Asymptotic performance of collaborative spectrum sensing under correlated log-normal shadowing," *IEEE Commun. Lett.*, vol. 11, no. 1, pp. 34–36, 2007.

[12] E. Larsson and G. Regnoli, "Primary system detection for cognitive radio: Does small-scale fading help?" *IEEE Commun. Lett.*, vol. 11, no. 10, pp. 799–801, 2007.

[13] J. Unnikrishnan and V. Veeravalli, "Cooperative sensing for primary detection in cognitive radio," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 18–27, Feb. 2008.

[14] D. Bhargavi and C. Murthy, "Performance comparison of energy, matched-filter and cyclostationarity-based spectrum sensing," in *Proc. SPAWC*, Jun. 2010, pp. 1–5.

[15] Y. R. Venugopalakrishna, C. R. Murthy, and D. N. Dutt, "Multiple transmitter

localization and communication footprint identification using energy measurements," *Physical Communication*, vol. 9, no. 0, pp. 184–192, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1874490712000717

[16] R. Thomas, L. DaSilva, and A. MacKenzie, "Cognitive networks," in *Proc. DyS-PAN*, Baltimore, USA, 2005, pp. 352–360.

[17] X. Jing, S.-C. Mau, D. Raychaudhuri, and R. Matyas, "Reactive cognitive radio algorithms for co-existence between IEEE 802.11b and 802.16a networks," in *Proc. Globecom*, vol. 5, St. Louis, USA, 2005, pp. 1–5.

[18] M. Gandetto and C. Regazzoni, "Spectrum sensing: A distributed approach for cognitive terminals," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 546–557, 2007.

[19] Y. Chen, G. Yu, Z. Zhang, H.-H. Chen, and P. Qiu, "On cognitive radio networks with opportunistic power control strategies in fading channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2752–2761, 2008.

[20] N. Devroye, P. Mitran, and V. Tarokh, "Cognitive multiple access networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Adelaide, Australia, 2005, pp. 57–61.

[21] K. Lee and A. Yener, "On the achievable rate of three-node cognitive hybrid wireless networks," in *International Conference on Wireless Networks, Communications and Mobile Computing*, vol. 2, Wuhan, China, 2005, pp. 1313–1318.

[22] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, USA, 2006, pp. 2413–2417.

[23] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, 2006.

[24] K. Nagananda, P. Mohapatra, C. R. Murthy, and S. Kishore, "Multiuser cognitive radio networks: an information-theoretic perspective," *International Journal of Advances in Engineering Sciences and Applied Mathematics*, vol. 5, no. 1, pp. 43–65, 2013. [Online]. Available: http://dx.doi.org/10.1007/s12572-013-0079-1

[25] A. Sharma and C. Murthy, "Group testing based spectrum hole search for cognitive radios," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1–12, Feb. 2014.

[26] J. Neel, R. Buehrer, J. Reed, and R. P. Gilles, "Game theoretic analysis of a network of cognitive radios," in *Proc. Midwest Symp. Circuits Syst.*, vol. 3, Oklahoma, USA, 2002, pp. 409–412.

[27] J. Neel, J. Reed, and R. Gilles, "Convergence of cognitive radio networks," in *Proc. WCNC*, vol. 4, Georgia, USA, 2004, pp. 2250–2255.

[28] C. Rieser, T. Rondeau, C. Bostian, and T. Gallagher, "Cognitive radio testbed: further details and testing of a distributed genetic algorithm based cognitive engine for programmable radios," in *Proc. MILCOM*, vol. 3, Monterey, USA, 2004, pp. 1437–1443.

[29] A. Fehske, J. Gaeddert, and J. Reed, "A new approach to signal classification using spectral correlation and neural networks," in *Proc. DySPAN*, Baltimore, USA, 2005, pp. 144–150.

[30] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive

radio applications," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 116–130, First Quarter 2009.

