Detection and Estimation Techniques in Cognitive Radio

A thesis submitted to The University of Manchester for the degree of

Doctor of Philosophy

in the Faculty of Engineering and Physical Sciences

2013

Juei-Chin Shen

School of Electrical and Electronic Engineering

CONTENTS

LIST OF FIGURES

NOMENCLATURE

AWGN Additive white Gaussian noise

BMRC Blind maximum ratio combining

CAF Cyclic autocorrelation function

CFD Cyclostationary feature detection

CFL Cycle-frequency-lag

CR Cognitive Radio

CSI Channel side information

CU Cognitive user

EGC Equal gain combining

FAR False alarm rate

FCC Federal Communications Commission

GLRT Generalized likelihood ratio test

i.i.d. Independent and identically distributed

MASS Multi-antenna spectrum sensing

MRC Maximum ratio combining

PU Primary user

RF Radio frequency

RMSs Received multi-antenna signals

ROC Receiver operating characteristic

SC Selection Combining

SINR Signal-to-interference-plus-noise ratio

SNR Signal-to-noise ratio

PDF Probability density function

χ^2_N Chi-square distribution of N degrees of freedom

$\chi^2_N(\gamma)$ Non-central chi-square distribution of N degrees of freedom with noncentrality γ

$\mathcal{H}_0$ Null hypothesis

$\mathcal{H}_1$ Alternative hypothesis

$\mathcal{N}(a, b)$ Normal distribution with mean a and variance b

$F_{\chi^2_N(\gamma)}(\cdot)$ Cumulative distribution of the $\chi^2_N(\gamma)$ variate

$F_{\chi^2_N}(\cdot)$ Cumulative distribution of the χ^2_N variate

$f_{ray}(x; \sigma)$ Rayleigh probability density function $\frac{x}{\sigma^2} e^{\frac{-x^2}{2\sigma^2}}$

P_d Probability of detection

P_f Probability of false alarm

$Q(x)$ Q-function $\frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(\frac{-u^2}{2})\, \mathrm{d}u$

$Q^{-1}(\cdot)$ Inverse Q-function

$Q_M(a, b)$ Generalized Marcum Q-function

Thesis Title: Detection and Estimation Techniques in Cognitive Radio
Doctor of Philosophy **The University of Manchester**
Juei-Chin Shen **September 2013**

Abstract

Faced with imminent spectrum scarcity largely due to inflexible licensed band arrangements, cognitive radio (CR) has been proposed to facilitate higher spectrum utilization by allowing cognitive users (CUs) to access the licensed bands without causing harmful interference to primary users (PUs). To achieve this without the aid of PUs, the CUs have to perform spectrum sensing reliably detecting the presence or absence of PU signals. Without reliable spectrum sensing, the discovery of spectrum opportunities will be inefficient, resulting in limited utilization enhancement.

This dissertation examines three major techniques for spectrum sensing, which are matched filter, energy detection, and cyclostationary feature detection. After evaluating the advantages and disadvantages of these techniques, we narrow down our research to a focus on cyclostationary feature detection (CFD). Our first contribution is to boost performance of an existing and prevailing CFD method. This boost is achieved by our proposed optimal and sub-optimal schemes for identifying best hypothesis test points. The optimal scheme incorporates prior knowledge of the PU signals into test point selection, while the sub-optimal scheme circumvents the need for this knowledge. The results show that our proposed can significantly outperform other existing schemes. Secondly, in view of multi-antenna deployment in CR networks, we generalize the CFD method to include the multi-antenna case. This requires effort to justify the joint asymptotic normality of vector-valued statistics and show the consistency of covariance estimates. Meanwhile, to effectively integrate the received multi-antenna signals, a novel cyclostationary feature based channel estimation is devised to obtain channel side information. The simulation results demonstrate that the errors of channel estimates can diminish sharply by increasing the sample size or the average signal-to-noise ratio. In addition, no research has been found that analytically assessed CFD performance over fading channels. We make a contribution to such analysis by providing tight bounds on the average detection probability

over Nakagami fading channels and tight approximations of diversity reception performance subject to independent and identically distributed Rayleigh fading.

For successful coexistence with the primary system, interference management in cognitive radio networks plays a prominent part. Normally certain average or peak transmission power constraints have to be placed on the CR system. Depending on available channel side information and fading types (fast or slow fading) experienced by the PU receiver, we derive the corresponding constraints that should be imposed. These constraints indicate that the second moment of interference channel gain is an important parameter for CUs allocating transmission power. Hence, we develop a cooperative estimation procedure which provides robust estimate of this parameter based on geolocation information. With less aid from the primary system, the success of this procedure relies on statistically correlated channel measurements from cooperative CUs. The robustness of our proposed procedure to the uncertainty of geolocation information is analytically presented. Simulation results show that this procedure can lead to better mean-square error performance than other existing estimates, and the effects of using inaccurate geolocation information diminish steadily with the increasing number of cooperative cognitive users.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Publications

Journal Contributions

1. J.-C. Shen and E. Alsusa, "An Efficient Multiple Lags Selection Method for Cyclostationary Feature Based Spectrum Sensing," *IEEE Signal Processing Letters*, vol. 20, pp. 133-136, Feb. 2013.

2. J.-C. Shen and E. Alsusa, "Cooperative Estimation of Path Loss in Interference Channels without Primary-user CSI Feedback," *IEEE Signal Processing Letters*, vol. 20, pp. 273-276, March 2013.

3. J.-C. Shen and E. Alsusa, "Joint Cycle Frequencies and Lags Utilization in Cyclostationary Feature Spectrum Sensing," *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5337-5346, Nov. 2013.

4. J.-C. Shen and E. Alsusa, "Cyclostationary Feature Based Multi-antenna Spectrum Sensing for Cognitive Radio," submitted to *IEEE Transactions on Signal Processing*.

Conference Contributions

1. J.-C. Shen, E. Alsusa, and D. K. So, "Performance Bounds on Cyclostationary Feature Detection over Fading Channels," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2013, pp. 2971-2975.

2. J.-C. Shen and E. Alsusa, "Post-Combining Based Cyclostationary Feature Detection for Cognitive Radio over Fading Channels," to be presented at the IEEE GLOBECOM, Atlanta, Dec. 2013.

3. J.-C. Shen and E. Alsusa, "Cyclostationary Feature Based MRC Multi-antenna Spectrum Sensing," submitted to IEEE WCNC 2014.

Acknowledgements

First and foremost, I thank my supervisor, Dr. Emad Alsusa, for his great support, encouragement and advice throughout my Ph.D. study. With his support, I had complete freedom to audit relevant lectures and acquainted myself with desired mathematical tools. With his encouragement, I submitted part of my thesis work to refereed journals or conferences and successfully published three journal articles and two conference papers. With his advice, I developed a positive yet practical attitude towards research. Looking back, I feel deeply indebted to him.

I also owe my sincerest gratitude to Dr. Daniel K.C. So for providing positive feedback on my first-year research report, to Dr. Zhirun Hu for scrutinizing the results of this thesis and facilitating my thesis defence, and to Dr. Kai-Kit Wong for challenging the contributions of this thesis and making outspoken comments. Because of their assistance, this thesis can be presented in a professional fashion.

This acknowledgment would not be complete without mentioning my colleagues: Ebtihal H. Gismalla, Wahyu Pramudito, Warit Prawatmuang, Khaled M. Rabie, Mohammad R. Mili, Javad Jafarian, Azwan Mahmud, Abubakar U. Makarfi, Fahimeh Jasbi, Tarla Abadi. Thank you for all the stimulating discussions and enjoyable time in the MACS group. I wish our paths will cross again in the future.

In closing, I would like to acknowledge the support of my family. I thank my parents, Ching-Shan and Su-Chao, my wife, Rui-Han, and my daughter, Sin-Yue, for their patience and love.

Dedicated to my parents

CHAPTER 1

INTRODUCTION

1.1 Overview

The conventional spectrum management method, allocating most radio spectrum of high economic value for licensed use, has caused a phenomenon, "spectrum scarcity". It made it seem as though there are no more available radio resources to keep up with the growing demand of high-data-rate wireless transmission. However, a report, conducted by the Federal Communications Commission (FCC), indicated that most licensed bands are not fully utilized. The inconsistency between scarcity and under-utilization results from inflexible spectrum management. Cognitive radio, the most promising technology of wireless communications, shows a liberal way out of this inconsistency by allowing unlicensed access in licensed bands in an opportunistic way. The accessible licensed bands, which are referred to as spectrum opportunities, can be categorized in different ways.

First, spectrum opportunities can arise when no licensed-system activities occur temporarily or geographically in the bands of interest. To discover these opportunities, a CR has to perform the task of spectrum sensing to make it aware of the surrounding radio environment. Cyclostationary feature detection, one of the methods for spectrum sensing, has shown its superiority over other methods. Although this method has existed for decades, there is still much room for innovation when it comes to applying CFD to CR. One purpose of this thesis is to review recent research into CFD and discover an efficient and high-performance CFD scheme.

The second kind of opportunities appears when the licensed transmission can tolerate certain levels of interference without performance degradation. In other words, CR transmission is permitted on the condition that the potential interference to the licensed system is not harmful.The success of this approach relies on appropriate inference constraints imposed on CR. Meanwhile, CR has to meet these constraints by exploiting channel side information (CSI). Therefore, there are two fundamental questions that need to be addressed. What interference constraints should be imposed on CR by a licensed system? How does CR acquire CSI and manage its radio resources to meet interference constraints? With this in mind, our second purpose is to answer these two questions.

1.2 Previous Work & Author's Contributions

This dissertation contributes to the topics of cyclostationary feature based spectrum sensing and interference management in cognitive radio.

Cyclostationary Spectrum Sensing

1. **Previous Work**: Cyclostationary feature detection is a preferred method for spectrum sensing under low signal-to-noise ratio (SNR) or/and noise uncertainty scenarios. To determine the presence, or otherwise, of PU signals, conventional CFD schemes tend to use test statistics over, either, multiple cycle frequencies for a fixed lag set, or, multiple lags for a fixed cycle frequency. These schemes are simple to implement, but carry with them the shortcoming of inefficient usage of cycle frequencies and lags.

   Contribution: This thesis proposed a new method that jointly utilizes cycle frequencies and lags to produce more reliable test statistics. As the optimal way to apply this joint utilization requires prior knowledge of the 4th-order cyclic cumulant, which can be challenging to obtain, an alternative sub-optimal scheme independent of this cumulant knowledge has been provided. It has been shown that, in the low SNR region, where it is most critical for CR applications, the proposed sub-optimal scheme can lead to similar detection performances as the optimal maximum likelihood technique. It has also been demonstrated that, compared to multi-cycle-frequency detection with selection combining, equal gain combining, or maximum ratio combining, the proposed provides superior performance.

2. **Previous Work**: Multi-antenna-assisted CFD has been applied to enhance spectrum sensing. However, the existing work fails to address essential prerequisites for forming asymptotic chi-square testing, such as joint asymptotic normality and consistency of estimates. Additionally, no attempt was made to justify the usage of maximum ratio combining in terms of maximizing detection performance. Another weakness is that there is no guarantee of analytical reliability of cyclostationary feature based channel estimation.

   Contribution: This thesis investigated multi-antenna-assisted cyclostationary feature based spectrum sensing. First, it has been shown that maximum ratio combining is the optimal linear pre-combining scheme for maximizing the performance metric. As maximum ratio combining requires CSI, we developed a blind channel estimation technique based on cyclostationary statistics for which analytical reliability is guaranteed. For post-combining, which requires no CSI, two practical schemes, joint combining and sum combining, have been examined. Especially, we rigorously established the desired joint asymptotic normality of vector-valued statistics and the consistency of the corresponding covariance estimates.

3. **Previous Work**: The relevant research to date tended to focus on analyzing energy detection over fading channels rather than CFD. Therefore, to obtain CFD performance in fading environments heavily relies on numerical simulations.

   Contribution: We provided an analytical expression of a tight upper bound on the average CFD detection performance. This is achieved by simplifying an argument which appears in the generalized Marcum Q-function. For diversity reception subject to independent and identically distributed Rayleigh fading, we were able to derive approximated detection probabilities in a series form for post addition combining and post selection combining. Provided numerical results demonstrate that the theoretical detection performance can be well approximated by our proposed method in the low average SNR region.

Interference Management

1. **Previous Work**: Generally, the licensed system will impose either a peak-power or an average-power constraint on CR transmitters, depending on its quality-of-service type. However, it is not clear what implications of using the peak-power and average-power constraints are from a physical layer perspective. Furthermore, it is also interesting to know what CSI is required corresponding to different constraints.

   Contribution: Our analysis has shown that what constraint to place depends on the channel fading type experienced by the licensed receiver. In slow fading environments, CR transmission has to satisfy the peak-power constraint with perfect CSI and the average-power constraint with partial CSI. This average-power constraint depends on certain outage probability limit. For the case of fast fading, a dynamic-power constraint has be imposed with perfect CSI, and a certain ergodic capacity limit has be achieved with partial CSI.

2. **Previous Work**: Estimation of channel knowledge has been a field well studied

in communications research. Particularly the need for cross-channel information with CR involved in a cooperative network setting is crucial. Many techniques in this regard have been suggested in literature, with path loss estimation being one of them. Two existing path loss estimation schemes are mainly based on the method of least squares.

Contribution: In this thesis, we have contributed the work on estimating path loss by adding another measurement that allows for smaller errors in path loss estimation, which is the de-correlation distance. For a single user system, this may not result in significant change, but the novelty lies in the performance achievable for a generally populated radio network. Moreover, the theoretical proof supports the improved simulation performance that guarantees a lower mean square error for the proposed estimation approach.

1.3 Description of the Chapters

This thesis has been organized in the following way.

- Chapter 2 begins by reviewing the background knowledge of CR. The most relevant spectrum-sensing studies and referred techniques will be examined.

- Chapter 3 provides a brief recapitulation of cyclostationary processes and the 2nd-order cyclostationary features. Based on these, a general statistical testing model in the cycle-frequency-lag domain is introduced. The derived asymptotic performance is further exploited to select test points. The optimal and sub-optimal schemes are proposed and discussed. For the purpose of comparison with our proposed method, the existing multi-cycle-frequency detection is succinctly reviewed. Asymptotically statistical properties of cyclostationary features of a linear modulated signal which is used as a simulation example are analytically presented.

- Chapter 4 gives a brief recapitulation of the kth-order almost cyclostationary processes and the system model for cyclostationary feature based multiple-antenna spectrum sensing. The pre-combining and post-combining (including joint combining and sum combining) schemes are analyzed. The sufficient conditions under which CU cochannel interference will not degrade our proposed multiple-antenna spectrum sensing performance are presented.

- Chapter 5 addresses analytical performance of CFD over fading channels. It first provides the upper and lower bounds on CFD performance over Nakagami fading channels. The case of diversity reception subject to independent and identically distributed Rayleigh fading is investigated based on a tight performance approximation.

- Chapter 6 discusses the suitable interference constraints to impose in a licensed-system perspective. A more realistic CSI assumption is proposed, this is, the statistical characterization of channels. The method of evaluating large-scale path losses between the CR and licensed systems is analytically presented.

- Chapter 7 gives conclusions of work already undertaken and their implications. Future work is also presented.

CHAPTER 2

LITERATURE REVIEW

2.1 Cognitive Radio

CR is a paradigm of integrating computational intelligence and wireless technology in which both radio environment and user needs are considered to provide personalized wireless communications services [1, 2]. The "radio environment" demonstrates enormous variety, ranging from artificial radio regulations to physical interference level—in other words all perceivable information describing the surrounding wireless environment. Hence, to allow CR to work properly, recognizing the radio environment, becomes an indispensable step. For instance, a remarkable application enabled by CR is spectrum pooling which permits spectrum renting between licensed and unlicensed users [3]. The radio environment that needs to be recognized by unlicensed users in this application is exploitable licensed bands.

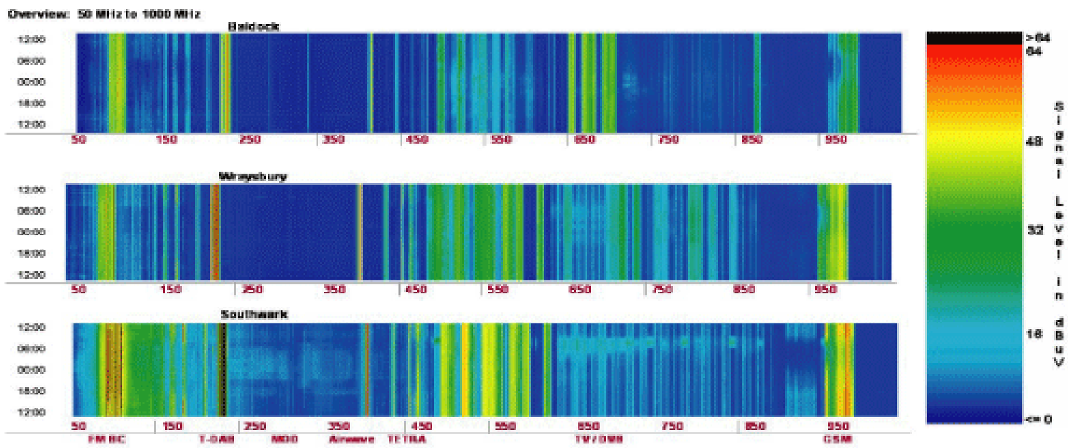


Figure 2.1: Spectrum occupancy measurements in a rural area (top), near Heathrow airport (middle) and in central London (bottom)

2.1.1 Remedy for Spectrum Scarcity

In November 2002, the Federal Communications Commission issued a report examining the spectrum utilization at the time and providing recommendations [4]. This report shed new light on the issue of spectrum scarcity, i.e., running out of usable radio frequencies. One informed comment stated that

> "In many bands, spectrum access is a more significant problem than physical scarcity of spectrum, in large part due to legacy command-and-control regulation that limits the ability of potential spectrum users to obtain such access."

In other words, the spectrum scarcity faced is largely due to inflexible spectrum access in untapped licensed bands which are identified in [5]. A similar phenomenon of inefficient spectrum utilization (see Fig. 2.1[1]) was also observed by the Office of Communications (Ofcom). As can be seen in this figure, a significant portion of

[1]Ofcom, Cognitive Radio Webpage, <http://stakeholders.ofcom.org.uk/market-data-research/technology-research/research/emerging-tech/cograd/>

radio resources was left unused especially in rural areas. This observation led to a remark, [6]

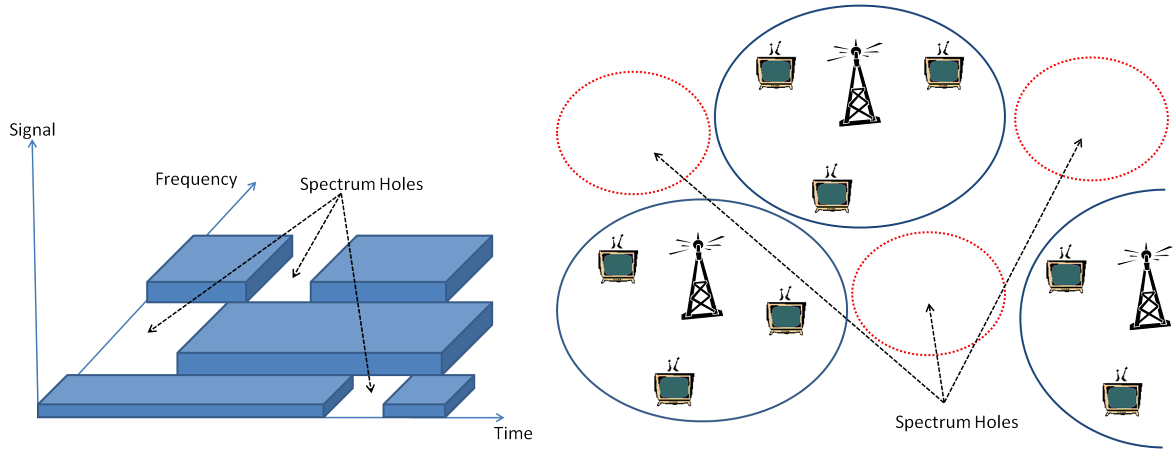
> "Such measurements indicate the potential gains that might be accrued from a system which was able to sense its radio environment and utilise it effectively,"

suggesting the adoption of CR-like technologies to enable the access to underutilized licensed bands. Television (TV) whitespaces, the frequency bands assigned to TV broadcasting services but not in use in a specific location, were the first licensed bands where unlicensed operation was approved by the FCC in November 2008. It was expected that a new wave of innovations would arrive in the coming years. Indeed, several novel whitespace applications were proposed such as an alternative rural link, wireless backhaul of a Wi-Fi hotspot, and local TV broadcasting. Later a CR-based standard, IEEE 802.22-2011, was finalized to enable rural broadband wireless access in TV whitespaces. The deployment of LTE in TV whitespaces with CR was also under consideration [7]. As a remedy for spectrum scarcity and the enabling technology of new services, CR has become an increasingly important research area in recent years.

2.1.2 Prerequisite of License-exempt Access

The prerequisite of successful license-exempt operation in licensed bands is no harmful interference to legacy (primary) systems. This is attained respectively in three principal CR network schemes: underlay, overlay, and interweave [8].

In the underlay scheme, the interference caused by cognitive transmitters to the



(a) Spectrum holes in time.

(b) Spectrum holes in space.

Figure 2.2: Spectrum holes.

receivers of legacy systems has to be restricted to an accepted level. For the purpose of evaluating interference level, a performance metric, interference temperature, has been proposed by FCC [9]. Based on this metric, the network capacity analysis, taking account of large-scale channel fading, has shown limited performance benefits [10]. When the knowledge of small-scale fading between a cognitive transmitter and a primary receiver is utilized, a dynamic transmission power allocation can result in a significant gain [11]. However, this knowledge is generally unavailable to the underlay CR system except when the channel reciprocity can be exploited.

The full interference immunity is achieved in the overlay scheme at the cost of requiring information on codebooks and messages of the primary transmission. By making use of more available information, an advanced precoding technique can be applied to completely mitigate interference seen by the receivers of the overlay CR system or the primary system. To provide an incentive for the primary system to share its information, the cognitive transmitter can partially serve as its relay node.

The interweave scheme allows a cognitive user to access temporarily or geographically unused licensed bands. These accessible spectrum opportunities either in time or in space are known as spectrum holes as shown in Fig. 2.2. To access spectrum in this opportunistic way, the interweave CR system is required to monitor primary user activities. There are three main approaches to collecting information about the PU activities, i.e., auxiliary beacon, geolocation database, and spectrum sensing. The auxiliary beacon approach is proposed to protect wireless microphones in TV whitespaces. At the PU location, enabling/disabling beacons have to be broadcast with stronger power to inform the CR system of the spectrum occupancy status [12]. The realistic design of beacon signaling and transmission is detailed in [13, 14]. The geolocation database can provide a list of available vacant channels to a GPS-equipped CU according to the CU's location, device type, and so on [15]. This approach has proved applicable to primary systems that have static or slowly varying channel usage patterns, such as TV broadcasting. A location-specific algorithm for the calculation of permitted CU transmission power in TV whitespaces is presented in [16]. Spectrum sensing requires the CR system to blindly detect the presence of the PU transmission in the vicinity using signal processing techniques. Thus, the burden of identifying vacant channels is largely shifted to the CU side. In addition, different from the beacon approach, spectrum sensing has to perform successfully even if the signal level of the PU transmission is very low. As the PU signals might be deeply faded in wireless channels, it raises the so-called hidden node problem. One way to overcome this problem which relies on collaboration among spatially distributed CU devices is known as cooperative sensing. Due to its applicability to dynamic PU channel usage patterns and placing less burden on the PU side, spectrum sensing is still an active research topic and is regarded as a promising long-term solution.

2.2 Spectrum Sensing

Borrowing terms from Computer Science and Cognitive Science, Simon Haykin, in 2005, gave a more engineering related definition of CR with two specific goals, that is, reliable communications and efficient spectrum utilization [17]. In order to interact with the radio frequency (RF) environment, the acquisition of knowledge of the surrounding environment, such as interference level, spectrum holes, and CSI, is involved in cognitive tasks. An alternative term of white spaces, "spectrum holes", was clearly defined in the same document,

> "A spectrum hole is a band of frequencies assigned to a primary user, but, at a particular time and specific geographic location, the band is not being utilized by that user."

In licensed bands, identifying these spectrum holes is significant in terms of increasing spectrum utilization. Nevertheless, using some band in the mistaken belief that a spectrum hole is in existence is having a serious effect, i.e., causing harmful interference to PUs. In fact, finding out spectrum holes is dual to being aware of the presence of PUs. The technique for accurately determining the existence of spectrum holes or PUs on a CR device (or by a cognitive user) is referred to as spectrum sensing.

Providing reliable communications within the CR system requires sufficiently accessible spectrum. In order to achieve this, CR will probably sense over a frequency span of several gigahertz-wide. It makes the high sensitivity to weak PU signals become a harsh requirement on an RF front-end of a CR device. Therefore, spectrum sensing cannot achieve high sensitivity without the aid of digital signal processing [18]. Depending on different knowledge levels of PU signal characteristics, there are three main digital signal processing techniques for spectrum sensing: matched filter, energy detection, and cyclostationary feature detection.

It is worth the effort of clarifying the difference between spectrum sensing and spectrum estimation. The latter focuses on accurately measuring spectrum in terms of unbiasedness, consistency, and high resolution [19]. Thus, spectrum estimation also provides a CR device the knowledge of the radio scene. However, spectrum sensing in this thesis emphasizes the accuracy of detecting active PUs. Here, the knowledge of the radio scene is referred to as the existence of PUs rather than the values of the spectrum. Though these values of the spectrum might be used to form a test statistic, the more important thing is what characteristic makes a PU signal stand out from background noise or interference. This characteristic might be a unique shape of a waveform or the spectrum, or an energy level. The three spectrum sensing techniques which are going to be introduced are exactly the detection strategies given different characteristic descriptions of a PU signal.

2.2.1 Formation of Detection Problems

Detecting the possible existence of the PU transmission is basically a problem of the binary hypothesis testing. Let the noise-only case be the null hypothesis $\mathcal{H}_0$ and the

PU-present case be the alternative hypothesis $\mathcal{H}_1$. The received discrete-time signal $\{Y_n\}_{n=0}^{\infty}$ under each hypothesis is expressed as

$$\begin{aligned}\mathcal{H}_0 &: \{Y_n\}_{n=0}^{\infty} = \{Z_n\}_{n=0}^{\infty}\,, \text{ noise only,} \\ \mathcal{H}_1 &: \{Y_n\}_{n=0}^{\infty} = \{X_n + Z_n\}_{n=0}^{\infty}\,, \text{ PU present,}\end{aligned} \tag{2.1}$$

where $\{Z_n\}_{n=0}^{\infty}$ denotes additive white Gaussian noise (AWGN) samples with the sample distribution $\mathcal{N}(0, \sigma_Z^2)$ and the sample variance σ_Z^2, and $\{X_n\}_{n=0}^{\infty}$ represents the samples of the PU transmitted signal. It is assumed that no interference comes from other PUs or CUs.

Based on this hypothesis model, two terms of interest can be defined, i.e., the probability of detection P_d and the probability of false alarm P_f. The probability of detection is the odds that the hypothesis $\mathcal{H}_1$ is regarded as true when it is true. The odds that the hypothesis $\mathcal{H}_1$ is regarded as true while it is not are named as the probability of false alarm. These two probabilities play an important role in evaluating the performance of a spectrum sensing technique. For the purpose of increasing spectrum utilization, the preliminary step is to discover potentially usable frequency bands. In the case of the binary testing, it is to make the right decision when the noise-only hypothesis is true, i.e. to make P_f as low as possible. While trying to accurately identify the available frequency band, CR has to avoid causing interference to the PU. Such disturbance could occur when CR chooses the null hypothesis in the presence of the PU and then decides to use that frequency band. Therefore, to maintain a reasonable high probability of detection P_d is also necessary.

It has been recognized that increasing P_d and lowering P_f are two conflicting goals for given observations $\mathbf{y} = [Y_1, Y_2, \ldots, Y_N]'$ of the fixed length N, where $'$ denotes the transpose. The next-best thing that can be achieved is maximizing P_d under the constraint of $P_f = \alpha$. This optimum criterion is known as Neyman-Pearson approach [20]. This criterion decides $\mathcal{H}_1$ if the likelihood ratio $\Lambda(\mathbf{y})$ is no less than a threshold λ, i.e.

$$\Lambda(\mathbf{y}) = \frac{f_{\mathbf{y}|\mathcal{H}_1}(\mathbf{y}|\mathcal{H}_1)}{f_{\mathbf{y}|\mathcal{H}_0}(\mathbf{y}|\mathcal{H}_0)} \geq \lambda, \tag{2.2}$$

where $f_{\mathbf{y}|\mathcal{H}_1}(\mathbf{y}|\mathcal{H}_1)$ stands for the joint distribution of the vector $\mathbf{y}$ under $\mathcal{H}_1$, and the threshold λ is derived from

$$P_f = \int_{\{\mathbf{y}:\Lambda(\mathbf{y})\geq\lambda\}} f_{\mathbf{y}|\mathcal{H}_0}(\mathbf{y}|\mathcal{H}_0)\, d\mathbf{y} = \alpha. \tag{2.3}$$

2.2.2 Matched Filter Detection

The matched filter is an optimum coherent detection method in terms of maximizing the SNR. This technique requires the knowledge of the PU signal at a CU. In fact there exist pilots, training symbols, or spreading codes in PU signals that can be exploited by the CU. For example, in the standard IEEE 802.16-2009, one or two predefined OFDM symbols are used for the initial acquisition. Because of prior knowledge, the CU is enabled to distinguish the PU's signal from interference and noise. Another advantage of the matched filter is its shorter time, due to requiring less observations, to satisfy the desired probability of false alarm than other schemes [18].

To do coherent detection involves the timing synchronization or channel estimation. In other words, the performance of matched filter is subject to imperfect synchronization [21]. Moreover, for sensing intended PUs of different systems, there should be different dedicated algorithms [22]. Hence the high implementation complexity is inevitable when the matched filter detection is applied to multiple-PU sensing.

In matched filter detection, the transmitted PU signal vector $\mathbf{x}=[X_1, X_2, \ldots, X_N]'$ with energy $E_{\mathbf{x}} \triangleq \mathbf{x}'\mathbf{x}$ is regarded as deterministic and known to the CU. The test statistic derived from the likelihood ratio is $\mathcal{T}(\mathbf{y}) \triangleq \mathbf{y}'\mathbf{x}$ [20]. The distribution of $\mathcal{T}(\mathbf{y})$ is given by

$$
\begin{aligned}
&\mathcal{H}_0 : \mathcal{T}(\mathbf{y}) \sim \mathcal{N}(0, \sigma_Z^2 E_{\mathbf{x}}), \\
&\mathcal{H}_1 : \mathcal{T}(\mathbf{y}) \sim \mathcal{N}(E_{\mathbf{x}}, \sigma_Z^2 E_{\mathbf{x}}).
\end{aligned} \tag{2.4}
$$

Given a false alarm probability constraint, the test threshold can be computed as $\lambda = \sqrt{\sigma_Z^2 E_{\mathbf{x}}} Q^{-1}(P_f)$, where $Q^{-1}(\cdot)$ denotes the inverse Q-function. Hence the detection probability P_d of this threshold is

$$
P_d = Q\left(\frac{\lambda - E_{\mathbf{x}}}{\sqrt{\sigma_Z^2 E_{\mathbf{x}}}}\right) = Q\left[Q^{-1}(P_f) - \sqrt{\mathrm{SNR}}\right], \tag{2.5}
$$

where $\mathrm{SNR} \triangleq E_{\mathbf{x}}/\sigma_Z^2$.

In Rayleigh fading channel, the received signal is redefined to be $Y_n = |h| X_n + Z_n$, where the channel gain $|h|$ is Rayleigh distributed $f_{ray}(x; \sigma_{ray}) = \frac{x}{\sigma_{ray}^2} \exp\left(\frac{-x^2}{2\sigma_{ray}^2}\right)$. The observation time of $\mathbf{y}$ is assumed to be less than the coherence time of the channel, so the channel gain does not vary significantly during this period. Therefore the distribution of the test statistic under the alternative hypothesis can be recast to be

$$\mathcal{H}_1 : \mathcal{T}\left(\mathbf{y}\right) \sim \mathcal{N}(\left|h\right| E_{\mathbf{x}}, \sigma_Z^2 E_{\mathbf{x}}). \tag{2.6}$$

With the same test threshold as the previous one for the same P_f, the probability of detection conditioned on the channel gain is given by

$$P_d\left(\left|h\right|\right) = Q\left(\frac{\lambda - \left|h\right| E_{\mathbf{x}}}{\sqrt{\sigma_Z^2 E_{\mathbf{x}}}}\right) = Q\left[Q^{-1}\left(P_f\right) - \left|h\right|\sqrt{\frac{E_{\mathbf{x}}}{\sigma_Z^2}}\right], \tag{2.7}$$

and then the average probability of detection can be expressed as

$$\bar{P}_d = \int_0^\infty P_d\left(\left|h\right|\right) f_{ray}\left(\left|h\right|; \sigma_{ray}\right) \,\mathrm{d}\left|h\right|. \tag{2.8}$$

Figure 2.3 presents the detection performance of the matched filter in an AWGN channel. The sample size is $N = 25$. The solid lines indicate the analytical results which are in agreement with the simulated results denoted by different marks. It can be seen that lowering the order of the desired P_f does significantly decrease the detection probability. The detection performance in a Rayleigh fading channel is shown in Fig. 2.4. The X-axis coordinate in this figure is the average SNR, ${}^{2\sigma_{ray}^2 E_{\mathbf{x}}}/{\sigma_Z^2}$. For $\sigma_{ray} = {}^1/\sqrt{2}$, the results show that the detection performance is seriously degraded by fading effects at higher SNR.

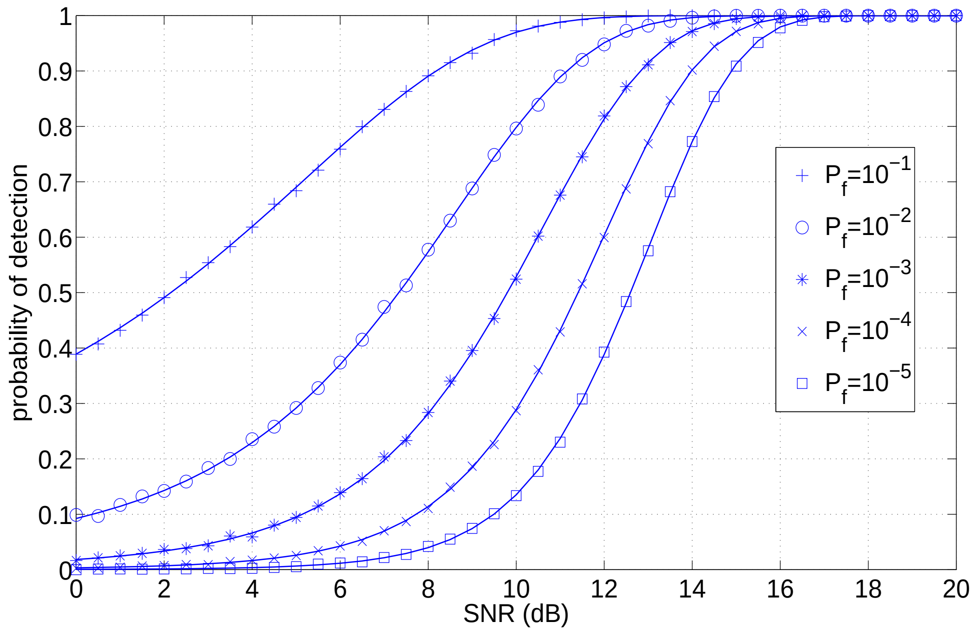


Figure 2.3: Matched filter detection in an AWGN channel.

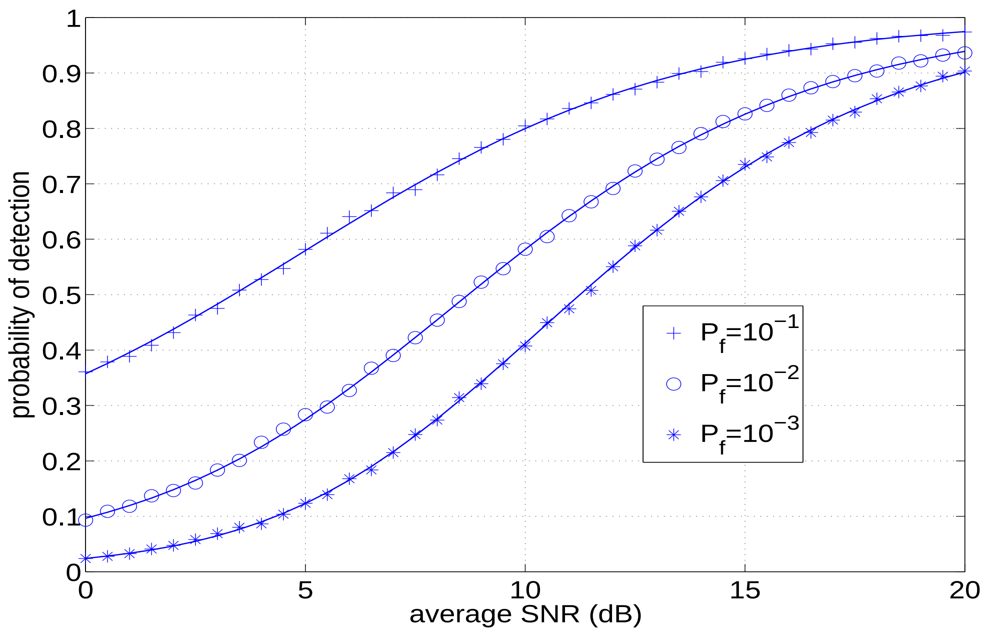


Figure 2.4: Matched filter detection in a Rayleigh fading channel.

2.2.3 Energy Detection

Different from the matched filter detection, energy detection is a noncoherent approach to spectrum sensing and in turn introduces less implementation complexity [18]. This approach just measures the collected energy of the received signal and compares it to the test threshold. Without ties with prior knowledge of the PU's signal, energy detection is more applicable to wideband spectrum sensing [21].

Although measuring energy is straightforward, energy detection does not work satisfactorily at the low SNR region [22], where the sensing ability of a CU should reach. For instance, spread spectrum signals with very low SNR per chip cannot be efficiently detected by energy measuring [18, 22]. Such shortcoming can be solved by more sophisticated spectrum sensing techniques or the cooperation with other CUs. However, it is still hard to distinguish the PU's signal from interference of the energy level which is higher than the noise floor [18, 22]. This is a consequence of having no knowledge of the PU's signal. Furthermore, in energy detection, the test threshold determined by the noise level is supposed to be known and is constant. In practice, there is a variation of noise floor, causing the poor performance [18, 21, 22]. To overcome this problem requires a reliable algorithm to monitor the noise level constantly.

Let's review this technique first and analyze its vulnerability to the noise variation in the next section. As its name suggests, energy detection uses the normalized energy of the received signal $\mathcal{T}(\mathbf{y}) = \mathbf{y}'\mathbf{y}/\sigma_Z^2$ as the test statistic whose probability density function (PDF) is given by [20]

$$\mathcal{H}_0 : \mathcal{T}(\mathbf{y}) \sim \chi_N^2,$$
$$\mathcal{H}_1 : \mathcal{T}(\mathbf{y}) \sim \chi_N^2(E_{\mathbf{x}}/\sigma_Z^2), \tag{2.9}$$

where χ_N^2 represents the chi-square distribution of N degrees of freedom and its cumulative distribution function is denoted by $F_{\chi_N^2}(\cdot)$, and $\chi_N^2(\gamma)$ denotes the noncentral chi-square distribution of N degrees of freedom with noncentrality γ and the corresponding cumulative distribution function is $F_{\chi_N^2(\gamma)}(\cdot)$. The test threshold for the given P_f is

$$\lambda = F_{\chi_N^2}^{-1}(1 - P_f), \tag{2.10}$$

where $F_{\chi_N^2}^{-1}(\cdot)$ is the inverse function of cumulative distribution of χ_N^2. The probability of detection can be computed as

$$P_d = 1 - F_{\chi_N^2(E_{\mathbf{x}}/\sigma_Z^2)}(\lambda) = Q_{N/2}\left(\sqrt{E_{\mathbf{x}}/\sigma_Z^2}, \sqrt{\lambda}\right), \tag{2.11}$$

where $Q_M(a, b)$ is the generalized Marcum Q-function.

In the case of a Rayleigh fading channel, the distribution of the test statistic under the alternative hypothesis is now given by

$$\mathcal{H}_1 : \mathcal{T}(\mathbf{y}) \sim \chi_N^2\left(|h|^2 E_{\mathbf{x}}/\sigma_Z^2\right). \tag{2.12}$$

The probability of detection conditioned on the channel gain is

$$P_d(|h|) = 1 - F_{\chi_N^2\left(|h|^2 E_{\mathbf{x}}/\sigma_Z^2\right)} = Q_{N/2}\left(\sqrt{|h|^2 E_{\mathbf{x}}/\sigma_Z^2}, \sqrt{\lambda}\right). \tag{2.13}$$

With this in mind, the average probability of detection is given by

$$\begin{aligned} \bar{P}_d &= \int_0^\infty P_d(|h|)\, f_{ray}\left(|h|; \sigma_{ray} = 1/\sqrt{2}\right) \mathrm{d}\,|h|, \\ &= \int_0^\infty Q_{N/2}\left(\sqrt{|h|^2 E_{\mathbf{x}}/\sigma_Z^2}, \sqrt{\lambda}\right) 2\,|h| \exp\left(-|h|^2\right) \mathrm{d}\,|h|, \end{aligned} \tag{2.14}$$

which can be computed through the formula [23]

$$\begin{aligned} &\int_0^\infty |h| \exp\left(\frac{-p^2\,|h|^2}{2}\right) Q_M\left(a\,|h|, b\right) \mathrm{d}\,|h|, \\ &= \frac{1}{p^2} \exp\left(\frac{-b^2}{2}\right) \left\{ \left(\frac{p^2 + a^2}{a^2}\right)^{M+1} \left[\exp\left(\frac{a^2 b^2}{2p^2 + 2a^2}\right) - \sum_{n=0}^{M-2} \frac{1}{n!} \left(\frac{a^2 b^2}{2p^2 + 2a^2}\right)^n\right] \right. \\ &\qquad \left. + \sum_{n=0}^{M-2} \frac{1}{n!} \left(\frac{b^2}{2}\right)^n \right\}. \end{aligned} \tag{2.15}$$

Figures 2.5 and 2.6 show the detection performance of energy detection in an AWGN channel and a Rayleigh fading channel respectively. The simulation parameters are the same as those in the previous section. As shown in these two figures, the analytical results (solid lines) are consistent with the simulated results (marks). Also, similar to what has been observed in the matched filter detection, the channel fading makes it harder to improve detection probability by increasing SNR.

Effect of Noise Uncertainty

As mentioned in the previous section, energy detection is subject to the variation of noise floor, also known as noise uncertainty. The noise uncertainty can be viewed

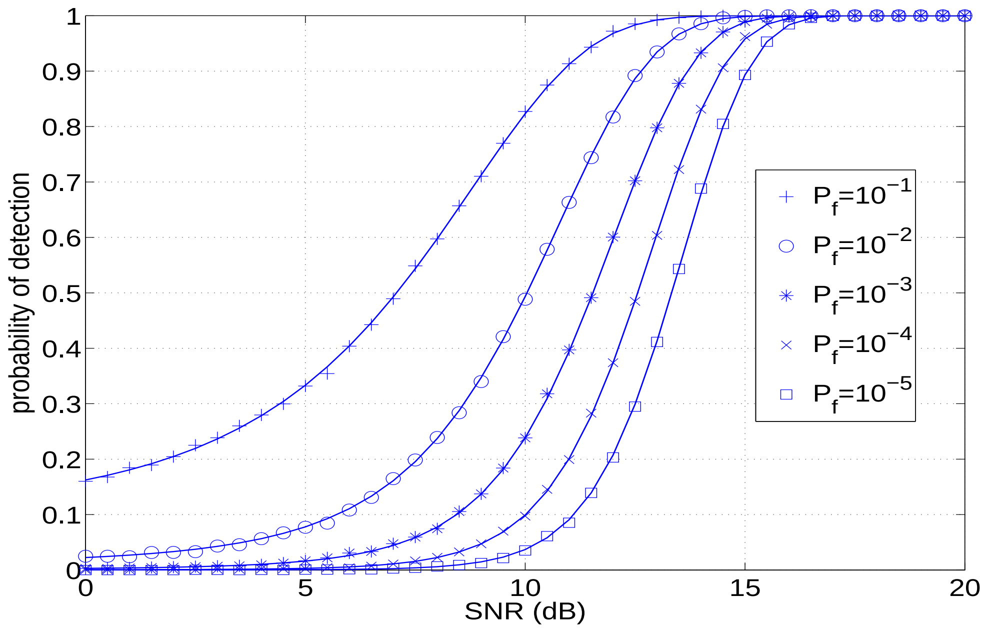


Figure 2.5: Energy detection in an AWGN channel.

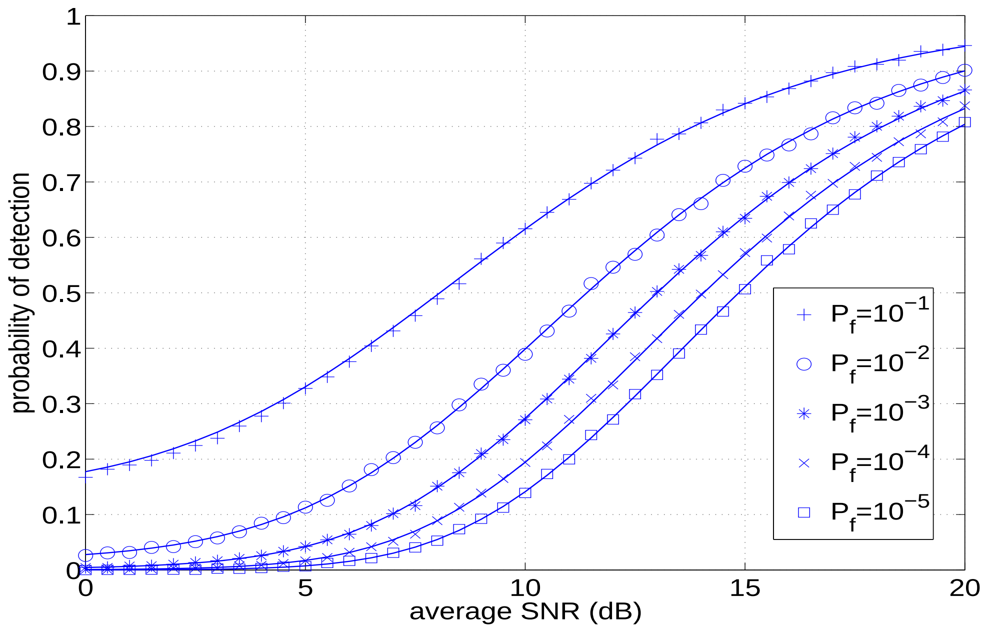


Figure 2.6: Energy detection in a Rayleigh fading channel.

as a confident interval of the estimate of noise power. There are several sources of noise uncertainty, such as calibration errors, changeable thermal noise, a time-varying low noise amplify (LNA) gain [24]. All of them contribute to non-ergodic characteristics of noise, which cannot be described by a stationary noise model. So, it is reasonable to assume that there is always some uncertainty level no matter how accurate the estimation technique is. In [24, 25], a rough range of noise uncertainty at a receiving end is shown to be ±1dB. When the SNR ratio is sufficiently high, the noise uncertainty does not cause problems. However, when it decreases to below some level, a performance bound will be created due to this noticeable noise uncertainty. The derivation of this performance bound can be found in [26, 27]. Here we derive this performance bound in our setting, following the argument in [26].

Under the assumption of noise uncertainty, the test statistic becomes $\mathcal{T}(\mathbf{y}) = \mathbf{y}'\mathbf{y}/\hat{\sigma}_Z^2$, where the variance $\hat{\sigma}_Z^2$ is the estimation result of actual noise power σ_Z^2. The real noise power might be time-varying and ranges between the interval $\left[\frac{1}{\rho}\hat{\sigma}_Z^2, \rho\hat{\sigma}_Z^2\right]$ in which ρ is called as the uncertainty gain. This uncertainty gain is assumed to be greater than one. Because of the time-varying characteristic, the underlying assumption of Gaussian distribution might not hold anymore. In spite of that, if the statistical deviation is relatively small and the uncertainty gain is not too large, the probability density function of the test statistic will keep the same form, that is,

$$\begin{aligned}
&\mathcal{H}_0 : \mathcal{T}(\mathbf{y}) \sim \chi_N^2, \\
&\mathcal{H}_1 : \mathcal{T}(\mathbf{y}) \sim \chi_N^2(NP_{\mathbf{x}}/\sigma_Z^2),
\end{aligned} \tag{2.16}$$

where $P_{\mathbf{x}} \triangleq {}^{E_{\mathbf{x}}}/_{N}$. If the sample size N is large enough, then by central limit theory, the distribution of the normalized test statistic $\tilde{\mathcal{T}}(\mathbf{y}) \triangleq (\hat{\sigma}_Z^2/N)\,\mathcal{T}(\mathbf{y})$ can be approximated by Gaussian distribution, i.e.,

$$\begin{aligned}
&\mathcal{H}_0 : \tilde{\mathcal{T}}(\mathbf{y}) \sim \mathcal{N}\left(\hat{\sigma}_Z^2, \frac{2\hat{\sigma}_Z^4}{N}\right), \\
&\mathcal{H}_1 : \tilde{\mathcal{T}}(\mathbf{y}) \sim \mathcal{N}\left(\hat{\sigma}_Z^2 + P_{\mathbf{x}}, \frac{2\hat{\sigma}_Z^4}{N} + \frac{4P_{\mathbf{x}}\hat{\sigma}_Z^2}{N}\right).
\end{aligned} \tag{2.17}$$

The probability of false alarm, as a function of the estimated noise power, is given by

$$P_f\left(\hat{\sigma}_Z^2\right) = Q\left(\frac{(\lambda/\hat{\sigma}_Z^2)}{\sqrt{2/N}}\right). \tag{2.18}$$

The Q-function is monotonically decreasing, so the possible false alarm probability will be in the range $[P_f(\hat{\sigma}_Z^2/\rho), P_f(\rho\hat{\sigma}_Z^2)]$. The robust setting of the test threshold, in terms of achieving the target probability of false alarm $P_f^{(\star)}$ under noise uncertainty, is to have $P_f(\rho\hat{\sigma}_Z^2) = P_f^{(\star)}$. It follows that

$$\lambda = \rho\hat{\sigma}_Z^2\left[1 + \sqrt{\frac{2}{N}}Q^{-1}\left(P_f^{(\star)}\right)\right]. \tag{2.19}$$

Similarly, the probability of detection will range between $[P_d(\hat{\sigma}_Z^2/\rho), P_d(\rho\hat{\sigma}_Z^2)]$. To achieve the target probability of detection $P_d^{(\star)}$ robustly, $P_d(\hat{\sigma}_Z^2/\rho)$ has to be made equal to $P_d^{(\star)}$, i.e.,

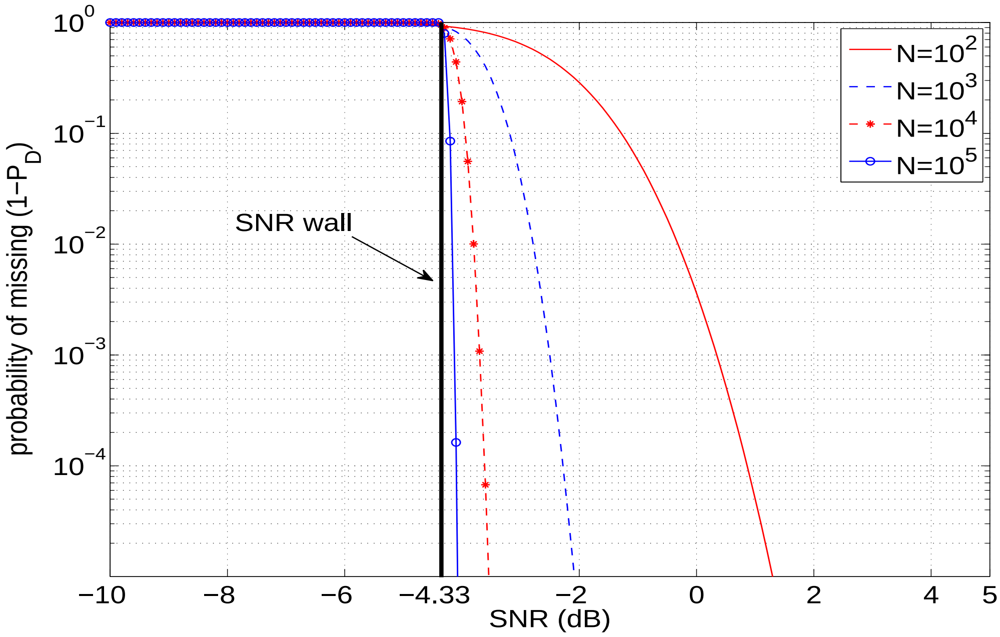


Figure 2.7: An illustrative simulation of SNR wall phenomenon.

$$\begin{aligned} P_d^{(\star)} &= Q\left[\frac{\lambda - \left(\frac{\hat{\sigma}_Z^2}{\rho} + P_{\mathbf{x}}\right)}{\sqrt{\left(\frac{2\hat{\sigma}_Z^4}{\rho^2} + \frac{4P_{\mathbf{x}}\hat{\sigma}_Z^2}{\rho}\right)/N}}\right], \\ &= Q\left[\frac{\rho^2 Q^{-1}\left(P_f^{(\star)}\right) + \sqrt{N/2}\,(\rho^2 - 1 - \rho^2\mathrm{SNR})}{\sqrt{1+2\rho\mathrm{SNR}}}\right]. \end{aligned} \tag{2.20}$$

Normally, any given target detection probability $P_d^{(\star)}$ can be reached by increasing the sampling number N. However, this requires the term $(\rho^2 - 1 - \rho^2\mathrm{SNR})$ to be less than zero, which implies that the Q-function can increase monotonically with the increasing number N. Thus, the minimum required SNR can be calculated and denoted as the SNR wall, i.e.,

$$\mathrm{SNR}_{\mathrm{wall}} = \frac{\rho^2 - 1}{\rho^2}. \tag{2.21}$$

If the value of SNR is less than $\mathrm{SNR}_{\mathrm{wall}}$, there will be some target detection probabilities, which cannot be achieved no matter how large is the sampling number N. The value of N has been assumed to be sufficiently large. Hence, the term $\sqrt{N/2}\,(\rho^2 - 1 - \rho^2\mathrm{SNR})$ becomes dominant and makes the detection probability approaching zero when increasing N. Figure 2.7 is a plot of (2.20) over different sample sizes, demonstrating the phenomenon of the SNR wall. The simulation parameters respectively are: $P_f = 0.1$, $\rho = 1.26$ and $\mathrm{SNR}_{\text{wall}} = -4.33$dB. As can been seen in this figure, the probability of missing $(1 - P_d)$ cannot be reduced further by increasing the sample size if the value of SNR is below $\mathrm{SNR}_{\text{wall}}$. In other words, making the sample size larger will not improve the detection performance once the operational SNR region is not beyond the SNR wall.

2.2.4 CFD

Most of modulated signals in communications, such as analog TV signals, AM and FM signals, and digital modulated signals, can be modeled as cyclostationary processes [28].

A zero-mean complex-valued discrete-time random process $\{X_n\}_{n=0}^{\infty}$ is said to be second-order almost-cyclostationary in the wide sense [29, 30] if its autocorrelation function $R_X[n;\tau] \triangleq \mathbb{E}[X_n X_{n+\tau}^*]$ for a given time lag τ exhibits periodicity for a collection of periods $\mathcal{T} = \{T_0, T_1, \ldots\}$, i.e.,

$$R_X[n;\tau] = R_X[n+T;\tau]\,,\; T \in \mathcal{T}, \tag{2.22}$$

where $\mathbb{E}$ is expectation and $*$ denotes complex conjugate. This periodic autocorrelation function accepts the Fourier-series expansion given by

$$R_X[n;\tau] = \sum_{\alpha\in\mathcal{A}} R_X^{\alpha}[\tau]\, e^{i\alpha n}, \tag{2.23}$$

where the set $\mathcal{A} = \{\alpha \in (-\pi, \pi];\ \alpha = 2\pi k/T,\ k \in \mathbb{Z}\}$ is a collection of cycle frequencies. The Fourier coefficients $R_X^{\alpha}[\tau]$ are known as the cyclic autocorrelation function (CAF) which can be evaluated by

$$R_X^{\alpha}[\tau] = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} R_X[n;\tau]\, e^{-i\alpha n}. \tag{2.24}$$

Several CAFs of standardized PU signals are provided in [31]. A counterpart of the CAF, the conjugate CAF, can be defined similarly. In this chapter, however, we confine the discussion to the CAF. A process is claimed to exhibit a 2^{nd}-order cyclostationary feature (or cyclostationarity) if its CAF $R_X^{\alpha}[\tau]$ is non-zero for some cycle frequency α and some lag τ. The CAFs of an orthogonal frequency division multiplexing (OFDM) signal and a linear modulated signal are respectively presented in Fig. 2.8 and Fig. 2.9. The patterns of these features depend on the modulation types, coding schemes, and so on. By utilizing these unique features, CFD is claimed to be robust to varying noise and able to distinguish different modulated signals with the same power spectrum densities [32].

To obtain cyclostationary features requires a sampling rate higher than Nyquist rate and a sufficient sensing time to achieve reliable estimates [33]. However, by exploiting the sparsity of cyclostationary features, the sub-Nyquist-sampling approach to performing CFD becomes feasible [33–35]. Moreover, a sequential CFD framework

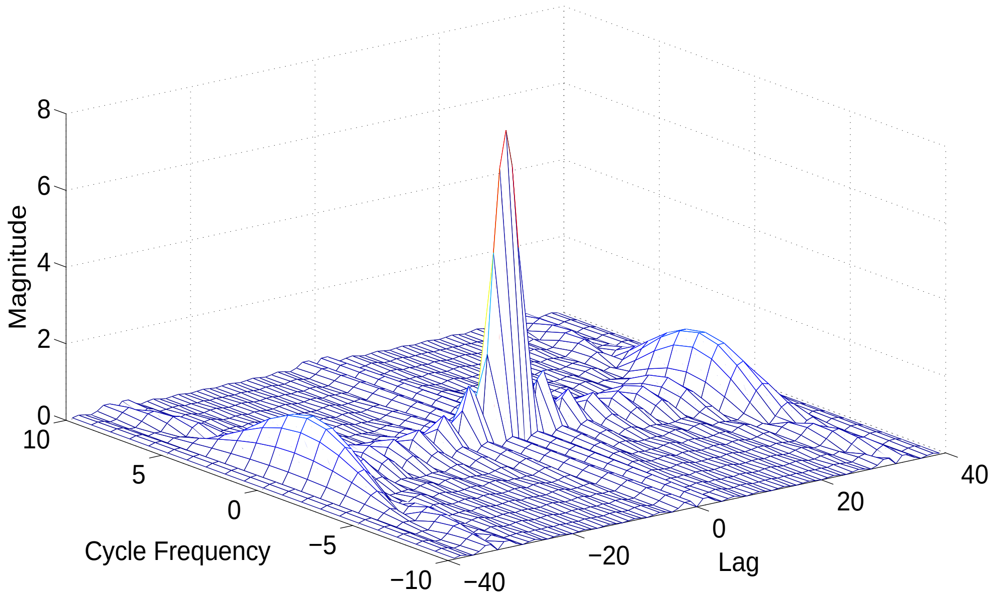


Figure 2.8: CAF of an OFDM signal.

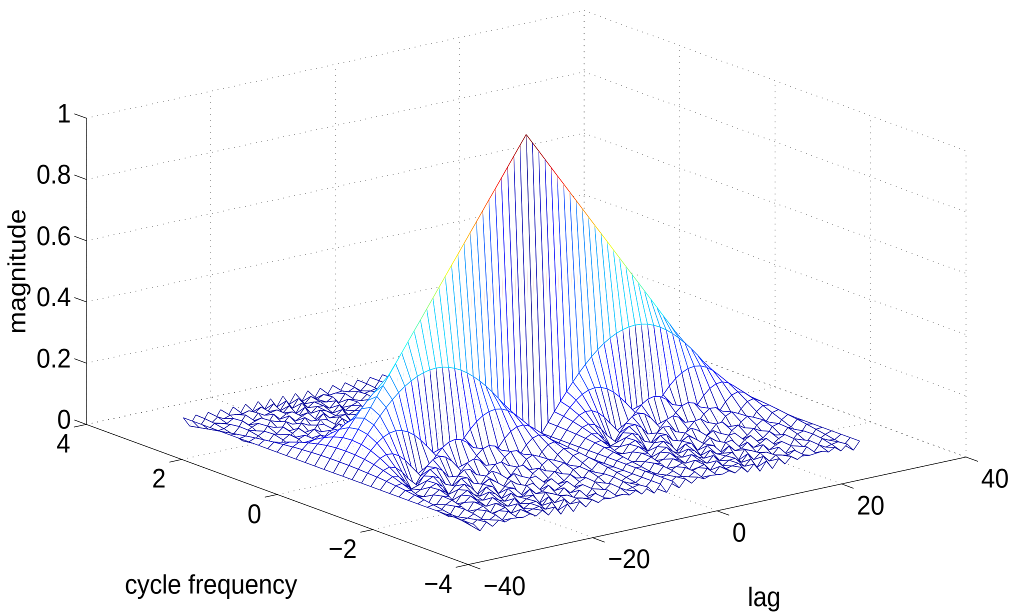


Figure 2.9: CAF of an linear modulated signal.

for average sensing time reduction has been proposed in [36].

2.3 Conclusions

In this chapter, we have introduced the concept of CR and its potential applications. As a remedy for spectrum scarcity, CR is required to identify accessible licensed bands while maintaining acceptable interference to licensed users. The issue of how to manage interference in three main CR network schemes has been briefly discussed. Then the focus shifts to spectrum sensing, a long-term solution to monitoring dynamic primary user activities in the interweave CR network.

Three major digital signal processing techniques for spectrum sensing have also been respectively examined. CFD stands out from the other two techniques, matched filter detection and energy detection. It is because that CFD is more robust to noise uncertainty compared to energy detection [37], and requires less prior knowledge of the PU signals compared to matched filter detection in which exact knowledge of PU signals is expected. The next chapter will continue to explore the potential usage of CFD.

CHAPTER 3

CYCLOSTATIONARY SPECTRUM SENSING

3.1 Introduction

Cyclostationary features appear in the common communication signals, such as OFDM signals and linear modulated signals. Based on these features, the binary hypothesis testing can be formed to determine the presence of PUs or otherwise. In the cyclic spectrum, the points selected for statistical testing are normally those of the maximum magnitude and usually the probability distribution of the formed test statistic is available under the null hypothesis [38, Eq.(10.26)] [39, Eq.(33)]. Due to lack of statistical descriptions under the alternative hypothesis, the likelihood-ratio test cannot be performed, thus the best detection performance for a given false alarm rate (FAR) is not guaranteed by using these test points. An alternative way to choose test points is provided in [40]. This however relies on extensive computer simulations to determine the test points that maximize the performance metric. In [29], Dandawate developed a generalized likelihood-ratio test (GLRT) by using CAFs. CAF involves two dimensions, the cycle-frequency domain and the lag domain. As this test was initiated for blindly detecting the existence of cycle frequencies, the test points are those in the lag domain for a given test cycle frequency. The criterion for selecting the test points (or equivalently the test multiple lags) is to use a single or two lags of maximum absolute cyclic autocorrelation [41, 42]. Methods for utilizing the cycle-frequency domain are proposed in [42, 43]. Basically, for a given set of lags, the test statistics, formed over multiple cycle frequencies, are combined by using one of three typical schemes, i.e., selection combining (SC), equal gain combing (EGC), and maximum ratio combining (MRC). As the detection performance can only be evaluated by simulation, the combination of the test lag set and the combining scheme that leads to the optimal detection performance is not analytically investigated. Another strategy for linearly combining test statistics over multiple cycle frequencies in some optimal sense is investigated in [44]. In this chapter, this category of statistical testing is referred to as multi-cycle-frequency detection.

3.2 CFD in Cycle-Frequency-Lag Domain

This section aims to provide an optimal scheme to select the best test points in the cycle-frequency-lag (CFL) domain. The optimality claimed is achieved when the asymptotic detection performance for a fixed FAR is maximized. Though the asymptotic optimality in the Neyman-Pearson sense is not guaranteed, the asymptotic detection performance for a fixed FAR provides a sensible benchmark for choosing test points in CAF. A similar approach can be found in [41,45], in which it was restricted to choosing a single test point. To reach our goal, we first generalize the one-dimension 2nd-order statistical testing model proposed by Dandawate to the CFL domain. Based on the derived asymptotic detection probability, we identify the required optimal test points (or the corresponding multiple cycle frequencies and lags) by an exhaustive search. Knowledge of the 4th-order cyclic cumulant of a PU signal, which is required in this optimal scheme, is not normally available a priori. Therefore, an alternative scheme that does not require knowing this cyclic cumulant is proposed. This scheme will be shown to lead to comparable detection performance achieved using the optimal scheme in the low signal-to-noise ratio region.

3.2.1 System Model

Consider that the PU signal is a zero-mean second-order almost-cyclostationary random process $\{X_n\}_{n=0}^{\infty}$. Based on prior knowledge of the CAFs of PU signals, we are able to formulate the statistical test for the presence of cyclostationarity over multiple cycle frequencies and lags. Let the received signal $\{Y_n\}_{n=0}^{\infty}$ under the noise-only and PU-present hypotheses be presented as

$$\begin{aligned} &\mathcal{H}_0 : \{Y_n\}_{n=0}^{\infty} = \{Z_n\}_{n=0}^{\infty}\,, \text{ noise-only}, \\ &\mathcal{H}_1 : \{Y_n\}_{n=0}^{\infty} = \{X_n + Z_n\}_{n=0}^{\infty}\,, \text{ PU-present}, \end{aligned} \tag{3.1}$$

where the PU signal $\{X_n\}_{n=0}^{\infty}$ is a zero-mean complex-valued second-order almost-cyclostationary process with variance σ_X^2, $\{Z_n\}_{n=0}^{\infty}$ is a circularly-symmetric white Gaussian process with variance σ_Z^2, and these two processes are assumed uncorrelated. The SNR is defined as ${}^{\sigma_X^2}/_{\sigma_Z^2}$. Given the CAF $R_X^{\alpha}[\tau]$ of the PU signal, M points in the CFL domain, chosen for statistical testing, are denoted by the test set $\mathcal{J} = \{(\alpha_1, \tau_1), \ldots, (\alpha_M, \tau_M)\}$. The coordinate of the i^{th} point is (α_i, τ_i) and its corresponding cyclic autocorrelation $R_X^{\alpha_i}[\tau_i]$ is assumed to be non-zero. Then a CAF-estimation vector can be established for testing the presence of cyclostationarity over these parameters based on N observations, that is,

$$\hat{\mathbf{r}}_Y^{\mathcal{J}}[N] \triangleq \begin{bmatrix} \hat{\mathbf{u}}_Y^{\mathcal{J}}[N] \\ \hat{\mathbf{v}}_Y^{\mathcal{J}}[N] \end{bmatrix}, \tag{3.2}$$

where

$$\hat{\mathbf{u}}_Y^{\mathcal{J}}[N] \triangleq \left[\mathrm{Re}\left\{\hat{R}_Y^{\alpha_1}[\tau_1; N]\right\}, \ldots, \mathrm{Re}\left\{\hat{R}_Y^{\alpha_M}[\tau_M; N]\right\}\right]', \tag{3.3}$$

and

$$\hat{\mathbf{v}}_Y^{\mathcal{J}}[N] \triangleq \left[\mathrm{Im}\left\{\hat{R}_Y^{\alpha_1}[\tau_1; N]\right\}, \ldots, \mathrm{Im}\left\{\hat{R}_Y^{\alpha_M}[\tau_M; N]\right\}\right]', \tag{3.4}$$

in which $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ represent the real and imaginary parts of an argument, and the superscript $'$ denotes the transpose. The consistent estimation $\hat{R}_Y^{\alpha}[\tau; N]$ of the received signal's CAF $R_Y^{\alpha}[\tau]$ is provided as

$$\hat{R}_Y^{\alpha}[\tau;N] \triangleq \frac{1}{N}\sum_{n=\max\{0,-\tau\}}^{N+\min\{0,-\tau\}-1} Y_n Y_{n+\tau}^* e^{-i\alpha n},$$
$$= R_Y^{\alpha}[\tau] + \varepsilon_{YY}^{\alpha}[\tau;N], \tag{3.5}$$

where $\varepsilon_{YY}^{\alpha}[\tau;N]$ denotes the estimation error. The notion of N will be omitted from now on for simplicity. Asymptotic distributions of the CAF-estimation vector $\hat{\mathbf{r}}_Y^{\mathcal{J}}$ under two hypotheses can be shown to be

$$\mathcal{H}_0 : \lim_{N\to\infty} \sqrt{N}\hat{\mathbf{r}}_{\mathrm{Y}}^{\mathcal{J}} \stackrel{\mathrm{D}}{=} \mathcal{N}(\mathbf{0}, \Sigma_Y),$$
$$\mathcal{H}_1 : \lim_{N\to\infty} \sqrt{N}\hat{\mathbf{r}}_{\mathrm{Y}}^{\mathcal{J}} \stackrel{\mathrm{D}}{=} \mathcal{N}\left(\sqrt{N}\mathbf{r}_Y^{\mathcal{J}}, \Sigma_Y\right), \tag{3.6}$$

where the symbol $\stackrel{\mathrm{D}}{=}$ denotes "the left hand side converges in distribution to" and $\mathcal{N}$ indicates a multivariate normal distribution. The vector $\mathbf{r}_Y^{\mathcal{J}}$ and the matrix Σ_Y respectively represent the limiting vector $\mathbf{r}_Y^{\mathcal{J}} = \lim_{N\to\infty}\mathbb{E}[\hat{\mathbf{r}}_Y^{\mathcal{J}}]$ and the limiting autocovariance matrix $\Sigma_Y = \lim_{N\to\infty}\mathrm{Cov}(\sqrt{N}\hat{\mathbf{r}}_Y^{\mathcal{J}}, \sqrt{N}\hat{\mathbf{r}}_Y^{\mathcal{J}})$, in which $\mathrm{Cov}(\mathbf{a},\mathbf{b}) \triangleq \mathbb{E}[(\mathbf{a}-\mathbb{E}[\mathbf{a}])(\mathbf{b}-\mathbb{E}[\mathbf{b}])^H]$ and the symbol "H" denotes Hermitian transpose. An alternative expression of the matrix Σ_Y, which will be used in Section 3.3, is provided here. Let's define the complex vector $\hat{\mathbf{w}}_Y^{\mathcal{J}} = \hat{\mathbf{u}}_Y^{\mathcal{J}} + i\hat{\mathbf{v}}_Y^{\mathcal{J}}$, the covariance matrix $\mathbf{C}_{\hat{\mathbf{w}}} = \lim_{N\to\infty}\mathrm{Cov}(\sqrt{N}\hat{\mathbf{w}}_Y^{\mathcal{J}}, \sqrt{N}\hat{\mathbf{w}}_Y^{\mathcal{J}})$, and the complementary covariance matrix $\bar{\mathbf{C}}_{\hat{\mathbf{w}}} = \lim_{N\to\infty}\mathrm{Cov}(\sqrt{N}\hat{\mathbf{w}}_Y^{\mathcal{J}}, \sqrt{N}(\hat{\mathbf{w}}_Y^{\mathcal{J}})^H)$. By making use of [46, Eq.(2.21-22)], the matrix Σ_Y can be recast as

$$\Sigma_Y = \begin{bmatrix} \frac{\mathrm{Re}(\mathbf{C}_{\hat{\mathbf{w}}}+\bar{\mathbf{C}}_{\hat{\mathbf{w}}})}{2} & \frac{\mathrm{Im}(-\mathbf{C}_{\hat{\mathbf{w}}}+\bar{\mathbf{C}}_{\hat{\mathbf{w}}})}{2} \\ \frac{\mathrm{Im}(\mathbf{C}_{\hat{\mathbf{w}}}+\bar{\mathbf{C}}_{\hat{\mathbf{w}}})}{2} & \frac{\mathrm{Re}(\mathbf{C}_{\hat{\mathbf{w}}}-\bar{\mathbf{C}}_{\hat{\mathbf{w}}})}{2} \end{bmatrix}. \tag{3.7}$$

The (i,j)th entries of $\mathbf{C}_{\hat{\mathbf{w}}}$ and $\bar{\mathbf{C}}_{\hat{\mathbf{w}}}$ are given respectively by

$$c_{i,j} = \lim_{N\to\infty} N\mathrm{Cov}(\hat{R}_Y^{\alpha_i}[\tau_i], \hat{R}_Y^{\alpha_j}[\tau_j]), \tag{3.8}$$

and

$$\bar{c}_{i,j} = \lim_{N\to\infty} N\mathrm{Cov}(\hat{R}_Y^{\alpha_i}[\tau_i], \hat{R}_Y^{\alpha_j *}[\tau_j]). \tag{3.9}$$

Due to (3.6), a GLRT can be formed, giving the final test statistic $\Lambda = N(\hat{\mathbf{r}}_Y^{\mathcal{J}})'\hat{\Sigma}_Y^{-1}\hat{\mathbf{r}}_Y^{\mathcal{J}}$ where $\hat{\Sigma}_Y^{-1}$ signifies a mean-square sense consistent estimation of the matrix Σ_Y^{-1}. The asymptotic probability density functions of this test statistic under two hypotheses are provided respectively,

$$\begin{aligned}
\mathcal{H}_0 &: \lim_{N\to\infty} \Lambda \overset{\mathrm{D}}{=} \chi^2_{2M}, \\
\mathcal{H}_1 &: \lim_{N\to\infty} \Lambda \overset{\mathrm{D}}{=} \chi^2_{2M}\left(N\left(\mathbf{r}_Y^{\mathcal{J}}\right)'\Sigma_Y^{-1}\mathbf{r}_Y^{\mathcal{J}}\right).
\end{aligned} \tag{3.10}$$

For a constant false alarm rate P_f, the asymptotic detection probability can be shown to be

$$P_d = 1 - F_{\chi^2_{2M}\left(N\left(\mathbf{r}_Y^{\mathcal{J}}\right)'\Sigma_Y^{-1}\mathbf{r}_Y^{\mathcal{J}}\right)}(\lambda), \tag{3.11}$$

where the detection threshold λ is chosen such that $P_f = 1 - F_{\chi^2_{2M}}(\lambda)$. The asymptotic normality exploited in (3.6) is a natural extension of the results in [29] when testing over multiple cycle frequencies. The explicit method for obtaining $\hat{\Sigma}_Y^{-1}$ is detailed in [29, 47]. The convergence in distribution under $\mathcal{H}_1$ in (3.10) is proved in Section 3.A. We will refer to this system model as joint-utilization detection.

3.2.2 Selecting Test Points for Joint-Utilization Detection

As the asymptotic detection performance in (3.11) is available, the asymptotically optimal set of test points for hypothesis testing might be chosen for any given false alarm rate. Unfortunately, it has been shown that, generally, no global optimum exists in terms of providing the best asymptotic detection performance over all SNR regions for a constant FAR [48]. In that analysis an alternative asymptotic distribution under $\mathcal{H}_1$ is used, i.e., when $N\left(\mathbf{r}_Y^{\mathcal{J}}\right)' \Sigma_Y^{-1}\mathbf{r}_Y^{\mathcal{J}} \gg M$,

$$\mathcal{H}_1 : \lim_{N\to\infty} \Lambda \sim \mathcal{N}\left(N\left(\mathbf{r}_Y^{\mathcal{J}}\right)' \Sigma_Y^{-1}\mathbf{r}_Y^{\mathcal{J}}, 4N\left(\mathbf{r}_Y^{\mathcal{J}}\right)' \Sigma_Y^{-1}\mathbf{r}_Y^{\mathcal{J}}\right), \tag{3.12}$$

where $\alpha_1 = \cdots = \alpha_M$ in $\mathcal{J}$. Although this distribution is conditionally true, it is sufficient to prove the absence of the global optimum when the more accurate asymptotic distribution in (3.10) is considered. As a result, a suitable lag set for testing should be a locally asymptotic optimum, depending on the SNR region within which a CU is working. This is especially realizable when the received SNR can be estimated at a CU end. However, to use (3.11) as a benchmark requires knowledge of the covariance matrix Σ_Y. The evaluation of this matrix involves 4^{th}-order cyclic cumulant calculation, which is normally hard to perform under the alternative hypothesis. While a consistent estimator of this matrix exists, it is not an efficient way to do this estimation for each possible combination of lags. In order to circumvent the difficulty of calculating the 4^{th}-order cyclic cumulant and offer an efficient evaluation of this covariance matrix, we restrict our test-set selection to the low SNR region. Certainly, this region where a CU is highly likely to cause harmful interference to a PU is critical for a CR system and it is well worth the effort to enhance detection performance by using a suitable test set in this region. On the other hand, doing hypothesis testing with optimal single point or points of maximum CAF magnitude, as shown in Section 3.4.3, could be sufficiently good for other SNR regions.

In our system model, the undecided test set $\mathcal{J} = \{(\alpha_1, \tau_1), \ldots, (\alpha_M, \tau_M)\}$ will be determined based on the asymptotic detection probability P_d for which prior knowledge of the cyclostationary process $\{X_n\}_{n=0}^{\infty}$ is required. Let $\hat{\mathcal{A}}_{\text{CFL}} = \{\hat{\alpha}_1, \cdots, \hat{\alpha}_K\}$ be a collection of cycle frequencies of interest and $\hat{\mathcal{L}}_{\text{CFL}} = \{\hat{\tau}_1, \cdots, \hat{\tau}_L\}$ be a set of feasible lags. The test set $\mathcal{J}$ will consist of test points in which $\alpha_i \in \hat{\mathcal{A}}_{\text{CFL}}$ and $\tau_i \in \hat{\mathcal{L}}_{\text{CFL}}$ for $1 \leq i \leq M \leq M_{\max}$ where $M_{\max}$ is an upper bound on the cardinality of $\mathcal{J}$.

Optimal Selection Scheme

As shown in (3.11), several parameters, such as the target FAR P_f, the sample size N, the SNR value, and the CAF of the PU signal, are required to specify P_d. Once these parameters are given, the optimal test set $\mathcal{J}_{opt}$ of length M_{opt} can be chosen by maximizing P_d, that is,

$$(\mathcal{J}_{opt}, M_{opt}) = \underset{\mathcal{J}, M < M_{\max}}{\arg} \max P_d. \tag{3.13}$$

This test set $\mathcal{J}_{opt}$ is said to be locally asymptotically optimal because it depends on the initial parameters and P_d is achieved for sufficiently large N. To check if there exists any simpler way to identify $\mathcal{J}_{opt}$, let's recast the detection probability as $P_d = Q_{2M}(\sqrt{N(\mathbf{r}_Y^{\mathcal{J}})'\Sigma_Y^{-1}\mathbf{r}_Y^{\mathcal{J}}}, \sqrt{\lambda})$. The monotonicity of $Q_m(a, b)$ indicates that P_d is strictly increasing in M and $N(\mathbf{r}_Y^{\mathcal{J}})'\Sigma_Y^{-1}\mathbf{r}_Y^{\mathcal{J}}$, and is strictly decreasing in λ [49]. However, it is impossible to increase M and decrease λ simultaneously due to $F(\lambda; 2M) = 1 - P_f$ for a fixed P_f. Moreover, there is no explicit description to show how the quantity $(\mathbf{r}_Y^{\mathcal{J}})'\Sigma_Y^{-1}\mathbf{r}_Y^{\mathcal{J}}$ varies with M. In other words, using more or less test points in terms of larger or smaller M does not necessarily lead to higher P_d. Even if the optimal length M_{opt} is known, finding out $\mathcal{J}_{opt}$ that maximizes $(\mathbf{r}_Y^{\mathcal{J}})'\Sigma_Y^{-1}\mathbf{r}_Y^{\mathcal{J}}$ will still rely on the exhaustive search over all possible sets $\mathcal{J} = \{(\alpha_1, \tau_1), \ldots, (\alpha_{M_{opt}}, \tau_{M_{opt}})\}$ unless the matrix Σ_Y^{-1} is an identity matrix.

Sub-optimal Selection Scheme

In the previous section, the asymptotic detection performance used for identifying proper test points involves a theoretical covariance matrix Σ_Y under $\mathcal{H}_1$. As shown in [48], to evaluate this matrix requires knowledge of the 4^{th}-order cyclic cumulants of the PU signal and the noise. When considering the white Gaussian noise, it is not hard to obtain its 4^{th}-order cyclic cumulant. Nevertheless, to the best of our knowledge, only a few cyclic cumulant expressions of discrete-time communication signals are available. To circumvent the difficulty of obtaining this cyclic cumulant, we provide an approximated inverse matrix $\tilde{\Sigma}_Y^{-1}$ of Σ_Y^{-1} under $\mathcal{H}_1$ in the low SNR region. Thus, an alternative approximate detection performance $\tilde{P}_d \triangleq 1 - P(\lambda; 2M, N(\mathbf{r}_Y^{\mathcal{J}})'\tilde{\Sigma}_Y^{-1}\mathbf{r}_Y^{\mathcal{J}})$ can be used to facilitate the identification of desired test points.

Let's define $\Sigma_Z \triangleq \Sigma_Y$ under $\mathcal{H}_0$ and decompose Σ_Y under $\mathcal{H}_1$ as

$$\mathcal{H}_1 : \Sigma_Y = \Sigma_Z + \Sigma_\Delta,$$

where $\Sigma_\triangle \triangleq \Sigma_Y - \Sigma_Z$. In the following proposition, we will show that the inverse Σ_Y^{-1} under $\mathcal{H}_1$ can be well approximated by Σ_Z^{-1} in the low SNR region. A special case of this proposition in which Σ_Z is an identity matrix has been shown in [50]. However, in our proposition, the matrix Σ_Z only has to be positive definite.

Proposition 1. *Given a test set $\mathcal{J} = \{(\alpha_1, \tau_1), \ldots, (\alpha_M, \tau_M)\}$ if the corresponding $2M \times 2M$ matrices Σ_Y and Σ_Z are positive definite. Then each entry of the matrix $\left(\Sigma_Y^{-1} - \Sigma_Z^{-1}\right)$ approaches zero as the SNR ratio goes to zero.*

Proof. See Section 3.B. □

With this proposition, we are eligible to use Σ_Z^{-1} as $\tilde{\Sigma}_Y^{-1}$ and replace P_d with $\tilde{P}_d$ in (3.13), yielding the sub-optimal test set $\mathcal{J}_{sub}$ of length M_{sub}, that is,

$$(\mathcal{J}_{sub}, M_{sub}) = \underset{\mathcal{J}, M < M_{\max}}{\arg \max} \tilde{P}_d. \tag{3.14}$$

Compared with the optimal scheme, this sub-optimal scheme requires less knowledge and leads to the similar performance in the low SNR region which will be shown in numerical results.

3.2.3 Multi-Cycle-Frequency Detection

In multi-cycle-frequency detection [42, 43], a couple of test statistics are formed over different cycle frequencies and then combined to be the final test statistic. Let's define $\hat{\mathcal{A}}_{\text{MD}} = \{\hat{\alpha}_1, \ldots, \hat{\alpha}_K\}$ and $\hat{\mathcal{L}}_{\text{MD}} = \{\hat{\tau}_1, \ldots, \hat{\tau}_L\}$ respectively as the sets of cycle frequencies and lags of interest. For any given cycle frequency $\alpha \in \hat{\mathcal{A}}_{\mathrm{MD}}$, a test set is given by $\mathcal{J}^{\alpha} = \{(\alpha, \tau_1), \ldots, (\alpha, \tau_m)\}$ with its corresponding test statistic $\Lambda_m(\alpha)$, which is obtainable by following the steps in Section 3.2.1. Typically, the lag set $\mathcal{L}_m^{\alpha} = \{\tau_1, \ldots, \tau_m\}$ of cardinality $m = 1$ or 2 in use is a collection such that $|R_X^{\alpha}[\tau]| = \max\{|R_X^{\alpha}[\rho]|\,|\rho \in \hat{\mathcal{L}}_{\mathrm{MD}}\}$ for $\tau \in \mathcal{L}_m^{\alpha}$. That is, $\mathcal{L}_1^{\alpha}$ includes a single lag of maximum absolute cyclic autocorrelation, and $\mathcal{L}_2^{\alpha}$ includes two. Thus, we can obtain a collection of test statistics over the cycle frequencies of interest $\{\Lambda_m(\alpha)\,|\alpha \in \hat{\mathcal{A}}_{\mathrm{MD}}\}$. Using any of the different diversity schemes, SC, EGC, and MRC, these test statistics are combined, yielding the final test statistics, i.e., for $m \in \{1, 2\}$,

$$\Lambda_m^{\mathrm{SC}} = \max_{\alpha \in \hat{\mathcal{A}}_{\mathrm{MD}}} \Lambda_m(\alpha), \tag{3.15}$$

$$\Lambda_m^{\mathrm{EGC}} = \sum_{\alpha \in \hat{\mathcal{A}}_{\mathrm{MD}}} \Lambda_m(\alpha), \tag{3.16}$$

and

$$\Lambda_m^{\mathrm{MRC}} = \sum_{k=1}^{K} w_k \Lambda_m(\hat{\alpha}_k), \tag{3.17}$$

where

$$w_k = \frac{\Lambda_m(\hat{\alpha}_k)}{\sqrt{\sum_{k=1}^{K} \Lambda_m^2(\hat{\alpha}_k)}}. \tag{3.18}$$

The asymptotic or approximated distributions of these final test statistics under the null hypothesis and their cumulative distributions are discussed in [42–44].

For the sake of completeness, it is worthwhile commenting on the complexity of our proposed joint-utilization detection relative to the multi-cycle-frequency detection technique. In the case of joint-utilization, the test sets selected by both optimal and sub-optimal schemes are parameter-dependent. In other words, the test sets are not necessarily the same over different initial parameter settings. In contrast, the test sets used in multi-cycle-frequency detection are fixed. Also, in our technique, it is required to have information, such as the fourth-order cyclic cumulants of the PU signal and the background noise, which is not a requirement in multi-cycle-frequency detection. Finally, while the requirement of the exhaustive search approach in joint-utilization detection is undesirable, as most PU signals in the licensed bands are standardized, the test sets to be exploited in joint-utilization detection can be determined in advance, hence eliminating the need for an exhaustive search. Furthermore, the joint-utilization detection promises superior performance in the low SNR region.

3.3 Case Study of Linear Modulated Signal

For illustration, we examine an example for which the theoretical CAF vector $\mathbf{r}_Y^{\mathcal{J}}$ and covariance matrix Σ_Y under $\mathcal{H}_1$ are both obtainable. Let the PU signal be the linear modulated signal with its samples $X_n = \sum_{k=-\infty}^{\infty} a_k p(nT_s - kT_{sym})$, where T_s is the sampling interval, T_{sym} is the symbol interval, a_k denotes independent and identically distributed zero-mean complex-valued symbols from a finite alphabet with $\mathbb{E}[a_n a_n^*] = \sigma_a^2$ and $\mathbb{E}[|a_n|^4] = \eta_a \sigma_a^2$, and $p\,(t)$ is a rectangular pulse of value 1 for $0 \leq t < T_{sym}$ and value 0 elsewhere. This PU signal X_n is second-order almost-cyclostationary with the cycle frequency $\alpha = {}^{2\pi\kappa}/_{N_{sym}}$ for some $\kappa \in \{k \in \mathbb{Z} | k \in ({}^{-N_{sym}}/_{2}, {}^{-N_{sym}}/_{2}]\}$ where $N_{sym} = {}^{T_{sym}}/_{T_s}$. We restrict the possible value of τ to $|\tau| < N_{sym}$. Given N received samples $\{Y_n\}_{n=0}^{N-1}$, the theoretical vector $\mathbf{r}_Y^{\mathcal{J}}$ and matrix Σ_Y under PU-present hypothesis are derived as follows. A property, to be used, is provided below.

Property 1. *For a given τ such that $|\tau| < N_{sym}$, let $p(nT_s - kT_{sym})\, p((n+\tau)T_s - lT_{sym})$ be defined as $F_{2p}(n;\tau;k,l)$, then $F_{2p}(n;\tau;k,l) = 1$ when $k = l$ and $kN_{sym} + \max\{0, -\tau\} \leq n < (k+1)N_{sym} + \min\{0, -\tau\}$.*

Proof. It can be easily shown that $p(t)\, p(t+\xi) = 1$ for $|\xi| < T_{sym}$ and $\max\{0, -\xi\} \leq t < T_{sym} + \min\{0, -\xi\}$. As $F_{2p}(n;\tau;k,l)$ can be expressed as $p(\hat{t})p(\hat{t}+\hat{\xi})$ where $\hat{t} = nT_s - kT_{sym}$ and $\hat{\xi} = \tau T_s + (k-l)T_{sym}$, we have the result that $F_{2p}(n;\tau;k,l) = 1$ for $|\tau + (k-l)N_{sym}| < N_{sym}$ and $kN_{sym} + \max\{0, -\hat{\xi}\} \leq n < (k+1)N_{sym} + \min\{0, -\hat{\xi}\}$. Due to the presumption $|\tau| < N_{sym}$, a necessary condition for $F_{2p}(n;\tau;k,l) = 1$ to be true is $k = l$. $\square$

Due to $\mathbf{r}_Y^{\mathcal{J}} = \lim_{N\to\infty} \mathbb{E}\left[\hat{\mathbf{r}}_Y^{\mathcal{J}}\right]$ and $\hat{\mathbf{w}}_Y^{\mathcal{J}}$ being the alternative form of $\hat{\mathbf{r}}_Y^{\mathcal{J}}$, we seek to evaluate the vector $\mathbf{w}_Y^{\mathcal{J}} \triangleq \lim_{N\to\infty} \mathbb{E}\left[\hat{\mathbf{w}}_Y^{\mathcal{J}}\right]$ instead of $\mathbf{r}_Y^{\mathcal{J}}$. The ith entry w_i of $\mathbf{w}_Y^{\mathcal{J}}$ can be presented as

$$\begin{aligned} w_i &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=\max\{0,-\tau_i\}}^{N+\min\{0,-\tau_i\}-1} \mathbb{E}\left[Y_n Y_{n+\tau_i}^*\right] e^{-i\alpha_i n}, \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=\max\{0,-\tau_i\}}^{N+\min\{0,-\tau_i\}-1} \mathbb{E}\left[X_n X_{n+\tau_i}^* + Z_n Z_{n+\tau_i}^*\right] e^{-i\alpha_i n}, \end{aligned}$$

by using the fact $\mathbb{E}\left[X_n Z_{n+\tau_i}^* + Z_n X_{n+\tau_i}^*\right] = 0$. Applying Property 1 to the first part of the previous limit yields

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=\max\{0,-\tau_i\}}^{N+\min\{0,-\tau_i\}-1}\mathbb{E}\left[X_nX_{n+\tau_i}^*\right]e^{-i\alpha_i n}$$
$$=\lim_{N\to\infty}\frac{\sigma_a^2}{N}\sum_{n=\max\{0,-\tau_i\}}^{N+\min\{0,-\tau_i\}-1}\sum_{k,l=-\infty}^{\infty}F_{2p}\left(n;\tau_i;k,l\right)e^{-i\alpha_i n},$$
$$=\lim_{K\to\infty}\frac{\sigma_a^2}{KN_{sym}}\sum_{n_q=0}^{K-1}\sum_{n_r=\max\{0,-\tau_i\}}^{N_{sym}+\min\{0,-\tau_i\}-1}e^{-i\alpha_i n_r},$$
$$=\frac{\sigma_a^2}{N_{sym}}\sum_{n_r=\max\{0,-\tau_i\}}^{N_{sym}+\min\{0,-\tau_i\}-1}e^{-i\alpha_i n_r}, \quad (3.19)$$

where n is replaced by $n_qN_{sym}+n_r$ for $0\le n_r<N_{sym}$ and N by KN_{sym}. The second part of the limit is given by

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=\max\{0,-\tau_i\}}^{N+\min\{0,-\tau_i\}-1}\mathbb{E}\left[Z_nZ_{n+\tau_i}^*\right]e^{-i\alpha_i n}=\sigma_Z^2\delta\left(\tau_i\right)\delta\left(\alpha_i\right), \quad (3.20)$$

where $\delta\left(\cdot\right)$ represents a function that maps zero to one and any other real number to zero. Finally, we evaluate w_i as the sum of (3.19) and (3.20).