[31] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Phys. Commun.*, vol. 4, no. 1, pp. 40–62, Mar. 2011.

[32] I. F. Akyildiz, L. Won Yeol, C. V. Mehmet, and M. Shantidev, "A survey on spectrum management in cognitive radio networks," *IEEE Commun. Mag.*, vol. 46, pp. 40–48, Apr. 2008.

[33] D. Ariananda, M. Lakshmanan, and H. Nikookar, "A survey on spectrum sensing techniques for cognitive radio," *Second International Workshop on Cognitive Radio and Advanced Spectrum Management (CogART)*, pp. 74–79, May 2009.

[34] "Spectrum Policy Task Force Report," Federal Communications Commission, US, Nov. 2002.

[35] C. Cordeiro, K. Challapali, D. Birru, and S. Shankar, "IEEE 802.22: An introduction to the first wireless standard based on cognitive radios," *Journal of Communications*, vol. 1, no. 1, pp. 38–47, Apr. 2006.

[36] H. L. Van Trees, *Detection, estimation, and modulation theory*, 1st ed. Wiley, Sep. 2001.

[37] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Department of EECS, MIT, Nov. 1984.

[38] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 4–17, 2008.

[39] A. Tkachenko, D. Cabric, and R. Brodersen, "Cognitive radio experiments using reconfigurable bee2," in *Proc. Asilomar Conf. on Signals, Syst., and Comput.*, CA, USA, 2006, pp. 2041–2045.

[40] N. Han, S. Shon, J.-H. Chung, and J.-M. Kim, "Spectral correlation based signal detection method for spectrum sensing in IEEE 802.22 WRAN systems," in *International Conference Advanced Communication Technology (ICACT)*, vol. 3, Korea, 2006, pp. 1770–1776.

[41] J. Lunden, S. Kassam, and V. Koivunen, "Nonparametric cyclic correlation based detection for cognitive radio systems," in *International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom)*, Singapore, 2008, pp. 1–6.

[42] Z. Quan, W. Zhang, S. Shellhammer, and A. Sayed, "Optimal spectral feature detection for spectrum sensing at very low SNR," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 201–212, 2011.

[43] F. F. Digham, M. S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 21–24, Jan. 2007.

[44] H. V. Poor, *An introduction to signal detection and estimation*, 2nd ed. Springer, New York, 1994.

[45] J. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems (MCSS)*, vol. 1, no. 2, pp. 167–182, 1988.

[46] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, no. 4, pp. 523–531, Apr. 1967.

[47] S. J. Shellhammer, N. S. Shankar, R. Tandra, and J. Tomcik, "Performance of power detector sensors of DTV signals in IEEE 802.22 WRANs," in *Proc. TAPAS*. New York, NY, USA: ACM, 2006.

[48] T. M. Cover and J. M. Thomas, *Elements of information theory*, 2nd ed. John Wiley and Sons, Inc., 2005.

[49] R. Blahut, "Hypothesis testing and information theory," *IEEE Trans. Inf. Theory*, vol. 20, no. 4, pp. 405–417, Jul. 1974.

[50] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[51] M. G. Kendall and A. Stuart, *The advanced theory of statistics*, 2nd ed. Hafner publishing company, 1961, vol. II.

[52] H. Wang, E.-H. Yang, Z. Zhao, and W. Zhang, "Spectrum sensing in cognitive radio using goodness of fit testing," *IEEE Trans. Wireless Commun.*, vol. 8, no. 11, pp. 5427–5430, Nov. 2009.

[53] L. Shen, H. Wang, W. Zhang, and Z. Zhao, "Blind spectrum sensing for cognitive

radio channels with noise uncertainty," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1721–1724, Jun. 2011.

[54] S. Rostami, K. Arshad, and K. Moessner, "Order-statistic based spectrum sensing for cognitive radio," *IEEE Commun. Lett.*, vol. 16, no. 5, pp. 592–595, May 2012.