To obtain Σ_Y requires knowledge of the entries $c_{i,j}$ and $\bar{c}_{i,j}$ of $\mathbf{C}_{\hat{\mathbf{w}}}$ and $\bar{\mathbf{C}}_{\hat{\mathbf{w}}}$. These two entries can be further expanded as

$$c_{i,j}=\lim_{N\to\infty}N\mathrm{Cov}\left(\hat{R}_X^{\alpha_i}\left[\tau_i\right],\hat{R}_X^{\alpha_j}\left[\tau_j\right]\right)+\lim_{N\to\infty}N\mathrm{Cov}\left(\hat{R}_Z^{\alpha_i}\left[\tau_i\right],\hat{R}_Z^{\alpha_j}\left[\tau_j\right]\right)$$
$$+\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\xi=-n}^{N-1-n}\left\{R_{XZZX}\left[n;\tau_i,\xi+\tau_j,\xi\right]+R_{ZXXZ}\left[n;\tau_i,\xi+\tau_j,\xi\right]\right\}e^{-j(\alpha_i-\alpha_j)n}e^{j\alpha_j\xi}, \quad (3.21)$$

and

$$\bar{c}_{i,j} = \lim_{N\to\infty} N\mathrm{Cov}\left(\hat{R}_X^{\alpha_i}[\tau_i], \hat{R}_X^{\alpha_j *}[\tau_j]\right) + \lim_{N\to\infty} N\mathrm{Cov}\left(\hat{R}_Z^{\alpha_i}[\tau_i], \hat{R}_Z^{\alpha_j *}[\tau_j]\right)$$
$$+\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\xi=-n}^{N-1-n}\left\{R_{XZZX}[n;\tau_i,\xi,\xi+\tau_j] + R_{ZXXZ}[n;\tau_i,\xi,\xi+\tau_j]\right\} e^{-j(\alpha_i+\alpha_j)n} e^{-j\alpha_j\xi}, \tag{3.22}$$

where $R_{ABCD}[n;\tau_1,\tau_2,\tau_3] \triangleq \mathbb{E}\left[A_n B^*_{n+\tau_1,} C_{n+\tau_2} D^*_{n+\tau_3}\right]$. The six limits which appear in (3.21) and (3.22) are evaluated in Section 3.C.

3.4 Numerical Results

3.4.1 Exploiting Multiple Lags

Let the PU signal be the 16-QAM linear modulated signal in which $N_{sym} = 4$. In this section, we evaluate the performance in the lag domain for a fixed cycle frequency, that is, $\hat{\mathcal{A}}_{\text{CFL}} = \{\alpha_0 = 2\pi/4\}$ and $\hat{\mathcal{L}}_{\text{CFL}} = \{-3,-2,-1,0,1,2,3\}$. The simulation parameters are chosen as follows. Three test sets used for illustration are given respectively by, $\mathcal{J}_1 = \{(\alpha_0, 3)\}$, $\mathcal{J}_2 = \{(\alpha_0,-3),(\alpha_0,3)\}$, and $\mathcal{J}_3 = \{(\alpha_0,-3),(\alpha_0,-2),(\alpha_0,2),(\alpha_0,3)\}$. The SNR region of interest is restricted to the range between -20dB and -8dB. The size of test samples is $N = 4000$.

Fig. 3.1 compares the analytical asymptotic detection performance P_d in (3.11) with the simulated detection performance in which the detection is performed by using the test statistic $\Lambda = N(\hat{\mathbf{r}}_Y^{\mathcal{J}})'\hat{\Sigma}_Y^{-1}\hat{\mathbf{r}}_Y^{\mathcal{J}}$. The simulated detection probability is obtained by averaging over 10000 Monte Carlo runs. It is apparent from this figure that the used sample size N is sufficiently large as the simulated performance approaches the

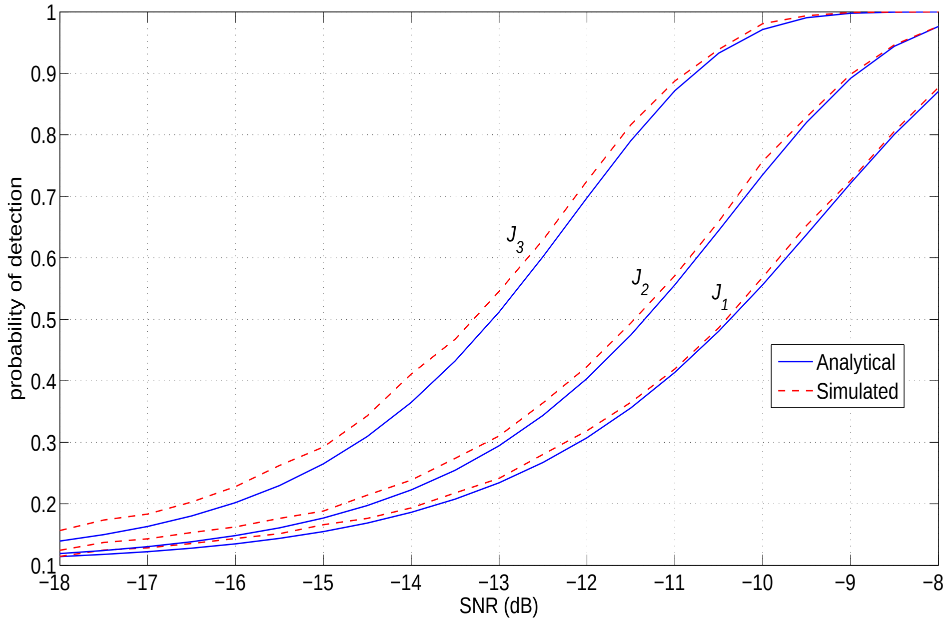


Figure 3.1: The comparison of the theoretical detection performance and the simulated detection performance over different lag sets, $\mathcal{J}_1$, $\mathcal{J}_2$, and $\mathcal{J}_3$, given $P_f = 0.1$.

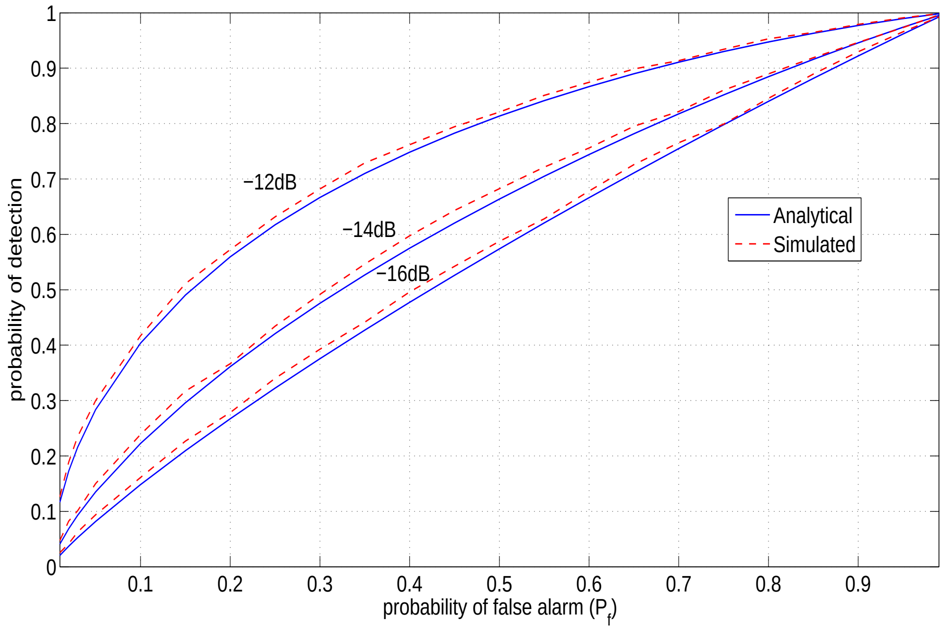


Figure 3.2: Detection probability vs. false alarm rate for the lag set $\mathcal{J}_2$ over different SNR ratios.

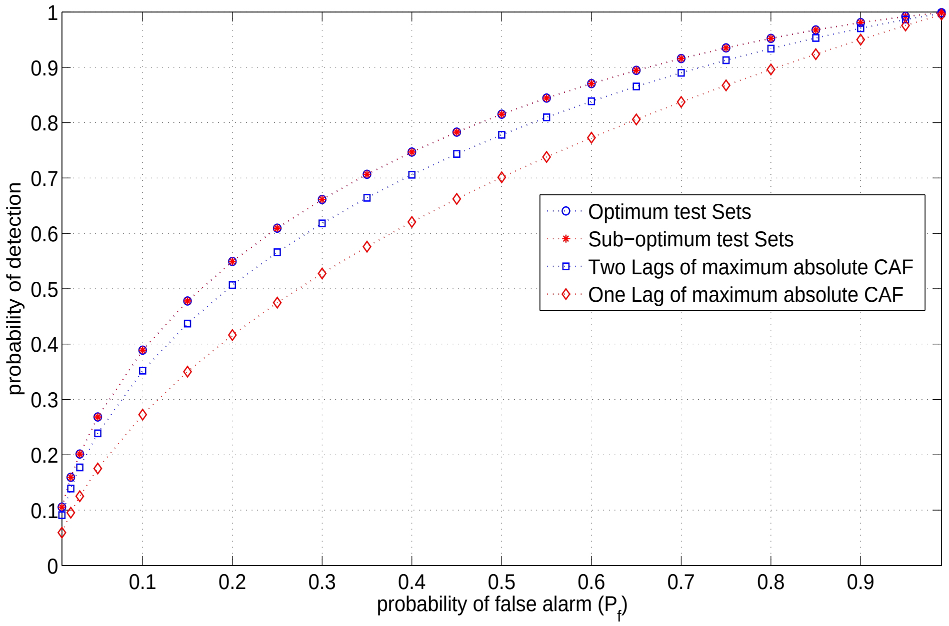


Figure 3.3: Theoretical detection probability vs. false alarm rate for different multiple lags selection schemes at SNR=-14dB.

analytical asymptotic performance. With the same sample size and fixed lag set $\mathcal{J}_2$, the corresponding receiver operating characteristic (ROC) curves of the analytical and simulated results are presented in Fig. 3.2. As can be seen from this figure, there is a significant improvement in detection performance from the lower SNR ratio to the higher. The same level of improvement can be observed from analytical results and simulated results. Because there is no notable difference between analytical and simulated results for the given parameter setting, the analytical performance P_d can be used to assess the effectiveness of our proposed test-point selection schemes.

For the given fixed SNR ratio -14dB, the analytical probability of detection versus false alarm rate for different test-point selection schemes is depicted in Fig. 3.3. This figure shows that, for any given FAR, the detector based on the proposed sub-optimal scheme outperforms that using test set $\{(\alpha_0, \bar{\tau}_1)\}$ or $\{(\alpha_0, \bar{\tau}_1), (\alpha_0, \bar{\tau}_2)\}$, in which $|R_X^{\alpha}[\tau]| = \max\{|R_X^{\alpha_0}[\rho]|\,|\rho \in \hat{\mathcal{L}}_{\text{CFL}}\}$ for $\tau \in \{\bar{\tau}_1, \bar{\tau}_2\}$. Furthermore, the proposed scheme leads to the best performance which is achieved by the optimal scheme for this given low SNR ratio. As claimed before, our proposed selection schemes give parameter-dependent test sets so they are likely to use different test sets at different simulated points over the performance curves. Thus, the explicit optimal and sub-optimal test sets, used at each simulated point, are not listed in the figure for compactness.

3.4.2 Jointly Exploiting Multiple Cycle Frequencies and Lags

The simulation parameters for comparing performance of our proposed schemes and multi-cycle-frequency detection are listed below. $\hat{\mathcal{A}}_{\text{CFL}} = \{-\alpha_0, \alpha_0, 2\alpha_0\}$, $M_{\max} = 6$, $\hat{\mathcal{A}}_{\text{MD}} = \{\alpha_0, 2\alpha_0\}$, $\hat{\mathcal{L}}_{\text{MD}} = \hat{\mathcal{L}}_{\text{CFL}}$, $\mathcal{L}_1^{\alpha=\alpha_0} = \{2\}$, and $\mathcal{L}_1^{\alpha=2\alpha_0} = \{1\}$. In $\hat{\mathcal{A}}_{\text{CFL}}$ and $\hat{\mathcal{A}}_{\text{MD}}$, the case of zero-valued cycle frequency is not considered because the test points for $\alpha = 0$ are subject to noise uncertainty and interference. Another three test sets with multiple cycle frequencies and lags are given by $\mathcal{J}_4 = \{(\alpha_0, 1), (2\alpha_0, 1)\}$, $\mathcal{J}_5 = \{(\alpha_0, 1), (2\alpha_0, 1), (\alpha_0, 3), (2\alpha_0, 3)\}$, and $\mathcal{J}_6 = \{(2\alpha_0, -3), (\alpha_0, -2), (\alpha_0, 1), (\alpha_0, 2)\}$. The parameters not listed are the same as those in the previous section.

The comparison between analytical and simulated performances over test sets with multiple cycle frequencies and lags, $\mathcal{J}_4$, $\mathcal{J}_5$, and $\mathcal{J}_6$, is depicted in Fig. 3.4. It shows that the analytical performance curve can match the simulated curve in the SNR region of interest for these general test sets as expected when the sample size is large. This verifies that the derived analytical CAF vector $\mathbf{r}_Y^{\mathcal{J}}$ and covariance matrix Σ_Y

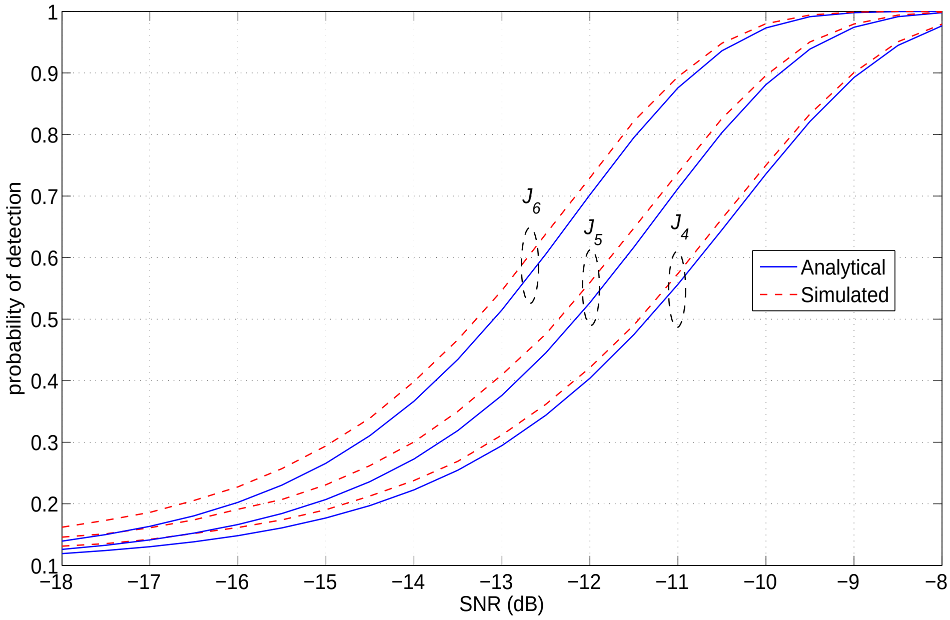


Figure 3.4: The comparison between analytical performance and simulated performance over different sets $\mathcal{J}_4$, $\mathcal{J}_5$, and $\mathcal{J}_6$, given $P_f = 0.1$.

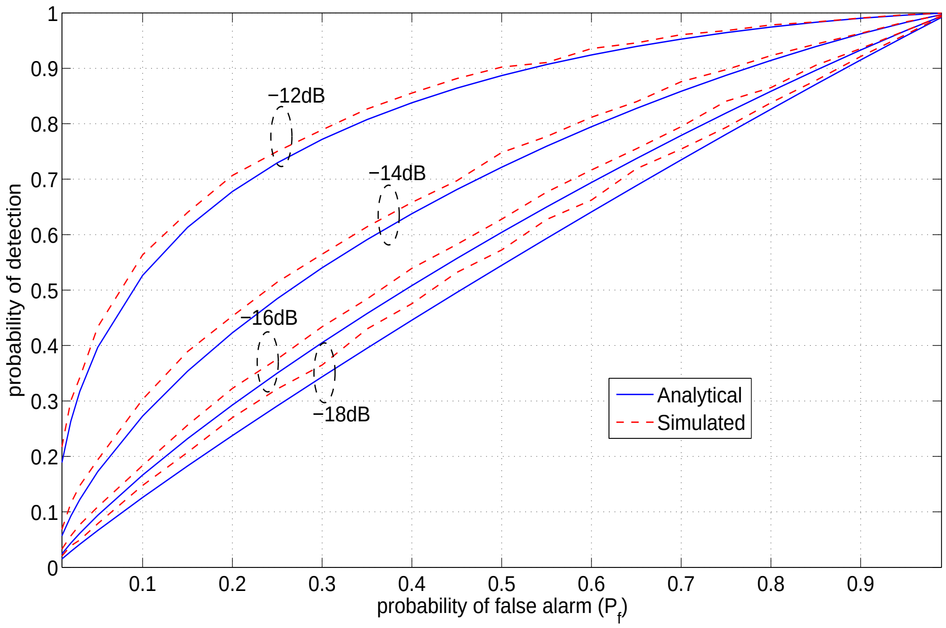


Figure 3.5: Detection probability vs. false alarm rate for the set $\mathcal{J}_5$ over different SNR ratios.

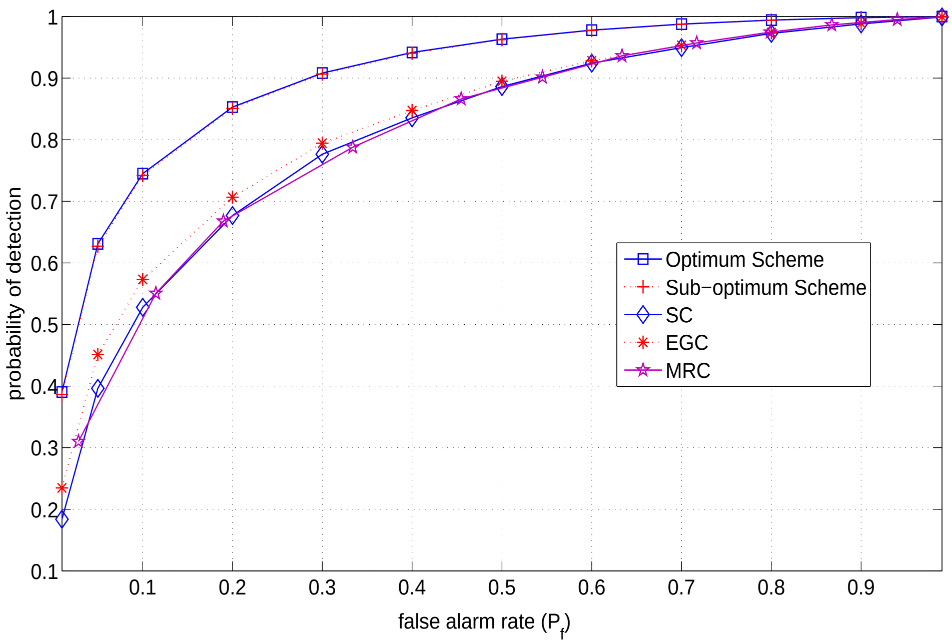


Figure 3.6: Probability of detection vs. false alarm rate for the optimal and sub-optimal selection schemes and multi-cycle-frequency detection at SNR=-12dB.

under $\mathcal{H}_1$ of the PU signal are both accurate. When restricted to the test set $\mathcal{J}_5$, the analytical ROC curves in Fig. 3.5 demonstrate that for any given FAR, the analytical value of P_d can approach the simulated value. This will justify the following part of using the analytical asymptotic performance as a benchmark, in which the analytical optimal results are compared to the sub-optimal results and multi-cycle-frequency detection results in terms of ROC curves.

Fig. 3.6 presents detection probability versus false alarm rate at SNR=-12dB over different test statistics formed by our proposed schemes or combining schemes in multi-cycle-frequency detection. The probability of detection corresponding to the optimal scheme or the sub-optimal scheme is obtained based on (3.11). For SC and EGC in multi-cycle-frequency detection, the detection performance is a simulation result averaged over 10000 Monte Carlo runs, where the detection thresholds are

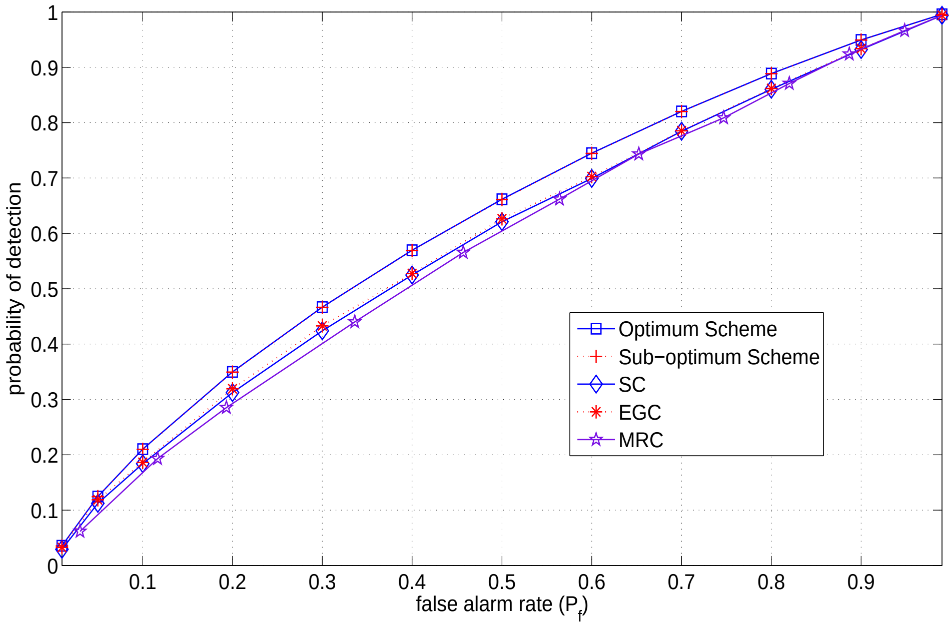


Figure 3.7: Probability of detection vs. false alarm rate for the optimal and sub-optimal selection schemes and multi-cycle-frequency detection at SNR=-16dB.

calculated based on the cumulative distributions provided in [42]. As the approximated distribution of the test statistic of MRC [43, 44] is not sufficiently accurate for our setting, the performance curve of MRC is acquired via extensive simulations over a range of thresholds to find the best detection probability and false alarm rate for comparison purposes. From this figure, we can see that our proposed schemes of jointly using cycle-frequencies and lags significantly outperform multi-cycle-frequency detection whether SC, EGC, or MRC are used. Simulation results at SNR=-16dB presented in Fig. 3.7 also show the comparative enhancement of our proposed.

3.4.3 Higher SNR Region and Smaller Sample Size

Previous analytical and simulated results focus on the low SNR region with sufficient samples. It is intriguing to know what will happen if the operation moves to the

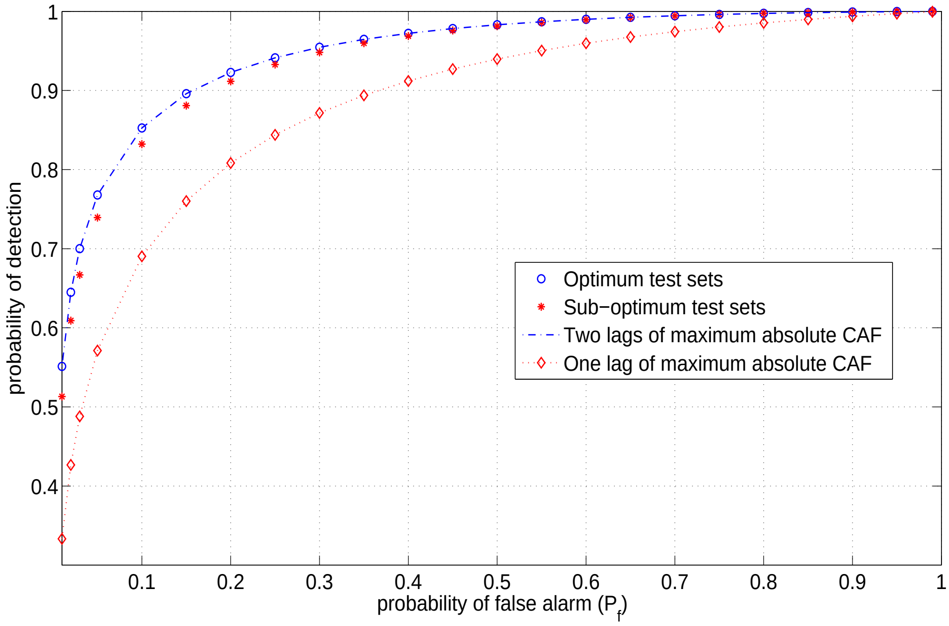


Figure 3.8: Theoretical detection probability vs. false alarm rate for different multiple lags selection schemes at SNR=-5dB.

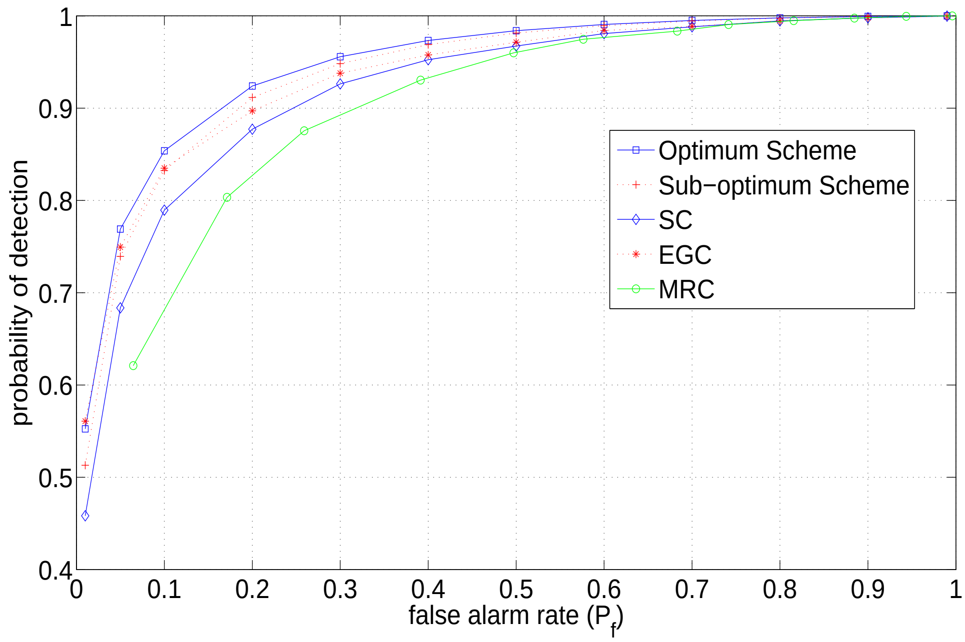


Figure 3.9: Probability of detection vs. false alarm rate for the optimal and sub-optimal selection schemes and multi-cycle-frequency detection at SNR=-5dB.

higher SNR region with smaller sample size. In Fig. 3.8 and Fig. 3.9, we reexamine the ROC curves at SNR=-5dB where the smaller sample size $N = 400$ is used. As can be seen from these two figures, the proposed sub-optimal scheme no longer achieves the optimal performance in this higher SNR region. Though this sub-optimal scheme can still provide decent performance, other existing schemes may outperform it.

3.5 Conclusions

This chapter has investigated cyclostationary feature spectrum sensing based on jointly utilizing multiple cycle frequencies and multiple lags. Using the generalized statistical testing model in the cycle-frequency-lag domain, an optimal scheme for selecting test points has been proposed. To reduce the required prior knowledge about PU signals, a practical method with comparable performance in the low SNR region was also developed. When restricted to a specific cycle frequency, using test points chosen by our proposed schemes, rather than using a single or two points of maximum absolute cyclic autocorrelation, can lead to notably superior performance. In the CFL domain, the results demonstrated superiority of our proposed over multi-cycle-frequency with SC, EGC, and MRC detection methods.

Preliminary results of using multiple lags have been published in IEEE Signal Processing Letters [51], and the work of the joint utilization approach has been accepted for publication by IEEE Transactions on Signal Processing [52].

3.A Proof of Asymptotic Distribution

In [47], it has been shown that, in the case of $\alpha_i = \alpha_j$ for $\alpha_i, \alpha_i \in \mathcal{J}$, the asymptotic distribution of the test statistic Λ under $\mathcal{H}_1$ can be approximated by the non-central chi-square distribution. However, a stronger statement can be made, that is, this test statistic converges in distribution to a non-central chi-square distributed random variable. The test statistic Λ is the squared norm of the vector $\sqrt{N}\hat{\Sigma}_Y^{-1/2}\hat{\mathbf{r}}_Y^{\mathcal{J}}$. By making use of (3.6) and [53, Corollary 1.7], it can be shown that this vector converges in distribution to a Gaussian random vector, that is, $\lim_{N\to\infty} \sqrt{N}\hat{\Sigma}_Y^{-1/2}\hat{\mathbf{r}}_Y^{\mathcal{J}} \overset{\mathrm{D}}{=} \mathcal{N}(\sqrt{N}\Sigma_Y^{-1/2}\mathbf{r}_Y^{\mathcal{J}}, \mathbf{I})$. Using the same corollary again and [53, Lemma 3.5], we reach a conclusion that the statistic Λ converges in distribution to a non-central chi-square distributed random variable with $2M$ degrees of freedom and a non-centrality parameter $N(\mathbf{r}_Y^{\mathcal{J}})'\Sigma_Y^{-1}\mathbf{r}_Y^{\mathcal{J}}$.

3.B Proof of Proposition 1

Due to Σ_Y being positive definite, there exists an orthonormal matrix $\mathbf{Q}$ such that $\Lambda_Y = \mathbf{Q}'\Sigma_Y\mathbf{Q}$ where Λ_Y denotes a diagonal matrix with diagonal entries $\{\lambda_{Y,n}\}_{n=1}^{2M}$. This matrix $\mathbf{Q}$ can also diagonalize Σ_Z and Σ_Δ. Thus we have diagonal matrices $\Lambda_Z = \mathbf{Q}'\Sigma_Z\mathbf{Q}$ and $\Lambda_\Delta = \mathbf{Q}'\Sigma_\Delta\mathbf{Q}$ with their diagonal entries $\{\lambda_{Z,n}\}_{n=1}^{2M}$ and $\{\lambda_{\Delta,n}\}_{n=1}^{2M}$ and the inequality $\lambda_{Z,n} + \lambda_{\Delta,n} = \lambda_{Y,n} > 0$.

The matrix Σ_Y^{-1} can be expanded as

$$\Sigma_Y^{-1} = \left(\Sigma_Z + \Sigma_\Delta\right)^{-1},$$
$$= \mathbf{QDQ}^{'},$$

where $\mathbf{D} = \left(\Lambda_Z + \Lambda_\Delta\right)^{-1}$. The nth diagonal element of $\mathbf{D}$ is given by $d_n = {}^1\!/\!\left(\lambda_{Z,n}+\lambda_{\Delta,n}\right)$. We recast d_n as

$$d_n = \frac{1}{\lambda_{Z,n}} - \frac{\lambda_{\Delta,n}}{\lambda_{Z,n}\left(\lambda_{Z,n}+\lambda_{\Delta,n}\right)}. \tag{3.23}$$

As each entry of $\Sigma_\triangle$ is a linear combination of terms such as $\mathbb{E}\left[X_i X_j X_k X_l\right]$ and $\mathbb{E}\left[X_i X_j Z_k Z_l\right]$, we can represent $\lambda_{\triangle,n}$ as $a_n\sigma_X^4 + b_n\sigma_X^2\sigma_Z^2$ for some $a_n, b_n \in \mathbb{R}^+$. Similarly, $\lambda_{Z,n}$ can be expressed as $c_n\sigma_Z^4$ for some $c_n \in \mathbb{R}^+$. Thus, d_n converges to ${}^1\!/\!\lambda_{Z,n}$ as the SNR ratio approaches zero due to

$$\begin{aligned}\left|\frac{\lambda_{\Delta,n}}{\lambda_{Z,n}\left(\lambda_{Z,n}+\lambda_{\Delta,n}\right)}\right| &= \frac{\left|a_n\sigma_X^4 + b_n\sigma_X^2\sigma_Z^2\right|}{c_n\sigma_Z^4\left(c_n\sigma_Z^4 + a_n\sigma_X^4 + b_n\sigma_X^2\right)},\\ &= \frac{\left|a_n\mathrm{SNR}^2 + b_n\mathrm{SNR}\right|}{c_n\sigma_Z^4\left(a_n\mathrm{SNR}^2 + b_n\mathrm{SNR} + c_n\right)}.\end{aligned}$$

Consequently, Σ_Y^{-1} converges to Σ_Z^{-1} and the required result follows.

3.C Evaluation of Limits

Here we only provide the derivation of the first limit in (3.21). The evaluation results of the other limits in (3.21) and (3.22) are just listed.

Let's represent $\lim_{N\to\infty} N\mathrm{Cov}\left(\hat{R}_X^{\alpha_i}\left[\tau_i\right], \hat{R}_X^{\alpha_j}\left[\tau_j\right]\right)$ as

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\xi=-n}^{N-1-n}\left\{R_{4X}\left[n;\tau_i,\xi+\tau_j,\xi\right]\right.$$

$$\left.-R_X\left[n;\tau_i\right]R_X^*\left[n+\xi;\tau_j\right]\right\}\times e^{-j(\alpha_i-\alpha_j)n}e^{j\alpha_j\xi}, \quad (3.24)$$

where $R_{4X}\left[\cdots\right]\triangleq R_{XXXX}\left[\cdots\right]$. The fourth-order moment $R_{4X}\left[n;\tau,\xi+\rho,\xi\right]$ can be further expanded as

$$\begin{aligned}
&R_{4X}\left[n;\tau,\xi+\rho,\xi\right]\\
&=\sum_{i=-\infty}^{\infty}\eta_a\sigma_a^4F_{4p}\left(n;\tau,\xi+\rho,\xi;i=j=k=l\right)\\
&\qquad+\sum_{i,k=-\infty}^{\infty}\sigma_a^4F_{4p}\left(n;\tau,\xi+\rho,\xi;i=j\neq k=l\right)\\
&\qquad\qquad+\sum_{i,k=-\infty}^{\infty}\sigma_a^4F_{4p}\left(n;\tau,\xi+\rho,\xi;i=l\neq j=k\right),
\end{aligned}\quad (3.25)$$

where

$$\begin{aligned}
&F_{4p}\left(n;\tau_1,\tau_2,\tau_3;i,j,k,l\right)\triangleq\\
&\qquad p\left(nT_s-iT_{sym}\right)p\left(\left(n+\tau_1\right)T_s-jT_{sym}\right)\\
&\qquad\qquad\times p\left(\left(n+\tau_2\right)T_s-kT_{sym}\right)p\left(\left(n+\tau_3\right)T_s-lT_{sym}\right).
\end{aligned}$$

Some properties for the case, $|\tau|<N_{sym}$ and $|\rho|<N_{sym}$, are provided below.

Property 2. $F_{4p}\left(n;\tau,\xi+\rho,\xi;i=j=k=l\right)=1$ *when* $i=n_q$, $\max\{0,-\tau\}\leq n_r<N_{sym}+\min\{0,-\tau\}$, *and* $-n_r+\max\{0,-\rho\}\leq\xi<N_{sym}-n_r+\min\{0,-\rho\}$.

Property 3. $F_{4p}\left(n;\tau,\xi+\rho,\xi;i=j\neq k=l\right)=1$ *when* $i=n_q$, $\max\{0,-\tau\}\leq n_r<N_{sym}+\min\{0,-\tau\}$, *and* $(k-i)N_{sym}-n_r+\max\{0,-\rho\}\leq\xi<(k-i+1)N_{sym}-$

$n_r + \min\{0, -\rho\}$.

Property 4. *$F_{4p}(n; \tau, \xi+\rho, \xi; i = l \neq j = k) = 1$ when $i = n_q$ and either (1) or (2), where (1) denotes $-N_{sym} < \tau < 0$, $-N_{sym} < \rho < 0$, $0 \leq n_r < -\tau$, and $-n_r \leq \xi < -n_r - \rho$, and (2) denotes $0 < \tau < N_{sym}$, $0 < \rho < N_{sym}$, $N_{sym} - \tau \leq n_r < N_{sym}$, and $N_{sym} - n_r - \rho \leq \xi < N_{sym} - n_r$.*

Property 5. *$F_{2p}(n; \tau; i = j)\, F_{2p}(n+\xi+\rho; -\rho; k = l) = 1$ when $i = n_q$, $\max\{0, -\tau\} \leq n_r < N_{sym} + \min\{0, -\tau\}$, and $(k-i)\, N_{sym} - n_r + \max\{0, -\rho\} \leq \xi < (k-i+1)\, N_{sym} - n_r + \min\{0, -\rho\}$.*

Table 3.1

	$R_{4X}[n; \tau, \xi+\rho, \xi] - R_X[n; \tau]\, R_X^*[n+\xi; \rho]$
$-N_{sym} < \tau < 0$ $-N_{sym} < \rho < 0$	$= (\eta_a - 1)\, \sigma_a^4$, if $-\tau \leq n_r < N_{sym}$ and $-n_r - \rho \leq \xi < N_{sym} - n_r$.
	$= \sigma_a^4$, if $0 \leq n_r < -\tau$ and $-n_r \leq \xi < -n_r - \rho$.
$0 < \tau < N_{sym}$ $0 < \rho < N_{sym}$	$= (\eta_a - 1)\, \sigma_a^4$, if $0 \leq n_r < N_{sym} - \tau$ and $-n_r \leq \xi < N_{sym} - n_r - \rho$.
	$= \sigma_a^4$, if $N_{sym} - \tau \leq n_r < N_{sym}$ and $N_{sym} - n_r - \rho \leq \xi < N_{sym} - n_r$.
$\tau\rho \leq 0$	$= (\eta_a - 1)\, \sigma_a^4$, if $\max\{0, -\tau\} \leq n_r < N_{sym} + \min\{0, -\tau\}$ and $-n_r + \max\{0, -\rho\} \leq \xi < N_{sym} - n_r + \min\{0, -\rho\}$.

Applying Properties 1-5 gives Table 3.1, the evaluation of

$$R_{4X}[n; \tau, \xi+\rho, \xi] - R_X[n; \tau]\, R_X^*[n+\xi; \rho] \tag{3.26}$$

for different combinations of the pair (τ, ρ). Making use of this table in (3.24), we obtain the required limit

$$\lim_{N\to\infty} N\mathrm{Cov}\left(\hat{R}_X^{\alpha_i}\left[\tau_i\right], \hat{R}_X^{\alpha_j}\left[\tau_j\right]\right)$$
$$= \frac{(\eta_a - 1)\,\sigma_a^4}{N_{sym}} \sum_{n_r=\max\{0,-\tau_i\}}^{N_{sym}+\min\{0,-\tau_i\}-1} \sum_{\xi=-n_r+\max\{0,-\tau_j\}}^{N_{sym}-n_r+\min\{0,-\tau_j\}-1} e^{-j(\alpha_i-\alpha_j)n_r} e^{j\alpha_j\xi}$$
$$+ \delta\left[\mathrm{sign}\left(\tau_i\tau_j\right) - 1\right]\delta\left[\mathrm{sign}\left(\tau_i\right) + 1\right] \frac{\sigma_a^4}{N_{sym}} \sum_{n_r=0}^{-\tau_i-1} \sum_{\xi=-n_r}^{-n_r-\tau_j-1} e^{-j(\alpha_i-\alpha_j)n_r} e^{j\alpha_j\xi}$$
$$+ \delta\left[\mathrm{sign}\left(\tau_i\tau_j\right) - 1\right]\delta\left[\mathrm{sign}\left(\tau_i\right) - 1\right] \frac{\sigma_a^4}{N_{sym}} \sum_{n_r=N_{sym}-\tau_i}^{N_{sym}-1} \sum_{\xi=N_{sym}-n_r-\tau_j}^{N_{sym}-n_r-1} e^{-j(\alpha_i-\alpha_j)n_r} e^{j\alpha_j\xi}, \tag{3.27}$$

where the function sign $(\cdot)$ extracts the sign of a real number.

The other limits in (3.21) and (3.22) are given as follows.

$$\lim_{N\to\infty} N\mathrm{Cov}\left(\hat{R}_X^{\alpha_i}\left[\tau_i\right], \hat{R}_X^{\alpha_j *}\left[\tau_j\right]\right)$$
$$= \frac{(\eta_a - 1)\,\sigma_a^4}{N_{sym}} \sum_{n_r=\max\{0,-\tau_i\}}^{N_{sym}+\min\{0,-\tau_i\}-1} \sum_{\xi=-n_r+\max\{0,-\tau_j\}}^{N_{sym}-n_r+\min\{0,-\tau_j\}-1} e^{-j(\alpha_i+\alpha_j)n_r} e^{-j\alpha_j\xi}$$
$$+ \delta\left[\mathrm{sign}\left(\tau_i\tau_j\right) + 1\right]\delta\left[\mathrm{sign}\left(\tau_i\right) + 1\right] \frac{\sigma_a^4}{N_{sym}} \sum_{n_r=0}^{-\tau_i-1} \sum_{\xi=-n_r-\tau_j}^{-n_r-1} e^{-j(\alpha_i+\alpha_j)n_r} e^{-j\alpha_j\xi}$$
$$+ \delta\left[\mathrm{sign}\left(\tau_i\tau_j\right) + 1\right]\delta\left[\mathrm{sign}\left(\tau_j\right) + 1\right] \frac{\sigma_a^4}{N_{sym}} \sum_{n_r=N_{sym}-\tau_i}^{N_{sym}-1} \sum_{\xi=N_{sym}-n_r}^{N_{sym}-n_r-\tau_j-1} e^{-j(\alpha_i+\alpha_j)n_r} e^{-j\alpha_j\xi}. \tag{3.28}$$

$$\lim_{N\to\infty} N\mathrm{Cov}\left(\hat{R}_Z^{\alpha_i}\left[\tau_i\right], \hat{R}_Z^{\alpha_j}\left[\tau_j\right]\right) = \sigma_Z^4 \delta\left(\alpha_i - \alpha_j\right)\delta\left(\tau_i - \tau_j\right). \tag{3.29}$$

$$\lim_{N\to\infty} N\mathrm{Cov}\left(\hat{R}_Z^{\alpha_i}\left[\tau_i\right], \hat{R}_Z^{\alpha_j *}\left[\tau_j\right]\right)$$
$$= \sigma_Z^4 \delta\left(\alpha_i + \alpha_j\right)\left\{\delta\left(\tau_i\right)\delta\left(\tau_i - \tau_j\right) + \left[1 - \delta\left(\tau_i\right)\right]\delta\left(\tau_i + \tau_j\right)e^{-j\alpha_j\tau_i}\right\}. \tag{3.30}$$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\xi=-n}^{N-1-n}\left\{R_{XZZX}\left[n;\tau_i,\xi+\tau_j,\xi\right]e^{-j(\alpha_i-\alpha_j)n}e^{j\alpha_j\xi}\right\}$$
$$=\frac{\sigma_a^2\sigma_z^2}{N_{sym}}e^{j\alpha_j(\tau_i-\tau_j)}\delta\left[\text{sign}\left(|\tau_i-\tau_j|-N_{sym}\right)+1\right]\sum_{n_r=\max\{0,-(\tau_i-\tau_j)\}}^{N_{sym}+\min\{0,-(\tau_i-\tau_j)\}-1}e^{-j(\alpha_i-\alpha_j)n_r}. \tag{3.31}$$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\xi=-n}^{N-1-n}\left\{R_{XZZX}\left[n;\tau_i,\xi,\xi+\tau_j\right]e^{-j(\alpha_i+\alpha_j)n}e^{-j\alpha_j\xi}\right\}$$
$$=\delta\left[\text{sign}\left(|\tau_i+\tau_j|-N_{sym}\right)+1\right]\frac{\sigma_a^2\sigma_z^2}{N_{sym}}e^{-j\alpha_j\tau_i}\sum_{n_r=\max\{0,-(\tau_i+\tau_j)\}}^{N_{sym}+\min\{0,-(\tau_i+\tau_j)\}-1}e^{-j(\alpha_i+\alpha_j)n_r}. \tag{3.32}$$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\xi=-n}^{N-1-n}\left\{R_{ZXXZ}\left[n;\tau_i,\xi+\tau_j,\xi\right]e^{-j(\alpha_i-\alpha_j)n}e^{j\alpha_j\xi}\right\}$$
$$\begin{aligned}
&=\frac{\sigma_a^2\sigma_z^2}{N_{sym}}\delta\left[\text{sign}\left(|\tau_i-\tau_j|-N_{sym}\right)+1\right]\Bigg\{\left\{\delta\left[\text{sign}\left(\tau_i-\tau_j\right)-1\right]+\delta\left[\text{sign}\left(\tau_i-\tau_j\right)\right]\right\}\\
&\quad\times\Bigg\{\sum_{n_r=\max\{0,-\tau_j\}}^{N_{sym}+\min\{0,-\tau_i\}-1}e^{-j(\alpha_i-\alpha_j)n_r}+\delta\left[\text{sign}\left(\tau_i\right)+1\right]\sum_{n_r=0}^{-\tau_i-1}e^{-j(\alpha_i-\alpha_j)n_r}\\
&\quad+\delta\left[\text{sign}\left(\tau_j\right)-1\right]\sum_{n_r=N_{sym}-\tau_j}^{N_{sym}-1}e^{-j(\alpha_i-\alpha_j)n_r}\Bigg\}+\delta\left[\text{sign}\left(\tau_i-\tau_j\right)+1\right]\\
&\quad\times\Bigg\{\sum_{n_r=\max\{0,-\tau_i\}}^{N_{sym}+\min\{0,-\tau_j\}-1}e^{-j(\alpha_i-\alpha_j)n_r}+\delta\left[\text{sign}\left(\tau_i\right)-1\right]\sum_{n_r=N_{sym}-\tau_i}^{N_{sym}-1}e^{-j(\alpha_i-\alpha_j)n_r}\\
&\quad+\delta\left[\text{sign}\left(\tau_j\right)+1\right]\sum_{n_r=0}^{-\tau_j-1}e^{-j(\alpha_i-\alpha_j)n_r}\Bigg\}\Bigg\}.
\end{aligned} \tag{3.33}$$

$$
\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\xi=-n}^{N-1-n} R_{XZZX}\left[n;\tau_i,\xi,\xi+\tau_j\right] e^{-j(\alpha_i+\alpha_j)n} e^{-j\alpha_j\xi} \\
&= \frac{\sigma_a^2\sigma_z^2}{N_{sym}} e^{j\alpha_j\tau_j}\delta\left[\text{sign}\left(|\tau_i+\tau_j|-N_{sym}\right)+1\right]\Bigg\{\left\{\delta\left[\text{sign}\left(\tau_i+\tau_j\right)-1\right]+\delta\left[\text{sign}\left(\tau_i+\tau_j\right)\right]\right\} \\
&\qquad\times\Bigg\{\sum_{n_r=\max\{0,\tau_j\}}^{N_{sym}+\min\{0,-\tau_i\}-1} e^{-j(\alpha_i+\alpha_j)n_r}+\delta\left[\text{sign}\left(\tau_i\right)+1\right]\sum_{n_r=0}^{-\tau_i-1} e^{-j(\alpha_i+\alpha_j)n_r} \\
&\qquad\quad+\delta\left[\text{sign}\left(\tau_j\right)+1\right]\sum_{n_r=N_{sym}+\tau_j}^{N_{sym}-1} e^{-j(\alpha_i+\alpha_j)n_r}\Bigg\}+\delta\left[\text{sign}\left(\tau_i+\tau_j\right)+1\right] \\
&\qquad\times\Bigg\{\sum_{n_r=\max\{0,-\tau_i\}}^{N_{sym}+\min\{0,\tau_j\}-1} e^{-j(\alpha_i+\alpha_j)n_r}+\delta\left[\text{sign}\left(\tau_i\right)-1\right]\sum_{n_r=N_{sym}-\tau_i}^{N_{sym}-1} e^{-j(\alpha_i+\alpha_j)n_r} \\
&\qquad\qquad\qquad\qquad+\delta\left[\text{sign}\left(\tau_j\right)-1\right]\sum_{n_r=0}^{\tau_j-1} e^{-j(\alpha_i+\alpha_j)n_r}\Bigg\}\Bigg\}. \qquad (3.34)
\end{aligned}
$$

CHAPTER 4

MULTI-ANTENNA SPECTRUM SENSING

4.1 Introduction

Multi-antenna techniques, which have been widely applied in wireless communications including CR either to increase diversity and throughput or to reduce the amount of interference to PUs, can also be exploited to significantly enhance the performance of spectrum sensing. As spectrum sensing can be performed in synchronous or asynchronous ways, multi-antenna spectrum sensing (MASS) can be considered in two scenarios.

In the first scenario, the CUs perform MASS in a quiet sensing period during which no CU transmission is allowed. In other words, the PU signals can be received without cochannel interference from the CUs. The conventional way of exploiting multiple receive antennas is to yield a power gain via receive beamforming, e.g., to maximize the output signal-to-interference-plus-noise ratio (SINR). Alternatively, one could also benefit from the quiet period if the received multi-antenna signals (RMSs) can be more manipulated to maximize the probability of detection. Normally, the Gaussian assumption is made about the received PU signals and the background noise so that the distribution of the RMSs becomes available under the noise-only or PU-present hypothesis. Moreover, a GLRT can be conducted, in which unknown parameters are estimated by the maximum likelihood method. The final test statistic is generally presented as the ratio of functions of eigenvalues of the sampling covariance matrix resulting from the RMSs [54–57]. The formation of the sampling covariance matrix depends on whether the spatial correlations or the temporal correlations of the RMSs are used. Even without making a Gaussian assumption about the PU signals, this sampling covariance matrix can still lead to a similar test statistic whose distribution is evaluated by random matrix theory [58]. Generally, all the methods which use eigenvalue-dependent test statistics are called the eigenvalue based approach. Another approach to utilizing this covariance matrix is exploiting the different matrix profiles under the noise-only and PU-present hypotheses [59]. Following the spirit of receive beamforming, a more conventional approach which combines RMSs by maximizing the output SNR and then applies energy detection can be found in [60]. The effect of the correlated RMSs, when using energy detection, has been investigated in [61].

When it comes to asynchronous spectrum sensing, in which some CUs may transmit at the same time, cochannel interference from these CUs has to be tackled. The previous eigenvalue based approach can apply to this scenario when interference is assumed to be Gaussian distributed [62]. For the general cochannel interference, there might be no mathematically tractable probabilistic models to describe the RMSs so that the GLRT can be applied. In view of this, a feasible and straightforward approach to MASS is to reduce cochannel interference via receive beamforming and then employ any detection technique on the combined signal. A considerable number of receive beamforming algorithms have been developed, in which a class of spectral self-coherence restoral (SCORE) algorithms is especially suitable for the purpose of spectrum sensing due to not requiring training signals and accurate manifold information or a structural constraint on the antenna array [63, 64]. Furthermore, cyclostationary features exploited by this algorithm to find the weighting vector, can immediately be utilized by cyclostationary feature detection to determine the presence of the PUs. In [65], a subspace approach is proposed to make the performance of the least-square SCORE algorithm comparable to the Cross-SCORE algorithm while having less computational complexity. An adaptive version of the Cross-SCORE algorithm for complexity reduction is investigated in [66]. Compared with the SCORE algorithms, another three cyclostationary feature based beamforming algorithms in [67] can lead to higher output SINR, less complexity, and faster convergence rates.

In this chapter, we explore the possible usage of the RMSs in cyclostaionary feature detection developed in [29]. The focus is on synchronous spectrum sensing for which per-combining and post-combining schemes are examined. For pre-combining, it will be shown that MRC, conventionally for improving output SNR, can also lead to the best detection performance in terms of maximizing a performance metric in the low SNR region. In addition, a blind channel estimation based on cyclostationary features will be presented for the practical MRC implementation. Our proposed channel estimation is based on [68]. In our modified technique, we can ensure that the estimation accuracy is asymptotically achieved. This MRC reception, also investigated in [69], assumes perfect channel side information and is only used for the purpose of improving output SNR. For post-combining, we investigate two possible schemes, the joint combining and the sum combining, which can be performed without need of the CSI. In [68], the pre-(post-)combining detectors are established based on the cyclic spectral coherence, while our proposed detectors are based on the cyclic autocorrelation. As cochannel interference can inherently be diminished when constructing cyclostationary statistics, we further indicate the conditions under which our proposed methods can work even with cochannel CU interferes.

4.2 Signal Model and Preliminary Results

Consider that spectrum sensing is performed by a CU equipped with L receive antennas. We assume that the received signal at each antenna is subject to independent and identically distributed (i.i.d.) slow fading and additive Gaussian noise. The observed N sample vectors under PU-absent and PU-present hypotheses are respectively given by, for $n = 0, \ldots, N-1$,

$$
\begin{aligned}
&\mathcal{H}_0 : \mathbf{y}_n = \mathbf{z}_n, \ \text{PU-absent}, \\
&\mathcal{H}_1 : \mathbf{y}_n = \mathbf{s}_n + \mathbf{z}_n, \ \text{PU-present},
\end{aligned} \tag{4.1}
$$

where $\mathbf{y}_n = [Y_n^{(1)}, Y_n^{(2)}, \ldots, Y_n^{(L)}]'$ denotes the nth sample vector of the RMSs, $\mathbf{s}_n = [h_1, h_2, \ldots, h_L]' X_n$ represents the sample of the received multi-antenna PU signals in which X_n is a sequence of discrete-time zero-mean complex-valued cyclostationary random variables with variance σ_X^2 and h_l denotes the i.i.d. complex-valued fading channel gains, and $\mathbf{z}_n = [Z_n^{(1)}, Z_n^{(2)}, \ldots, Z_n^{(L)}]'$ is a sequence of the i.i.d. circularly-symmetric Gaussian noise sample vectors with distribution $\mathbf{z}_n \sim \mathcal{CN}(\mathbf{0}, \sigma_Z^2 \mathbf{I}_L)$. The noise samples are suitably assumed to be independent of the PU signal. The instantaneous SNR is defined as $\gamma_l \triangleq |h_l|^2 \, \sigma_X^2/\sigma_Z^2$ and σ_X^2/σ_Z^2 is assumed to be one throughout the paper. Let's first give the definition of being k**th-order almost-cyclostationary**.

Definition. *The sequence X_n is claimed to be kth-order almost-cyclostationary [29] if its kth-order cumulant $c_{kX}[n;\boldsymbol{\tau},\diamondsuit] \triangleq \mathrm{cum}\{X_n^{\diamondsuit_0}, X_{n+\tau_1}^{\diamondsuit_1}, \ldots, X_{n+\tau_{k-1}}^{\diamondsuit_{k-1}}\}$ complies with the Fourier-series expansion,*

$$c_{kX}\left[n;\boldsymbol{\tau},\diamondsuit\right] = \sum_{\alpha \in \mathcal{A}_k[\boldsymbol{\tau},\diamondsuit]} C_{kX}^{\alpha}\left[\boldsymbol{\tau},\diamondsuit\right] e^{j\alpha n}, \tag{4.2}$$

where the time lag $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_{k-1})$, $\diamondsuit_i \in \{*, \bar{*}\}$ indicates an optional complex conjugate, $\bar{*}$ denotes that there is no conjugate operation, $\diamondsuit = (\diamondsuit_0, \ldots, \diamondsuit_{k-1})$, $\mathcal{A}_k\left[\boldsymbol{\tau},\diamondsuit\right] = \{\alpha \in (-\pi,\pi];\, C_{kX}^{\alpha}\left[\boldsymbol{\tau},\diamondsuit\right] \neq 0\}$ is a set of cycle frequencies, and the Fourier coefficient

$$C_{kX}^{\alpha}\left[\boldsymbol{\tau},\diamondsuit\right] \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} c_{kX}\left[n;\boldsymbol{\tau},\diamondsuit\right] e^{-j\alpha n}. \tag{4.3}$$

Several standardized PU signals have been shown to be second-order almost cyclostationary with the Fourier coefficients $C_{2\mathrm{X}}^{\alpha}[\tau, \diamondsuit = (\bar{*}, *)]$ and $C_{2\mathrm{X}}^{\alpha(*)}[\tau, \diamondsuit = (\bar{*}, \bar{*})]$, known as the cyclic autocorrelation and the conjugate cyclic autocorrelation [31, 70]. For simplicity, the notation $\diamondsuit$ in these two coefficients will be omitted from now on.

Let the RMSs $\mathbf{y}_n$ be linearly combined, yielding the sequence $Y_n^{\mathrm{LC}} = \mathbf{w}'\mathbf{y}_n$ where

$\mathbf{w}' = [w_1, \ldots, w_L]$ is a complex-valued weight vector with its 2-norm $\|\mathbf{w}\|_2 = 1$. A mixing condition that facilitates the derivation of asymptotic normality is given by

$$\mathbf{A1} \quad \sum_{\boldsymbol{\tau}=-\infty}^{\infty} \sup_n |\tau_i| \left| \operatorname{cum}\left\{ X_n^{\diamondsuit_0}, X_{n+\tau_1}^{\diamondsuit_1}, \ldots, X_{n+\tau_k}^{\diamondsuit_k} \right\} \right| < \infty, \text{ for } 1 \le i \le k, \ \forall k \in \mathbb{N}_0, \tag{4.4}$$

where $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_k)$. The next lemma will show the relationship between X_n and $Y_n^{(l)}$(or Y_n^{LC}).

Lemma 1. *If the sequence X_n is kth-order almost-cyclostationary and satisfies the condition* **A1**, *then the same statistical properties are inherited by the sequences $Y_n^{(l)} = h_l X_n + Z_n^{(l)}$ for $1 \le l \le L$ and Y_n^{LC} under $\mathcal{H}_1$.*

Proof. See Section 4.A. □

Moreover, it can be easily shown that the sequences $Y_n^{(l)}$ and Y_n^{LC} exhibit cyclic autocorrelations given by

$$C_{2\text{Y}^{(l)}}^{\alpha}[\tau] = |h_l|^2 C_{2\text{X}}^{\alpha}[\tau], \tag{4.5}$$

and

$$C_{2\text{Y}^{\text{LC}}}^{\alpha}[\tau] = |\mathbf{w}'\mathbf{h}|^2 C_{2\text{X}}^{\alpha}[\tau], \tag{4.6}$$

where $\mathbf{h} = [h_1, h_2, \ldots, h_L]'$. This indicates that $C_{2\text{X}}^{\alpha}[\tau]$ is scaled up to a real number after applying linear combination to the RMSs $\mathbf{y}_n$.

Next, the pre-combining and the post-combining MASS will respectively exploit

$C^{\alpha}_{2\mathrm{Y}^{(l)}}[\tau]$ and $C^{\alpha}_{2\mathrm{Y}^{\mathrm{LC}}}[\tau]$ to form a chi-square test statistic by following the approach in [29]. Let $\mathcal{L} = \{\tau_1, \tau_2, \ldots \tau_M; M \in \mathbb{N}\}$ be a set of lags and α be a non-zero cycle frequency of interest. Although only single cycle frequency is considered here, the utilization of multiple cycle frequencies is also feasible. For pre-combining, a cyclic autocorrelation vector estimate can be constructed as

$$\begin{aligned}\hat{\mathbf{r}}_Y^{\mathrm{LC}} \triangleq\ & \left[\mathrm{Re}\left\{\hat{C}^{\alpha}_{2\mathrm{Y}^{\mathrm{LC}}}[\tau_1]\right\}, \ldots, \mathrm{Re}\left\{\hat{C}^{\alpha}_{2\mathrm{Y}^{\mathrm{LC}}}[\tau_M]\right\},\right. \\ & \left.\mathrm{Im}\left\{\hat{C}^{\alpha}_{2\mathrm{Y}^{\mathrm{LC}}}[\tau_1]\right\}, \ldots, \mathrm{Im}\left\{\hat{C}^{\alpha}_{2\mathrm{Y}^{\mathrm{LC}}}[\tau_M]\right\}\right]'. \end{aligned} \tag{4.7}$$

Similarly, for post-combining, we establish the vector estimate $\hat{\mathbf{r}}_Y^{(l)}$ based on $Y_n^{(l)}$, that is

$$\begin{aligned}\hat{\mathbf{r}}_Y^{(l)} \triangleq\ & \left[\mathrm{Re}\left\{\hat{C}^{\alpha}_{2\mathrm{Y}^{(l)}}[\tau_1]\right\}, \ldots, \mathrm{Re}\left\{\hat{C}^{\alpha}_{2\mathrm{Y}^{(l)}}[\tau_M]\right\},\right. \\ & \left.\mathrm{Im}\left\{\hat{C}^{\alpha}_{2\mathrm{Y}^{(l)}}[\tau_1]\right\}, \ldots, \mathrm{Im}\left\{\hat{C}^{\alpha}_{2\mathrm{Y}^{(l)}}[\tau_M]\right\}\right]'. \end{aligned} \tag{4.8}$$

$\hat{C}^{\alpha}_{2\mathrm{Y}^{\mathrm{LC}}}[\tau]$ and $\hat{C}^{\alpha}_{2\mathrm{Y}^{(l)}}[\tau]$ are respectively the consistent estimates of $C^{\alpha}_{2\mathrm{Y}^{\mathrm{LC}}}[\tau]$ and $C^{\alpha}_{2\mathrm{Y}^{(l)}}[\tau]$, that is, given N observations,

$$\hat{C}^{\alpha}_{2\mathrm{Y}^{\mathrm{LC}}}[\tau] \triangleq \frac{1}{N} \sum_{n=\max\{0,-\tau\}}^{N+\min\{0,-\tau\}-1} Y_n^{\mathrm{LC}} Y_{n+\tau}^{\mathrm{LC}*} e^{-j\alpha n}, \tag{4.9}$$

and

$$\hat{C}^{\alpha}_{2\mathrm{Y}^{(l)}}[\tau] \triangleq \frac{1}{N} \sum_{n=\max\{0,-\tau\}}^{N+\min\{0,-\tau\}-1} Y_n^{(l)} Y_{n+\tau}^{(l)*} e^{-j\alpha n}. \tag{4.10}$$

Making use of Lemma 1 and asymptotic normality shown in [29] yields the asymptotic distributions of $\sqrt{N}\hat{\mathbf{r}}_Y^{\mathrm{LC}}$ and $\sqrt{N}\hat{\mathbf{r}}_Y^{(l)}$, i.e., $\mathcal{N}(\sqrt{N}|\mathbf{w}'\mathbf{h}|^2\mathbf{r}_X, \Sigma_Y^{\mathrm{LC}})$ and $\mathcal{N}(\sqrt{N}|h_l|^2\mathbf{r}_X, \Sigma_Y^{(l,l)})$, respectively, where $h_l \triangleq 0$ under $\mathcal{H}_0$, Σ_Y^{LC} and $\Sigma_Y^{(l,l)}$ are the limiting autocovariance matrices of $\sqrt{N}\hat{\mathbf{r}}_Y^{\text{LC}}$ and $\sqrt{N}\hat{\mathbf{r}}_Y^{(l)}$, and

$$\begin{aligned}\mathbf{r}_X &= [\text{Re}\{C_{2\text{x}}^{\alpha}[\tau_1]\}, \ldots, \text{Re}\{C_{2\text{x}}^{\alpha}[\tau_M]\}, \\ &\quad \text{Im}\{C_{2\text{x}}^{\alpha}[\tau_1]\}, \ldots, \text{Im}\{C_{2\text{x}}^{\alpha}[\tau_M]\}]'. \end{aligned} \tag{4.11}$$

The method of obtaining the consistent estimates $\hat{\Sigma}_Y^{\text{LC}}$ and $\hat{\Sigma}_Y^{(l,l)}$ of Σ_Y^{LC} and $\Sigma_Y^{(l,l)}$ is detailed in [29, 42]. Based on $\hat{\mathbf{r}}_Y^{\text{LC}}$ and $\hat{\Sigma}_Y^{\text{LC}}$, we can perform the GRLT for the pre-combining MASS as follows

$$\begin{aligned} L_G &= \frac{f\left(\sqrt{N}\hat{\mathbf{r}}_Y^{\text{LC}}; |\mathbf{w}'\mathbf{h}|^2\mathbf{r}_X = \hat{\mathbf{r}}_Y^{\text{LC}}, \Sigma_Y^{\text{LC}} = \hat{\Sigma}_Y^{\text{LC}}\right)}{f\left(\sqrt{N}\hat{\mathbf{r}}_Y^{\text{LC}}; |\mathbf{w}'\mathbf{h}|^2\mathbf{r}_X = \mathbf{0}, \Sigma_Y^{\text{LC}} = \hat{\Sigma}_Y^{\text{LC}}\right)}, \\ &= \frac{\exp\left[-\frac{1}{2}N\left(\hat{\mathbf{r}}_Y^{\text{LC}} - \hat{\mathbf{r}}_Y^{\text{LC}}\right)'\hat{\Sigma}_Y^{\text{LC}-1}\left(\hat{\mathbf{r}}_Y^{\text{LC}} - \hat{\mathbf{r}}_Y^{\text{LC}}\right)\right]}{\exp\left[-\frac{1}{2}N\left(\hat{\mathbf{r}}_Y^{\text{LC}}\right)'\hat{\Sigma}_Y^{\text{LC}-1}\left(\hat{\mathbf{r}}_Y^{\text{LC}}\right)\right]} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \end{aligned} \tag{4.12}$$

which leads to the final test statistic

$$T_{\text{LC}} = N\hat{\mathbf{r}}_Y^{\text{LC}'}\left(\hat{\Sigma}_Y^{\text{LC}}\right)^{-1}\hat{\mathbf{r}}_Y^{\text{LC}}. \tag{4.13}$$

The derivation of the test statistics for the post-combining MASS is postponed to Section 4.4.

Another lemma concerning the product sequence $\psi_{n,\tau}^{(l)} \triangleq Y_n^{(l)}Y_{n+\tau}^{(l)*}$ for fixed τ and $1 \le l \le L$ is provided below.

Lemma 2. *If the sequence X_n is almost-cyclostationary up to fourth order and satisfies the condition* **A1***, then the sequence $\psi_{n,\tau}^{(l)}$ under $\mathcal{H}_1$ is second-order almost-cyclostationary and satisfies another mixing condition,*

A2 $\sum_{\boldsymbol{\xi}=-\infty}^{\infty} \sup_n |\xi_i| \left| \mathrm{cum}\left\{ \psi_{n,\tau_0}^{(l_0)\diamondsuit_0}, \psi_{n+\xi_1,\tau_1}^{(l_1)\diamondsuit_1}, \ldots, \psi_{n+\xi_k,\tau_k}^{(l_k)\diamondsuit_k} \right\} \right| < \infty,$

$$\textit{for } 1 \leq i \leq k, \ \forall k \in \mathbb{N}_0, \quad (4.14)$$

where $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_k)$.

Proof. See Section 4.B. □

Due to this lemma, another cyclic statistic, the cyclic spectrum of $Y_n^{(l)}$, can be defined as [71]

$$S_{\psi_{\tau,\rho}^{(l_1,l_2)}}(\alpha;\omega) \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{\xi=-\infty}^{\infty} \mathrm{cum}\left\{ \psi_{n,\tau}^{(l_1)}, \psi_{n+\xi,\rho}^{(l_2)} \right\} e^{-j\alpha n} e^{-j\omega\xi}. \quad (4.15)$$

Similarly, the definition of the conjugate cyclic spectrum is given by

$$\tilde{S}_{\psi_{\tau,\rho}^{(l_1,l_2)}}(\alpha;\omega) \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{\xi=-\infty}^{\infty} \mathrm{cum}\left\{ \psi_{n,\tau}^{(l_1)}, \psi_{n+\xi,\rho}^{(l_2)*} \right\} e^{-j\alpha n} e^{-j\omega\xi}. \quad (4.16)$$

Later on, these cyclic spectra will be shown to be related to the evaluation of $\Sigma_Y^{(l,l)}$.

4.3 Pre-Combining Scheme

In the pre-combining scheme, the RMSs are first linearly combined and further utilized to obtain the vector $\hat{\mathbf{r}}_Y^{\mathrm{LC}}$ and the matrix $\hat{\Sigma}_Y^{\mathrm{LC}}$. One possible choice of the weight vector $\mathbf{w}$ is using MRC when CSI is avaliable. Conventionally, MRC can provide either diversity gains or power gains, supporting reliable communications [72]. In the case of spectrum sensing, this combining technique can improve the detection performance by maximizing the output SNR, $\mathrm{SNR}_o \triangleq \mathrm{Var}(\mathbf{w}'\mathbf{s}_n)/\mathrm{Var}(\mathbf{w}'\mathbf{z}_n)$, under $\mathcal{H}_1$

without changing the statistical properties of noise under $\mathcal{H}_0$. So it has been applied to either energy-based or cyclostationary-feature-based MASS [69, 73]. The following discussion will justify the usage of MRC in our proposed pre-combining scheme from another viewpoint. Meanwhile, CSI, required for implementing MRC, will be acquired by our proposed estimation procedure.

4.3.1 Usage of MRC

The distribution of the test statistic T_{LC} resulting from the linearly combined sequence Y_n^{LC} is provided below [29, 47]

$$\begin{aligned} \mathcal{H}_0: \quad T_{\text{LC}} \sim & \ \chi^2_{2M}, \\ \mathcal{H}_1: \quad T_{\text{LC}} \sim & \ \chi^2_{2M}\left(N|\mathbf{w}'\mathbf{h}|^4 \mathbf{r}'_X \left(\Sigma_Y^{\text{LC}}\right)^{-1} \mathbf{r}_X\right). \end{aligned}$$

where χ^2_{2M} denotes the central chi-square distribution with $2M$ degrees of freedom and $\chi^2_{2M}(\varsigma(T_{\text{LC}}))$ represents the non-central chi-square distribution with $2M$ degrees of freedom and the noncentrality parameter $\varsigma(T_{\text{LC}}) = N|\mathbf{w}'\mathbf{h}|^4\mathbf{r}'_X(\Sigma_Y^{\text{LC}})^{-1}\mathbf{r}_X$. The weight vector $\mathbf{w}$ is said to optimize the detection performance in the sense of maximizing a modified deflection coefficient (MDC) [74]

$$\begin{aligned} d_m^2\left(T_{\text{LC}}\right) & \triangleq \frac{\left[\mathbb{E}\left(T_{\text{LC}}|\mathcal{H}_1\right) - \mathbb{E}\left(T_{\text{LC}}|\mathcal{H}_0\right)\right]^2}{\text{Var}\left(T_{\text{LC}}|\mathcal{H}_1\right)}, \\ & = \frac{\varsigma\left(T_{\text{LC}}\right)^2}{2\left[2M + 2\varsigma\left(T_{\text{LC}}\right)\right]}, \end{aligned} \tag{4.17}$$

which is equivalent to maximizing $\varsigma(T_{\text{LC}})$. This MDC has been shown as a generalized SNR and a good detection performance measure [75, 76]. Because of [51, Proposition], the matrix $(\Sigma_Y^{\mathrm{LC}})^{-1}$ in the low SNR region can be viewed as being independent of $\mathbf{w}$. Hence, the final metric to be maximized is $|\mathbf{w}'\mathbf{h}|^2$ whose maximum is achieved by using $\mathbf{w}'_{\mathrm{opt}} = {e^{j\theta}\mathbf{h}^H}/{\|\mathbf{h}\|_2}\ \forall\theta \in [0, 2\pi)$. The MRC weight vector, $\mathbf{w}'_{\mathrm{MRC}} = {\mathbf{h}^H}/{\|\mathbf{h}\|_2}$, is one of optimal choices. This justifies the usage of MRC in the most critical SNR region where our proposed pre-combining scheme is intended for enhancing performance. On the other hand, the matrix $(\Sigma_Y^{\mathrm{LC}})^{-1}$ in the high SNR region can be well approximated by $|\mathbf{w}'\mathbf{h}|^{-4}\Sigma_X^{-1}$ where the matrix Σ_X, resulting from the sequence X_n, is independent of $\mathbf{w}$. This makes $\varsigma(T_{\mathrm{LC}}) = N\mathbf{r}'_X\Sigma_X^{-1}\mathbf{r}_X$ being constant no matter what weight vector $\mathbf{w}$ is used. It implies that choosing the proper wight vector becomes less important as the SNR increases.

4.3.2 Blind Channel Estimation

Perfect CSI with which the presence or the absence of PUs becomes unambiguous is not a practical assumption for spectrum sensing. Without the aid of PUs, two approaches have been proposed for blindly estimating channel information based on cyclostationary features. The first approach [77] utilizes spectral correlations to acquire phase information and performs equal gain combining in the frequency domain. The second [68] directly estimates the channel gains up to a phase rotation by using the cyclic crosscorrelations. Here, we recast this second approach with moderate modification and keep its original name as blind MRC (BMRC).

Let's define the cyclic cross-correlation of the received signals as

$$\begin{aligned} C^{\alpha}_{Y^{(l_1,l_2)}}[\tau] &\triangleq \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\text{cum}\left\{Y_n^{(l_1)},Y_{n+\tau}^{(l_2)*}\right\}e^{-j\alpha n}, \\ &= h_{l_1}h_{l_2}^{*}C^{\alpha}_{2X}[\tau]. \end{aligned} \tag{4.18}$$

Since $C^{\alpha}_{2X}[\tau]$ is known a priori, the quantity $h_{l_1}h^*_{l_2}$ can be adequately evaluated by utilizing a consistent estimate $\hat{C}^{\alpha}_{Y^{(l_1,l_2)}}[\tau] \triangleq \frac{1}{N}\sum_{n=\max\{0,-\tau\}}^{N+\min\{0,-\tau\}-1} Y_n^{(l_1)}Y_{n+\tau}^{(l_2)*}e^{-j\alpha n}$. Consider a $M(L^2+L)\times 1$ estimated vector

$$\begin{aligned} \hat{\mathbf{r}}_Y^{\text{CE}} \triangleq\ & \left[\text{Re}\left\{\hat{\mathbf{r}}^{\text{CE}}_{Y,1}\right\},\text{Re}\left\{\hat{\mathbf{r}}^{\text{CE}}_{Y,2}\right\},\ldots,\text{Re}\left\{\hat{\mathbf{r}}^{\text{CE}}_{Y,M}\right\}\right. \\ & \left.\text{Im}\left\{\hat{\mathbf{r}}^{\text{CE}}_{Y,1}\right\},\text{Im}\left\{\hat{\mathbf{r}}^{\text{CE}}_{Y,2}\right\},\ldots,\text{Im}\left\{\hat{\mathbf{r}}^{\text{CE}}_{Y,M}\right\}\right]', \end{aligned} \tag{4.19}$$

where

$$\begin{aligned} \hat{\mathbf{r}}^{\text{CE}}_{Y,i} =\ & \left[\hat{C}^{\alpha}_{Y^{(1,1)}}[\tau_i],\hat{C}^{\alpha}_{Y^{(1,2)}}[\tau_i],\ldots,\hat{C}^{\alpha}_{Y^{(1,L)}}[\tau_i],\right. \\ & \hat{C}^{\alpha}_{Y^{(2,2)}}[\tau_i],\hat{C}^{\alpha}_{Y^{(2,3)}}[\tau_i],\ldots,\hat{C}^{\alpha}_{Y^{(2,L)}}[\tau_i],\ldots, \\ & \left.\hat{C}^{\alpha}_{Y^{(L-1,L-1)}}[\tau_i],\hat{C}^{\alpha}_{Y^{(L-1,L)}}[\tau_i],\hat{C}^{\alpha}_{Y^{(L,L)}}[\tau_i]\right]_{1\times\frac{(L^2+L)}{2}}. \end{aligned}$$

We decompose it as

$$\hat{\mathbf{r}}^{\text{CE}}_Y = \mathbf{R}^{\text{CE}}\mathbf{h}_{\text{CP}} + \boldsymbol{\epsilon}, \tag{4.20}$$

where the $(L^2+L)\times 1$ cross-product channel gain vector $\mathbf{h}_{\text{CP}} = [\text{Re}\{\mathbf{h}_{\text{cp}}\},\text{Im}\{\mathbf{h}_{\text{cp}}\}]'$ with the sub-vector

$$\mathbf{h}_{\text{cp}} = [h_{1,1},\ldots,h_{1,L},h_{2,2},\ldots,h_{2,L},\ldots,h_{L-1,L-1},h_{L-1,L},h_{L,L}]_{1\times\frac{(L^2+L)}{2}},$$

$$h_{l_1,l_2} \triangleq h_{l_1}h^*_{l_2},$$

the residual vector $\boldsymbol{\epsilon} = (\hat{\mathbf{r}}^{\text{CE}}_Y - \mathbf{R}^{\text{CE}}\mathbf{h}_{\text{CP}})$, and the $M(L^2+L)\times(L^2+L)$ matrix

$\mathbf{R}^{\text{CE}} = [\mathbf{R}^{\text{CE}}_{1,1}; \mathbf{R}^{\text{CE}}_{1,2}; \cdots; \mathbf{R}^{\text{CE}}_{1,M}; \mathbf{R}^{\text{CE}}_{2,1}; \mathbf{R}^{\text{CE}}_{2,2}; \cdots; \mathbf{R}^{\text{CE}}_{2,M}]$ where

$$\mathbf{R}^{\text{CE}}_{1,i} = \begin{bmatrix} \operatorname{Re}\left\{\mathbf{R}^{\text{CE}}_{i,\mathbb{C}}\right\} & -\operatorname{Im}\left\{\mathbf{R}^{\text{CE}}_{i,\mathbb{C}}\right\} \end{bmatrix}_{\frac{(L^2+L)}{2}\times(L^2+L)}, \tag{4.21}$$

$$\mathbf{R}^{\text{CE}}_{2,i} = \begin{bmatrix} \operatorname{Im}\left\{\mathbf{R}^{\text{CE}}_{i,\mathbb{C}}\right\} & \operatorname{Re}\left\{\mathbf{R}^{\text{CE}}_{i,\mathbb{C}}\right\} \end{bmatrix}_{\frac{(L^2+L)}{2}\times(L^2+L)}, \tag{4.22}$$

and

$$\mathbf{R}^{\text{CE}}_{i,\mathbb{C}} = C^{\alpha}_{2X}\left[\tau_i\right]\mathbf{I}_{\frac{(L^2+L)}{2}}. \tag{4.23}$$

In view of this decomposition, we have the following channel estimation procedure.

Step 1 Obtain the estimated vector $\hat{\mathbf{r}}^{\text{CE}}_{Y}$.

Step 2 Compute the least-squares estimate of $\mathbf{h}_{\text{CP}}$ given by

$$\hat{\mathbf{h}}_{\text{CP}} = [(\mathbf{R}^{\text{CE}})'\,\mathbf{R}^{\text{CE}}]^{-1}\,(\mathbf{R}^{\text{CE}})'\,\hat{\mathbf{r}}^{\text{CE}}_{Y}. \tag{4.24}$$

Step 3 Form a matrix estimate $\hat{\mathbf{H}}$ of $\mathbf{H} = \mathbf{h}\mathbf{h}^H$ by using the available estimates of real and image parts of $h_{l_1}h^*_{l_2}$, that is

$$\hat{\mathbf{H}} = \begin{bmatrix} \operatorname{Re}\{\hat{h}_{1,1}\} & \hat{h}_{1,2} & \cdots & \hat{h}_{1,L} \\ \hat{h}^*_{1,2} & \operatorname{Re}\{\hat{h}_{2,2}\} & \cdots & \hat{h}_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{h}^*_{1,L} & \hat{h}^*_{2,L} & \cdots & \operatorname{Re}\{\hat{h}_{L,L}\} \end{bmatrix}. \tag{4.25}$$

Step 4 Let $\hat{\mathbf{h}}$ be the normalized eigenvector corresponding to the maximum eigenvalue of $\hat{\mathbf{H}}$.

Then, $\hat{\mathbf{h}}$ will be an estimate of ${}^{\mathbf{h}}/{\|\mathbf{h}\|_2}$ up to some phase rotation, i.e., ${}^{e^{j\theta}\mathbf{h}^H}/{\|\mathbf{h}\|_2}$ for some $\theta \in [0, 2\pi)$. According to the previous discussion, using $\mathbf{w} = \hat{\mathbf{h}}$ will achieve the desired result.

The matrix $\mathbf{H}$ is Hermitian and rank-one, so there exists a decomposition $\mathbf{H} = \sum_{i=1}^{L} \lambda_i \mathbf{u}_i \mathbf{u}_i^H$ where λ_i are eigenvalues and $\mathbf{u}_i$ are orthonormal eigenvectors. Let's define the first eigenpair $(\lambda_1, \mathbf{u}_1) = (\|\mathbf{h}\|_2^2, \mathbf{h}/\|\mathbf{h}\|_2)$ and $\lambda_{i\neq 1} = 0$. The matrix $\hat{\mathbf{H}}$ with its eigenvalues $\tilde{\lambda}_i$ can be presented as $\mathbf{H} + \delta_{\mathbf{H}}$ where $\delta_{\mathbf{H}}$ is a perturbation matrix. The eigenvector corresponding to $\tilde{\lambda}_{\max} = \max\{\tilde{\lambda}_i, 1 \leq i \leq L\}$ is denoted by $\tilde{\mathbf{u}}_{\max}$ when the algebraic multiplicity of $\tilde{\lambda}_{\max}$ is one. The following lemma is going to validate the proposed estimation procedure by showing that the perturbed eigenpair $(\tilde{\lambda}_{\max}, \tilde{\mathbf{u}}_{\max})$ asymptotically approaches $(\lambda_1, \mathbf{u}_1)$. This convergence property is based on the facts that both $\mathbf{H}$ and $\delta_{\mathbf{H}}$ are Hermitian, and the eigenpair $(\lambda_1, \mathbf{u}_1)$ is simple.

Proposition 2. *As the sample size N goes to infinity, the algebraic multiplicity of $\tilde{\lambda}_{\max}$ is one and both quantities $|\tilde{\lambda}_{\max} - \lambda_1|$ and $||\tilde{\mathbf{u}}_{\max} - \mathbf{u}_1||_2$ decrease toward zero.*

Proof. See Section 4.C. □

The quality of the estimated CSI $\hat{\mathbf{h}}$ can be quantified by the angle distance between $\hat{\mathbf{h}}$ and $\mathbf{h}$ [78, 79], that is

$$\angle\left(\hat{\mathbf{h}}, \mathbf{h}\right) = \cos^{-1} \frac{\left|\hat{\mathbf{h}}^H \mathbf{h}\right|}{\|\mathbf{h}\|_2}. \tag{4.26}$$

This angle distance $\angle(\hat{\mathbf{h}}, \mathbf{h})$ is zero if $\hat{\mathbf{h}}$ is equal to $\mathbf{h}/\|\mathbf{h}\|_2$ up to a phase rotation. In simulation results, $\angle(\hat{\mathbf{h}}, \mathbf{h})$ will be used to illustrate the effectiveness of the proposed channel estimation procedure.

4.4 Post-Combining Scheme

Unlike the pre-combining scheme, the received signal $Y_n^{(l)}$ at the lth antenna branch is directly utilized to form a vector estimate $\hat{\mathbf{r}}_Y^{(l)}$ in the post-combining scheme. Moreover, CSI is not necessary in this scheme. The following will provide two possible approaches to exploiting $\hat{\mathbf{r}}_Y^{(l)}$ for $1 \leq l \leq L$.

4.4.1 Joint Combining

Let's arrange the L cyclic autocorrelation vector estimates $\hat{\mathbf{r}}_Y^{(l)}$ into one vector $\hat{\mathbf{r}}_Y^{\mathrm{JC}} \triangleq [\hat{\mathbf{r}}_Y^{(1)'}, \hat{\mathbf{r}}_Y^{(2)'}, \ldots, \hat{\mathbf{r}}_Y^{(L)'}]'$. Though joint asymptotic normality for each vector $\hat{\mathbf{r}}_Y^{(l)}$ has been established, it needs to be verified that the same property is held by $\hat{\mathbf{r}}_Y^{\mathrm{JC}}$.

Proposition 3. *If the second-order almost-cyclostationary sequence X_n satisfies the mixing condition **A1**, then the scaled vector $\sqrt{N}\hat{\mathbf{r}}_Y^{JC}$ is asymptotically normally distributed $\mathcal{N}\left(\sqrt{N}\mathbf{R}_X\underline{\gamma}, \Sigma_Y^{JC}\right)$, where $\underline{\gamma} \triangleq [\gamma_1, \gamma_2, \ldots, \gamma_L]'$,*

$$\mathbf{R}_X \triangleq \begin{bmatrix} \mathbf{r}_X & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{r}_X & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{r}_X \end{bmatrix}_{2ML \times L}, \tag{4.27}$$

and

$$\Sigma_Y^{JC} \triangleq \begin{bmatrix} \Sigma_Y^{(1,1)} & \Sigma_Y^{(1,2)} & \cdots & \Sigma_Y^{(1,L)} \\ \Sigma_Y^{(2,1)} & \Sigma_Y^{(2,2)} & \cdots & \Sigma_Y^{(2,L)} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_Y^{(L,1)} & \Sigma_Y^{(L,2)} & \cdots & \Sigma_Y^{(L,L)} \end{bmatrix}_{2ML\times 2ML}. \tag{4.28}$$

The $2M \times 2M$ limiting covariance matrix $\Sigma_Y^{(l_1,l_2)} \triangleq \lim_{N\to\infty} \mathrm{Cov}(\sqrt{N}\hat{\mathbf{r}}_Y^{(l_1)}, \sqrt{N}\hat{\mathbf{r}}_Y^{(l_2)})$ can be written as

$$\Sigma_Y^{(l_1,l_2)} = \begin{bmatrix} Re\left\{\frac{\mathbf{Q}^{(l_1,l_2)}+\tilde{\mathbf{Q}}^{(l_1,l_2)}}{2}\right\} & Im\left\{\frac{\mathbf{Q}^{(l_1,l_2)}-\tilde{\mathbf{Q}}^{(l_1,l_2)}}{2}\right\} \\ Im\left\{\frac{\mathbf{Q}^{(l_1,l_2)}+\tilde{\mathbf{Q}}^{(l_1,l_2)}}{2}\right\} & Re\left\{\frac{-\mathbf{Q}^{(l_1,l_2)}+\tilde{\mathbf{Q}}^{(l_1,l_2)}}{2}\right\} \end{bmatrix}, \tag{4.29}$$

where $\mathbf{Q}^{(l_1,l_2)}$ and $\tilde{\mathbf{Q}}^{(l_1,l_2)}$ are two $M \times M$ matrices with their (i,j)th entries,

$$\mathbf{Q}_{i,j}^{(l_1,l_2)} = S_{\psi_{\tau_i,\tau_j}^{(l_1,l_2)}}(2\alpha;\alpha), \tag{4.30}$$

$$\tilde{\mathbf{Q}}_{i,j}^{(l_1,l_2)} = \tilde{S}_{\psi_{\tau_i,\tau_j}^{(l_1,l_2)}}(0;-\alpha). \tag{4.31}$$

Proof. See Section 4.D. □

Note that this proposition generalizes [29, Theorem 1] which takes account of the single receive antenna case ($L = 1$). Corresponding to this generalization, the general estimates of $\mathbf{Q}_{i,j}^{(l_1,l_2)}$ and $\tilde{\mathbf{Q}}_{i,j}^{(l_1,l_2)}$ are provided below.

Proposition 4. *Assume that the sequence X_n is almost-cyclostationary up to fourth order and satisfies the condition* **A1**. *The mean-square-sense consistent estimates of the unconjugated and conjugate cyclic spectra of $\psi_{n,\tau}^{(l)}$ are given by*

$$\hat{S}_{\psi_{\tau,\rho}^{(l_1,l_2)}}(\alpha;\omega) = \frac{1}{NS}\sum_{s=-(S-1)/2}^{(S-1)/2} G(s)\, F_{\tau}^{(l_1)}\left(\alpha-\omega+\frac{2\pi s}{N}\right) F_{\rho}^{(l_2)}\left(\omega-\frac{2\pi s}{N}\right), \quad (4.32)$$

and

$$\hat{\tilde{S}}_{\psi_{\tau,\rho}^{(l_1,l_2)}}(\alpha;\omega) \;=\; \frac{1}{NS}\sum_{s=-(S-1)/2}^{(S-1)/2} G(s)\, F_{\tau}^{(l_1)}\left(\alpha-\omega+\frac{2\pi s}{N}\right) F_{\rho}^{(l_2)*}\left(-\omega+\frac{2\pi s}{N}\right), \quad (4.33)$$

where $F_{\tau}^{(l)}(\omega) = \sum_{n=0}^{N-1}(\psi_{n,\tau}^{(l)} - \mathbb{E}[\psi_{n,\tau}^{(l)}])e^{-j\omega n}$ *and* $G(s)$ *is an odd-length* (S) *smoothing window.*

Proof. The sequences $\psi_{n,\tau}^{(l)}$ for $1 \le l \le L$ can be viewed as the vector-valued time series, having cyclic spectra defined by (4.15) and (4.16). As this vector-valued time series satisfies [71, Assumption 1] due to Lemma 2, the consistency of the proposed cyclic spectrum estimates $\hat{S}_{\psi_{\tau,\rho}^{(l_1,l_2)}}(\alpha;\omega)$ and $\hat{\tilde{S}}_{\psi_{\tau,\rho}^{(l_1,l_2)}}(\alpha;\omega)$ has been shown in [71] for the real-valued $\psi_{n,\tau}^{(l)}$. To show the consistency for the complex-valued case requires some alterations which can be found in [80]. $\square$

An immediate corollary of this proposition is:

Corollary. *The consistent estimates of* $\mathbf{Q}_{i,j}^{(l_1,l_2)}$ *and* $\tilde{\mathbf{Q}}_{i,j}^{(l_1,l_2)}$ *are given by*

$$\hat{\mathbf{Q}}_{i,j}^{(l_1,l_2)} = \hat{S}_{\psi_{\tau_i,\tau_j}^{(l_1,l_2)}}(2\alpha;\alpha), \quad (4.34)$$

$$\hat{\tilde{\mathbf{Q}}}_{i,j}^{(l_1,l_2)} = \hat{\tilde{S}}_{\psi_{\tau_i,\tau_j}^{(l_1,l_2)}}(0;-\alpha). \quad (4.35)$$

Consider that prior knowledge of the conjugate cyclic autocorrelation vector $\mathbf{r}_X$ is acquired by the CU, while the instantaneous SNR vector $\underline{\gamma}$ is unknown. Performing

GLRT by using Proposition 3 can yield the test statistic

$$T_{\text{JC}} = N\hat{\mathbf{r}}_Y^{\text{JC}\prime} \left(\hat{\Sigma}_Y^{\text{JC}}\right)^{-1} \hat{\mathbf{r}}_Y^{\text{JC}}, \tag{4.36}$$

where $\hat{\Sigma}_Y^{\text{JC}}$ is the estimate of Σ_Y^{JC} by exploiting (4.34) and (4.35). The corresponding asymptotic distribution of T_{JC} is given by

$$\begin{aligned} \mathcal{H}_0 : \quad T_{\text{JC}} \sim \quad & \chi^2_{2ML}, \\ \mathcal{H}_1 : \quad T_{\text{JC}} \sim \quad & \chi^2_{2ML}\left(N\underline{\gamma}'\mathbf{R}_X' \left(\Sigma_Y^{\text{JC}}\right)^{-1} \mathbf{R}_X\underline{\gamma}\right). \end{aligned} \tag{4.37}$$

4.4.2 Sum Combining

Let's add cyclic autocorrelation vector estimates $\hat{\mathbf{r}}_Y^{(l)}$ for $1 \leq l \leq L$ to be $\hat{\mathbf{r}}_Y^{\text{SUM}} \triangleq \sum_{l=1}^{L} \hat{\mathbf{r}}_Y^{(l)}$. By utilizing [53, Corollary 1.7], [81, Lemma 2.3.2], and the relationship $\hat{\mathbf{r}}_Y^{\text{SUM}} = \mathbf{D}\hat{\mathbf{r}}_Y^{\text{JC}}$ where $\mathbf{D} = [\mathbf{I}_{2M}, \mathbf{I}_{2M}, \ldots, \mathbf{I}_{2M}]_{2M\times 2ML}$, the asymptotic distribution of $\sqrt{N}\hat{\mathbf{r}}_Y^{\text{SUM}}$ can be given by $\mathcal{N}(\sqrt{N}\gamma^{\text{SUM}}\mathbf{r}_X, \Sigma_Y^{\text{SUM}})$ where $\gamma^{\text{SUM}} = \sum_{l=1}^{L}\gamma_l$ and $\Sigma_Y^{\text{SUM}} = \sum_{l_1,l_2=1}^{L}\Sigma_Y^{(l_1,l_2)}$. Hence, the GRLT statistic can be presented as

$$T_{\text{SUM}} = N\hat{\mathbf{r}}_Y^{\text{SUM}\prime} \left(\hat{\Sigma}_Y^{\text{SUM}}\right)^{-1} \hat{\mathbf{r}}_Y^{\text{SUM}}, \tag{4.38}$$

where $\hat{\Sigma}_Y^{\text{SUM}}$ is an estimate of Σ_Y^{SUM}. The distribution of T_{SUM} is provided as

$$\begin{aligned} \mathcal{H}_0 : \quad T_{\text{SUM}} \sim \quad & \chi^2_{2M}, \\ \mathcal{H}_1 : \quad T_{\text{SUM}} \sim \quad & \chi^2_{2M}\left(N\gamma^{\text{SUM}\prime}\mathbf{r}_X' \left(\Sigma_Y^{\text{SUM}}\right)^{-1} \mathbf{r}_X\gamma^{\text{SUM}}\right). \end{aligned} \tag{4.39}$$

When Σ_Y^{SUM} is decomposed as in (4.29) in which the superscript (l_1, l_2) is replaced by SUM, the corresponding consistent estimate of $\mathbf{Q}_{i,j}^{\text{SUM}}$ is given by $\hat{\mathbf{Q}}_{i,j}^{\text{SUM}} = \hat{S}_{\psi_{\tau_i,\tau_j}^{\text{SUM}}}(2\alpha; \alpha)$ where $\psi_{n,\tau}^{\text{SUM}} \triangleq \sum_{l=1}^{L} Y_n^{(l)} Y_{n+\tau}^{(l)*}$ and

$$\hat{S}_{\psi_{\tau,\rho}^{\text{SUM}}}(\alpha; \omega) = \frac{1}{NS} \sum_{s=-(S-1)/2}^{(S-1)/2} G(s)\, F_\tau^{\text{SUM}}\left(\alpha - \omega + \frac{2\pi s}{N}\right) F_\rho^{\text{SUM}}\left(\omega - \frac{2\pi s}{N}\right), \quad (4.40)$$

in which

$$F_\tau^{\text{SUM}}(\omega) = \sum_{n=0}^{N-1} \left(\psi_{n,\tau}^{\text{SUM}} - \mathbb{E}\left[\psi_{n,\tau}^{\text{SUM}}\right]\right) e^{-j\omega n}. \quad (4.41)$$

The other estimate $\hat{\tilde{\mathbf{Q}}}_{i,j}^{\text{SUM}}$ of $\tilde{\mathbf{Q}}_{i,j}^{\text{SUM}}$ can be defined in the same way. Thus, the evaluation of $\hat{\Sigma}_Y^{\text{SUM}}$ can be performed.