[55] D. Middleton, "Non-gaussian noise models in signal processing for telecommunications: new methods an results for class A and class B noise models," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.

[56] S. Gurugopinath, C. R. Murthy, and V. Sharma, "Error exponent analysis of energy-based Bayesian spectrum sensing under fading channels," in *Proc. Globecom*, Houston, USA, Dec. 2011, pp. 1–5.

[57] G. Sanjeev, K. V. K. Chaythanya, and C. R. Murthy, "Bayesian decentralized spectrum sensing in cognitive radio networks," in *Int. Conf. on Sig. Proc. and Commun. (SPCOM)*, Bangalore, India, Jul. 2010, pp. 1–5.

[58] S. Gurugopinath, "Near-optimal detection thresholds for bayesian spectrum sensing under fading," in *Proc. SPCOM*, Bangalore, India, Jul. 2014, pp. 1–6.

[59] S. Gurugopinath, R. Akula, C. R. Murthy, R. Prasanna, and B. Amruthur, "Spectrum sensing with a frequency-hopping primary: from theory to practice," in *Proc. ICC*, Sydney, Australia, Jun. 2014, pp. 1–6.

[60] S. Gurugopinath, C. R. Murthy, and C. S. Seelamantula, "Zero-crossings based spectrum sensing under noise uncertainties," in *Proc. NCC*, Kanpur, India, Mar. 2014, pp. 1–6.

[61] S. Gurugopinath, "Multi-dimensional goodness-of-fit tests for spectrum sensing based on stochastic distances," in *Proc. SPCOM*, Bangalore, India, Jul. 2014, pp. 1–6.

[62] S. Haykin, "Cognitive radio: Brain empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 201–220, Feb. 2005.

[63] E. Visotsky, S. Kuffner, and R. Peterson, "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," in *Proc. DySPAN*, Baltimore, USA, 2005, pp. 338–345.

[64] S. Mishra, A. Sahai, and R. Brodersen, "Cooperative sensing among cognitive radios," in *Proc. ICC*, vol. 4, Istanbul, Turkey, 2006, pp. 1658–1663.

[65] Y.-C. Liang, K.-C. Chen, G. Li, and P. Mahonen, "Cognitive radio networking and communications: an overview," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3386–3407, Sep. 2011.

[66] M. López-Benítez and F. Casadevall, "Versatile, accurate, and analytically tractable approximation for the Gaussian Q-function," *IEEE Trans. Commun.*, vol. 59, no. 4, pp. 917–922, Apr. 2011.

[67] A. Makarfi and K. Hamdi, "Interference analysis of energy detection for spectrum sensing," *IEEE Trans. Veh. Technol.*, vol. 62, no. 6, pp. 2570–2578, Jul. 2013.

[68] P. Sofotasios, E. Rebeiz, L. Zhang, T. Tsiftsis, D. Cabric, and S. Freear, "Energy detection based spectrum sensing over $\kappa-\mu$ and $\kappa-\mu$ extreme fading channels," *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1031–1040, Mar. 2013.

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

[70] Y. Lee and Y. Sung, "Generalized Chernoff information for mismatched Bayesian detection and its application to energy detection," *IEEE Signal Process. Lett.*, vol. 19, no. 11, pp. 753–756, Nov. 2012.

[71] P. K. Varshney, *Distributed detection and data fusion*, 1st ed.   Springer-Verlag New York, Inc., 1996.

[72] Y. Zeng, Y. C. Liang, and R. Zhang, "Blindly combined energy detection for spectrum sensing in cognitive radio," *IEEE Signal Process. Lett.*, vol. 15, pp. 649–652, 2008.

[73] P. Wang, J. Fang, N. Han, and H. Li, "Multiantenna-assisted spectrum sensing for cognitive radio," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1791–1800, May 2010.

[74] A. Pandharipande and J. Linnartz, "Performance analysis of primary user detection in a multiple antenna cognitive radio," in *Proc. ICC*, Glasgow, UK, Jun. 2007, pp. 6482–6486.

[75] S. Atapattu, C. Tellambura, and H. Jiang, "Energy detection based cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. on Wireless Commun.*, vol. 10, no. 4, pp. 1232 –1241, Apr. 2011.

[76] W. Zhang, R. Mallik, and K. Letaief, "Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5761–5766, Dec. 2009.

[77] H. Hashemi, "The indoor radio propagation channel," *Proc. IEEE*, vol. 81, no. 7, pp. 943–968, Jul. 1993.

[78] A. Iyer, K. K. Chintalapudi, V. Navda, R. Ramjee, V. Padmanabhan, and C. R. Murthy, "Spectrum sensing sans frontières," in *8th USENIX Symposium on Networked Systems Design and Implementation (NSDI '11)*, Boston, USA, Mar. 2011.

[79] S. Zheng, P.-Y. Kam, Y.-C. Liang, and Y. Zeng, "Spectrum sensing for digital primary signals in cognitive radio: A Bayesian approach for maximizing spectrum utilization," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1774–1782, Apr. 2013.

[80] D. Cabric, A. Tkachenko, and R. Brodersen, "Spectrum sensing measurements of pilot, energy, and collaborative detection," in *Proc. MILCOM*, Washington D.C., USA, Oct. 2006, pp. 1–7.

[81] A. Sonnenschein and P. M. Fishman, "Radiometric detection of spread-spectrum signals in noise of uncertain power," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, pp. 654–660, Jul. 1992.

[82] P. Barsocchi, "Channel models for terrestrial wireless communications: A survey," *available at http://puma.isti.cnr.it/dfdownload.php?ident=/cnr.isti/2006-TR-16*, 2006.

[83] Z. Li, F. Yu, and M. Huang, "A distributed consensus-based cooperative spectrum-sensing scheme in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 383–393, Jan. 2010.

[84] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, no. 1, pp. 98–101, Jan. 1986.

[85] J. Filho and M. Yacoub, "Simple precise approximations to Weibull sums," *IEEE Commun. Lett.*, vol. 10, no. 8, pp. 614–616, Aug. 2006.

[86] M. Di Renzo, F. Graziosi, and F. Santucci, "Further results on the approximation of log-normal power sum via Pearson type IV distribution: A general formula for log-moments computation," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 893–898, Apr. 2009.

[87] R. Srinivasan, *Importance sampling: Applications in communications and detection.* Springer Verlag, 2002.

[88] S. Astaneh and S. Gazor, "Cooperative spectrum sensing over mixture-nakagami channels," *IEEE Commun. Lett.*, vol. 2, no. 3, pp. 259–262, 2013.

[89] E. Axell, G. Leus, E. Larsson, and H. Poor, "Spectrum sensing for cognitive radio : State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101–116, May 2012.

[90] X. Li and X.-B. Li, "Common fallacies in applying hypothesis testing," in *11th Int. Conf. on Information Fusion*, Cologne, Germany, Jul. 2008, pp. 1–8.

[91] M.-S. Alouini and M. K. Simon, "Performance of generalized selection combining over Weibull fading channels: Research articles," *Wireless. Commun. and Mobile Comput.*, vol. 6, pp. 1077–1084, Dec. 2006.

[92] "Coverage prediction for mobile radio systems operating in the 800/900 MHz frequency range," *IEEE Trans. on Veh. Tech.*, vol. 37, no. 1, pp. 3–72, 1988.

[93] S. Y. D. Dubey, "Normal and weibull distributions," *Naval Research Logistics*, no. 14, pp. 69–79, 1967.

[94] F. Hernandez and R. A. Johnson, "The large-sample behavior of transformations to normality," *Journal of the American Statistical Association*, vol. 75, no. 372, pp. 855–861, 1980.