Compared with joint combining, sum combining requires less computational complexity due to the smaller size of covariance matrix Σ_Y^{SUM} that needs to be estimated. In addition, sum combining dose not necessarily lead to larger modified deflection coefficient $d_m^2(T_{\text{SUM}})$. For instance, let $\Sigma_Y^{(l_1,l_2)} = \mathbf{I}_{2M}$ if $l_1 = l_2$ and $\Sigma_Y^{(l_1,l_2)} = \mathbf{0}$ if $l_1 \neq l_2$. This is the case with the low SNR region and $\tau_i \neq 0$ for $1 \leq i \leq M$. Then the inequality,

$$d_m^2(T_{\text{SUM}}) \quad = \quad \frac{\varsigma(T_{\text{SUM}})^2}{2\left[2M + 2\varsigma(T_{\text{SUM}})\right]} \quad \geq \quad \frac{\varsigma(T_{\text{JC}})^2}{2\left[2ML + 2\varsigma(T_{\text{JC}})\right]} \quad = \quad d_m^2(T_{\text{JC}}), \quad (4.42)$$

where $\varsigma(T_{\text{SUM}}) = {}^N\!/_L(\sum_{l=1}^{L}\gamma_l)^2 \mathbf{r}_X' \mathbf{r}_X$ and $\varsigma(T_{\text{JC}}) = N(\sum_{l=1}^{L}\gamma_l^2)\mathbf{r}_X'\mathbf{r}_X$, does not always hold because of $(\sum_{l=1}^{L}\gamma_l)^2 \leq L(\sum_{l=1}^{L}\gamma_l^2)$. In this sense, better detection performance is not guaranteed by using either the sum combining or the joint combining.

4.5 Cochannel Interference Immunity

In this section, we are interested in conditions under which the previous analysis will not alter in the presence of cochannel interference from other CUs. Let's assume that there are two CU cochannel interference sources $X_n^{(I1)}$ and $X_n^{(I2)}$, which are statistically independent of each other, PU signals, and noise. The hypotheses to be considered are

$$\begin{aligned}
&\mathcal{H}_0 : \mathbf{y}_n = \mathbf{z}_n, \text{ noise-only},\\
&\mathcal{H}_1 : \mathbf{y}_n = \mathbf{s}_n^{(I1)} + \mathbf{s}_n^{(I2)} + \mathbf{z}_n, \text{ interferece-only},\\
&\mathcal{H}_2 : \mathbf{y}_n = \mathbf{s}_n^{(I0)} + \mathbf{s}_n^{(I1)} + \mathbf{s}_n^{(I2)} + \mathbf{z}_n, \text{ PU plus interferece},\\
&\mathcal{H}_3 : \mathbf{y}_n = \mathbf{s}_n^{(I0)} + \mathbf{z}_n, \text{ only PU present},
\end{aligned}$$

where $\mathbf{s}_n^{(I0)} = \mathbf{h}X_n^{(I0)}$ replaces the notation $\mathbf{s}_n = \mathbf{h}X_n$, $\mathbf{s}_n^{(Ii)} = \mathbf{h}X_n^{(Ii)}$ for $i \in \{1, 2\}$ represents the sample of the received multi-antenna cochannel interference, and $X_n^{(Ii)}$ denotes a zero-mean complex sequence from the ith interference source. This case is adequate for the purpose of illustration, although it can be generalized to include more interference sources or more hypothese such as only-first-interference-source-present hypothesis.

Pre-combining is taken as an example. If we can show that the asymptotic distribution of $\sqrt{N}\hat{\mathbf{r}}_Y^{\text{LC}}$ under $\mathcal{H}_1$ is the same as that under $\mathcal{H}_0$, and likewise the distribution under $\mathcal{H}_2$ the same as that under $\mathcal{H}_3$, this multiple hypotheses test can be simplified to be the original binary hypothesis test shown in (4.1). The following proposition will give sufficient conditions under which this simplification becomes feasible.

Proposition 5. *The asymptotic distribution of* $\sqrt{N}\hat{\mathbf{r}}_Y^{LC}$ *under* $\mathcal{H}_1$ *is the same as that under* $\mathcal{H}_0$*, and likewise the distribution under* $\mathcal{H}_2$ *the same as that under* $\mathcal{H}_3$*, if the interference sequences* $X_n^{(Ii)}$ *for* $i \in \{1,2\}$ *satisfy the following conditions, for any* $\{\diamondsuit_0, \diamondsuit_1, \diamondsuit_2, \diamondsuit_3\}$ *and* $\{\tau_0, \tau_1, \tau_2, \tau_3\}$*,*

1.

$$C^{\alpha}_{2X^{(Ii)}}\left[\boldsymbol{\tau}, \diamondsuit\right] = 0, \tag{4.43}$$

2.

$$S_{2X^{(Ii)}}\left(2\alpha;\alpha\right) = S_{2X^{(Ii)}}\left(0;-\alpha\right) = 0, \tag{4.44}$$

3.

$$S_{2X^{(I1)}2X^{(I2)}}\left(2\alpha;\alpha\right) = S_{2X^{(I1)}2X^{(I2)}}\left(0;-\alpha\right) = 0, \tag{4.45}$$

4.

$$S_{2X^{(I0)}2X^{(Ii)}}\left(2\alpha;\alpha\right) = S_{2X^{(I0)}2X^{(Ii)}}\left(0;-\alpha\right) = 0, \tag{4.46}$$

where

$$S_{2X^{(Ii)}}\left(\alpha;\omega\right) \triangleq \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\xi=-\infty}^{\infty} \mathrm{cum}\left\{X^{(Ii)\diamondsuit_0}_{n+\tau_0}, X^{(Ii)\diamondsuit_1}_{n+\xi+\tau_1}\right\} e^{-j\alpha n}e^{-j\omega\xi},$$

and

$$\begin{aligned} S_{2X^{(Ii)}2X^{(Ij)}}\left(\alpha;\omega\right) \triangleq \lim_{N\to\infty}\frac{1}{N}&\sum_{n=0}^{N-1}\sum_{\xi=-\infty}^{\infty} \\ &\times \mathrm{cum}\left\{X^{(Ii)\diamondsuit_0}_{n+\tau_0}X^{(Ii)\diamondsuit_1}_{n+\tau_1}, X^{(Ij)\diamondsuit_2}_{n+\xi+\tau_2}X^{(Ij)\diamondsuit_3}_{n+\xi+\tau_3}\right\} e^{-j\alpha n}e^{-j\omega\xi}. \end{aligned}$$

Proof. See Section 4.E. □

The constraints 2-3 in this proposition indicate that the second-order and fourth-order uncorrelatedness between the cochannel CU signals is required.

Although the pre-combining MASS is proposed without considering cochannel interference, full immunity from CU cochannel interference is achievable as indicated in this proposition. In other words, the asynchronous cyclostationary feature based MASS can achieve comparable performance to the synchronous one once the required conditions are satisfied.

4.6 Numerical Results

In this section, we establish the simulation by modeling the PU signal as a linear modulated signal with its sample sequence $X_n = \sum_{k=-\infty}^{\infty} a_k p\left(nT_s - kT_{sym}\right)$, where T_s is the sampling interval, T_{sym} is the symbol interval, a_k denotes i.i.d. zero-mean complex-valued symbols from a finite alphabet, and $p\left(t\right)$ is a rectangular pulse of value 1 for $0 \leq t < T_{sym}$ and value 0 elsewhere. This sample sequence X_n is second-order almost-cyclostationary, having the conjugated cyclic auto-correlation given by

$$R_{2X}^{\alpha}\left[\tau\right] = \frac{e^{-j\alpha \max\{0,-\tau\}}\left[1 - e^{-j\alpha(N_{sym}-|\tau|)}\right]}{\left(1 - e^{-j\alpha}\right)}, \tag{4.47}$$

where $N_{sym} = {}^{T_{sym}}/_{T_s}$ and the cycle frequency $\alpha = {}^{2\pi\kappa}/_{N_{sym}}$ for some $\kappa \in \{k \in \mathbb{Z} | k \in \left({}^{-N_{sym}}/_{2}, {}^{-N_{sym}}/_{2}\right]\}$. The probability density function of the instantaneous SNR γ_l is given by $f_{\text{Ray}}(\gamma_l) = \frac{1}{\overline{\gamma}}\exp(-\frac{\gamma_l}{\overline{\gamma}})$ where $\overline{\gamma}$ is the average SNR. The simulation parameters are respectively given by $N_{sym} = 6$, $\alpha = {}^{2\pi}/_{6}$, and the lag set $\mathcal{L} = \{\tau_1 = -3, \tau_2 = 3\}$. The sample size, proportional to the spectrum sensing time, is given by $N = N_{sym} \times N_s$

where N_s denotes the number of received data symbols.

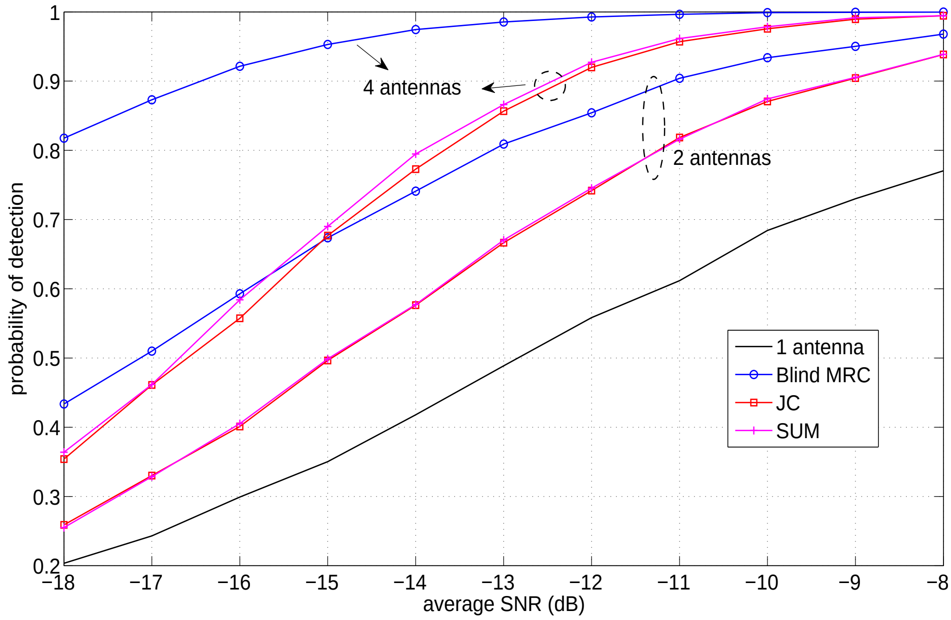


Figure 4.1: Probability of detection versus average SNR over multiple antennas using BMRC, joint combining, or sum combining.

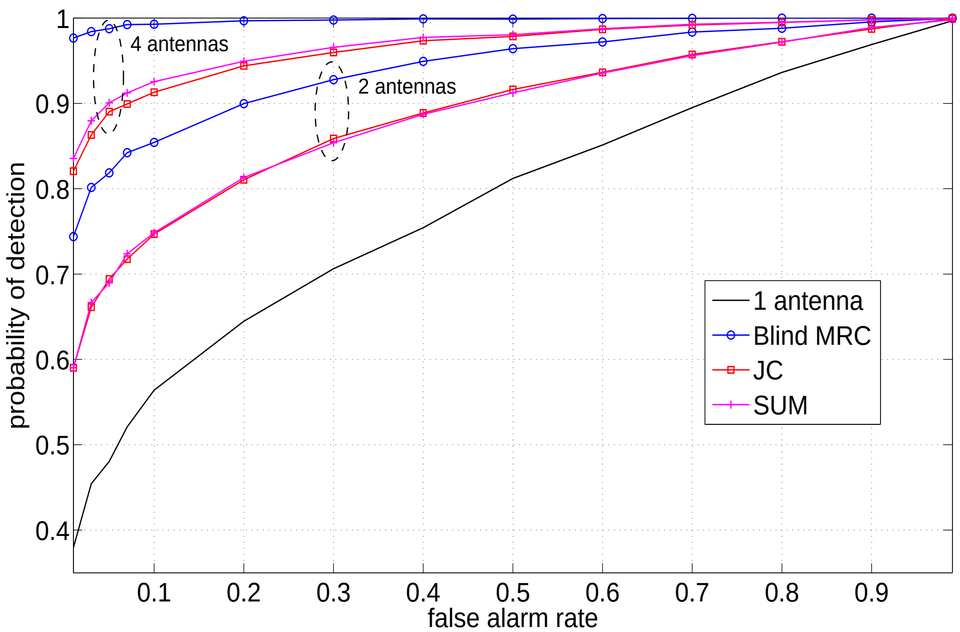


Figure 4.2: Probability of detection versus false alarm rate over multiple antennas using BMRC, joint combining, or sum combining at $\overline{\gamma} = -12$ dB.

Figure 4.1 shows the detection performance of cyclostationary feature based multi-antenna spectrum sensing over different average SNRs given the false alarm rate $P_{FA} = 0.1$ and $N_s = 10^3$. The operational average SNR region is between -18dB and -8dB. It can be seen from this figure that increasing the number of antennas does improve the probability of detection whether using BMRC, joint combining, or sum combining. Compared with joint combining and sum combining, BMRC, taking advantage of estimated CSI, results in notablely better detection performance. Moreover, joint combining and sum combining lead to comparable performance resutls in the sense that neither of them can significantly outperform the other over all average SNR region.

The receiver operating characteristic curves of our proposed GLRTs over different numbers of receive antennas are presented in Fig. 4.2. The curves are obtained with $\overline{\gamma} = -12$dB. It is apparent that the performance is improved when more receive antennas are utilized. Also from this figure, we can see that the probability of detection corresponding to BMRC appers to be higher than that of post-combining for any given false alarm rate.

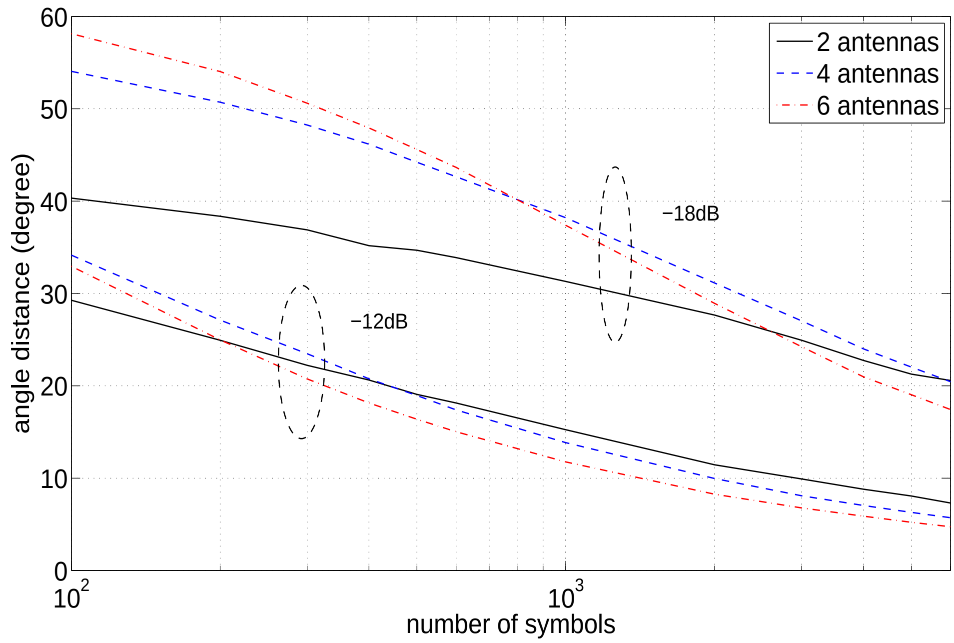


Figure 4.3: Average angle distance versus number of symbols over multiple antennas at the average SNR=-12 dB and -18dB.

Figure 4.3 presents how the average angle distance between $\hat{\mathbf{h}}$ and $\mathbf{h}$ varies with the number of received data symbols N_s for $\overline{\gamma} = -18$dB and -12dB. The average curves are obtained from 10^4 Monte Carlo runs. As shown in this figure, the average angle distance declines rapidly with the order of the number N_s. This result demonstrates the effectiveness of our proposed blind channel estimation and supports the claim in Proposition 2. It can also be observed that performing channel estimation in the higher average SNR region can result in higher average angle distance. In addition, increasing the number of antennas does not lead to significantly greater average angle distance when the sample size is sufficiently large.

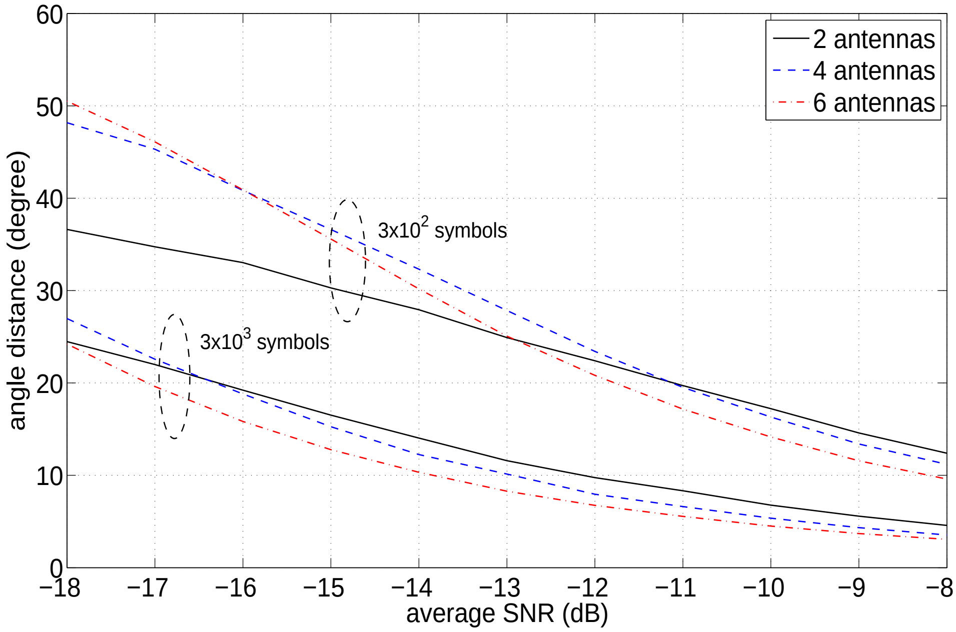


Figure 4.4: Average angle distance versus average SNR over multiple antennas with the symbol size 3×10^2 or 3×10^3.

The depiction of how the average angle distance varies with the average SNR over multiple antennas is shown in Fig. 4.4. The sizes of symbols utilized for acquisition of cyclic statistics are $N_s = 3 \times 10^2$ and 3×10^3. As can be seen in this figure, there is a negative correlation between the average angle distance and the average SNR. This negative correlation becomes stronger when larger symbol size is used.

4.7 Conclusions

This chapter has given an account of cyclostationary feature based multi-antenna spectrum sensing with its two possible schemes, pre-combining and post-combining. We have shown that MRC is the optimal linear per-combining strategy which maximizes the detection performance metric in the low SNR region. The asymptotic performance guarantee of our proposed cyclic crosscorrelation based channel estimation has been analytically examined. For post-combining, the joint asymptotic normality required in joint combining and sum combining has been validated once a mixing condition is met. Concerning the presence of CU cochannel interference in asynchronous spectrum sensing, we present the sufficient conditions which should be satisfied by CU transmission signals such that the full interference immunity can be achieved.

The results presented in this chapter have been submitted to IEEE Transactions on Signal Processing and IEEE Wireless Communications and Networking Conference 2014.

4.A Proof of Lemma 1

First, we show that $c_{kY^{(l)}}[n;\boldsymbol{\tau},\diamondsuit]$ can be presented in the required form, that is

$$\begin{aligned} c_{kY^{(l)}}[n;\boldsymbol{\tau},\diamondsuit] &= \operatorname{cum}\left\{Y_n^{(l)\diamondsuit_0}, Y_{n+\tau_1}^{(l)\diamondsuit_1}, \ldots, Y_{n+\tau_{k-1}}^{(l)\diamondsuit_{k-1}}\right\}, \\ &= h_l^{\diamondsuit_0} h_l^{\diamondsuit_1} \cdots h_l^{\diamondsuit_{k-1}} \operatorname{cum}\left\{X_n^{\diamondsuit_0}, X_{n+\tau_1}^{\diamondsuit_1}, \ldots, X_{n+\tau_{k-1}}^{\diamondsuit_{k-1}}\right\} \\ &\quad + \operatorname{cum}\left\{Z_n^{(l)\diamondsuit_0}, Z_{n+\tau_1}^{(l)\diamondsuit_1}, \ldots, Z_{n+\tau_{k-1}}^{(l)\diamondsuit_{k-1}}\right\}, \quad (4.48) \\ &= \sum_{\alpha\in\mathcal{A}_k[\boldsymbol{\tau},\diamondsuit]} \left\{h_l^{\diamondsuit_0} h_l^{\diamondsuit_1} \cdots h_l^{\diamondsuit_{k-1}} C_{kX}^{\alpha}[\tau,\diamondsuit] \right. \\ &\quad \left. + C_{kZ}^{\alpha}[\tau,\diamondsuit] \right\} e^{j\alpha n}, \quad (4.49) \end{aligned}$$

where the properties [82, eq. (6a) (6e)] are applied in the first equality and

$$C_{kZ}^{\alpha}[\boldsymbol{\tau},\diamondsuit] = \begin{cases} \sigma_Z^2 \delta(\alpha), & \text{for } k=2, \boldsymbol{\tau}=(0), \text{ and } \diamondsuit = (\bar{*}, *) \text{ or } (*, \bar{*}), \\ 0, & \text{elsewhere,} \end{cases} \quad (4.50)$$

where the fact that the joint cumulants (order greater than 2) of jointly Gaussian random variables are all zero [83] is used. Therefore, it follows that $Y_n^{(l)}$ is kth-order almost-cyclostationary.

By exploiting the expansion (4.48), it can be easily verified that the sequence $Y_n^{(l)}$ satisfies **A1** for $k \geq 3$. This statement is also true for $1 \leq k \leq 2$ because of

$$\sum_{\tau_1=-\infty}^{\infty} \sup_n |\tau_1| \left|\operatorname{cum}\left\{Z_n^{(l)\diamondsuit_0}, Z_{n+\tau_1}^{(l)\diamondsuit_1}\right\}\right| = 0, \quad (4.51)$$

and

$$\mathrm{cum}\left\{Z_n^{(l)\diamondsuit_0}\right\} = 0. \tag{4.52}$$

The same conclusions can be reached when it comes to the sequence Y_n^{LC}.

4.B Proof of Lemma 2

Lemma 3. *For* $1 \leq i \leq k$, $\forall k \in \mathbb{N}_0$, *and* $1 \leq l_0, \ldots, l_k \leq L$,

$$\sum_{\xi_1,\ldots,\xi_k=-\infty}^{\infty} \sup_n |\xi_i| \left|\mathrm{cum}\left\{Y_n^{(l_0)\diamondsuit_0}, Y_{n+\xi_1}^{(l_1)\diamondsuit_1}, \ldots, Y_{n+\xi_k}^{(l_k)\diamondsuit_k}\right\}\right| \;<\; \infty. \tag{4.53}$$

Proof. The proof is similar to that of Lemma 1. □

For simplicity, let's discard the superscript (l) for a moment, i.e., using $\psi_{n,\tau}$ and Y_n rather than $\psi_{n,\tau}^{(l)}$ and $Y_n^{(l)}$. By the definition of the second-order cumulant, we have

$$\begin{aligned}
&c_{2\psi_{n,\tau}}[n;\xi,\diamondsuit] = \mathrm{cum}\left\{\psi_{n,\tau}^{\diamondsuit_0}, \psi_{n+\xi,\tau}^{\diamondsuit_1}\right\},\\
&= \mathbb{E}\left[Y_n^{\diamondsuit_0} Y_{n+\tau}^{\diamondsuit_0 *} Y_{n+\xi}^{\diamondsuit_1} Y_{n+\xi+\tau}^{\diamondsuit_1 *}\right] - \mathbb{E}\left[Y_n^{\diamondsuit_0} Y_{n+\tau}^{\diamondsuit_0 *}\right] \mathbb{E}\left[Y_{n+\xi}^{\diamondsuit_1} Y_{n+\xi+\tau}^{\diamondsuit_1 *}\right],\\
&= c_{4Y}\left[n; \boldsymbol{\tau}_1 = (\tau, \xi, \xi+\tau), \diamondsuit_1 = (\diamondsuit_0, \diamondsuit_0 *, \diamondsuit_1, \diamondsuit_1 *)\right]\\
&\quad + c_{2Y}\left[n; \boldsymbol{\tau}_2 = (\xi), \diamondsuit_2 = (\diamondsuit_0, \diamondsuit_1)\right] c_{2Y}\left[n+\xi+\tau; \boldsymbol{\tau}_3 = (-\xi), \diamondsuit_3 = (\diamondsuit_1 *, \diamondsuit_0 *)\right]\\
&+ c_{2Y}\left[n; \boldsymbol{\tau}_4 = (\xi+\tau), \diamondsuit_4 = (\diamondsuit_0, \diamondsuit_1 *)\right] c_{2Y}\left[n+\tau; \boldsymbol{\tau}_5 = (\xi-\tau), \diamondsuit_5 = (\diamondsuit_0 *, \diamondsuit_1)\right],
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\alpha \in \mathcal{A}_4[\boldsymbol{\tau}_1, \diamondsuit_1]} C_{4Y}^{\alpha} [\boldsymbol{\tau}_1, \diamondsuit_1] e^{i\alpha n} \\
&\quad + \sum_{\alpha \in \mathcal{A}_2[\boldsymbol{\tau}_2, \diamondsuit_2]} C_{2Y}^{\alpha} [\boldsymbol{\tau}_2, \diamondsuit_2] e^{j\alpha n} \sum_{\beta \in \mathcal{A}_2[\boldsymbol{\tau}_3, \diamondsuit_3]} C_{2Y}^{\beta} [\boldsymbol{\tau}_3, \diamondsuit_3] e^{j\beta(n+\xi+\tau)} \\
&\quad + \sum_{\alpha \in \mathcal{A}_2[\boldsymbol{\tau}_4, \diamondsuit_4]} C_{2Y}^{\alpha} [\boldsymbol{\tau}_4, \diamondsuit_4] e^{j\alpha n} \sum_{\beta \in \mathcal{A}_2[\boldsymbol{\tau}_5, \diamondsuit_5]} C_{2Y}^{\beta} [\boldsymbol{\tau}_5, \diamondsuit_5] e^{j\beta(n+\tau)}, \\
&= \sum_{\alpha \in \mathcal{A}_4[\boldsymbol{\tau}_1, \diamondsuit_1] \cup \mathcal{A}_2} C_{2\psi_{n,\tau}}^{\alpha} [\xi, \diamondsuit] e^{j\alpha n}, \qquad (4.54)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{A}_2 = \{\alpha_1 + \beta_1, \alpha_2 + \beta_2 | \alpha_1 \in \mathcal{A}_2 [\boldsymbol{\tau}_2, \diamondsuit_2], \beta_1 \in \mathcal{A}_2 [\boldsymbol{\tau}_3, \diamondsuit_3], \\
\alpha_2 \in \mathcal{A}_2 [\boldsymbol{\tau}_4, \diamondsuit_4], \beta_2 \in \mathcal{A}_2 [\boldsymbol{\tau}_5, \diamondsuit_5]\}, \qquad (4.55)
\end{aligned}
$$

$$
\begin{aligned}
C_{2\psi_{n,\tau}}^{\alpha} [\xi, \diamondsuit] &= C_{4Y}^{\alpha} [\boldsymbol{\tau}_1, \diamondsuit_1] + \\
&+ \sum_{\{\alpha_1 \in \mathcal{A}_2[\boldsymbol{\tau}_2, \diamondsuit_2], \beta_1 \in \mathcal{A}_2[\boldsymbol{\tau}_3, \diamondsuit_3] | \alpha_1 + \beta_1 = \alpha\}} C_{2Y}^{\alpha_1} [\boldsymbol{\tau}_2, \diamondsuit_2] C_{2Y}^{\beta_1} [\boldsymbol{\tau}_3, \diamondsuit_3] e^{j\beta_1(\xi+\tau)} \\
&+ \sum_{\{\alpha_1 \in \mathcal{A}_2[\boldsymbol{\tau}_4, \diamondsuit_4], \beta_1 \in \mathcal{A}_2[\boldsymbol{\tau}_5, \diamondsuit_5] | \alpha_1 + \beta_1 = \alpha\}} C_{2Y}^{\alpha_1} [\boldsymbol{\tau}_4, \diamondsuit_4] C_{2Y}^{\beta_1} [\boldsymbol{\tau}_5, \diamondsuit_5] e^{j\beta_1 \tau}, \qquad (4.56)
\end{aligned}
$$

[29, eq. (89)] is applied in the third equality and the property, Y_n being almost-cyclostationary up to fourth order due to Lemma 1, is used in the fourth equality. This Fourier-series expansion (4.54) indicates that $\psi_{n,\tau}^{(l)}$ is second-order almost-cyclostationary.

Moreover,

$$\begin{aligned}
&\mathrm{cum}\left\{\psi_{n,\tau_0}^{(l_0)\diamondsuit_0}, \psi_{n+\xi_1,\tau_1}^{(l_1)\diamondsuit_1}, \ldots, \psi_{n+\xi_k,\tau_k}^{(l_k)\diamondsuit_k}\right\} \\
&\qquad = \mathrm{cum}\left\{Y_n^{(l_0)\diamondsuit_0} Y_{n+\tau_0}^{(l_0)\diamondsuit_0 *}, Y_{n+\xi_1}^{(l_1)\diamondsuit_1} Y_{n+\xi_1+\tau_1}^{(l_1)\diamondsuit_1 *}, \ldots, Y_{n+\xi_k}^{(l_k)\diamondsuit_k} Y_{n+\xi_k+\tau_k}^{(l_k)\diamondsuit_k *}\right\}, \\
&\qquad = \sum_{\boldsymbol{\nu}} \mathrm{cum}\left\{\nu_1\right\} \cdots \mathrm{cum}\left\{\nu_p\right\}, \quad (4.57)
\end{aligned}$$

where [84, Theorem 2.3.2] is used in the second equality in which $\nu_1 \cup \cdots \cup \nu_p$ forms an indecomposable partition of the following $(k+1) \times 2$ array:

$$\begin{array}{cc}
Y_n^{(l_0)\diamondsuit_0} & Y_{n+\tau_0}^{(l_0)\diamondsuit_0 *} \\
Y_{n+\xi_1}^{(l_1)\diamondsuit_1} & Y_{n+\xi_1+\tau_1}^{(l_1)\diamondsuit_1 *} \\
\vdots & \vdots \\
Y_{n+\xi_k}^{(l_k)\diamondsuit_k} & Y_{n+\xi_k+\tau_k}^{(l_k)\diamondsuit_k *}
\end{array}, \tag{4.58}$$

each ν_i is a subset of entries of this array, and $\boldsymbol{\nu}$ represents the collection of all indecomposable partitions. Applying Lemma 3 gives, for $1 \leq i \leq p$,

$$\sum_{\boldsymbol{\xi}_i=-\infty}^{\infty} \sup_n \left|\xi_{i_m}\right| \left|\mathrm{cum}\left\{\nu_i\right\}\right| < \infty \tag{4.59}$$

where

$$\boldsymbol{\xi}_i \triangleq \left(\xi_{i_m} \text{ for } 1 \leq i_m \leq k \,\middle|\, Y_{n+\xi_{i_m}}^{(l_{i_m})\diamondsuit_{i_m}} \text{ or } Y_{n+\xi_{i_m}+\tau_{i_m}}^{(l_{i_m})\diamondsuit_{i_m} *} \in \nu_i\right). \tag{4.60}$$

Thus, we have

$$\begin{aligned}
&\sum_{\boldsymbol{\xi}=-\infty}^{\infty} \sup_n |\xi_i| \left|\text{cum}\left\{\psi_{n,\tau_0}^{(l_0)\diamondsuit_0}, \psi_{n+\xi_1,\tau_1}^{(l_1)\diamondsuit_1}, \ldots, \psi_{n+\xi_k,\tau_k}^{(l_k)\diamondsuit_k}\right\}\right| \\
&\leq \sum_{\boldsymbol{\nu}} \left\{ \sum_{\boldsymbol{\xi}=-\infty}^{\infty} \sup_n |\xi_i| \left|\text{cum}\left\{\nu_1\right\}\right| \cdots \left|\text{cum}\left\{\nu_p\right\}\right| \right\}, \\
&\leq \sum_{\boldsymbol{\nu}} \left\{ \sum_{\boldsymbol{\xi}=-\infty}^{\infty} \sup_n \left[\max\left\{\left(\sup_{1_m} |\xi_{1_m}|\right), 1\right\} \left|\text{cum}\left\{\nu_1\right\}\right| \right.\right. \\
&\qquad\qquad \left.\left. \cdots \max\left\{\left(\sup_{p_m} |\xi_{p_m}|\right), 1\right\} \left|\text{cum}\left\{\nu_p\right\}\right| \right] \right\}, \\
&\leq \sum_{\boldsymbol{\nu}} \left\{ \sum_{\boldsymbol{\xi}_1=-\infty}^{\infty} \sup_n \left[\max\left\{\left(\sup_{1_m} |\xi_{1_m}|\right), 1\right\} \left|\text{cum}\left\{\nu_1\right\}\right| \right] \right. \\
&\qquad\qquad \left. \cdots \sum_{\boldsymbol{\xi}_p=-\infty}^{\infty} \sup_n \left[\max\left\{\left(\sup_{p_m} |\xi_{p_m}|\right), 1\right\} \left|\text{cum}\left\{\nu_p\right\}\right| \right] \right\}, \\
&\leq \sum_{\boldsymbol{\nu}} \left\{ \left[\sum_{\boldsymbol{\xi}_1=-\infty}^{\infty} \sup_n \left[\left(\sup_{1_m} |\xi_{1_m}|\right) \left|\text{cum}\left\{\nu_1\right\}\right| \right] + \sum_{\boldsymbol{\xi}_1=(0,\ldots,0)} \sup_n \left|\text{cum}\left\{\nu_1\right\}\right| \right] \cdots \right. \\
&\left. \left[\sum_{\boldsymbol{\xi}_p=-\infty}^{\infty} \sup_n \left[\left(\sup_{p_m} |\xi_{p_m}|\right) \left|\text{cum}\left\{\nu_p\right\}\right| \right] + \sum_{\boldsymbol{\xi}_1=(0,\ldots,0)} \sup_n \left|\text{cum}\left\{\nu_1\right\}\right| \right] \right\} < \infty, \quad (4.61)
\end{aligned}$$

where the factorization (4.57) is applied in the first inequality and the condition (4.59) in the last inequality, showing that the condition **A2** is met by the sequence $\psi_{n,\tau}^{(l)}$.

The same conclusions can be drawn about the sequence Y_n^{LC}.

4.C Proof of Proposition 2

By replacing $\hat{\mathbf{r}}_Y^{\text{CE}}$ in (4.24) with $\mathbf{R}^{\text{CE}}\mathbf{h}_{\text{CP}}+\boldsymbol{\epsilon}$, it can be easily shown that $\hat{\mathbf{h}}_{\text{CP}}$ approaches $\mathbf{h}_{\text{CP}}$ asymptotically. Due to $||\delta_{\mathbf{H}}||_F \leq 2||\hat{\mathbf{h}}_{\text{CP}} - \mathbf{h}_{\text{CP}}||_2$, $\hat{\mathbf{H}}$ approaches $\mathbf{H}$ as N goes to infinity.

Let $\mathbf{U} = [\mathbf{u}_2\, \mathbf{u}_3\, \ldots\, \mathbf{u}_L]$ and assume that $4||\delta_{\mathbf{H}}||^2 < (\lambda_1 - |\mathbf{u}_1^H \delta_{\mathbf{H}} \mathbf{u}_1| - ||\mathbf{U}^H \delta_{\mathbf{H}} \mathbf{U}||)^2$

for $N > N_1$. As a result of [78, Theorem (3.11)], there exist a scalar φ and a vector $\mathbf{p}$ such that $(\tilde{\lambda}, \tilde{\mathbf{u}}) = (\lambda_1 + \varphi, \mathbf{u}_1 + \mathbf{U}\mathbf{p})$ where $(\tilde{\lambda}, \tilde{\mathbf{u}})$ is an eigenpair of $\hat{\mathbf{H}}$. The quantities $|\varphi|$ and $||\mathbf{U}\mathbf{p}||_2$ have been shown to be of order $O(||\delta_{\mathbf{H}}||_2)$. If we can show that the algebraic multiplicity of $\tilde{\lambda}_{\max}$ is one and $\tilde{\lambda} = \tilde{\lambda}_{\max}$, then the required results will follow.

Let's define an index set $I_{\max} = \{1 \leq i \leq L \,|\, \tilde{\lambda}_i = \tilde{\lambda}_{\max}\}$. Assume that the cardinality of $I_{\max}$ is greater than one. Applying [85, Theorem (4.3.1)] gives $|\tilde{\lambda}_i - \lambda_1| \leq ||\delta_{\mathbf{H}}||_2$ for $i \in I_{\max}$ and $|\tilde{\lambda}_j| \leq ||\delta_{\mathbf{H}}||_2$ for $j \neq i \in I_{\max}$. Let N_2 be an integer such that $||\delta_{\mathbf{H}}||_2 < {}^{\lambda_1}\!/_2$ for $N > N_2$. Given $N > N_2$, there exists $\tilde{\lambda}_i \neq \tilde{\lambda}_j$ for $i, j \in I_{\max}$, which is a contradiction. Therefore, the cardinality of $I_{\max}$ is one for $N > N_2$ and so is the algebraic multiplicity of $\tilde{\lambda}_{\max}$.

Assume that $\tilde{\lambda} = \tilde{\lambda}_i$ for some $i \notin I_{\max}$. Let $N_3 > \max\{N_1, N_2\}$ be an integer such that $|\varphi| = |\tilde{\lambda}_i - \lambda_1| < {}^{\lambda_1}\!/_2$ for $N > N_3$. Given $N > N_3$, using [85, Theorem (4.3.1)] again gives $|\tilde{\lambda}_i| \leq ||\delta_{\mathbf{H}}||_2 < {}^{\lambda_1}\!/_2$, which is a contradiction. Hence, $\tilde{\lambda} = \tilde{\lambda}_{\max}$ for $N > N_3$.

4.D Proof of Proposition 3

The mean vector $\sqrt{N}\mathbf{R}_X \underline{\gamma}$ is easily obtained by using the asymptotic property of $\sqrt{N}\hat{\mathbf{r}}_Y^{(l)}$. To show the joint asymptotic normality of $\sqrt{N}\hat{\mathbf{r}}_Y$, we need to verify that cumulants of $\sqrt{N}\hat{C}^{\alpha}_{2Y^{(l)}}[\tau]$ with order greater than 2 approach to zero asymptotically, that is, for $1 \leq l_i \leq L$ and $m \geq 2$,

$$\lim_{N\to\infty} \operatorname{cum}\left\{\sqrt{N}\hat{C}^{\alpha\diamondsuit_0}_{2Y^{(l_0)}}[\tau_0], \ldots, \sqrt{N}\hat{C}^{\alpha\diamondsuit_m}_{2Y^{(l_m)}}[\tau_m]\right\} = 0. \tag{4.62}$$

Exploiting the multilinearity of cumulants [84] gives

$$\begin{aligned}
&N^{m+1} \operatorname{cum}\left\{\hat{C}^{\alpha\diamondsuit_0}_{2Y^{(l_0)}}[\tau_0], \ldots, \hat{C}^{\alpha\diamondsuit_m}_{2Y^{(l_m)}}[\tau_m]\right\} \\
&= \sum_{n_0,\ldots,n_m=0}^{N-1} \operatorname{cum}\left\{\psi^{(l_0)\diamondsuit_0}_{n_0,\tau_0}, \ldots, \psi^{(l_m)\diamondsuit_m}_{n_m,\tau_m}\right\} e^{-j\alpha(\pm n_0 \pm \cdots \pm n_m)}, \\
&= \sum_{\xi_1,\ldots,\xi_m=-(N-1)}^{(N-1)} \sum_{n=n_a}^{n_b} \operatorname{cum}\left\{\psi^{(l_0)\diamondsuit_0}_{n,\tau_0}, \ldots, \psi^{(l_m)\diamondsuit_m}_{n+\xi_m,\tau_m}\right\} e^{-j\alpha[\pm n \pm (n+\xi_1)\cdots \pm (n+\xi_m)]},
\end{aligned} \tag{4.63}$$

where $\pm$ to be plus or minus depends on $\diamondsuit_i$, the setting, $n_0 \triangleq n$, $\xi_i \triangleq n_i - n$ for $1 \leq i \leq m$, $n_a \triangleq -\min(0, \xi_1, \ldots, \xi_m)$, and $n_b \triangleq N - 1 - \max(0, \xi_1, \ldots, \xi_m)$, is used in the second equality. By using Lemma 2, we have

$$\begin{aligned}
&\left|\operatorname{cum}\left\{\hat{C}^{\alpha}_{2Y^{(l_0)}}[\tau_0], \ldots, \hat{C}^{\alpha}_{2Y^{(l_m)}}[\tau_m]\right\}\right| \\
&\qquad \leq N^{-(m+1)} \sum_{\boldsymbol{\xi}=-(N-1)}^{(N-1)} \sum_{n=n_a}^{n_b} \left|\operatorname{cum}\left\{\psi^{(l_0)\diamondsuit_0}_{n,\tau_0}, \ldots, \psi^{(l_m)\diamondsuit_m}_{n+\xi_m,\tau_m}\right\}\right|, \\
&\qquad \leq N^{-m} \sum_{\boldsymbol{\xi}=-(N-1)}^{(N-1)} \sup_n \left|\operatorname{cum}\left\{\psi^{(l_0)\diamondsuit_0}_{n,\tau_0}, \ldots, \psi^{(l_m)\diamondsuit_m}_{n+\xi_m,\tau_m}\right\}\right|, \\
&\qquad = O\left(N^{-m}\right),
\end{aligned} \tag{4.64}$$

where $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)$. Therefore, the limit (4.62) immediately follows.

Following closely the derivation in [29, eq. (86-88)], we can also obtain the results of (4.30) and (4.31).

4.E Proof of Proposition 5

Assume that the asymptotic normality of $\sqrt{N}\hat{\mathbf{r}}_Y^{\mathrm{LC}}$ under any hypothesis is true, i.e., $\lim_{N\to\infty}\sqrt{N}\hat{\mathbf{r}}_Y^{\mathrm{LC}} \sim \mathcal{N}(\mathbf{r}_{Y,\mathcal{H}_i}^{\mathrm{LC}}, \Sigma_{Y,\mathcal{H}_i}^{\mathrm{LC}})$ under $\mathcal{H}_i$ for $0 \leq i \leq 3$. Applying Condition 1 can lead to the conclusion that $\mathbf{r}_{Y,\mathcal{H}_0}^{\mathrm{LC}} = \mathbf{r}_{Y,\mathcal{H}_1}^{\mathrm{LC}}$ and $\mathbf{r}_{Y,\mathcal{H}_2}^{\mathrm{LC}} = \mathbf{r}_{Y,\mathcal{H}_3}^{\mathrm{LC}}$. As shown in [29], the entries of $\Sigma_{Y,\mathcal{H}_i}^{\mathrm{LC}}$ are the real or imaginary parts of the linear combinations of

$$S_{\psi_{\tau,\rho}^{\mathrm{LC}}}(2\alpha;\alpha) \triangleq \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\xi=-\infty}^{\infty}\mathrm{cum}\left\{\psi_{n,\tau}^{\mathrm{LC}}, \psi_{n+\xi,\rho}^{\mathrm{LC}}\right\}e^{-j2\alpha n}e^{-j\alpha\xi}, \tag{4.65}$$

and

$$\tilde{S}_{\psi_{\tau,\rho}^{\mathrm{LC}}}(0;-\alpha) \triangleq \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{\xi=-\infty}^{\infty}\mathrm{cum}\left\{\psi_{n,\tau}^{\mathrm{LC}}, \psi_{n+\xi,\rho}^{\mathrm{LC}*}\right\}e^{j\alpha\xi}, \tag{4.66}$$

where $\psi_{n,\tau}^{\mathrm{LC}} = Y_n^{\mathrm{LC}}Y_{n+\tau}^{\mathrm{LC}*}$. Hence, we need to show that $S_{\psi_{\tau,\rho}^{\mathrm{CE}}}(2\alpha;\alpha)$ and $\tilde{S}_{\psi_{\tau,\rho}^{(l_1,l_2)}}(0;-\alpha)$ under $\mathcal{H}_0$ are respectively the same as those under $\mathcal{H}_1$, and likewise under $\mathcal{H}_2$ and $\mathcal{H}_3$. For the sake of simplicity, only the case of $S_{\psi_{\tau,\rho}^{\mathrm{LC}}}(2\alpha;\alpha)$ being the same under $\mathcal{H}_0$ and $\mathcal{H}_1$ is proved here.