[95] N. Shepherd, "Radio wave loss deviation and shadow loss at 900 MHz," *IEEE Trans. on Veh. Tech.*, vol. 26, no. 4, pp. 309–313, 1977.

[96] A. Coulson, A. Williamson, and R. Vaughan, "Improved fading distribution for mobile radio," *IEE Proc. Commun.*, vol. 145, no. 3, pp. 197–202, Jun. 1998.

[97] D.-C. Oh and Y.-H. Lee, "Low complexity fft based spectrum sensing in bluetooth system," in *Proc. VTC*, 2009, pp. 1–5.

[98] G. R. Cooper, "Detection of frequency-hop signals," *IEEE MILCOM*, vol. 1, pp. 1–5, Oct. 1986.

[99] L. Miller, J. Lee, and D. Torrieri, "Frequency-hopping signal detection using partial band coverage," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, pp. 540–553, Apr. 1993.

[100] R. A. Dillard and G. M. Dillard, "Likelihood-ratio detection of frequency-hopped signals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, pp. 543–553, April 1996.

[101] Z. Chen, N. Guo, and R. Qiu, "Demonstration of real-time spectrum sensing for cognitive radio," in *Proc. MILCOM*, Oct 2010, pp. 323–328.

[102] Y. Hur, J. Park, K. Kim, J. Lee, K. Lim, C. H. Lee, S. H. Kim, and J. Laskar, "A cognitive radio (cr) testbed system employing a wideband multi-resolution spectrum sensing (mrss) technique," *Proc. VTC*, pp. 1–5, Sep. 2006.

[103] Z. Peng, Q. Robert, and N. Guo, "Demonstration of spectrum sensing with blindly learned features," *IEEE Commun. Lett.*, vol. 15, pp. 548–550, May 2011.

[104] D. Kun and N. A. Morgan, "A new low-cost cfar detector for spectrum sensing with cognitive radio systems," *IEEE Aerospace Conference*, pp. 1–8, Mar. 2008.

[105] L. U. Ancarani and G. Gasaneo, "Derivatives of any order of the hypergeometric function $_2F_1(a, b; c; z)$ with respect to the parameters a, b and c," *Journal of Physics A: Mathematical and Theoretical*, vol. 43, no. 8, p. 085210, 2010.

[106] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, pp. 523–531, Apr. 1967.

[107] P. B. Patnaik, "The non-central χ^2- and F-distribution and their applications," *Biometrika*, vol. 36, no. 1/2, pp. 202–232, Jun. 1949.

[108] G. Chavali, "Signal detection and modulation classification in non-gaussian noise environments," Ph.D. dissertation, Virginia Polytechnic Institute and State University, Jul. 2012.

[109] K. S. Vastola, "Threshold detection in narrow-band non-gaussian noise," *IEEE Trans. Commun.*, vol. 32, no. Feb., pp. 134–139, 1984.

[110] J. Miller and J. Thomas, "The detection of signals in impulsive noise modeled as a mixture process," *IEEE Trans. Commun.*, vol. 24, no. 5, pp. 559–563, May 1976.

[111] B. Aazhang and H. Poor, "Performance of DS/SSMA communications in impulsive channels–part I: Linear correlation receivers," *IEEE Trans. Commun.*, vol. 35, no. 11, pp. 1179–1188, Nov. 1987.

[112] K. Jitvanichphaibool, Y.-C. Liang, and Y. Zeng, "Spectrum sensing using multiple antennas for spatially and temporally correlated noise environments," in *2010 IEEE Symposium on New Frontiers in Dynamic Spectrum*, Singapore, Apr. 2010, pp. 1–7.

[113] V. Pohl, F. Suratman, A. Zoubir, and H. Boche, "Spectrum sensing for cognitive radio architectures based on sub-nyquist sampling schemes," in *2011 International ITG Workshop on Smart Antennas (WSA)*, Aachen, Germany, Feb. 2011, pp. 1–8.