The cumulant $\mathrm{cum}\left\{\psi_{n,\tau}^{\mathrm{LC}}, \psi_{n+\xi,\rho}^{\mathrm{LC}}\right\}$ under $\mathcal{H}_1$ can be expanded as

$$\begin{aligned}
&\mathrm{cum}\left\{\zeta_n^{\mathrm{LC}}\zeta_{n+\tau}^{\mathrm{LC}*}, \zeta_{n+\xi}^{\mathrm{LC}}\zeta_{n+\xi+\rho}^{\mathrm{LC}*}\right\}\\
&\quad+\left|\mathbf{w}'\mathbf{h}\right|^4\left[\mathrm{cum}\left\{X_n^{(I1)}X_{n+\tau}^{(I1)*}, X_{n+\xi}^{(I1)}X_{n+\xi+\rho}^{(I1)*}\right\} + \mathrm{cum}\left\{X_n^{(I2)}X_{n+\tau}^{(I2)*}, X_{n+\xi}^{(I2)}X_{n+\xi+\rho}^{(I2)*}\right\}\right.
\end{aligned}$$

$$+\operatorname{cum}\left\{X_n^{(I1)}X_{n+\tau}^{(I2)*}, X_{n+\xi}^{(I1)}X_{n+\xi+\rho}^{(I2)*}\right\}+\operatorname{cum}\left\{X_n^{(I1)}X_{n+\tau}^{(I2)*}, X_{n+\xi}^{(I2)}X_{n+\xi+\rho}^{(I1)*}\right\}$$
$$+\operatorname{cum}\left\{X_n^{(I2)}X_{n+\tau}^{(I1)*}, X_{n+\xi}^{(I2)}X_{n+\xi+\rho}^{(I1)*}\right\}+\operatorname{cum}\left\{X_n^{(I2)}X_{n+\tau}^{(I1)*}, X_{n+\xi}^{(I1)}X_{n+\xi+\rho}^{(I2)*}\right\}\Big]$$
$$+\left|\mathbf{w}'\mathbf{h}\right|^2\sigma_Z^2\left\{\delta\left(\xi+\rho\right)\left[\operatorname{cum}\left\{X_{n+\tau}^{(I1)*}, X_{n+\xi}^{(I1)}\right\}+\operatorname{cum}\left\{X_{n+\tau}^{(I2)*}, X_{n+\xi}^{(I2)}\right\}\right]\right.$$
$$\left.+\delta\left(\tau-\xi\right)\left[\operatorname{cum}\left\{X_n^{(I1)}, X_{n+\xi+\rho}^{(I1)*}\right\}+\operatorname{cum}\left\{X_n^{(I2)}, X_{n+\xi+\rho}^{(I2)*}\right\}\right]\right\}, \quad (4.67)$$

where $\zeta_n^{\mathrm{LC}}=\mathbf{w}'\mathbf{z}_n$ is a sequence of the i.i.d. circularly-symmetric complex Gaussian noise samples with variance σ_Z^2. Replacing the cumulant in (4.65) with (4.67) and utilizing Condition 2 and 3 simplify $S_{\psi_{\tau,\rho}^{\mathrm{LC}}}(2\alpha;\alpha)$ under $\mathcal{H}_1$ to be that under $\mathcal{H}_0$.

CHAPTER 5

SPECTRUM SENSING OVER FADING CHANNELS

5.1 Introduction

The performance of spectrum sensing is subject to wireless channel uncertainty, including large-scale fading and small-scale fading. It is highly desirable to have analytical expressions of detection and false alarm probabilities, taking into account of various fading distributions. These analytical expressions make performance over fading channels predictive and facilitate further system analysis in cognitive radio networks.

There is a sizable literature on the study of energy detection over fading channels. The closed-from expressions that describe the average detection probability over Nakagami and Rician fading channels have been presented in [86]. In the same paper, diversity reception (using equal gain combining, selection combining, and switch and stay combining (SSC)) subject to i.i.d. Rayleigh fading has also been investigated. The performance of two low-complexity reception schemes, square-law combining (SLC) and square-law selection (SLS), was subsequently analyzed in [87]. A more general $\kappa - \mu$ fading distribution has been recently discussed in [88]. When energy detection is implemented in the spectral domain, the corresponding performance analysis based on different spectral density estimates has been addressed in [89, 90]. While the above analyses mainly focus on small-scale fading, the composite effect of small-scale and large-scale fading has its importance. To circumvent the issue of complicated composite fading distributions, such as chi-square-gamma or Nakagami-lognormal distributions, alternative distribution approximations have to be applied to get the closed-form results [91,92]. In CR networks, the scenarios of relay-assisted and cooperative energy detection are considered in [93].

Cyclostationary featuer detection plays an important role in the performance enhancement of spectrum sensing. However, far too little attention has been paid to the analytical analysis of CFD over fading channels. It is partly because some CFD approaches are lacking in analytical descriptions of detection probability conditioned on the channel gain. The existing analysis of detection probability relies on numerical simulation [94]. In view of the shortage of analytical results, this chapter seeks to analytically examine the detection probability of CFD proposed in [29] over fading channels. In fading environments, the asymptotic detection performance conditioned on the fading channel gain can be shown to be a generalized Marcum Q-function. In the first part, we aim to provide analytical expressions of detection performance bounds without requiring integrating over the Marcum Q-function. The difficulty of obtaining this kind of analytical expressions arises from a complicated argument in the generalized Marcum Q-function. Inspired by [95], we seek alternative expressions of this argument, i.e., its upper and lower bounds. By further exploiting monotonicity of the generalized Marcum Q-function, the upper and lower bounds of detection performance averaged over the probability density function of the fading channel gain can be obtained. The second part will analyze the detection performance of post-combining over i.i.d. Rayleigh fading channels. This analysis is based on two tight approximations of detection performance over Nakagami fading channels at low average SNR. Two post-combining techniques to be examined are post addition combining (PAC) and post selection combining (PSC), which are first proposed for exploiting multiple cycle frequencies in [42].

5.2 Performance Bounds over Nakagami Fading Channels

Let's introduce the complex-valued fading channel gain h in the binary hypothesis testing model presented in (3.1), giving the modified model,

$$\begin{aligned}
&\mathcal{H}_0 : \{Y_n\}_{n=0}^{\infty} = \{Z_n\}_{n=0}^{\infty}\,, \text{ noise only,}\\
&\mathcal{H}_1 : \{Y_n\}_{n=0}^{\infty} = \{hX_n + Z_n\}_{n=0}^{\infty}\,, \text{ feature-present.} \qquad (5.1)
\end{aligned}$$

The instantaneous SNR is defined as $\gamma \triangleq |h|^2 \, {}^{\sigma_X^2}/{}_{\sigma_Z^2}$ and ${}^{\sigma_X^2}/{}_{\sigma_Z^2} = 1$ is assumed throughout this chapter. The probability density function of γ is given by

$$f_{\text{Nak}}(\gamma; l) = \frac{l^l}{\Gamma(l)\,\overline{\gamma}^l}\gamma^{l-1}\exp\left(-\frac{l}{\overline{\gamma}}\gamma\right), \tag{5.2}$$

where $\overline{\gamma}$ is the average SNR, l is the fading parameter, and $\Gamma(\cdot)$ is the Gamma function. To simplify the notation, we restrict our analysis to the case of single cycle frequency, i.e., $\alpha_1 = \cdots = \alpha_M = \alpha$ in the test set $\mathcal{J}$ defined in Chapter 3, and denote the lag set $\mathcal{L} = \{\tau_1, \tau_2, \ldots \tau_M\}$. As the counterparts of the CAF-vector estimate $\hat{\mathbf{r}}_Y^{\mathcal{J}}[N]$ in (3.2) and the test statistic Λ in Section 3.2.1, we have

$$\begin{aligned}\hat{\mathbf{r}}_Y^{\alpha} \triangleq\ & \Big[\mathrm{Re}\left\{\hat{R}_Y^{\alpha}[\tau_1; N]\right\}, \ldots, \mathrm{Re}\left\{\hat{R}_Y^{\alpha}[\tau_M; N]\right\}, \\ & \mathrm{Im}\left\{\hat{R}_Y^{\alpha}[\tau_1; N]\right\}, \ldots, \mathrm{Im}\left\{\hat{R}_Y^{\alpha}[\tau_M; N]\right\}\Big]', \end{aligned} \tag{5.3}$$

and

$$\mathcal{T} = N\left(\hat{\mathbf{r}}_Y^{\alpha}\right)' \hat{\Sigma}_Y^{-1}\hat{\mathbf{r}}_Y^{\alpha}. \tag{5.4}$$

For a constant false alarm rate, the asymptotic detection performance $P_{d|\gamma}$ conditioned on the SNR is given by

$$P_{d|\gamma} = Q_M\left(\gamma\sqrt{N\left(\mathbf{r}_X^{\alpha}\right)' \Sigma_Y^{-1}\mathbf{r}_X^{\alpha}}, \sqrt{\lambda}\right), \tag{5.5}$$

where λ is a detection threshold and $\gamma\mathbf{r}_X^{\alpha}$ is the limiting vector of $\hat{\mathbf{r}}_Y^{\alpha}$ under the feature-present hypothesis. The matrix Σ_Y under $\mathcal{H}_1$ can be written as $\Sigma_Y = \gamma^2\Sigma_{4X} + \gamma\bar{\Sigma}_{XZ} + \Sigma_{4Z}$ where

$$\bar{\Sigma}_{XZ} = \Sigma_{XXZZ} + \Sigma_{ZZXX} + \Sigma_{XZXZ} + \Sigma_{ZXZX} + \Sigma_{XZZX} + \Sigma_{ZXXZ}, \tag{5.6}$$

and

$$\Sigma_{ABCD} = \begin{bmatrix} \mathrm{Re}\left\{\frac{\mathbf{C}_{ABCD}+\tilde{\mathbf{C}}_{ABCD}}{2}\right\} & \mathrm{Im}\left\{\frac{\mathbf{C}_{ABCD}-\tilde{\mathbf{C}}_{ABCD}}{2}\right\} \\ \mathrm{Im}\left\{\frac{\mathbf{C}_{ABCD}+\tilde{\mathbf{C}}_{ABCD}}{2}\right\} & \mathrm{Re}\left\{\frac{-\mathbf{C}_{ABCD}+\tilde{\mathbf{C}}_{ABCD}}{2}\right\} \end{bmatrix}. \tag{5.7}$$

$\mathbf{C}_{ABCD}$ and $\tilde{\mathbf{C}}_{ABCD}$ are two $M \times M$ covariance matrices with their (i,j)th entries,

$$\mathbf{C}^{i,j}_{ABCD} \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{\xi=-\infty}^{\infty} \mathrm{Cov}\left(A_n B^*_{n+\tau_i}, C_{n+\xi} D^*_{n+\xi+\tau_j}\right) e^{-j\alpha\xi} e^{-j2\alpha n}, \tag{5.8}$$

and

$$\tilde{\mathbf{C}}^{i,j}_{ABCD} \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{\xi=-\infty}^{\infty} \mathrm{Cov}\left(A_n B^*_{n+\tau_i}, C^*_{n+\xi} D_{n+\xi+\tau_j}\right) e^{j\alpha\xi}. \tag{5.9}$$

Without loss of generality, the covariance matrices Σ_{4Z} and Σ_Y are assumed positive-definite as the pair $(\alpha, \mathcal{L})$ can always be modified to make this assumption ture. Therefore, there exists an orthogonal matrix $\mathbf{Q}$ such that $\Sigma_Y = \mathbf{Q}\Lambda_Y\mathbf{Q}'$ where $\Lambda_Y = \mathrm{diag}\left(\lambda_{y,i}; 1 \le i \le 2M\right)$. The eigenvalue $\lambda_{y,i}$ is equal to $(\gamma^2\lambda_{x,i} + \gamma\lambda_{xz,i} + \lambda_{z,i})$ in which $\lambda_{x,i}$, $\lambda_{xz,i}$, and $\lambda_{z,i}$ are respectively the diagonal entries of matrices $\Lambda_X = \mathbf{Q}'\Sigma_{4X}\mathbf{Q}$, $\bar{\Lambda}_{XZ} = \mathbf{Q}'\bar{\Sigma}_{XZ}\mathbf{Q}$, and $\Lambda_Z = \mathbf{Q}'\Sigma_{4Z}\mathbf{Q}$. Let's define a vector as $\mathbf{v} \triangleq \mathbf{Q}'\mathbf{r}^{\alpha}_{X} = [v_1, v_2, \ldots, v_{2M}]'$. The average probability of detection can be presented as

$$\begin{aligned} P_d &= \int_0^\infty P_{d|\gamma} f_{\mathrm{Nak}}(\gamma; l)\, \mathrm{d}\gamma, \\ &= \frac{l^l}{\Gamma(l)\,\overline{\gamma}^l} \int_0^\infty Q_M\left(\sqrt{N f(\gamma)}, \sqrt{\lambda}\right) \xi^{l-1} e^{-\frac{l}{\overline{\gamma}}\gamma}\, \mathrm{d}\gamma, \end{aligned} \tag{5.10}$$

where

$$f(\gamma) = \sum_{i=1}^{2M} \left(\frac{v_i^2 \gamma^2}{\lambda_{x,i}\gamma^2 + \lambda_{xz,i}\gamma + \lambda_{z,i}} \right). \tag{5.11}$$

As the closed form of (5.10) is not obtainable, we seek for upper and lower performance bounds. To do this, first, some properties of the function $f(\gamma)$ and the generalized Marcum Q-function are provided. Then we exploit some available series expansion and exponential-type bound of $Q_m(a,b)$ to acquire the wanted analytic performance bounds.

5.2.1 Upper and Lower Bounds of $f(\gamma)$

The function $f(\gamma)$ involves a sum of rational functions of different types resulting from the facts that Σ_{4X} is positive semi-definite, $\bar{\Sigma}_{XZ}$ is not necessarily positive semi-definite, and Σ_{4Z} and Σ_Y are positive definite. The possible types of rational functions involved are given by $f_1(\gamma; a_1, b_1, c_1, d_1) = \frac{a_1\gamma^2}{b_1\gamma^2 + c_1\gamma + d_1}$, $a_1, b_1, c_1, d_1 > 0$, $f_2(\gamma; a_2, b_2, d_2) = \frac{a_2\gamma^2}{b_2\gamma^2 + d_2}$, $a_2, b_2, d_2 > 0$, $f_3(\gamma; a_3, c_3, d_3) = \frac{a_3\gamma^2}{c_3\gamma + d_3}$, $a_3, c_3, d_3 > 0$, and $f_4(\gamma; a_4, b_4, c_4, d_4) = \frac{a_4\gamma^2}{b_4\gamma^2 + c_4\gamma + d_4}$, $a_4, b_4, d_4 > 0$, $-\sqrt{4b_4d_4} < c_4 < 0$. Thus, without loss of generality, the function $f(\gamma)$ can be expanded as

$$\sum_{i\in E_1} f_1\left(\gamma; v_i^2, \lambda_{x,i}, \lambda_{xz,i}, \lambda_{z,i}\right) + \sum_{i\in E_2} f_2\left(\gamma; v_i^2, \lambda_{x,i}, \lambda_{z,i}\right) + \sum_{i\in E_3} f_3\left(\gamma; v_i^2, \lambda_{xz,i}, \lambda_{z,i}\right) + \sum_{i\in E_4} f_4\left(\gamma; v_i^2, \lambda_{x,i}, \lambda_{xz,i}, \lambda_{z,i}\right),$$

where $E_i \subset \{1, 2, \cdots, 2M\}$ and $E_i \cap E_j = \emptyset$ for $i \neq j$. Some upper and lower bounds of $f(\gamma)$ are provided below.

Property 6. *$f(\gamma) \leq s_1\gamma^2$ for $\gamma \geq 0$, where*

$$s_1 = \sum_{i\in\left\{\cup_{n=1}^{3}E_n\right\}} \left(\frac{v_i^2}{\lambda_{z,i}}\right) + \sum_{i\in E_4}\left(\frac{4\lambda_{x,i}v_i^2}{4\lambda_{x,i}\lambda_{z,i}-\lambda_{xz,i}^2}\right). \tag{5.12}$$

Proof. It can be easily shown that $f_n(\gamma) \le \frac{v_i^2}{\lambda_{z,i}}\gamma^2$ where $i \in E_n$ for $\gamma \ge 0$ when $n = 1, 2, 3$. When $i \in E_4$,

$$\lambda_{x,i}\gamma^2 + \lambda_{xz,i}\gamma + \lambda_{z,i} \ge \frac{4\lambda_{x,i}\lambda_{z,i}-\lambda_{xz,i}^2}{4\lambda_{x,i}}, \text{ for } \gamma \ge 0. \tag{5.13}$$

Therefore,

$$f_4(\gamma) \le \frac{4\lambda_{x,i}v_i^2}{4\lambda_{x,i}\lambda_{z,i}-\lambda_{xz,i}^2}\gamma^2, \text{ for } i \in E_4 \text{ and } \gamma \ge 0. \tag{5.14}$$

□

Property 7. *$f(\gamma) \le s_2\gamma$ for $\gamma \ge 0$, where*

$$s_2 = \sum_{i\in\left\{\cup_{n=1}^{2}E_n\right\}} \left[f_n\left(\sqrt{\tfrac{\lambda_{z,i}}{\lambda_{x,i}}}\right)\Big/\sqrt{\tfrac{\lambda_{z,i}}{\lambda_{x,i}}}\right] + \sum_{i\in E_3}\left(\frac{v_i^2}{\lambda_{xz,i}}\right) + \sum_{i\in E_4}\left(\frac{v_i^2}{\lambda_{xz,i}+\sqrt{4\lambda_{x,i}\lambda_{z,i}}}\right).$$

Proof. Let $g_n(\gamma) = \frac{f_n(\gamma)}{\gamma}$ for $\gamma \ge 0$ and $g_n(\gamma) = 0$ elsewhere.

For $n = 1, 2$, $g_n'(\gamma) = 0$ when $\gamma = \sqrt{\frac{\lambda_{z,i}}{\lambda_{x,i}}}$, and $g_n''\left(\sqrt{\frac{\lambda_{z,i}}{\lambda_{x,i}}}\right) < 0$. Thus, $g_n(\gamma) \le g_n\left(\sqrt{\frac{\lambda_{z,i}}{\lambda_{x,i}}}\right)$.

For $n = 3$, it can be easily shown that $g_3(\gamma) \le \frac{v_i^2}{\lambda_{xz,i}}$.

For $n = 4$. It can be verified that $\lambda_{x,i}\gamma^2 + \lambda_{xz,i}\gamma + \lambda_{z,i} \ge \left(\lambda_{xz,i} + \sqrt{4\lambda_{x,i}\lambda_{z,i}}\right)\gamma$ for $\gamma \ge 0$. Thus, $g_4(\gamma) \le \frac{v_i^2}{\lambda_{xz,i}+\sqrt{4\lambda_{x,i}\lambda_{z,i}}}$. □

Property 8. *$f(\gamma) \le s_3\gamma + t_3$ for $\gamma \ge 0$, where $s_3 = \sum_{i\in E_3}\left(\frac{v_i^2}{\lambda_{xz,i}}\right)$ and*

$$t_3 = \sum_{i\in\left\{\cup_{n=1}^{2}E_n\right\}} \left(\frac{v_i^2}{\lambda_{x,i}}\right) + \sum_{i\in E_4}\left(\frac{4\lambda_{z,i}v_i^2}{4\lambda_{z,i}\lambda_{x,i}-\lambda_{xz,i}^2}\right). \tag{5.15}$$

Proof. It can be easily shown that $f_n(\gamma) \le \frac{v_i^2}{\lambda_{x,i}}$ where $i \in E_n$ for $\gamma \ge 0$ when $n = 1, 2$.

For $n = 4$. It can be shown that $\lambda_{x,i}\gamma^2 + \lambda_{xz,i}\gamma + \lambda_{z,i} \ge \left(\lambda_{x,i} - \frac{\lambda_{xz,i}^2}{4\lambda_{z,i}}\right)\gamma^2$ for $\gamma \ge 0$. Thus, $f_4(\gamma) \le \frac{4\lambda_{z,i}v_i^2}{\left(4\lambda_{z,i}\lambda_{x,i} - \lambda_{xz,i}^2\right)}$.

By using the fact that $f_3(\gamma) \le \frac{v_i^2}{\lambda_{xz,i}}\gamma$ for $i \in E_3$, the desired result can be obtained.

□

Property 9. *$f(\gamma) \ge s_4\gamma^2$ for $\gamma \in [0, \gamma_0]$, where*

$$s_4 = \sum_{i \in \left\{\cup_{n=1}^{3} E_n\right\}} \frac{f_n(\gamma_0)}{\gamma_0^2} + \sum_{i \in E_4} \frac{v_i^2}{\lambda_{x,i}\gamma_0^2 + \lambda_{z,i}}. \tag{5.16}$$

Proof. For $\gamma \in [0, \gamma_0]$, it can be easily shown that $f_n(\gamma) - \frac{f_n(\gamma_0)}{\gamma_0^2}\gamma^2 \ge 0$ when $n = 1, 2, 3$.

When $n = 4$, $\lambda_{x,i}\gamma^2 + \lambda_{xz,i}\gamma + \lambda_{z,i} \le \lambda_{x,i}\gamma^2 + \lambda_{z,i}$. Therefore, $f_4(\gamma) \ge \frac{v_i^2\gamma^2}{\lambda_{x,i}\gamma^2 + \lambda_{z,i}} \ge \frac{v_i^2\gamma^2}{\lambda_{x,i}\gamma_0^2 + \lambda_{z,i}}$ for $\gamma \in [0, \gamma_0]$. □

Property 10. *$f(\gamma) \ge s_5\gamma + t_5$ for $\gamma \in [\gamma_0, \infty)$, where $s_5 = \sum_{i \in E_3} f_3'(\gamma_0)$ and*

$$t_5 = \sum_{i \in \left\{\cup_{n=1}^{3} E_n\right\}} f_n(\gamma_0) + \sum_{i \in E_4} \frac{v_i^2\gamma_0^2}{\lambda_{x,i}\gamma_0^2 + \lambda_{z,i}} - \left[\sum_{i \in E_3} f_3'(\gamma_0)\right]\gamma_0. \tag{5.17}$$

Proof. When $n = 1, 2$, it can be shown that $f_n(\gamma)$ is monotonically increasing for $\gamma \in [0, \infty)$. Therefore, $f_n(\gamma) \ge f_n(\gamma_0)$ for $\gamma \in [\gamma_0, \infty)$.

When $n = 3$, $f_3(\gamma) \ge f_3(\gamma_0) + f_3'(\gamma_0)(\gamma - \gamma_0)$ for $\gamma \in [\gamma_0, \infty)$ as $f_3'(\gamma) > 0$ and $f_3''(\gamma) > 0$ for $\gamma \in [0, \infty)$.

When $n = 4$, it has been indicated that $f_4(\gamma) \ge \frac{v_i^2\gamma^2}{\lambda_{x,i}\gamma^2 + \lambda_{z,i}}$ for $\gamma \in [0, \infty)$. Thus, $f_4(\gamma) \ge \frac{v_i^2\gamma_0^2}{\lambda_{x,i}\gamma_0^2 + \lambda_{z,i}}$ for $\gamma \in [\gamma_0, \infty)$. □

5.2.2 Upper Bounds on the Average Detection Probability

To apply the upper and lower bounds in the previous section requires a monotonicity property of the generalized Marcum Q-function, which is provided below.

Property 11. *$Q_m(a, b)$ is greater than zero and monotonically increasing on $a \in (0, \infty)$ for each $b > 0$, $m \in \mathbb{N}$.*

Proof. Employing [96, Eq.(4.44)], we can show that for each $b > 0$, $m \in \mathbb{N}$,

$$Q_m(0, b) = \frac{\Gamma\left(m, {}^{b^2}/_{2}\right)}{\Gamma(m)} > 0. \tag{5.18}$$

Differentiating $Q_m(a, b)$ with respect to a by using [97, Eq.(16)] yields

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}a} Q_m(a, b) &= -a\left[Q_m(a, b) - Q_{m+1}(a, b)\right], \\ &> 0, \end{aligned}$$

where the inequality holds because $Q_r(a, b)$ is strictly increasing on $r \in (0, \infty)$ for each $a \geq 0$, $b > 0$ [98]. □

By making use of Property 6 and Property 11, the detection performance can be upper bounded by

$$\begin{aligned} & P_d \leq \frac{l^l}{\Gamma(l)\,\overline{\gamma}^l} \int_0^\infty Q_M\left(\sqrt{N s_1 \gamma^2}, \sqrt{\lambda}\right) \gamma^{l-1} e^{-\frac{l}{\overline{\gamma}}\gamma}\, \mathrm{d}\gamma, \\ & = \frac{l^l}{\Gamma(l)\,\overline{\gamma}^l} \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{N s_1}{2}\right)^n \left[\sum_{k=0}^{n+M-1} e^{-\frac{\lambda}{2}} \frac{\left(\frac{\lambda}{2}\right)^k}{k!}\right] G_{01}\left(0, \infty; 2n + l - 1; -\frac{N s_1}{2}, -\frac{l}{\overline{\gamma}}, 0\right), \\ & \triangleq P_{d,\mathrm{UB01}}, \end{aligned} \tag{5.19}$$

where the generalized Marcum Q-function is replaced by its series expansion [96, Eq.(4.47)]. The function $G_{01}(u, v; n; a, b, c)$ is defined below

$$\int_u^v x^n \exp\left(ax^2 + bx + c\right) \mathrm{d}x, \text{ for } n \geq 0 \text{ and } a < 0,$$

$$= \exp\left(\frac{4ac - b^2}{4a}\right) \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} \left(-\frac{b}{2a}\right)^{n-k} \left\{ \left. \frac{-\Gamma\left[\frac{k+1}{2}, -ay^2\right]}{2\left(-a\right)^{(k+1)/2}} \right|_{y=u+\frac{b}{2a}}^{v+\frac{b}{2a}} \right\},$$

$$\triangleq G_{01}\left(u, v; n; a, b, c\right),$$

where [99, Eq.(2.33.10)] is applied. Similarly, applying Property 7 and [87, Eq.(7,8)] yields another upper bound, namely

$$P_{d,\mathrm{UB02}} \triangleq \mathrm{A}_1 + \beta^l e^{-\frac{\lambda}{2}} \sum_{n=1}^{M-1} \frac{\left(\frac{\lambda}{2}\right)^n}{n!} {}_1F_1\left(l; n+1; \frac{\lambda\left(1-\beta\right)}{2}\right), \tag{5.20}$$

where $\beta = \frac{(2l)}{(2l + N s_2 \overline{\gamma})}$, ${}_1F_1$ denotes the confluent hypergeometric function, and A_1 is given by

$$\mathrm{A}_1 = e^{-\frac{\lambda\beta}{2}} \left[\beta^{l-1} L_{l-1}\left(\frac{-\lambda\left(1-\beta\right)}{2}\right) + \left(1-\beta\right) \sum_{n=0}^{l-2} \beta^n L_n\left(\frac{-\lambda\left(1-\beta\right)}{2}\right)\right], \tag{5.21}$$

where L_n stands for Laguerre polynomial of degree n.

Subsequently, making use of Property 8 will lead to the third upper bound. In this section, this third upper bound is not discussed because it is effective at high average SNR.

5.2.3 Lower Bound on the Average Detection Probability

A lower bound [100, Eq.(11,12)] on the generalized Marcum Q-function $Q_m\left(a, b\right)$ to be used is presented below,

$$1-\frac{1}{2}\exp\left[-\frac{(a-b)^2}{2}\right]+\left[\sum_{k=1}^{m-1}\frac{\left(\frac{b}{2}\right)^k}{k!}+\frac{1}{2}\right]\exp\left[-\frac{(a+b)^2}{2}\right],\ \text{for } 0<b<a. \quad (5.22)$$

By making use of Property 9 and Property 10, we can define a lower bound $f_{\text{LB}}(\gamma)$ of $f(\gamma)$ as

$$f_{\text{LB}}(\xi)=\begin{cases} s_4\gamma^2 & \text{if } 0\leq\gamma<\gamma_0, \\ s_5\gamma+t_5 & \text{if } \gamma_0\leq\gamma<\infty, \end{cases} \quad (5.23)$$

where $\gamma_0 \triangleq \gamma_{ab}$ and γ_{ab} is the point such that $f(\gamma_{ab})=\frac{\lambda}{N}$. With this in mind, a lower bound on P_d can be given by

$$\begin{aligned} P_{d,\text{LB}} \triangleq \frac{l^l}{\Gamma(l)\,\overline{\gamma}^l}\Bigg\{&\int_0^{\gamma_{ab}} Q_M\left(\sqrt{Ns_4\gamma^2},\sqrt{\lambda}\right)\gamma^{l-1}e^{-\frac{l}{\overline{\gamma}}\gamma}\,\mathrm{d}\gamma \\ &+\int_{\gamma_{ab}}^{\infty} Q_M\left(\sqrt{N(s_5\gamma+t_5)},\sqrt{\lambda}\right)\gamma^{l-1}e^{-\frac{l}{\overline{\gamma}}\gamma}\,\mathrm{d}\gamma\Bigg\}. \quad (5.24)\end{aligned}$$

Applying the same technique used in (5.19) to the first integral of (5.24), we obtain its lower bound

$$\begin{aligned}&\int_0^{\gamma_{ab}} Q_M\left(\sqrt{Ns_4\gamma^2},\sqrt{\lambda}\right)\gamma^{l-1}e^{-\frac{l}{\overline{\gamma}}\gamma}\,\mathrm{d}\gamma \\ &\geq \sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{Ns_1}{2}\right)^n G_{01}\left(0,\xi_{ab};2n+l-1;-\frac{Ns_4}{2},-\frac{l}{\overline{\gamma}},0\right)\sum_{k=0}^{n+M-1}e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^k}{k!}. \quad (5.25)\end{aligned}$$

If $s_5\neq 0$, we rewrite the second integral in (5.24) as

$$
\begin{aligned}
&\int_{\gamma_{ab}}^{\infty} Q_M\left(\sqrt{N\left(s_5\gamma+t_5\right)},\sqrt{\lambda}\right)\gamma^{l-1}e^{-\frac{l}{\overline{\gamma}}\gamma}\,\mathrm{d}\gamma \\
&\qquad = 2\mathrm{exp}\left(\frac{lt_5}{\overline{\gamma}s_5}\right)\sum_{k=0}^{l-1}\left(\begin{array}{c} l-1 \\ k \end{array}\right)\left(-\frac{t_5}{s_5}\right)^{l-1-k} \\
&\qquad\qquad \times\int_{\psi_0}^{\infty} Q_M\left(\sqrt{Ns_5\psi^2},\sqrt{\lambda}\right)\psi^{2l-1}e^{-\frac{l}{\overline{\gamma}}\psi^2}\,\mathrm{d}\psi. \quad (5.26)
\end{aligned}
$$

where $\psi=\sqrt{\gamma+\frac{t_5}{s_5}}$ and $\psi_0\triangleq\sqrt{\gamma_{ab}+\frac{t_5}{s_5}}$. By using (5.22), the integral in (5.26) can be lower bounded by

$$
\begin{aligned}
&\int_{\psi_0}^{\infty} Q_M\left(\sqrt{Ns_5\psi^2},\sqrt{\lambda}\right)\psi^{2l-1}e^{-\frac{l}{\overline{\gamma}}\psi^2}\,\mathrm{d}\psi \\
&\geq G_{02}\left(\psi_0,\infty;2l-1,2;\frac{l}{\overline{\gamma}}\right)+\left[\sum_{k=1}^{M-1}\frac{\left(\frac{\sqrt{\lambda}}{2}\right)^k}{k!}+\frac{1}{2}\right] \\
&\qquad \times G_{01}\left(\psi_0,\infty;2l-1;\frac{-Ns_5}{2}-\frac{l}{\overline{\gamma}},-\sqrt{\lambda Ns_5},\frac{-\lambda}{2}\right) \\
&\qquad\qquad -\frac{1}{2}G_{01}\left(\psi_0,\infty;2l-1;\frac{-Ns_5}{2}-\frac{l}{\overline{\gamma}},\sqrt{\lambda Ns_5},\frac{-\lambda}{2}\right), \quad (5.27)
\end{aligned}
$$

where applying [99, Eq.(2.33.10)] yields

$$
\begin{aligned}
G_{02}\left(u,v;l,n;a\right) &\triangleq \int_u^v x^{l-1}e^{-ax^n}\,\mathrm{d}x,\ \text{for } a\neq 0,\ n\neq 0, \\
&= \frac{\Gamma\left(\frac{l}{n},au^n\right)-\Gamma\left(\frac{l}{n},av^n\right)}{na^{\frac{l}{n}}}.
\end{aligned}
$$

If $s_5=0$, we have

$$
\int_{\gamma_{ab}}^{\infty} Q_M\left(\sqrt{Nt_5},\sqrt{\lambda}\right)\gamma^{l-1}e^{-\frac{l}{\overline{\gamma}}\gamma}\,\mathrm{d}\gamma = Q_M\left(\sqrt{Nt_5},\sqrt{\lambda}\right)G_{02}\left(\gamma_{ab},\infty;l,1;\frac{l}{\overline{\gamma}}\right). \quad (5.28)
$$

With the above results (5.24)~(5.28), the required lower bound $P_{d,\mathrm{LB}}$ can be obtained.

5.3 Post-Combining over IID Rayleigh Fading Channels

In this section, we examine the detection performance of J-branch post-combining where each branch is subject to i.i.d. Rayleigh fading. Based on the same tested cycle frequency α and lag set $\mathcal{L}$, each branch generates a statistic $\mathcal{T}_j = N f_j(\gamma_j)$ where j is the branch index and the PDF of γ_j is given by

$$f_{\text{Ray}}(\gamma_j) = \frac{1}{\bar{\gamma}} \exp\left(-\frac{\gamma_j}{\bar{\gamma}}\right). \tag{5.29}$$

The statistics resulting from J branches are integrated by using either post addition combining or post selection combining. Different from two post-combining schemes introduced in Section 4.4, the statistics $\mathcal{T}_j$ rather than the CAF-vector estimates $\hat{\mathbf{r}}^{\alpha}_{Y,j}$ for $1 \leq j \leq J$ are exploited.

Instead of providing performance bounds, we will first introduce two approximated detection performance over Nakagami fading channels and utilize them to derive approximated closed-form results. Numerical results will show the tightness of the approximated performance.

Two approximations of $f(\gamma)$ to be used are given below.

Approximation 1. *$f(\gamma) \approx s_6 \gamma^2$ at low SNR, where*

$$s_6 = \sum_{i \in \left\{\cup_{n=1}^{4} E_n\right\}} \left(v_i^2 / \lambda_{z,i}\right). \tag{5.30}$$

Approximation 2. *$f(\gamma) \approx s_7\gamma + t_7$ at high SNR, where $s_7 = \sum_{i\in E_3} \left(v_i^2/\lambda_{xz,i}\right)$ and*

$$t_7 = \sum_{i\in\{E_1\cup E_2\cup E_4\}} \left(v_i^2/\lambda_{x,i}\right). \tag{5.31}$$

By making use of Approximation 1, the average detection performance over Nakagami fading channels can be approximated by

$$\begin{aligned} P_d &\approx \frac{l^l}{\Gamma(l)\,\overline{\gamma}^l}\int_0^\infty Q_M\left(\sqrt{N s_6\gamma^2}, \sqrt{\lambda}\right)\gamma^{l-1}e^{-\frac{l}{\overline{\gamma}}\gamma}\,\mathrm{d}\gamma, \\ &\triangleq \widetilde{P}_{d,01}, \end{aligned} \tag{5.32}$$

where the closed form of the integral can be obtained with the similar way shown in (5.19). As Approximation 1 is used, $\widetilde{P}_{d,01}$ is expected to serve as a good approximation at low average SNR. On the other hand, once Approximation 2 is applied, we expect to approximate the detection performance well at high average SNR. For CR applications, the interesting operational region is at low SNR, so we focus on the first approximation in this section.

Let's define an approximation $\widetilde{f}(\gamma)$ of $f(\gamma)$ as

$$\widetilde{f}(\gamma) = \begin{cases} s_6\gamma^2 & \text{if } 0 \le \gamma < \gamma_0, \\ f_{max} & \text{if } \gamma_0 \le \gamma < \infty, \end{cases} \tag{5.33}$$

where $\gamma_0 = \sqrt{\lambda/N s_1}$, $f_{max} = t_7$ if $s_7 = 0$, and $f_{max} = \infty$ if $s_7 \neq 0$. With this in mind, another approximated detection probability can be given by

$$P_d \approx \frac{l^l}{\Gamma(l)\,\overline{\gamma}^l}\left\{\int_0^{\gamma_0} Q_M\left(\sqrt{Ns_6\gamma^2},\sqrt{\lambda}\right)\gamma^{l-1}e^{-\frac{l}{\overline{\gamma}}\gamma}\,\mathrm{d}\gamma\right.$$
$$\left.+\int_{\gamma_0}^{\infty} Q_M\left(\sqrt{Nf_{max}},\sqrt{\lambda}\right)\gamma^{l-1}e^{-\frac{l}{\overline{\gamma}}\gamma}\,\mathrm{d}\gamma\right\} \triangleq \widetilde{P}_{d,02}. \quad (5.34)$$

After some manipulation, the second approximation can be represented by

$$\widetilde{P}_{d,02} = \frac{l^l}{\Gamma(l)\,\overline{\gamma}^l}\left\{\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{Ns_1}{2}\right)^n\right.$$
$$\times\left[\sum_{k=0}^{n+M-1} e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^k}{k!}\right] G_{01}\left(0,\gamma_0;2n+l-1;-\frac{Ns_1}{2},-\frac{l}{\overline{\gamma}}\right)$$
$$\left.+Q_M\left(\sqrt{Nf_{max}},\sqrt{\lambda}\right)G_{02}\left(\gamma_0,\infty;l,1;\frac{l}{\overline{\gamma}}\right)\right\}. \quad (5.35)$$

It will be shown using computer simulation that the series in $\widetilde{P}_{d,02}$ converges faster than that in $\widetilde{P}_{d,01}$.

5.3.1 Post Addition Combining

The PAC combiner simply produces the sum of statistics from each branch, i.e., $\mathcal{T}_{\text{PAC}} = \sum_{j=1}^{J}\mathcal{T}_j$. This sum is noncentral chi-square distributed with $2JM$ DOFs and the noncentrality parameter $N\sum_{j=1}^{J} f_j(\gamma_j)$. Thus, the corresponding conditional detection probability can be given by

$$P_{d|\{\gamma_j\}_{j=1}^J} = Q_{JM}\left(\sqrt{N\sum_{j=1}^{J} f_j(\gamma_j)},\sqrt{\lambda}\right). \quad (5.36)$$

Applying Approximation 1 to (5.36) gives

$$P_{d|\{\gamma_j\}_{j=1}^J} \approx Q_{JM}\left(\sqrt{Ns_1\gamma_{\text{PAC}}^2},\sqrt{\lambda}\right), \quad (5.37)$$

where $\gamma_{\text{PAC}} \triangleq \sqrt{\sum_{j=1}^{J} \gamma_j^2}$. The closed-form PDF of γ_{PAC} is unknown, so we seek for its simple approximation. By using the method proposed in [101], we can obtain a precisely approximated PDF of γ_{PAC}^2 in terms of an $\alpha-\mu$ distribution, and then further approximate the PDF of γ_{PAC} with a gamma distribution, namely

$$f_{\text{PAC}}\left(\gamma_{\text{PAC}}\right) \approx \frac{1}{\theta^{\kappa}\Gamma\left(\kappa\right)}\gamma_{\text{PAC}}^{(\kappa-1)}\exp\left[-\frac{\gamma_{\text{PAC}}}{\theta}\right], \tag{5.38}$$

where the method of obtaining κ and θ is presented in Section 5.A. By employing the similar technique in the previous section, two approximations of the average detection probability of J-branch PAC can be given respectively by

$$\begin{aligned}\widetilde{P}_{d,\text{PAC1}} = & \frac{1}{\theta^{\kappa}\Gamma\left(\kappa\right)}\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{Ns_1}{2}\right)^n\left[\sum_{k=0}^{n+JM-1}e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^k}{k!}\right] \\ & \times G_{01}\left(0,\infty;2n+\kappa-1;-\frac{Ns_1}{2},-\frac{1}{\theta}\right),\end{aligned} \tag{5.39}$$

and

$$\begin{aligned}\widetilde{P}_{d,\text{PAC2}} = & \frac{1}{\theta^{\kappa}\Gamma\left(\kappa\right)}\left\{\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{Ns_1}{2}\right)^n\left[\sum_{k=0}^{n+JM-1}e^{-\frac{\lambda}{2}}\right.\right. \\ & \left.\times\frac{\left(\frac{\lambda}{2}\right)^k}{k!}\right]G_{01}\left(0,\gamma_0;2n+\kappa-1;-\frac{Ns_1}{2},-\frac{1}{\theta}\right) \\ & \left.+Q_{JM}\left(\sqrt{Nf_{max}},\sqrt{\lambda}\right)G_{02}\left(\gamma_0,\infty;\kappa,1;-\frac{1}{\theta}\right)\right\}.\end{aligned} \tag{5.40}$$

Without performing multiple integration which is required for obtaining the average detection probability, two approximations can be attained.

5.3.2 Post Selection Combining

The PSC combiner selects the branch with the largest test statistic, i.e., $\mathcal{T}_{\text{PSC}} = \max\{\mathcal{T}_1, \cdots, T_J\}$. The corresponding SNR γ_{PSC} is $\max\{\gamma_1, \cdots, \gamma_J\}$ with its PDF given by

$$f_{\text{PSC}}(\gamma_{\text{PSC}}) = J\sum_{i=0}^{J-1}\begin{pmatrix} J-1 \\ i \end{pmatrix}\frac{(-1)^i}{i+1}\frac{1}{\overline{\gamma}/(i+1)}\exp\left[-\frac{\gamma_{\text{PSC}}}{\overline{\gamma}/(i+1)}\right], \tag{5.41}$$

which is a linear combination of exponential PDFs with the parameter $\overline{\gamma}/(i+1)$. Therefore, the approximations of the average detection probability for the PSC scheme can be presented as

$$\widetilde{P}_{d,\text{PSC1}} = J\sum_{i=0}^{J-1}\begin{pmatrix} J-1 \\ i \end{pmatrix}\frac{(-1)^i}{i+1}\widetilde{P}_{d,01}\left(1, \frac{\overline{\gamma}}{i+1}\right), \tag{5.42}$$

and

$$\widetilde{P}_{d,\text{PSC2}} = J\sum_{i=0}^{J-1}\begin{pmatrix} J-1 \\ i \end{pmatrix}\frac{(-1)^i}{i+1}\widetilde{P}_{d,02}\left(1, \frac{\overline{\gamma}}{i+1}\right), \tag{5.43}$$

where $\widetilde{P}_{d,01}\left(1, \overline{\gamma}/(i+1)\right)$ and $\widetilde{P}_{d,02}\left(1, \overline{\gamma}/(i+1)\right)$ are respectively $\widetilde{P}_{d,01}$ and $\widetilde{P}_{d,02}$ in which $l = 1$ and $\overline{\gamma}$ is replaced by $\overline{\gamma}/(i+1)$.

5.4 Numerical Results

In this section, we examine an example for which the analytical CAF vector $\mathbf{r}_Y^\alpha$ and the covariance matrix Σ_Y under $\mathcal{H}_1$ are both obtainable. Let's define a communication signal $X_n = \sum_{k=-\infty}^{\infty} a_k p\left(nT_s - kT_{sym}\right)$, where T_s is the sampling interval, T_{sym} is

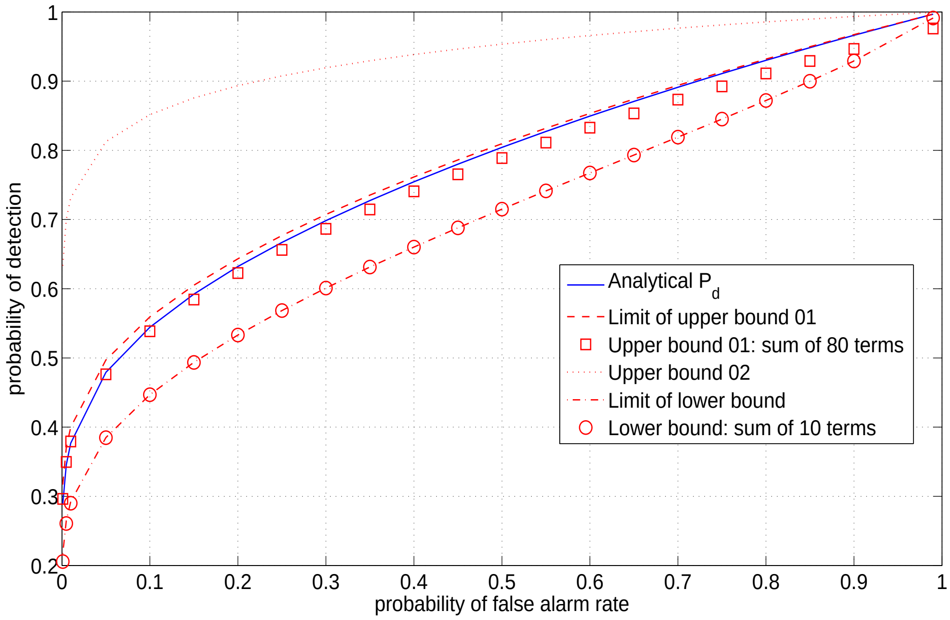


Figure 5.1: Upper and lower bounds of the ROC curve for a Nakagami fading channel (N =6000, l=1, $\bar{\gamma} = -12$dB)

the symbol interval, a_k denotes identically and independently distributed zero-mean complex-valued 16-QAM symbols, and $p(t)$ is a rectangular pulse with value 1 for $0 \le t < T_{sym}$ and value 0 elsewhere. This signal exhibits cyclostationary features at cyclic frequencies $\alpha = {}^{2\pi k}/_{N_{sym}}$, where $k \in \mathbb{Z}$ and $N_{sym} = {}^{T_{sym}}/_{T_s}$. Assuming that $N_{sym} = 6$. The cycle frequency and the lag set for testing are given respectively by $\alpha = {}^{2\pi}/_{6}$ and $\mathcal{L} = \{-3, 3\}$.

5.4.1 Performance Bounds

Fig. 5.1 presents the receiver operating characteristic curve and its upper and lower bounds over a Nakagami fading channel ($l = 1$) with low average SNR $\bar{\gamma} = -12$dB. The number of used samples is $N = 6000$. The ROC curve results from the numerical integration of (5.10). As shown in this figure, the first upper bound $P_{d,\text{UB01}}$, the sum

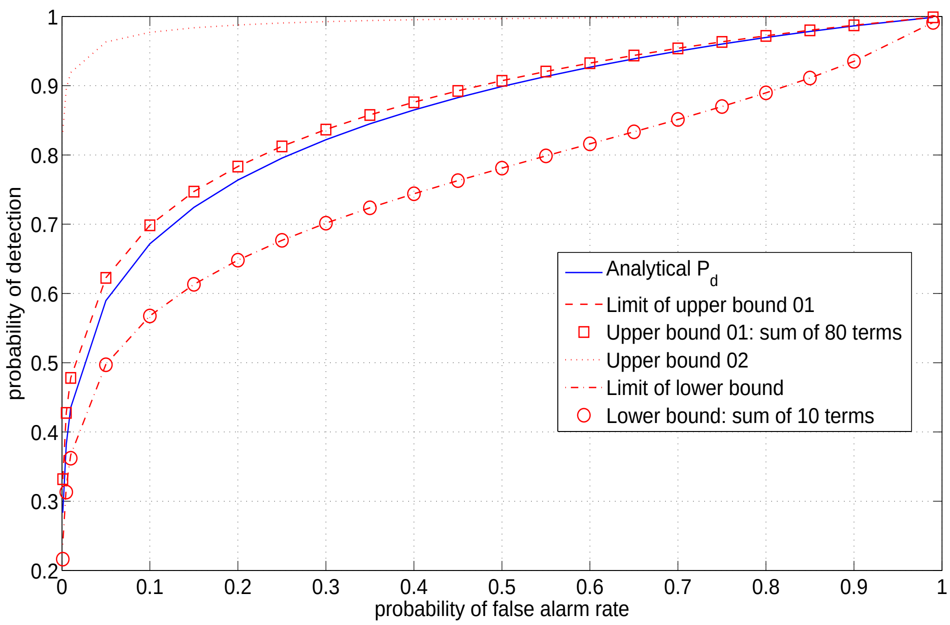
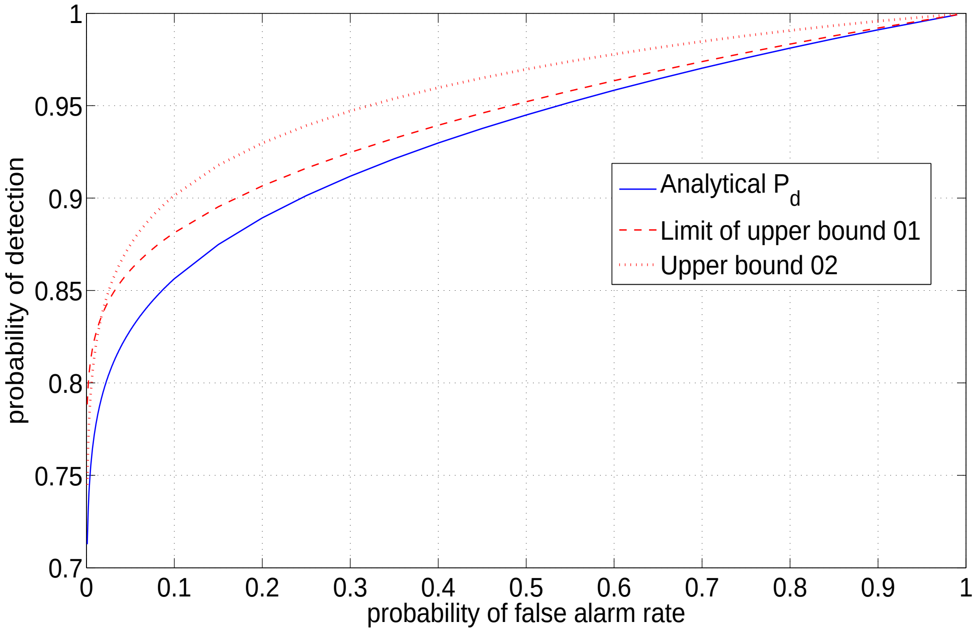


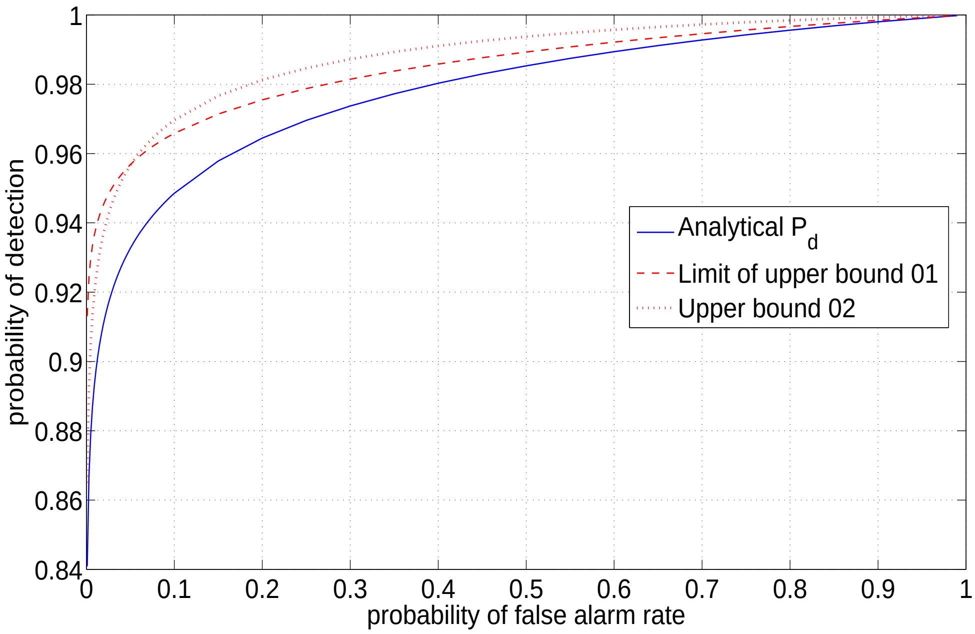
Figure 5.2: Upper and lower bounds of the ROC curve for a Nakagami fading channel (N =6000, l=4, $\bar{\gamma} = -12$dB)

of first 80 terms in the series expression (5.19), approaches its limiting values and can serve as a tight approximation. The explanation for $P_{d,\text{UB01}}$ being a tight performance bound is that the upper bound in Property 6 captures the behavior of $f(\gamma)$ very well at low average SNR when the set E_4 is empty. It is apparent that $P_{d,\text{UB02}}$ does not provide a satisfactory performance bound. For the lower bound, $P_{d,\text{LB}}$ can properly bound the ROC curve from below to some extent. The case of $l = 4$ is shown in Fig. 5.2. It can be seen that the upper bound $P_{d,\text{UB02}}$ is still loose, while $P_{d,\text{UB01}}$ keeps as a tight upper bound.

The ROC curve and its upper bounds at higher average SNR $\bar{\gamma} = 0$dB with sample size $N = 600$ over different Nakagami fading settings (l=1 and 2) are plotted in Fig. 5.3. The curve due to the upper bound $P_{d,\text{UB01}}$ is presented by using limiting values. As can be seen in Fig.5.3(a) and Fig.5.3(b), the upper bound $P_{d,\text{UB02}}$ becomes tighter



(a) l=1



(b) l=2

Figure 5.3: Upper bounds of the ROC curve for Nakagami fading channels (N =600, $\bar{\gamma}$ = 0dB)

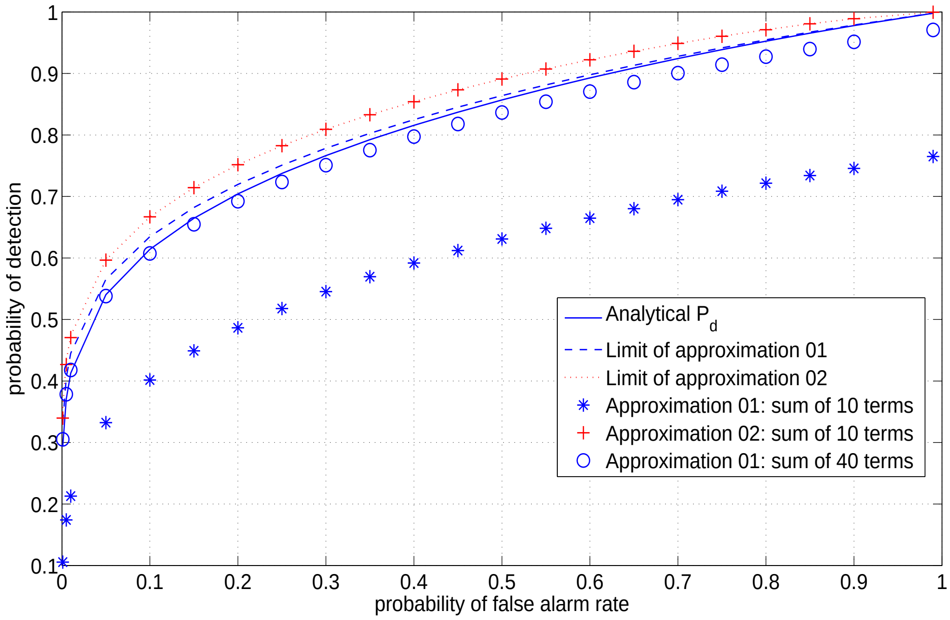


Figure 5.4: Approximated ROC curve for a Nakagami fading channel (l=2).

at this higher average SNR and can be tighter than $P_{d,\mathrm{UB01}}$ for small false alarm rates. This is partly because the upper bound in Property 7 captures well the behavior of $f(\gamma)$ at higher average SNR.

5.4.2 Post-Combining

Fig. 5.4 illustrates the receiver operating characteristic curve and its two approximations, resulting from $\widetilde{P}_{d,01}$ and $\widetilde{P}_{d,02}$, over a Nakagami fading channel with the parameter $l = 2$. As shown in this figure, the ROC curve due to the limiting values of $\widetilde{P}_{d,01}$ is tighter than the curve due to $\widetilde{P}_{d,02}$. Moreover, the ROC curve resulting from the partial sum of first 10 terms in the series of $\widetilde{P}_{d,02}$ has almost converged to its limiting curve. On the other hand, the ROC curves resulting from first 10 or 40 terms in the series of $\widetilde{P}_{d,01}$ obviously have not converged to its limiting curve. In other

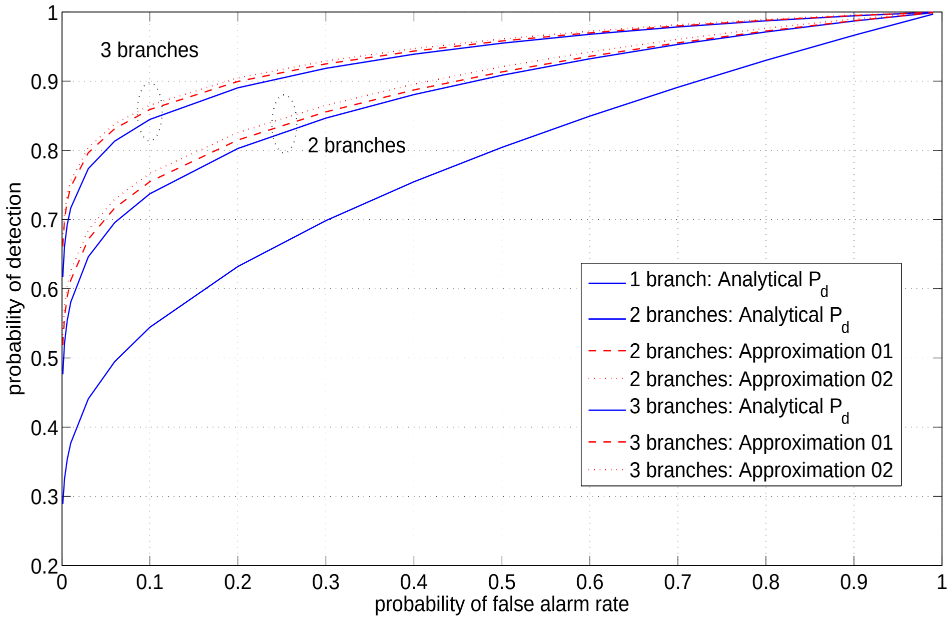


Figure 5.5: Approximated ROC curves for D-branch PAC.

words, the sequence of partial sums due to $\widetilde{P}_{d,02}$ converges to its limit faster than the sequence due to $\widetilde{P}_{d,01}$.

Fig. 5.5 and Fig. 5.6 respectively present the ROC curves and the complementary ROC curves for J-branch PAC over i.i.d. Rayleigh fading channels. In these two figures, only the limiting curves due to $\widetilde{P}_{d,\text{PAC1}}$ and $\widetilde{P}_{d,\text{PAC2}}$ are plotted. It can be seen that the ROC curves or their complementary curves due to our proposed approximations, $\widetilde{P}_{d,\text{PAC1}}$ and $\widetilde{P}_{d,\text{PAC2}}$, can closely approach the curves due to the analytical P_d or $1 - P_d$.

The numerical results for J-branch PSC can be found in Fig. 5.7 and Fig. 5.8. In both figures, it is apparent that both the ROC curves and the complementary ROC curves can be tightly approximated by our proposed.

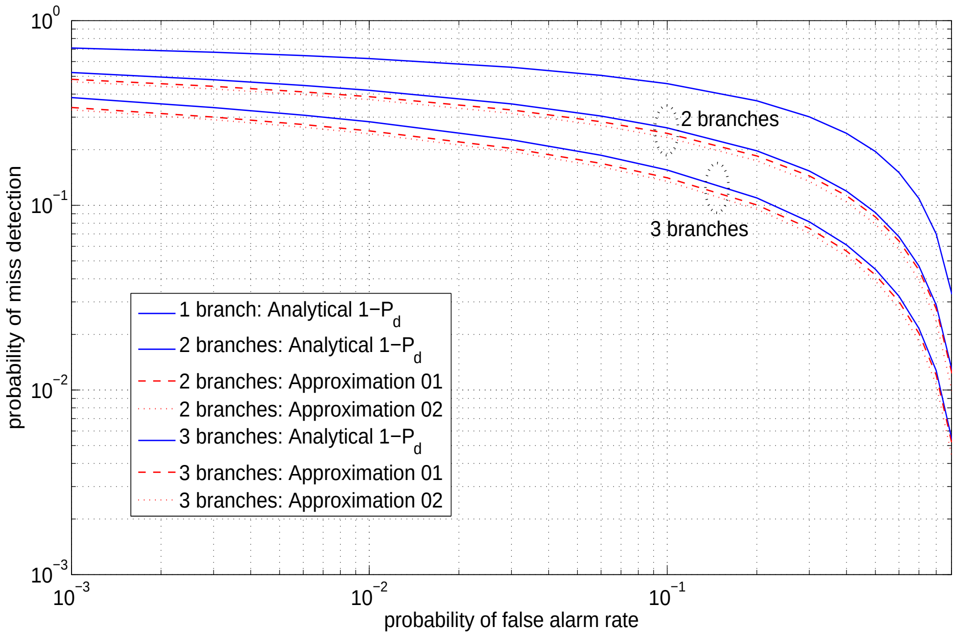


Figure 5.6: Approximated complementary ROC curves for D-branch PAC.

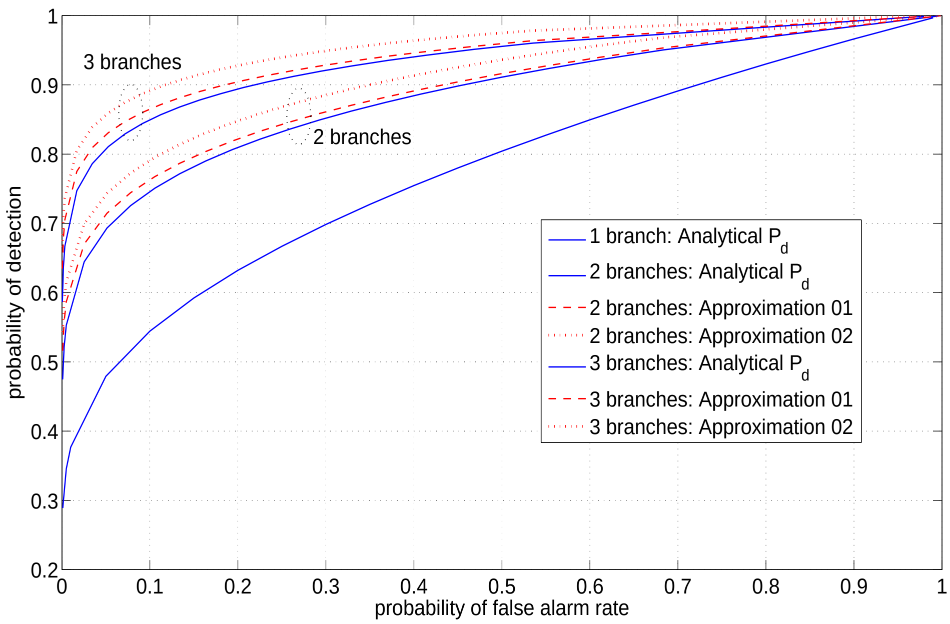


Figure 5.7: Approximated ROC curves for J-branch PSC.

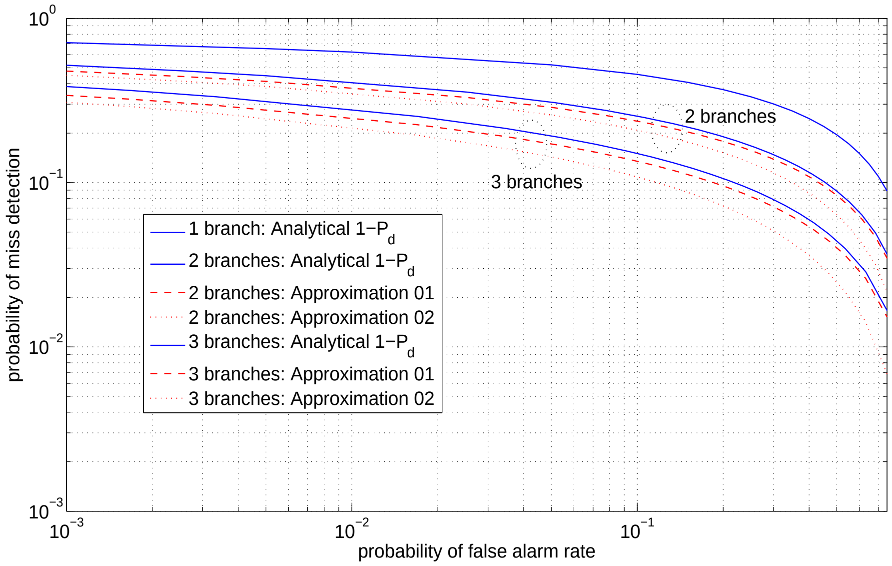


Figure 5.8: Approximated complementary ROC curves for J-branch PSC.

5.5 Conclusions

An analytic investigation of the second-order cyclostationary feature detection has been carried out in this chapter. A tight upper bound of average detection performance at low average SNR is presented as an infinite series. For the higher average SNR region, the proposed closed-form upper bound can serve as a reasonable approximation of detection performance. In addition, approximated detection performances in a series form are analytically presented for post addition combining and post selection combining. These approximations are especially effective at low average SNR which is the region of interest for cognitive radio applications.

The results of performance bounds over fading channels have been presented in IEEE Wireless Communications and Networking Conference (WCNC) [102]. The analysis of post-combining schemes is going to be presented in IEEE Global Communications Conference (GLOBECOM) 2013 [103].

5.A Derivation of (5.38)

As shown in (5.29), the variate γ_j is exponentially distributed. So its square $\psi_j = \gamma_j^2$ is Weibull distributed, i.e., $f_{\text{Wei}}\left(\psi_j\right) = {}^1\!/_{2\bar{\gamma}\sqrt{\psi_j}} \exp\left(-\sqrt{\psi_j}/\bar{\gamma}\right)$. Let's define ψ_{PAC} as γ_{PAC}^2, the sum of i.i.d. Weibull variates. In [101], an $\alpha - \mu$ distribution used to precisely approximate the PDF of the variate ψ_{PAC} is given by

$$f\left(\psi_{\text{PAC}}\right) \approx \frac{\beta\mu^{\mu}\psi_{\text{PAC}}^{\beta\mu-1}}{\Omega^{\mu}\Gamma\left(\mu\right)} \exp\left(-\frac{\mu\psi_{\text{PAC}}^{\beta}}{\Omega}\right), \tag{5.44}$$

where the parameters, β, μ, and Ω, are determined through a moments matching method. Employing the similar idea, we try to approximate the PDF of γ_{PAC} with a gamma distribution as shown in (5.38). The reason for using the gamma distribution is that it is a generalized exponential distribution in which two parameters, κ and θ, need to be determined. Another reason is that this distribution is mathematically tractable for evaluating the average detection probability. Matching the first and the second moments, $\mathbb{E}\left[\gamma_{\text{PAC}}\right] = \mathbb{E}\left[\sqrt{\psi_{\text{PAC}}}\right]$ and $\mathbb{E}\left[\gamma_{\text{PAC}}^2\right] = \mathbb{E}\left[\psi_{\text{PAC}}\right]$, yields

$$\kappa = \frac{\zeta_1^2}{\zeta_2 - \zeta_1^2}, \tag{5.45}$$

and

$$\theta = \frac{\zeta_2 - \zeta_1^2}{\zeta_1}, \tag{5.46}$$

where $\zeta_1 = \frac{\beta\mu^{\mu}}{\Omega^{\mu}\Gamma(\mu)} G_{02}\left(0, \infty; \beta\mu + \frac{1}{2}, \beta; \frac{\mu}{\Omega}\right)$ and $\zeta_2 = \frac{\beta\mu^{\mu}}{\Omega^{\mu}\Gamma(\mu)} G_{02}\left(0, \infty; \beta\mu + 1, \beta; \frac{\mu}{\Omega}\right)$.

CHAPTER 6

INTERFERENCE CHANNEL ESTIMATION

6.1 Introduction

The conventional strategies for interference management, such as transmission power limit, spectral masking, and transmitter location arrangement, are transmitter-centric [104, 105]. All these strategies are intended for minimizing interference to receivers without knowing the actual level of interference experienced by them. It means that the transmitter-centric strategies are less adaptive to the real interference level. If applying these strategies to the CR network, it will lead to less efficient CR communications.

In 2002, FCC proposed a new metric, interference temperature, measuring the RF interference received by a PU receiver with which the CR network has to coexist. Since then, the past decade has seen a rapid development of receiver-centric CR interference management. The immediate benefit of using this metric is that it provides the criterion to judge if harmful interference has been introduced in the band of interest. At the same time, the maximum tolerable interference level, which should not be exceeded by cumulative interference from the CR network, can also be estimated. Hence, a more flexible and accurate cap can be derived and placed on CU transmission power. CR communications can be performed whenever the amount of interference to each PU receiver is within the limit.

The scheme to utilize interference temperature can be presented as follows. First, each PU receiver measures its own interference temperature and decides its own interference limit based on its receiver sensitivity. Being informed of this PU-side information by the PU system, the CR system should transmit signals within a power constraint. In addition, computing this power constraint requires knowledge of the channel gains between an intended CU transmitter and each PU receiver. It is because the amount of increased interference temperature at PU receivers is directly related to these channel gains. For clarity, channel gains between the PU system and the CU system are referred to as interference channel gains (or interference gains). With this CSI, the CU system can reliably predict the increased interference temperature at the PU receivers based on its transmission power. Moreover, the CU system can allocate more power for the transmission causing less interference, especially when the interference gain is significantly low.

A considerable amount of literature has been published on power control in CR networks. Perfect CSI including interference gains is assumed in most studies. Knowledge of interference gains can be obtained either from the primary system or from a

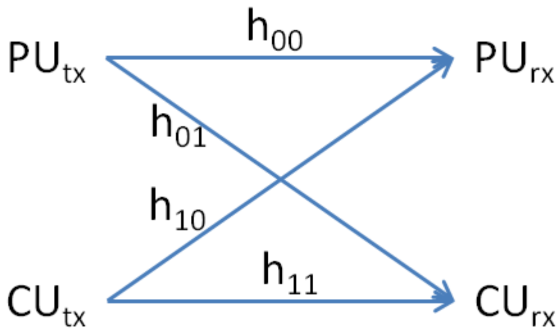


Figure 6.1: Interference channel.

band manager [106]. In [107], a method of evaluating interference gains is suggested. It involves close cooperation with the primary system in terms of the PU receiver feedbacking its measured SNR and SINR to the CR system. In this chapter, we consider a more realistic scenario in which perfect CSI is not assumed and interference gains are estimated with less primary-system-side aid, i.e., without CSI feedback from the PU system. First, with partial CSI, the issue of what interference constraints should be imposed on the CR system is addressed. Furthermore, a novel approach to obtaining this partial CSI is presented. Our proposed method will rely heavily on cooperation within the CR network and be based on knowledge such as geolocations of PUs and CUs, and the cross-correlation between two interference gains. Although only partial CSI is acquired, it serves as a key parameter for the CR network to satisfy peak or average interference power constraints.

6.2 Interference Constraints with Partial CSI

In Fig. 6.1, the interference channel has two pairs of the transmitter and the receiver, $\{\mathrm{PU_{tx}}, \mathrm{PU_{rx}}\}$ and $\{\mathrm{CU_{tx}}, \mathrm{CU_{rx}}\}$ from primary and CR systems respectively. Let h_{ij}

denote the channel gain between $\varsigma(i)_{\mathrm{tx}}$ and $\varsigma(i)_{\mathrm{rx}}$ where $\varsigma(0) = \mathrm{PU}$ and $\varsigma(1) = \mathrm{CU}$. These channel gains are assumed to be statistically independent. The transmission from $\mathrm{CU}_{\mathrm{tx}}$ depends on h_{10} and h_{11} , and has to satisfy its own transmission-power constraint and the interference-power constraint imposed by the primary system. Under the assumption of perfect information of h_{10} at $\mathrm{CU}_{\mathrm{tx}}$, a peak-interference-power constraint or an average-interference-power constraint is normally placed on $\mathrm{CU}_{\mathrm{tx}}$. This section will examine what interference-power constraints the PU system should impose, assuming that $\mathrm{CU}_{\mathrm{tx}}$ only knows the statistical distribution of h_{10} instead of the exact value h_{10}. This issue is discussed in two different scenarios, both channels $\{h_{00}, h_{10}\}$ being slow fading or fast fading. Some common assumptions in both scenarios are given as follows. Let the received SINR at $\mathrm{PU}_{\mathrm{rx}}$ denoted by $\mathrm{SINR}_{\mathrm{PU}} = P_0 |h_{00}|^2 / \left(P_1 |h_{10}|^2 + I_{ex} + N_0\right)$ where P_0 and P_1 are transmission powers of $\mathrm{PU}_{\mathrm{tx}}$ and $\mathrm{CU}_{\mathrm{tx}}$ respectively, I_{ex} denotes the stable interference from other sources, and N_0 is the power spectral density of AWGN at $\mathrm{PU}_{\mathrm{rx}}$. It is further assumed that the overall interference $\left(P_1 |h_{10}|^2 + I_{ex}\right)$ behaves like white Gaussian noise.

6.2.1 Slow Fading Channel

In the slow fading situation, the absolute values $\{|h_{00}|, |h_{10}|\}$ remain constant for the period of interest but follow some fading distribution with the second moment $\mu_{2,h_{ij}} = E\left[|h_{ij}|^2\right]$. The maximum reliable transmission rate $\log(1 + \mathrm{SINR}_{\mathrm{PU}})$ bits/s/Hz is a function of $\{|h_{00}|, |h_{10}|\}$. $\mathrm{PU}_{\mathrm{rx}}$ is said to be in outage whenever $\log(1 + \mathrm{SINR}_{\mathrm{PU}}) < R_{\mathrm{PU}}$ where R_{PU} is the data transmission rate from $\mathrm{PU}_{\mathrm{tx}}$ [72]. The outage probability is defined as $P_{out}(R_{\text{PU}}) \triangleq \mathbb{P}\{\log(1+\text{SINR}_{\text{PU}}) < R_{\text{PU}}\}$.

When PU_{rx} has already been in outage (i.e., $\log(1+\text{SINR}_{\text{PU}}) < R_{\text{PU}}$ with $P_1 = 0$), no interference power constraint is needed to be imposed on CU_{tx}. So CU_{tx} might transmit signals at any power levels without consideration of $|h_{10}|$. On the other hand, if PU_{rx} is not in outage (i.e., $\log(1+\text{SINR}_{\text{PU}}) > R_{\text{PU}}$ with $P_1 = 0$), the opportunity to access the licensed band is still available. This time a reasonable interference constraint should be placed such as

$$P_1 |h_{10}|^2 \leq \hbar, \tag{6.1}$$

where $\hbar \triangleq \left[P_0 |h_{00}|^2 / 2^{(R_{\text{PU}}-1)} - (I_{ex} + N_0)\right]$. Hence the inequality $\log(1+\text{SINR}_{\text{PU}}) \geq R_{\text{PU}}$ can be maintained. However, this can be achieved only when the knowledge of $|h_{10}|$ is available at CU_{tx}.

What if CU_{tx} knows the fading distribution of $|h_{10}|$ instead of its value? Obviously, whatever transmission power $P_1 \neq 0$ is used, the interference constraint (6.1) is no longer guaranteed. In other words, an outage can occur with probability $\mathbb{P}\left\{P_1 |h_{10}|^2 > \hbar\right\}$. Therefore, the overall outage probability is increased by

$$\left[1 - P_{out}(R_{\text{PU}}; P_1 = 0)\right] \mathbb{P}\left\{P_1 |h_{10}|^2 > \hbar\right\}. \tag{6.2}$$

From the perspective of the primary system, this increase in the outage probability should be upper bounded, which suggests another constraint, i.e.,

$$\mathbb{P}\left\{P_1 |h_{10}|^2 > \hbar\right\} \leq \epsilon, \text{ for some } 0 < \epsilon < 1. \tag{6.3}$$

For illustration, let $|h_{10}|$ be Rayleigh distributed $f_{ray}\left(|h_{10}|\,;\sigma_{h_{10}}=\sqrt{\mu_{2,h_{ij}}/2}\right)$. To satisfy (6.3), $\mathrm{CU}_{\mathrm{tx}}$ should follow the peak-power limit

$$P_1 \leq \hbar/\left[2\sigma_{h_{10}}^2 \ln\left(1/\epsilon\right)\right]. \tag{6.4}$$

Now the probabilistic constraint (6.3) has been translated into a practical power limit which can be placed on $\mathrm{CU}_{\mathrm{tx}}$. Compared to the power limit in (6.1), this peak-power limit in (6.4) has a high chance of being more stringent. However, the knowledge of the exact value $|h_{10}|$ is no longer required in the new peak-power limit.

6.2.2 Fast Fading Channel

In the case of fast fading channels, the channel gains $\{h_{00}, h_{10}\}$ are assumed to be stationary and ergodic. With CSI about h_{00} at $\mathrm{PU}_{\mathrm{rx}}$, an achievable ergodic capacity is given by $C_{eg}\left(h_{00}\right)=\mathbb{E}_{h_{00}}\left[\log\left(1+\mathrm{SINR}_{\mathrm{PU}}\right)\right]$ in which the received interference power has to be kept constant $P_1\left|h_{10}\right|^2$ [72]. This scheme is feasible when $\mathrm{CU}_{\mathrm{tx}}$ knows the interference channel gain h_{10} and dynamically adjusts its transmission power P_1. The value $P_1\left|h_{10}\right|^2$ to be maintained should be derived from the inequality $C_{eg}\left(h_{00}\right)\geq R_{\mathrm{PU}}$. Take the capacity at low SINR (i.e., $\left[P_0/\left(P_1\left|h_{10}\right|^2+I_{ex}+N_0\right)\right]\ll 1$) as an example. The ergodic capacity can be approximated by [72]

$$\begin{aligned} C_{eg}\left(h_{00}\right) &\approx \left(\log_2 e\right)\mathbb{E}_{h_{00}}\left[\frac{P_0\left|h_{00}\right|^2}{\left(P_1\left|h_{10}\right|^2+I_{ex}+N_0\right)}\right], \\ &= \frac{\left(\log_2 e\right)P_0\mu_{2,h_{00}}}{\left(P_1\left|h_{10}\right|^2+I_{ex}+N_0\right)}, \end{aligned} \tag{6.5}$$

which implies the interference constraint

Table 6.1: Interference constraints for different scenarios.

	Slow fading channel	Fast fading channel
CSI (channel realization h_{10}) at a CU transmitter	$P_1 \|h_{10}\|^2 \leq \hbar$	$P_1 \|h_{10}\|^2 = k$ such that $C_{eg}(h_{00}) \geq R_{\text{PU}}$
CSI (statistical characterization $f_{h_{10}}$) at a CU transmitter	$\mathbb{P}\left\{P_1 \|h_{10}\|^2 > \hbar\right\} \leq \epsilon$	$C_{eg}(h_{00}, h_{10}) \geq R_{\text{PU}}$

$$P_1 |h_{10}|^2 \leq \left[\frac{(\log_2 e) P_0 \mu_{2,h_{00}}}{R_{\text{PU}}} - (I_{ex} + N_0)\right]. \tag{6.6}$$

Similarly, with information about $\{h_{00}, h_{10}\}$ at PU_{rx}, the ergodic capacity $C_{eg}(h_{00}, h_{10}) = \mathbb{E}_{h_{00},h_{10}}[\log(1 + \text{SINR}_{\text{PU}})]$ is an average over two independent variables $\{h_{00}, h_{10}\}$. The inequality $C_{eg}(h_{00}, h_{10}) \geq R_{\text{PU}}$ implicitly determines the required interference constraint. At low SINR,

$$C_{eg}(h_{00}, h_{10}) \approx \frac{(\log_2 e) P_0 \mu_{2,h_{00}}}{(P_1 \mu_{2,h_{10}} + I_{ex} + N_0)}, \tag{6.7}$$

which suggests the constraint

$$P_1 \mu_{2,h_{10}} \leq \left[\frac{(\log_2 e) P_0 \mu_{2,h_{00}}}{R_{\text{PU}}} - (I_{ex} + N_0)\right]. \tag{6.8}$$

In this constraint, no prior knowledge of the channel realization h_{10} but its statistical characterization is required at CU_{tx}. Although the interference constraint can still be imposed on CU_{tx} without knowing exact channel realization, it is done at the cost of requiring more CSI at PU_{rx}. The summary of results is given in Table 6.1.

6.2.3 Remarks

The previous discussion indicates the possibility of placing interference constraint on the CR transmitter with information of channel statistical characterization rather than the channel realization. The next question is how to obtain the statistical information of the interference channel between the primary system and the CR system. A key statistical parameter which has been shown above is the second moment of the channel gain. This second moment is taken account of in an empirical large-scale path-loss model as follows. In a wireless channel model, the channel power gain $|h_{ij}|^2$ in decibels consists of a large-scale path loss and a small-scale fading

$$|h_{ij}|^2_{\mathrm{dB}} = -PL_{ij\mathrm{dB}} + 10\log_{10}\left(|g_{ij}|^2\right), \tag{6.9}$$

where the path loss $PL_{ij\mathrm{dB}} = -10\log_{10}\left(\mu_{2,h_{ij}}\right)$ and the normalized small-scale fading gain $g_{ij} = h_{ij}/\sqrt{\mu_{2,h_{ij}}}$. To evaluate the second moment $\mu_{2,h_{ij}}$ is equivalent to evaluate the large-scale path loss. A method for estimating this large-scale path loss within the CR network will be presented in the next section.

In the case of slow fading, the interference constraint is derived based on the outage probabiliy which is acceptable to the primary system. If a primary receiver undergoes an outage even without interfernce, then any CR transmission power can be allowed. However, if this outage is due to CR transmission, CR should have some mechanism to detect this event and stop its own transmission. Otherwise, this outage can last for an intolerable duration.

6.3 Cooperative Interference Channel Estimation

Underlay CR systems allow a CU to access a licensed band provided that the introduced interference at any PU receiver is below a certain threshold. It is possible for a CU to transmit data at high power when an interference channel, the link between a CU transmitter and a PU receiver, experiences deep fades. In [11], the capacity of such systems has been analyzed under the assumption that perfect CSI, including large-scale and small-scale propagation effects, is available a priori. However, in practice, CSI feedback delay and channel time variations introduce channel errors, thus only partial and outdated CSI at a CU could be available. It has been shown that the capacity loss increases notably with decreasing correlation between the outdated and instantaneous channels [108]. Hence, in [109], a power control scheme based on mean-value CSI instead of perfect CSI has been developed, putting large-scale-only CSI to practical use. This mean-value CSI is referred to here as the large-scale path loss.

For clarification, an example is given below. Let's denote the instantaneous interference channel power gain by g and the outdated one by $\widehat{g}$, where $\widehat{g}$ is exponentially distributed with mean $\overline{g}$. The probability density function of g given $\widehat{g}$ can be expressed as [110]

$$f_{g|\widehat{g}}\left(g|\widehat{g}\right) = \frac{\overline{g}}{\left(1-\rho^2\right)} \exp\left[\frac{-\overline{g}\left(g+\rho^2\widehat{g}\right)}{\left(1-\rho^2\right)}\right] I_0\left(\frac{2\overline{g}\rho}{1-\rho^2}\sqrt{\widehat{g}g}\right), \tag{6.10}$$

where $I_0\left(\cdot\right)$ denotes the zeroth-order modified Bessel function and ρ is a correlation coefficient. We are interested in two mean square errors (MSEs), $\mathbb{E}\left[\left(g-\widehat{g}\right)^2\right]$ and

$\mathbb{E}\left[(g-\overline{g})^2\right]$, where $\mathbb{E}$ denotes expectation. When the former is greater than the latter, it implies that using $\overline{g}$ rather than $\widehat{g}$ can lead to better performance. The difference between them, that is the former minus the latter, is given by $\left(1-2\rho^2\right)\overline{g}^2$, which is greater than zero as $\rho < {}^{1}/_{\sqrt{2}}$. In [109], the value ρ for real systems such WiMAX and 3GPP LTE can be less than ${}^{1}/_{\sqrt{2}}$, which justifies the usage of the mean-value CSI.

In this section, we propose a cooperative scheme among CUs for estimating the large-scale path loss in interference channels. The novel contributions of our scheme are two-fold. First, unlike conventional estimation schemes such as channel-reciprocity-based estimation and PU-aid channel estimation [77, 111], no assumption about the PU-system duplex mode is made, nor do we assume that the signal-to-interference-and-noise ratio measurements at the PU receiver are fed back to a CU transmitter. Secondly, by making use of the geolocation information of CUs and the PU receiver, path loss model parameters can be inferred by using maximum likelihood estimation instead of the least squares (LS) estimation proposed in [112, 113]. Moreover, cross-correlations among shadow fading factors are exploited to estimate the shadow fading factor at the PU receiver using the minimum MSE (MMSE) criterion. Analytical performance of our scheme is presented in terms of the MSE. Compared to LS based estimators, it will be shown that the proposed method offers a better estimate of the path-loss component of the channel between a PU and a CU. The robustness of our proposed method is verified via matching analytical and simulation results.

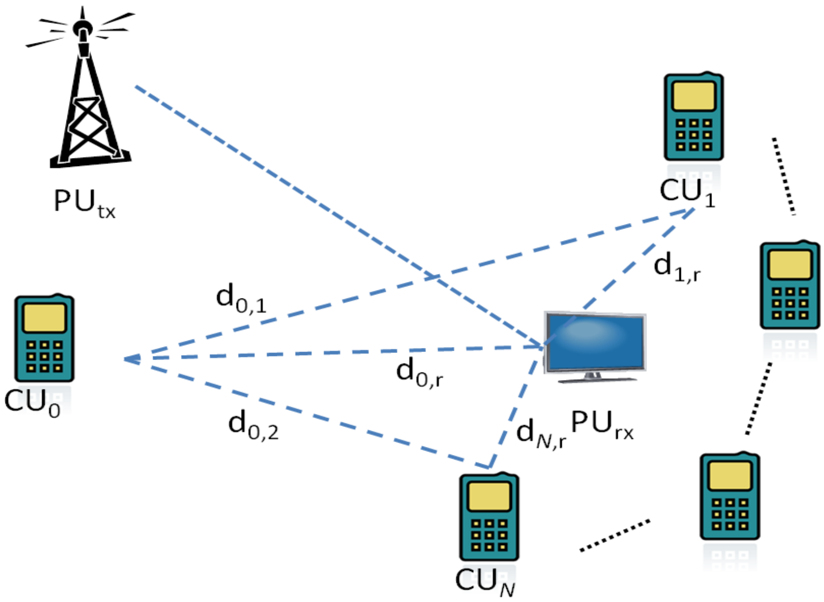


Figure 6.2: A primary receiver appears in a CR network.

6.3.1 System Model

Fig. 6.2 presents a centralized CR network in which a PU receiver $\mathrm{PU_{rx}}$ is situated at the center of a disc of radius r_{dis}, and N CUs ($\mathrm{CU}_1 \sim \mathrm{CU}_N$) are uniformly distributed over this disc. The aim is to estimate the large-scale path loss between the CU base station CU_0 and $\mathrm{PU_{rx}}$. We denote the relative distance between CU_i and CU_j by $d_{i,j}$, and the distance between CU_i and $\mathrm{PU_{rx}}$ by $d_{i,r}$. The log-normal path loss (in dB units) between CU_0 and another user, which could be either a CU or a PU, is given by [114]

$$PL\left(d_{0,j}\right)_{\mathrm{dB}} = PL\left(d_0\right) + 10\nu \log\left(d_{0,j}/d_0\right) + \psi_j, \tag{6.11}$$

for $j = 1 \sim N$ or r, where d_0 is the close-in reference distance, ν is the path loss exponent, and ψ_i is a shadow fading factor with a normal distribution $\mathcal{N}(0, \sigma_\psi^2)$. This propagation model has been adopted for power control in CR networks [115]. It has been shown that this model matches empirical data from urban and suburban environments [116]. Taking the random vector $[\psi_1, \psi_2, \ldots, \psi_N, \psi_r]$ to be jointly Gaussian distributed, the cross-correlation between the shadow fading factors follows the empirical formula $\mathbb{E}[\psi_i \psi_j] = \sigma_\psi^2 \gamma_{i,j}$ where $\gamma_{i,j} = \exp(-d_{i,j}/d_c)$ and d_c is the de-correlation distance which is determined empirically. This empirical formula was first proposed under the assumption that correlated shadow fading factors are measured along a straight line [116]. Here, we assume that $\mathrm{CU}_1 \sim \mathrm{CU}_N$ and the PU receiver $\mathrm{PU}_{\mathrm{rx}}$ do not need to be aligned along a straight line. However, since these CUs are geographically close to $\mathrm{PU}_{\mathrm{rx}}$, their corresponding coefficients ν, $\mathcal{N}(0, \sigma_\psi^2)$ and d_c tend to be correlated. The coefficients ν and σ_ψ^2 are unknown, while the de-correlation distance d_c is obtainable from the geolocation database.

6.3.2 Path Loss Estimation

Estimating the path loss $PL(d_{0,r})$ is equivalent to estimating the values of two coefficients ν and ψ_r. We propose that this is to be performed in two steps. First, obtain estimates $\hat{\nu}$ and $\hat{\sigma}_\psi^2$ with respect to ν and σ_ψ^2 based on the observations $\{PL(d_0), PL(d_{0,i}); i = 1, \ldots, N\}$. Then, an MMSE estimator of ψ_r can be formed by assuming $\psi_i = PL(d_{0,i}) - PL(d_0) - 10\hat{\nu}\log(d_{0,i}/d_0)$ and $\mathbb{E}[\psi_i \psi_r] = \hat{\sigma}_\psi^2 \gamma_{i,r}$ for $i = 1, \ldots, N$. Here are the detailed steps of our proposed method:

Step 1: estimating the path loss exponent and the shadow fading factor variance.

For simplicity, let's define $X_i = PL(d_{0,i}) - PL(d_0)$, which is Gaussian distributed

$\mathcal{N}(\nu\delta_i, \sigma_\psi^2)$ with $\delta_i = 10\log(d_{0,i}/d_0)$, $\mathbf{X} = [X_1, X_2, \ldots, X_N]^{'}$, where the superscript $'$ denotes the transpose, $\boldsymbol{\mu}_X = \nu\mathbf{d}$, and $\mathbf{d} = [\delta_1, \delta_2, \ldots, \delta_N]^{'}$. The relative distances between users within the CR system can be estimated by using the method proposed in [117]. A possible scheme to obtain information of path losses $PL(d_{0,i})$ for $i = 1 \sim N$ is described below. When measuring the channel gains between CUs in the licensed band, the CR system can transmit at a conservative power level such that the introduced interference seen by the PU receiver is acceptable, that is, the SINR at a PU receiver being above a certain threshold. At the same time, the SINR requirement should be met at each CU receiver. Due to the expected SINR at a CU receiver being higher than that at the PU receiver, a CU receiver should be more sensitive compared to a PU receiver. These SINR values can be evaluated using geolocation information, the known PU transmission power, and prior statistical knowledge of channels. Moreover, the effects of small-scale fadings at a CU can be averaged out if each CU is equipped with multiple antennas or has a certain level of mobility such that independent observations can be collected [112]. Therefore, the vector $\mathbf{X}$ is obtainable and its distribution is given by

$$P_{\mathbf{X}} = \frac{1}{\sqrt{(2\pi)^N |\Sigma_{\mathbf{X}}|}} \exp[\frac{-(\mathbf{X} - \boldsymbol{\mu}_X)' \Sigma_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_X)}{2}], \tag{6.12}$$

where $|\cdot|$ denotes the matrix determinant, $\Sigma_{\mathbf{X}} = \sigma_\psi^2\mathbf{H}$, and $\mathbf{H}$ is an N-by-N matrix with the (i,j)-th entry $\gamma_{i,j}$. Generally, the symmetric and positive-valued matrix is not positive-semidefinite, so it is not an eligible covariance matrix. However, it has be shown that the matrix $\Sigma_{\mathbf{X}}$ of this type is positive-definite by the following proposition.