[114] T. W. Anderson and D. A. Darling, "Asymptotic theory of certain "Goodness of Fit" criteria based on stochastic processes," *Ann. Math. Statist.*, vol. 23, no. 2, pp. 193–212, Jun. 1952.

[115] B. Kedem and E. Slud, "Time series discrimination by higher order crossings," *The Annals of Statistics*, vol. 10, no. 3, pp. 786–794, Sep. 1982.

[116] I. Gradshteyn and I. Ryzhik, *Tables of integrals, series and products*, 7th ed. Academic Press, 2007.

[117] M. A. Stephens, "EDF statistics for goodness of fit and some comparisons," *Journal of the American Statistical Association*, vol. 69, no. 347, pp. 730–737, Sep. 1974.

[118] B. Kedem, "Spectral analysis and discrimination by zero-crossings," *Proc. IEEE*, vol. 74, no. 11, pp. 1477–1493, Nov. 1986.

[119] B. L. Welch, "On linear combinations of several variances," *Journal of the American Statistical Association*, vol. 51, no. 273, pp. 132–148, Mar. 1956.

[120] M. West, "On scale mixtures of normal distributions," *Biometrika*, vol. 74, no. 3, pp. 646–648, 1987.

[121] R. S. Blum, Y. Zhang, B. M. Sadler, and R. J. Kozick, "On the approximation of correlated non-gaussian noise pdfs using gaussian mixture models," in *Conference on the Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics*, Jun. 1999.

[122] S. Shellhammer, "Spectrum sensing in IEEE 802.22," in *First Workshop on Cognitive Info. Process.(CIP 2008)*, Santorini, Greece, 2008.

[123] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: challenges and solutions," *EURASIP J. Adv. Signal Process.*, vol. 2010, Jan. 2010.

[124] T. Feng, T. Field, and S. Haykin, "Stochastic differential equation theory applied to wireless channels," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1478–1483, 2007.

[125] S. Krusevac, P. Rapajic, and R. Kennedy, "Channel capacity estimation for MIMO systems with correlated noise," in *Proc. Globecom*, vol. 5, St. Louis, USA, Dec. 2005, pp. 1–5.

[126] V. Aalo and R. Viswanathan, "Asymptotic performance of a distributed detection system in correlated gaussian noise," *IEEE Trans. Signal Process.*, vol. 40, no. 1, pp. 211–213, Jan. 1992.

[127] R. Zhang, T. J. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 84–88, Jan. 2010.

[128] S. Kritchman and B. Nadler, "Non-parametric detection of the number of signals: Hypothesis testing and random matrix theory," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3930–3941, Oct. 2009.

[129] P. Bianchi, M. Debbah, M. Maida, and J. Najim, "Performance of statistical tests for single-source detection using random matrix theory," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2400–2419, Apr. 2011.

[130] S. John, "The distribution of a statistic used for testing sphericity of normal distributions," *Biometrika*, vol. 59, no. 1, pp. 169–173, Apr. 1972.

[131] Y. Zeng and Y. C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1784–1793, Jun. 2009.

[132] F. Penna, R. Garello, and M. Spirito, "Cooperative spectrum sensing based on the limiting eigenvalue ratio distribution in wishart matrices," *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 507–509, Jul. 2009.

[133] L. Wei and O. Tirkkonen, "Spectrum sensing in the presence of multiple primary users," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1268–1277, May 2012.

[134] J. W. Mauchly, "Significance test for sphericity of a normal n-variate distribution," *The Annals of Mathematical Statistics*, vol. 11, no. 2, pp. 204–209, Jun. 1940.

[135] R. Bartoszynski, D. K. Pearl, and J. Lawrence, "A multidimensional goodness-of-fit test based on interpoint distances," *Journal of the American Statistical Association*, vol. 92, no. 438, pp. 577–586, Jun. 1997.

[136] M. Salicru, D. Morales, M. Menendez, and L. Pardo, "On the applications of divergence type measures in testing statistical hypotheses," *Journal of Multivariate Analysis*, vol. 51, no. 2, pp. 372–391, Nov. 1994.

[137] J.-F. Maa, D. K. Pearl, and R. Bartoszynski, "Reducing multidimensional two-sample data to one-dimensional interpoint comparisons," *The Annals of Statistics*, vol. 24, no. 3, pp. 1069–1074, Jun. 1996.

[138] A. Nielsen, H. Skriver, and K. Conradsen, "Complex wishart distribution based analysis of polarimetric synthetic aperture radar data," in *International Workshop on the Analysis of Multi-temporal Remote Sensing Images (MultiTemp)*, Provinciehuis Leuven, Belgium, 2007, pp. 1–6.

[139] A. C. Frery, A. D. C. Nascimento, and R. J. Cintra, "Information theory and image understanding: An application to polarimetric SAR imagery," *Chilean Journal of Statistics*, vol. 2, no. 2, pp. 81–101, Sep. 2011.

[140] F. Yilmaz and M.-S. Alouini, "Sum of weibull variates and performance of diversity systems." in *Proc. IWCMC*, Leipzig, Germany, Jun. 2009, pp. 247–252.

[141] A. Winkelbauer, "Moments and absolute moments of the normal distribution," *Arxiv preprint arXiv:1209.4340*, Sep. 2012.

[142] R. L. Winkler, G. M. Roodman, and R. R. Britney, "The determination of partial moments," *Management Science*, vol. 19, no. 3, pp. pp. 290–296, 1972.

[143] S. Gezici, Z. Sahinoglu, and H. Poor, "On the optimality of equal gain combining for energy detection of unknown signals," *IEEE Commun. Lett.*, vol. 10, no. 11, pp. 772–774, Nov. 2006.

[144] I. Hoballah and P. Varshney, "Distributed bayesian signal detection," *IEEE Trans. Inf. Theory*, vol. 35, no. 5, pp. 995–1000, Sep. 1989.

[145] I. Csiszar and J. Korner, *Information theory coding theorems for discrete memoryless systems*, 1st ed. Academic Press, Inc. (London) Ltd., 1981.

[146] N. C. Sagias and G. S. Tombras, "On the cascaded Weibull fading channel model," *Journal of the Franklin Institute*, vol. 344, no. 1, pp. 1–11, 2007.

[147] C. Withers and S. Nadarajah, "A generalized Suzuki distribution," *Wireless Pers. Commun.*, pp. 1–24, 2010, 10.1007/s11277-010-0095-4.

[148] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. Wiley & Sons, 1995, vol. 2.

[149] A. Gok, S. Joshi, J. Villasenor, and C. Danijela., "Estimating the number of frequency hopping interferers using spectral sensing with time and frequency offset measurements," *IEEE MILCOM*, pp. 1–7, Oct. 2009.

[150] L. J. Tick and P. Shaman, "Sampling rates and appearance of stationary gaussian processes," *Technometrics*, vol. 8, no. 1, pp. 91–106, Feb. 1966.

[151] S. O. Rice, "Mathematical analysis of random noise," *Bell System Tech. J.*, vol. 23, no. 3, pp. 282–332, Jul. 1944.

[152] W. Hoeffding, "On a theorem of V. M. Zolotarev," *Thoer. Probability Appl.*, vol. 9, pp. 89–92, 1964.

[153] M. A. Stephens, "Asymptotic results for goodness-of-fit statistics with unknown parameters," *The Annals of Statistics*, vol. 4, no. 2, pp. 357–369, Mar. 1976.

[154] Student, "The probable error of a mean," *Biometrika*, vol. 6, no. 1, pp. 1–25, Mar. 1908.

[155] K. Pearson, "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *Philosophical Magazine Series*, vol. 50, p. 157175, 1908.

[156] A. Harri and K. H. Coble, "Normality testing: two new tests using L-moments," *Journal of Applied Statistics*, vol. 38, no. 7, pp. 1369–1379, 2011.

[157] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 3, no. 52, Dec. 1965.

[158] R. B. D'Agostino, "An omnibus test of normality for moderate and large size samples," *Biometrika*, vol. 58, no. 2, pp. 341–348, Aug. 1971.

[159] ——, "Small sample probability points for the D test of normality," *Biometrika*, vol. 59, no. 1, pp. 219–221, Apr. 1972.

[160] P. Royston, "Approximating the shapiro-wilk W-test for non-normality," *Statistics and Computing*, vol. 2, no. 3, pp. 117–119, Oct. 1992.

[161] L. Chen and S. S. Shapiro, "An alernative test for normality based on normalized spacings," *Journal of Statistical Computation and Simulation*, vol. 53, no. 12, pp. 269–287, Dec. 1995.

[162] P. Zhang, "Omnibus test of normality using the q statistic," *Journal of Applied Statistics*, vol. 26, no. 4, pp. 519–528, Oct. 1999.

[163] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *The Annals of Mathematical Statistics*, vol. 19, no. 2, pp. 279–281, 1948.

[164] R. B. D'Agostino, A. Belanger, and J. D'Agostino, Ralph B., "A suggestion for using powerful and informative tests of normality," *The American Statistician*, vol. 44, no. 4, pp. 316–321, Nov. 1990.

[165] E. Seier, "Comparison of tests for univariate normality," *Statistics on the Internet*, pp. 1–17, Jan. 2002.

[166] J. R. M. Hosking, "L-moments: Analysis and estimation of distributions using linear combinations of order statistics," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 52, no. 1, pp. 105–124, 1990.

⋮

Nicolaus (Mikoaj Kopernik) Copernicus (1499)

Georg Joachim von Leuchen Rheticus (1535)

Moritz Valentin Steinmetz (1550, 67)

Christoph Meurer (1582)

Friedrich Leibniz (1622)

Philipp Mller (1604)

Jakob Thomasius (1643)

Erhard Weigel (1650)

Christian Huygens (1647,55)

Otto Mencke (1665)

Gottfried Wilhelm Leibniz (1672)

Nicolas Malebranche (1672)

Jacob Bernoulli (1676,84)

Johann Wichmannshausen (1685)

Johann Bernoulli (1694)

Christian Hausen (1713)

Leonhard Euler (1726)

Jean LeRond d'Alembert

Johann Elert Bode

Abraham Kastner (1739)

Joseph Louis Lagrange (1754)

Pierre Simon Laplace

Johann Friedrich Pfaff (1786)

Carl Friedrich Gauss (1799)

Jean-Bapiste Joseph Fourier

Simeon Denis Poisson (1800)

Martin Ohm (1811)

Christian Ludwig Gerling (1812)

Julius Plucker (1823)

Gustav Peter Lejeune Dirichlet (1827)

Ernst Eduard Kummer (1831)

Rudolf Otto Sigismund Lipschitz (1853)

Leo Pochhammer (1863)

C. L. Ferdinand (Carl Louis) Lindemann (1873)

David Hilbert (1885)

Hermann Minkowski (1885)

Martin Kutta (1900)

Arnold Johannes Wilhelm Sommerfeld (1891)

Ernst Adolph Guillemin (1926)

Robert Mario Fano (MIT 1947)

John McReynolds Wozencraft (MIT 1957)

Thomas Kailath (MIT 1961)

David Forney Jr. (MIT 1965)

Jimmy K. Omura (MIT 1966)

Sun-Yuan Kung (Stanford 1977)

Ali H. Sayed (MIT 1992)

Babak Hassibi (MIT 1996)

Bhaskar D. Rao (USC 1983)

KVS Hari (UCSD 1990)

Chandra Ramabhadra Murthy (UCSD 2006)

David Wipf (UCSD)

Sanjeev Gurugopinath (IISc 2015)

⋮