Proposition 6. *Suppose that there are N points $\{x_i\}_{i=1}^N$ on a plane and any two points of them do not overlap. The relative distance between any two points, x_i and x_j, is denoted by $d_{i,j}$. The N-by-N matrix $\mathbf{A}_N = [a_{i,j}]_{N\times N}$ with the i-th row and j-th column entry $a_{i,j} = exp\left(-d_{i,j}\right)$ is positive-definite for $N \in \mathbb{N}$.*

By exploiting [81, Theorem 3.2.1], the maximum likelihood estimators of these two parameters are respectively given by

$$\hat{\nu} = \frac{\mathbf{d}'\mathbf{H}^{-1}\mathbf{X}}{\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}}, \tag{6.13}$$

and

$$\hat{\sigma}_\psi^2 = \frac{1}{N}\left(\mathbf{X} - \hat{\nu}\mathbf{d}\right)'\mathbf{H}^{-1}\left(\mathbf{X} - \hat{\nu}\mathbf{d}\right). \tag{6.14}$$

The estimator $\hat{\nu}$ of the path loss exponent is an unbiased estimate of uniformly minimum variance and its distribution is $\mathcal{N}(\nu, \sigma_\psi^2/(\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}))$. It can also be shown (see proof below) that the estimator $\hat{\sigma}_\psi^2$ is a weighted sum of chi-square distributed random variables, i.e.,

$$\hat{\sigma}_\psi^2 = \frac{1}{N}\sum_{m=1}^{M}\lambda_m\chi_1^2\left(\mu_{Y_m}^2\right). \tag{6.15}$$

Step 2: estimating the shadowing factor and the interference path loss. With the estimates $\hat{\nu}$ and $\hat{\sigma}_\psi^2$, we can compute the shadow fading factor $\hat{\psi}_i = X_i - \delta_i\hat{\nu}$ and assume that the cross-correlation $\mathbb{E}\left[\hat{\psi}_i\hat{\psi}_r\right] = \hat{\sigma}_\psi^2\gamma_{i,r}$ holds for $1 \le i \le N$. To evaluate $\gamma_{i,r}$ requires the information of the relative distances between a CU and a PU receiver. This information is obtainable to the CR system from the geolocation database [118]. Making use of [119, Eq.(4.52)] yields the MMSE estimator of ψ_r, that is

$$\hat{\psi}_r = \boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\hat{\boldsymbol{\psi}}_{CU}, \tag{6.16}$$

where $\boldsymbol{\gamma}_r = [\gamma_{1,r}, \gamma_{2,r}, \ldots, \gamma_{N,r}]^{'}$ and $\hat{\boldsymbol{\psi}}_{CU} = \left[\hat{\psi}_1, \hat{\psi}_2, \ldots, \hat{\psi}_N\right]^{'} = \mathbf{X} - \hat{\nu}\mathbf{d}$. Finally, the desired estimate of the path loss $PL(d_{0,r})$ follows as

$$\hat{PL}(d_{0,r}) = PL(d_0) + 10\hat{\nu}\log(d_{0,r}/d_0) + \hat{\psi}_r. \tag{6.17}$$

The MSE of the estimation $\hat{PL}(d_{0,r})$ can be expressed as

$$\begin{aligned} MSE &= \mathbb{E}\left\{\left[\hat{PL}(d_{0,r}) - PL(d_{0,r})\right]^2\right\}, \\ &= \sigma_\psi^2\left[\frac{\left(\delta_r - \mathbf{d}^{'}\mathbf{H}^{-1}\boldsymbol{\gamma}_r\right)^2}{\mathbf{d}^{'}\mathbf{H}^{-1}\mathbf{d}} + 1 - \boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\boldsymbol{\gamma}_r\right], \end{aligned} \tag{6.18}$$

where $\delta_r = 10\log(d_{0,r}/d_0)$. The derivation of (6.18) is provided in Section 6.C.

6.3.3 Robustness in an Asymptotic Sense

In practice, the parameter of the de-correlation distance in use might not be accurate. However, we are going to show that even using inaccurate de-correlation distance $\tilde{d}_c$ will asymptotically lead to the same performance of using the accurate one. That is, the MSE of the proposed path loss estimator $MSE\left(\tilde{d}_c\right)$ due to $\tilde{d}_c$ will converge to MSE in (6.18) as the number of cooperative users N increases. The expression of $MSE\left(\tilde{d}_c\right)$ is given by

$$MSE\left(\tilde{d}_c\right) = \sigma_\psi^2 \left[\frac{\mathbf{d}'\tilde{\mathbf{H}}^{-1}\mathbf{H}\tilde{\mathbf{H}}^{-1}\mathbf{d}\left(\delta_r - \mathbf{d}'\tilde{\mathbf{H}}^{-1}\tilde{\boldsymbol{\gamma}}_r\right)^2}{\left(\mathbf{d}'\tilde{\mathbf{H}}^{-1}\mathbf{d}\right)^2} + \tilde{\boldsymbol{\gamma}}_r'\tilde{\mathbf{H}}^{-1}\left(\mathbf{H}\tilde{\mathbf{H}}^{-1}\tilde{\boldsymbol{\gamma}}_r - 2\boldsymbol{\gamma}_r\right) \right.$$

$$\left. + \frac{2\delta_r\mathbf{d}'\tilde{\mathbf{H}}^{-1}\left(\mathbf{H}\tilde{\mathbf{H}}^{-1}\tilde{\boldsymbol{\gamma}}_r - \boldsymbol{\gamma}_r\right)}{\mathbf{d}'\tilde{\mathbf{H}}^{-1}\mathbf{d}} + 1 + \frac{2\tilde{\boldsymbol{\gamma}}_r'\tilde{\mathbf{H}}^{-1}\mathbf{d}\mathbf{d}'\tilde{\mathbf{H}}^{-1}\left(\boldsymbol{\gamma}_r - \mathbf{H}\tilde{\mathbf{H}}^{-1}\tilde{\boldsymbol{\gamma}}_r\right)}{\mathbf{d}'\tilde{\mathbf{H}}^{-1}\mathbf{d}} \right], \tag{6.19}$$

where $\tilde{\mathbf{H}}$ and $\tilde{\boldsymbol{\gamma}}_r$ are inaccurate versions of $\mathbf{H}$ and $\boldsymbol{\gamma}_r$ in which $\gamma_{a,b}$ is replaced by $\tilde{\gamma}_{a,b} = \exp\left(-d_{a,b}/\tilde{d}_c\right)$. The following proposition will show that our assertion is true for a one-dimensional scenario.

Proposition 7. *Assume that N cooperative CUs are aligned along a line and the relative distance between CU_i and CU_j is defined as $d_{i,j} = \frac{2r_{dis}}{N-1}|i-j|$ for $1 \leq i, j \leq N$. PU_{rx} is situated at the center of these CUs. Then, $\lim_{N\to\infty} MSE\left(\tilde{d}_c\right) = MSE$.*

Proof. The matrix $\mathbf{H}$ corresponding to this scenario is a symmetric Toeplitz matrix with its tri-diagonal inverse [120]

$$\mathbf{H}^{-1} = \frac{1}{1-\gamma_0^2}\begin{bmatrix} 1 & -\gamma_0 & 0 & \cdots & 0 \\ -\gamma_0 & 1+\gamma_0^2 & -\gamma_0 & \cdots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \cdots & -\gamma_0 & 1+\gamma_0^2 & -\gamma_0 \\ 0 & \cdots & 0 & -\gamma_0 & 1 \end{bmatrix}, \tag{6.20}$$

where $\gamma_0 = \exp\left[\frac{-2r_{\text{dis}}}{(N-1)d_c}\right]$. Similarly, $\tilde{\mathbf{H}}^{-1}$ is $\mathbf{H}^{-1}$ in which γ_0 is replaced by $\tilde{\gamma}_0 = \exp\left[\frac{-2r_{\text{dis}}}{(N-1)\tilde{d}_c}\right]$. It can be easily shown that $\lim_{N\to\infty}\tilde{\mathbf{H}}^{-1} = \frac{\tilde{d}_c}{d_c}\mathbf{H}^{-1}$ because of $\lim_{N\to\infty}\frac{1-\gamma_0^2}{1-\tilde{\gamma}_0^2} = \frac{\tilde{d}_c}{d_c}$, $\lim_{N\to\infty}\frac{1+\tilde{\gamma}_0^2}{1+\gamma_0^2} = 1$, and $\lim_{N\to\infty}\frac{\tilde{\gamma}_0}{\gamma_0} = 1$. Likewise, $\lim_{N\to\infty}\tilde{\boldsymbol{\gamma}}_r = \frac{d_c}{\tilde{d}_c}\boldsymbol{\gamma}_r$. Replacing $\tilde{\mathbf{H}}^{-1}$and $\tilde{\boldsymbol{\gamma}}_r$ in (6.19) with their limits results in $MSE\left(\tilde{d}_c\right) = MSE$. □

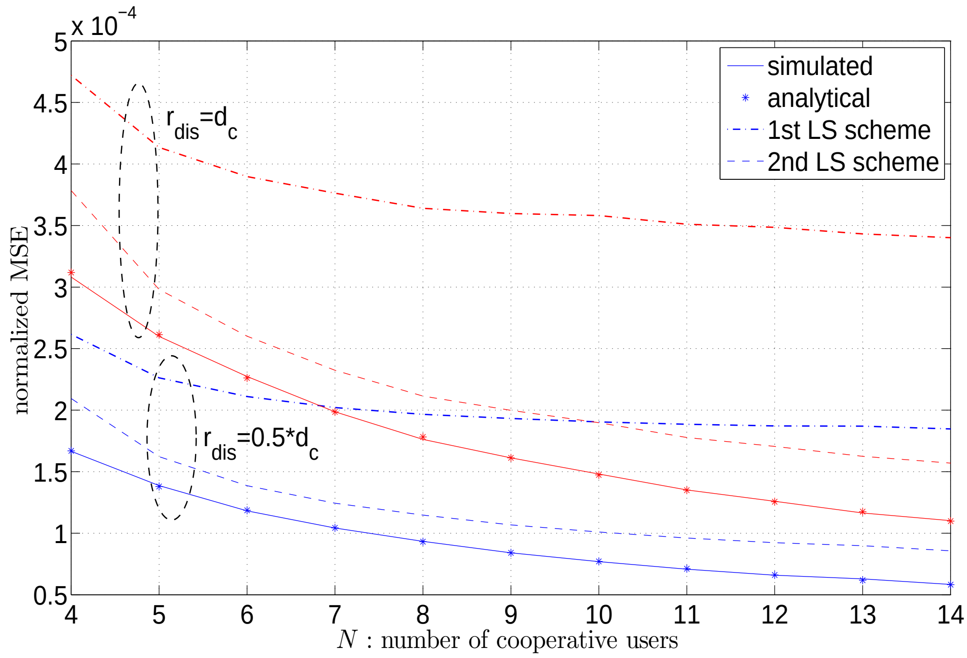


Figure 6.3: MSE of path loss estimation using different schemes versus the number of cooperative CUs over different sizes of discs.

This proposition also points that the same asymptotic property, $\lim_{N\to\infty} MSE\left(\tilde{d}_c\right) = MSE$, is held in the two-dimensional scenario. As N is sufficiently large, it is very likely to have a subset of CUs approximately arranged as the mentioned one-dimensional scenario. Meanwhile, the path loss estimation is mainly determined by reports from this subset of CUs. In other words, the MSE performance depends strongly on observations from this subset of CUs. Thus, we can argue that our assertion holds in a general two-dimension scenario.

6.4 Simulation Results

To verify the proposed path loss estimator, we consider an outdoor path loss model established using measurements in urban environments [121]. The relevant param-

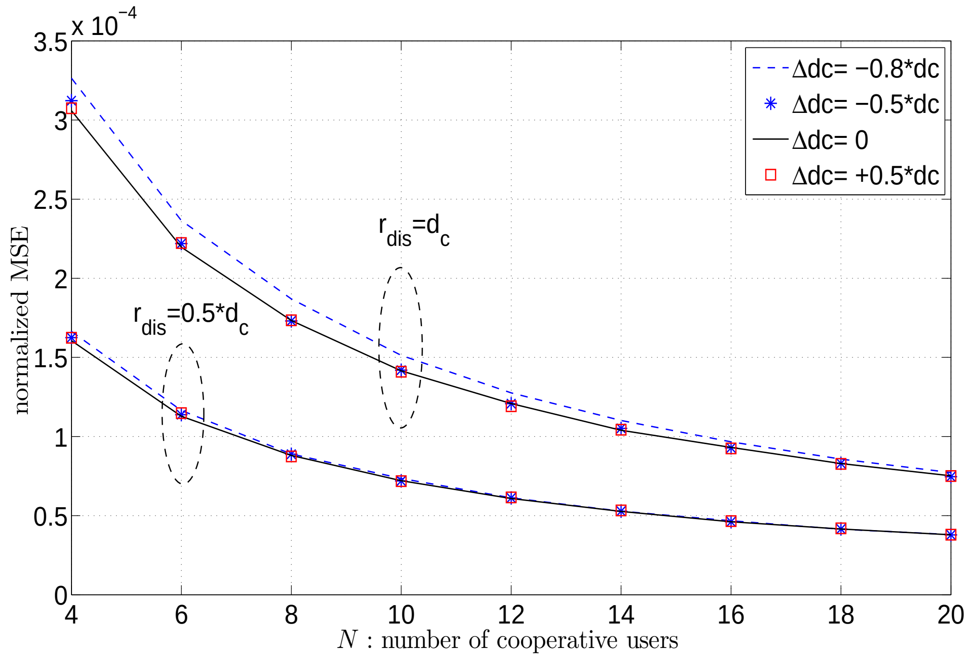


Figure 6.4: MSE of path loss estimation using inaccurate de-correlation distances versus the number of cooperative CUs over different sizes of discs.

eters are given by $\nu = 3$, $\sigma_\psi\,(\text{dB}) = 7.9$, $d_c = 55\text{m}$, and $PL\,(d_0)\,(\text{dB}) = 31.7$ at $d_0 = 1\text{m}$. The distance between CU_0 and PU_{rx} is 500 meters. To evaluate the effect of MSE, we will use the normalized MSE as the performance measure, i.e.,

$$\overline{MSE} = \frac{MSE}{\mathbb{E}^2\left[PL\left(d_{0,r}\right)\right]}. \tag{6.21}$$

Fig. 6.3 shows that the mean square error of the proposed path loss estimator decreases significantly when the density of the cooperative CUs within a given disc increases. The simulated results perfectly match the averaged analytical values given by (6.18) over 100,000 different CUs' location realizations. Compared to the MSE curve due to the first LS based scheme [112], in which only the mean-value part $PL\,(d_0) + 10\nu \log\left(d_{0,r}/d_0\right)$ is evaluated, our proposed estimator can lead to lower

MSE. Moreover, our proposed method is also superior to the second LS based scheme [113], in which ψ_r is inferred by linear estimation. The reason for this is that more information, such as the de-correlation distance, has been exploited in our scheme.

In Fig. 6.4, the impact of using inaccurate de-correlation distances $\tilde{d}_c$ has been depicted in terms of increases in MSEs when compared to the performance curve of using accurate information d_c. For the curve with $r_{dis} = d_c$ and $\Delta d_c = -0.8d_c$, where $\Delta d_c \triangleq \tilde{d}_c - d_c$, the increase in MSEs caused by using $\tilde{d}_c$ diminishes with the increased N. This illustrates the robustness mentioned in Section 6.3.3. It is also apparent from this figure that the absolute deviation $|\Delta d_c| = 0.5d_c$ does not cause significant increases in MSEs for $r_{\text{dis}} = d_c \, \text{or} \, 0.5d_c$. Numerically a higher level of robustness is guaranteed in the sense that the MSE increase is negligible as $|\Delta d_c|$ below some threshold.

6.5 Conclusions

The proper interference constraint, corresponding to what CSI is at the CU transmitter and at the PU receiver, has been explored. Conventional peak and average transmission-power constraints are imposed, depending on whether the quality of service of the primary system is delay sensitive or not. Nevertheless, our results indicate that what constraint to place in different scenarios is according to outage probability or ergodic capacity of the primary system.

The second part of this chapter proposed and analyzed a new method for estimating the large scale path loss of a CR interference channel without primary-user feedback. This method is performed mainly through cooperation among CUs. Both analytical and simulation results are in agreement and indicate that the proposed method can provide excellent performance relative to existing techniques. This part of work has been published in IEEE SPL [122].

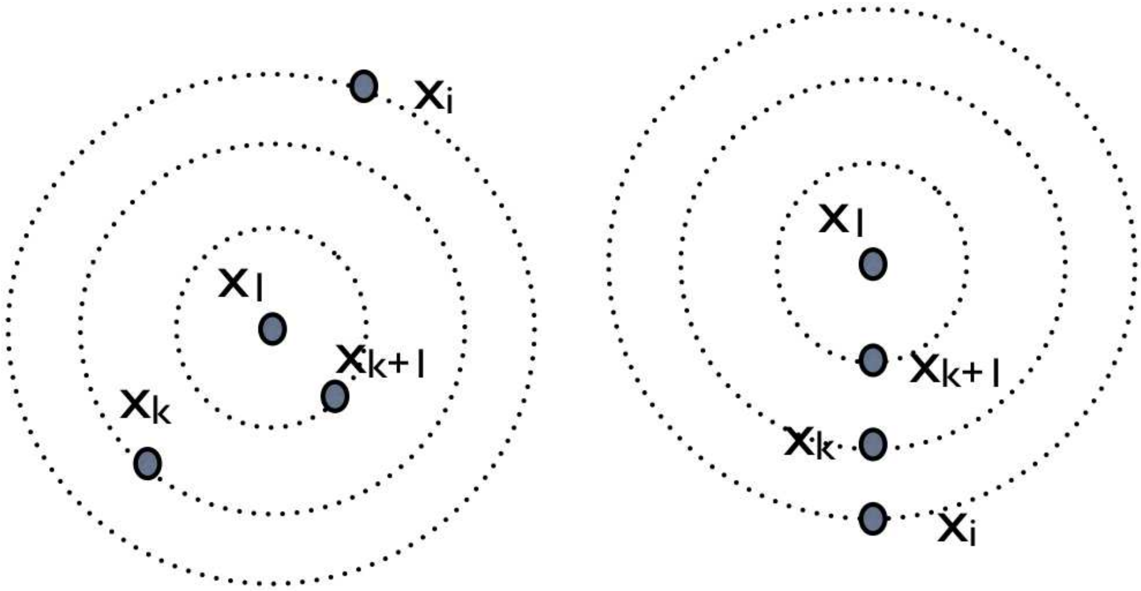
6.A Proof of Proposition 6

When $N = 1$ or 2, the statement can be easily shown to be true. Assume that the statement is true for $N = k$, we must show that the matrix $\mathbf{A}_{k+1}$ is positive-definite. Let's express $\mathbf{A}_{k+1}$ as

$$\mathbf{A}_{k+1} = \begin{bmatrix} \mathbf{A}_k & \mathbf{y} \\ \mathbf{y}' & 1 \end{bmatrix}, \tag{6.22}$$

where $\mathbf{y}' = [a_{1,k+1}, a_{2,k+1}, \cdots, a_{k,k+1}]$. By Sylvester's criterion, if the determinant of $\mathbf{A}_{k+1}$, $|\mathbf{A}_{k+1}|$, is greater than zero, then $\mathbf{A}_{k+1}$ must be positive-definite. With the expansion $|\mathbf{A}_{k+1}| = \left(1 - \mathbf{y}'\mathbf{A}_k^{-1}\mathbf{y}\right)|\mathbf{A}_k|$, $|\mathbf{A}_{k+1}|$ is greater than zero as $\mathbf{y}'\mathbf{A}_k^{-1}\mathbf{y} < 1$. The rest of work is to show $\mathbf{y}'\mathbf{A}_k^{-1}\mathbf{y} < 1$.

Let $a_{1,k+1}$ and $a_{1,k}$ respectively be the first and the second largest values in the set $\{a_{1,i}\}_{i=2}^{k+1}$. The matrix $\mathbf{A}_{k+1}$ can be turned into this form by performing row exchanges and matrix transposes without changing the value $|\mathbf{A}_{k+1}|$. As shown in Fig. 6.5(a), it means that the points x_{k+1} and x_k are respectively the first and the second closest points to x_1, that is, $0 < d_{1,k+1} < d_{1,k} < d_{1,i}$ for $2 \leq i \leq k-2$. The case of two or more points at the same distance from x_1 is excluded here. However, the proof presented below can be easily extended to include this case. By making use of the triangle inequality, we can get the possible ranges of the elements in the vector $\mathbf{y}$, that is, $a_{1,i}a_{1,k+1} < a_{i,k+1} < \frac{a_{1,i}}{a_{1,k+1}}$ for $2 \leq i \leq k$. These ranges are further extended to be $a_{1,i}a_{1,k+1} \leq a_{i,k+1} \leq \frac{a_{1,i}}{a_{1,k+1}}$ for $2 \leq i \leq k$ because of the possibility of the three points, x_1, x_{k+1}, and x_i, lying on the same line. All these ranges together define a



(a) scattered points

(b) aligned points

Figure 6.5: Geometric illustration

closed compact region $\mathcal{D}(a_{1,k+1})$ which depends on $a_{1,k+1}$.

The quadratic-form function $f(\mathbf{y}) = \mathbf{y}'\mathbf{A}_k^{-1}\mathbf{y}$ is concave up because of $\mathbf{A}_k^{-1}$ being positive-definite. Therefore, the maximum value of $f(\mathbf{y})$, $\forall \mathbf{y} \in \mathcal{D}(a_{1,k+1})$, will appear at the upper boundary of $\mathcal{D}(a_{1,k+1})$, that is, $a_{i,k+1} = \frac{a_{1,i}}{a_{1,k+1}}$ for $2 \le i \le k$. When $\mathbf{y}$ approaches this upper boundary, it forces all the points $\{x_i\}_{i=1}^{k+1}$ to lie on the same line and the point x_{k+1} to lie between x_i and x_1 for $2 \le i \le k$. This is demonstrated in Fig. 6.5(b). Let's expand the function $f(\mathbf{y})$ on this upper boundary as

$$f(a_{1,k+1}) = \mathbf{y}'\frac{\mathbf{C}_k}{|\mathbf{A}_{k+1}|}\mathbf{y} = \frac{1}{|\mathbf{A}_{k+1}|}\left\{C_{1,1}a_{1,k+1}^2 \right.$$
$$\left. +2\sum_{i=2}^{k} C_{1,i}a_{1,i} + \left(\textstyle\sum_{i,j=2}^{k} C_{i,j}a_{1,i}a_{1,j}\right)/a_{1,k+1}^2 \right\},$$

where $\mathbf{C}_k = [C_{i,j}]$ denotes the k-by-k cofactor matrix with the cofactor $C_{i,j}$ of the (i,j) entry of $\mathbf{A}_k$ and the possible range of $a_{1,k+1}$ is given by $a_{1,k} < a_{1,k+1} < 1$.

The second derivative $f''(a_{1,k+1}) = 2C_{1,1} + 6\left(\sum_{i,j=2}^{k} C_{i,j}a_{1,i}a_{1,j}\right)a_{1,k+1}^{-4}$ is strictly positive for $a_{1,k} < a_{1,k+1} < 1$. Thus, the function $f(a_{1,k+1})$ is concave up and has the maximum at either $a_{1,k+1} = a_{1,k}$ or $a_{1,k+1} = 1$. When $a_{1,k+1}$ approaches $a_{1,k}$, it means that the point x_{k+1} is getting close to x_k. Thus, $a_{i,k+1}$ is forced to become $a_{i,k}$ for $2 \leq i \leq k$. The value $f(a_{1,k+1} = a_{1,k})$ can be shown to be one by using the fact $\mathbf{C}_k\mathbf{A}_k = |\mathbf{A}_{k+1}|\,\mathbf{I}_k$ where $\mathbf{I}_k$ denotes the identity matrix of size k. With similar arguments, $f(a_{1,k+1} = 1)$ can also be shown to be one. Hence, $\mathbf{y}'\mathbf{A}_k^{-1}\mathbf{y}$ is strictly less than one and the statement in the proposition holds.

6.B Derivation of Equation (6.15)

The derivation of this proof follows the idea in [123]. First, rewrite the estimator $\hat{\sigma}_\psi^2$ as

$$\hat{\sigma}_\psi^2 = \frac{1}{N}\mathbf{X}'\mathbf{D}'\mathbf{H}^{-1}\mathbf{D}\mathbf{X}, \text{ where } \mathbf{D} \triangleq \left(\mathbf{I} - \frac{\mathbf{d}\mathbf{d}'\mathbf{H}^{-1}}{\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}}\right). \tag{6.23}$$

Let $\mathbf{Q}$ be an orthonormal matrix such that $\mathbf{\Lambda} \triangleq \mathbf{Q}\Sigma_{\mathbf{X}}^{1/2}\mathbf{A}\Sigma_{\mathbf{X}}^{1/2}\mathbf{Q}' = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$ where $\mathbf{A} \triangleq \mathbf{D}'\mathbf{H}^{-1}\mathbf{D}$. If the rank of the matrix $\mathbf{A}$ is M, then we can assume that $\lambda_m = 0$ for $m > M$. Let's define $\mathbf{Y} = \mathbf{Q}\Sigma_{\mathbf{X}}^{-1/2}\mathbf{X}$, then its distribution is $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{Y}}, \mathbf{I})$ where $\boldsymbol{\mu}_{\mathbf{Y}} = [\mu_{\mathbf{Y}_i}]_{N\times 1} = \nu\mathbf{Q}\Sigma_{\mathbf{X}}^{-1/2}\mathbf{d}$. As a result, we can recast the estimator $\hat{\sigma}_\psi^2$ as

$$\begin{aligned}\hat{\sigma}_{\psi}^{2} &= \frac{1}{N}\mathbf{X}^{'}\mathbf{A}\mathbf{X},\\ &= \frac{1}{N}\mathbf{X}^{'}\Sigma_{\mathbf{X}}^{-1/2}\Sigma_{\mathbf{X}}^{1/2}\mathbf{A}\Sigma_{\mathbf{X}}^{1/2}\Sigma_{\mathbf{X}}^{-1/2}\mathbf{X},\\ &= \frac{1}{N}\mathbf{X}^{'}\Sigma_{\mathbf{X}}^{-1/2}\mathbf{Q}^{'}\mathbf{\Lambda}\mathbf{Q}\Sigma_{\mathbf{X}}^{-1/2}\mathbf{X},\\ &= \frac{1}{N}\mathbf{Y}^{'}\mathbf{\Lambda}\mathbf{Y} = \frac{1}{N}\sum_{m=1}^{M}\lambda_m\chi_1^2\left(\mu_{\mathbf{Y}_i}^2\right). \end{aligned} \tag{6.24}$$

The mean and variance of the estimator are respectively given by

$$\mu_{\hat{\sigma}_{\psi}^{2}} = \frac{1}{N}\left(\sum_{m=1}^{M}\lambda_m + \sum_{m=1}^{M}\lambda_m\mu_{\mathbf{Y}_m}^2\right), \tag{6.25}$$

and

$$\sigma_{\hat{\sigma}_{\psi}^{2}}^{2} = \frac{2}{N^2}\left(\sum_{m=1}^{M}\lambda_m^2 + 2\sum_{m=1}^{M}\lambda_m^2\mu_{\mathbf{Y}_m}^2\right). \tag{6.26}$$

6.C Derivation of Equation (6.18)

The MSE of the estimated path loss $\hat{PL}\left(d_{0,r}\right)$ is given by

$$\begin{aligned} MSE &= \mathbb{E}\left[\left(\left(\delta_r\hat{\nu} + \hat{\psi}_r\right) - \left(\delta_r\nu + \psi_r\right)\right)^2\right],\\ &= \mathbb{E}\left[\delta_r^2\left(\hat{\nu}-\nu\right)^2 + 2\delta_r\left(\hat{\nu}-\nu\right)\left(\hat{\psi}_r - \psi_r\right) + \left(\hat{\psi}_r - \psi_r\right)^2\right]. \end{aligned} \tag{6.27}$$

The expectation in (6.27) can be evaluated in three parts. The first two are given by

$$\mathbb{E}\left[\delta_r^2\left(\hat{\nu}-\nu\right)^2\right] = \delta_r^2\frac{\sigma_{\psi}^2}{\mathbf{d}^{'}\mathbf{H}^{-1}\mathbf{d}}, \tag{6.28}$$

and

$$\begin{aligned}\mathbb{E}\left[2\delta_r\left(\hat{\nu}-\nu\right)\left(\hat{\psi}_r-\psi_r\right)\right] &= 2\delta_r\mathbb{E}\left[\hat{\nu}\hat{\psi}_r-\hat{\nu}\psi_r-\nu\hat{\psi}_r+\nu\psi_r\right], \\ &= -2\delta_r\sigma_\psi^2\frac{\mathbf{d}'\mathbf{H}^{-1}\boldsymbol{\gamma}_r}{\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}}, \end{aligned} \tag{6.29}$$

where

$$\begin{aligned}\mathbb{E}\left[\hat{\nu}\hat{\psi}_r\right] &= \mathbb{E}\left[\hat{\nu}\boldsymbol{\gamma}_r'\mathbf{H}^{-1}\hat{\boldsymbol{\psi}}_{CU}\right], \\ &= \mathbb{E}\left[\hat{\nu}\boldsymbol{\gamma}_r'\mathbf{H}^{-1}\left(\mathbf{X}-\hat{\nu}\mathbf{d}\right)\right], \\ &= \boldsymbol{\gamma}_r'\mathbf{H}^{-1}\mathbb{E}\left[\hat{\nu}\mathbf{X}-\hat{\nu}^2\mathbf{d}\right], \\ &= \boldsymbol{\gamma}_r'\mathbf{H}^{-1}\left\{\mathbb{E}\left[\hat{\nu}\mathbf{X}\right]-\mathbb{E}\left[\hat{\nu}^2\mathbf{d}\right]\right\}, \\ &= \boldsymbol{\gamma}_r'\mathbf{H}^{-1}\left\{\frac{\mathbb{E}\left[\mathbf{X}\mathbf{X}'\right]\mathbf{H}^{-1}\mathbf{d}}{\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}}-\mathbb{E}\left[\hat{\nu}^2\right]\mathbf{d}\right\}, \\ &= \boldsymbol{\gamma}_r'\mathbf{H}^{-1}\left\{\frac{\left(\sigma_\psi^2\mathbf{H}+\nu^2\mathbf{d}\mathbf{d}'\right)\mathbf{H}^{-1}\mathbf{d}}{\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}}-\left(\nu^2+\frac{\sigma_\psi^2}{\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}}\right)\mathbf{d}\right\}=0;\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left[\hat{\nu}\psi_r\right] &= \mathbb{E}\left[\frac{\mathbf{d}'\mathbf{H}^{-1}\mathbf{X}}{\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}}\psi_r\right], \\ &= \mathbb{E}\left[\frac{\mathbf{d}'\mathbf{H}^{-1}\left(\nu\mathbf{d}+\boldsymbol{\psi}_{CU}\right)}{\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}}\psi_r\right], \\ &= \mathbb{E}\left[\nu\psi_r\right]+\frac{\mathbf{d}'\mathbf{H}^{-1}\mathbb{E}\left[\boldsymbol{\psi}_{CU}\psi_r\right]}{\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}},\ \boldsymbol{\psi}_{CU}=\left[\psi_1,\psi_2,\ldots,\psi_N\right]', \\ &= \frac{\sigma_\psi^2\mathbf{d}'\mathbf{H}^{-1}\boldsymbol{\gamma}_r}{\mathbf{d}'\mathbf{H}^{-1}\mathbf{d}},\ \because \mathbb{E}\left[\boldsymbol{\psi}_{CU}\psi_r\right]=\sigma_\psi^2\boldsymbol{\gamma}_r;\end{aligned}$$

$$\mathbb{E}\left[\nu\hat{\psi}_r\right]=\mathbb{E}\left[\nu\boldsymbol{\gamma}_r'\mathbf{H}^{-1}\left(\mathbf{X}-\hat{\nu}\mathbf{d}\right)\right]=0;$$

$$\mathbb{E}\left[\nu\psi_r\right]=\nu\mathbb{E}\left[\psi_r\right]=0.$$

For the third,

$$\begin{aligned}\mathbb{E}\left[\left(\hat{\psi}_r - \psi_r\right)^2\right] &= \mathbb{E}\left[\hat{\psi}_r^2 - 2\hat{\psi}_r\psi_r + \psi_r^2\right], \\ &= \sigma_\psi^2\left[1 - \boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\boldsymbol{\gamma}_r + \frac{\boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\mathbf{d}\mathbf{d}^{'}\mathbf{H}^{-1}\boldsymbol{\gamma}_r}{\mathbf{d}^{'}\mathbf{H}^{-1}\mathbf{d}}\right], \end{aligned} \tag{6.30}$$

where

$$\begin{aligned}\mathbb{E}\left[\hat{\psi}_r^2\right] &= \boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\mathbb{E}\left[\left(\mathbf{X} - \hat{\nu}\mathbf{d}\right)\left(\mathbf{X} - \hat{\nu}\mathbf{d}\right)^{'}\right]\mathbf{H}^{-1}\boldsymbol{\gamma}_r, \\ &= \boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\left[\sigma_\psi^2\mathbf{H} - \sigma_\psi^2\frac{\mathbf{d}\mathbf{d}^{'}}{\mathbf{d}^{'}\mathbf{H}^{-1}\mathbf{d}}\right]\mathbf{H}^{-1}\boldsymbol{\gamma}_r, \\ &= \sigma_\psi^2\left[\boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\boldsymbol{\gamma}_r - \frac{\boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\mathbf{d}\mathbf{d}^{'}\mathbf{H}^{-1}\boldsymbol{\gamma}_r}{\mathbf{d}^{'}\mathbf{H}^{-1}\mathbf{d}}\right];\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left[\hat{\psi}_r\psi_r\right] &= \mathbb{E}\left[\boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\left(\mathbf{X} - \hat{\nu}\mathbf{d}\right)\psi_r\right], \\ &= \boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\mathbb{E}\left[\mathbf{X}\psi_r\right] - \boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\mathbf{d}\mathbb{E}\left[\hat{\nu}\psi_r\right], \\ &= \boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\mathbb{E}\left[\boldsymbol{\psi}_{CU}\psi_r\right] - \boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\mathbf{d}\left(\frac{\sigma_\psi^2\mathbf{d}^{'}\mathbf{H}^{-1}\boldsymbol{\gamma}_r}{\mathbf{d}^{'}\mathbf{H}^{-1}\mathbf{d}}\right), \\ &= \sigma_\psi^2\left(\boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\boldsymbol{\gamma}_r - \frac{\boldsymbol{\gamma}_r^{'}\mathbf{H}^{-1}\mathbf{d}\mathbf{d}^{'}\mathbf{H}^{-1}\boldsymbol{\gamma}_r}{\mathbf{d}^{'}\mathbf{H}^{-1}\mathbf{d}}\right);\end{aligned}$$

$$\mathbb{E}\left[\psi_r^2\right] = \sigma_\psi^2.$$

As a result of (6.27), (6.28), (6.29), and (6.30), Equation (6.18) follows.

CHAPTER 7

CONCLUSIONS

7.1 Final Remarks

The research presented in this thesis has investigated cyclostationary spectrum sensing, multi-antenna spectrum sensing, detection performance over fading channels, and interference channel estimation. This work first contributes to existing multi-cycle-frequency detection by ensuring better utilization of PU cyclostationary features and in turn achieving relatively higher detection probability. As the counterpart of an existing single-antenna scheme, multi-antenna cyclostationary spectrum sensing has been proposed and rigorously examined. The required mixing conditions which establish joint asymptotic normality of the test statistics, have been analytically verified. Moreover, the performance bounds and approximations of cyclostationary spectrum sensing over fading channels are provided to make CR system performance more predictable. Finally, a novel cooperative estimate of the path loss between the primary and CR systems is proposed. This vital information about the path loss can be used to facilitate efficient CR communications.

A number of caveats need to be noted regarding the present study. First, the current investigation was limited by the assumption of centralized networks. Due to this, the proposed detection and estimation methods are not immediately applicable in distributed CR networks. To make our proposed methods distributed might require application of principles in Game Theory. Another limitation of this study is that the proposed spectrum sensing methods were restricted to single frequency band detection. To sense over multiple frequency bands will inevitably increase the computational complexity. It is not surprising that there will be some trade-off between computational complexity and efficient utilization involved in the multiband spectrum sensing. Thirdly, the study did not evaluate the use of indoor path loss models in interference channel estimation. This makes our proposed path loss estimation only applicable for outdoor environments.

7.2 Future Works

In this thesis, several detection and estimation techniques have been explored for cognitive radio applications. The broadness and depth of this work can be made greater if the following tasks are carried out.

- Chapter 3 presents the optimal and sub-optimal schemes to identify test points in the cycle-frequency-lag domain. Though our sub-optimal scheme requires less prior knowledge of PU signals, it still relies on the exhaustive search. A computationally efficient scheme might exist in the field of Mathematical Optimization. It means that the problem of identifying test points might be formulated in some canonical form for which an efficient solver exists. Another possible extension could be to provide analytical expression of the fourth-order cumulant of OFDM signals which are commonly seen in the primary system. The derivation of this analytical expression could be based on our work of deriving the 4th-order cyclic cumulant of linear modulated signals.

- Multi-antenna spectrum sensing discussed in Chapter 4 mainly focuses on the synchronous scenario. As asynchronous spectrum sensing can occur in the distributed CR networks, it is well worth the effort to address this scenario, in which it may arise the issue such as using multiple antennas to minimize CR interference while maintaining the received PU signal level. It is also intriguing to come up with some eigenvalue-based methods for tackling interference issues. Most existing eigenvalue-based methods only work on the assumption of no interference.

- The detection performance analysis of cyclostationary spectrum sensing over fading channels has been presented in Chapter 5. As finding closed-form expressions of average detection probability is generally difficult, some alternative close-form performance bounds or series-expansion approximations are provided. The tightness of these bounds and approximations has not yet been quantitatively expressed. On top of this, several diversity reception schemes, such as maximum ratio combining and switch and stay combining, have not been analytically discussed.

- In Chapter 6 we investigate the interference constraints which should be imposed on the CR system with perfect or partial channel side information. There are some scenarios that can be considered in the future such as the PU transmission undergoing slow fading while the CR interference is fast faded, and vice versa. In addition, it is interesting to know what will happen without the assumption that interference from other sources seen by the PU receiver remains stationary.

- In our proposed cooperative path loss estimation, the geographical information plays an important role. However, the sensitivity of our proposed method to relative distance uncertainty has not been analyzed. Moreover, the simulation results have shown an interesting phenomenon that the accuracy of the parameter of the de-correlation distance is not very important. The quantitative analysis of how this inaccuracy affects our proposed method should be further addressed.

REFERENCES

[1] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, pp. 13–18, Aug. 1999.

[2] I. Mitola, J., "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Royal Institute of Technology (KTH), 2000.

[3] T. Weiss and F. Jondral, "Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency," *IEEE Communications Magazine*, vol. 42, no. 3, pp. S8–14, 2004.

[4] "Spectrum policy task force report," Federal Communications Commission, Tech. Rep. ET Docket no. 02-135, Nov. 2002.

[5] G. Staple and K. Werbach, "The end of spectrum scarcity [spectrum allocation and utilization]," *IEEE Spectrum*, vol. 41, no. 3, pp. 48–52, March 2004.

[6] A. Shukla, "Cognitive radio technology - a study for ofcom," Office of Communications, Tech. Rep. QINETIQ/06/00420, Feb. 2007.

[7] J. Xiao, R. Hu, Y. Qian, L. Gong, and B. Wang, "Expanding LTE network spectrum with cognitive radios: From concept to implementation," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 12–19, 2013.

[8] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[9] P. J. Kolodzy, "Interference temperature: a metric for dynamic spectrum utilization," *International Journal of Network Management*, vol. 16, pp. 103–113, 2006.

[10] T. Clancy, "Dynamic spectrum access using the interference temperature model," *annals of telecommunications - annales des télécommunications*, vol. 64, no. 7-8, pp. 573–592, 2009.

[11] A. Ghasemi and E. S. Sousa, "Fundamental limits of spectrum-sharing in fading environments," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 649–658, Feb. 2007.

[12] M. Nekovee, T. Irnich, and J. Karlsson, "Worldwide trends in regulation of secondary access to white spaces using cognitive radio," *IEEE Wireless Communications*, vol. 19, no. 4, pp. 32–40, 2012.

[13] G. Buchwald, S. Kuffner, L. Ecklund, M. Brown, and E. Callaway, "The design and operation of the IEEE 802.22.1 disabling beacon for the protection of TV whitespace incumbents," in *3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2008, pp. 1–6.

[14] Z. Lei and F. Chin, "A reliable and power efficient beacon structure for cognitive radio systems," in *IEEE International Conference on Communications*, 2008, pp. 2038–2042.

[15] K. Shin, H. Kim, A. Min, and A. Kumar, "Cognitive radios for dynamic spectrum access: from concept to reality," *IEEE Wireless Communications*, vol. 17, no. 6, pp. 64–74, 2010.

[16] H. Karimi, "Geolocation databases for white space devices in the UHF TV bands: Specification of maximum permitted emission levels," in *IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2011, pp. 443–454.

[17] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 201–220, Feb. 2005.

[18] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, Monterey, CA, Nov. 2004, pp. 772–776.

[19] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications.* Cambridge University Press, 1993, vol. 1.

[20] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory.* Prentice Hall, 1998.

[21] J. Ma, G. Li, and B. H. Juang, "Signal processing in cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 805–823, May 2009.

[22] 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, 2009.

[23] A. Nuttall, "Some integrals involving the Q_M function (corresp.)," *IEEE Transactions on Information Theory*, vol. 21, no. 1, pp. 95–96, Jan. 1975.

[24] S. Shellhammer and R. Tandra, "Performance of the power detector with noise uncertainty," Tech. Rep. IEEE Std. 802.22-06/0134r0, Jul. 2006.

[25] S. Shellhammer and G. Chouinard, "Spectrum sensing requirements summary," Tech. Rep. IEEE 802.22-06/0089r1, 2006.

[26] A. Sonnenschein and P. Fishman, "Radiometric detection of spread-spectrum signals in noise of uncertain power," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 654–660, Jul 1992.

[27] R. Tandra and A. Sahai, "Fundamental limits on detection in low snr under noise uncertainty," in *International Conference on Wireless Networks, Communications and Mobile Computing*, vol. 1, June 2005, pp. 464–469.

[28] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *IEEE Signal Processing Magazine*, vol. 8, pp. 14–36, Apr. 1991.

[29] A. V. Dandawate and G. B. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Transactions on Signal Processing*, vol. 42, pp. 2355–2369, Sept. 1994.

[30] W. A. Gardner, *Introduction to random processes : with applications to signals and systems.* New York: MacMillan, 1986.

[31] M. Oner and F. Jondral, "Cyclostationarity based air interface recognition for software radio systems," in *IEEE Radio and Wireless Conference*, Atlanta, GA, Sept. 2004, pp. 263–266.

[32] W. A. Gardner and C. M. Spooner, "Signal interception: performance advantages of cyclic-feature detectors," *IEEE Transactions on Communications*, vol. 40, pp. 149–159, Jan. 1992.

[33] Z. Tian, Y. Tafesse, and B. Sadler, "Cyclic feature detection with sub-nyquist sampling for wideband spectrum sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 1, pp. 58–69, 2012.

[34] D. Cohen, E. Rebeiz, V. Jain, Y. Eldar, and D. Cabric, "Cyclostationary feature detection from sub-nyquist samples," in *4th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, 2011, pp. 333–336.

[35] E. Rebeiz, V. Jain, and D. Cabric, "Cyclostationary-based low complexity wideband spectrum sensing using compressive sampling," in *IEEE International Conference on Communications*, 2012, pp. 1619–1623.

[36] K. W. Choi, W. S. Jeon, and D. G. Jeong, "Sequential detection of cyclostationary signal for cognitive radio systems," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4480–4485, 2009.

[37] L. Izzo, L. Paura, and M. Tanda, "Signal interception in non-gaussian noise," *IEEE Transactions on Communications*, vol. 40, no. 6, pp. 1030–1037, 1992.

[38] H. L. Hurd and A. Miamee, *Periodically Correlated Random Sequences: Spectral Theory and Practice.* Hoboken, NJ: Wiley-Interscience, 2007.

[39] H.-S. Chen, W. Gao, and D. Daut, "Spectrum sensing using cyclostationary properties and application to ieee 802.22 wran," in *IEEE Global Telecommunications Conference*, Nov. 2007, pp. 3133–3138.

[40] Z. Ye, J. Grosspietsch, and G. Memik, "Spectrum sensing using cyclostationary spectrum density for cognitive radios," in *IEEE Workshop on Signal Processing Systems*, Oct. 2007, pp. 1–6.

[41] H. Feng, Y. Wang, and S. Li, "Statistical test based on finding the optimum lag in cyclic autocorrelation for detecting free bands in cognitive radios," in *the 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, Singapore, May 2008, pp. 1–6.

[42] J. Lunden, V. Koivunen, A. Huttunen, and H. V. Poor, "Spectrum sensing in cognitive radios based on multiple cyclic frequencies," in *the 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, Orlando, FL, Aug. 2007, pp. 37–43.

[43] M. Kim, P. Kimtho, and J.-i. Takada, "Performance enhancement of cyclostationarity detector by utilizing multiple cyclic frequencies of ofdm signals," in *IEEE Symposium on New Frontiers in Dynamic Spectrum*, Apr. 2010, pp. 1–8.

[44] H. Sadeghi, P. Azmi, and H. Arezumand, "Optimal multi-cycle cyclostationarity-based spectrum sensing for cognitive radio networks," in *19th Iranian Conference on Electrical Engineering*, May 2011, pp. 1–6.

[45] D. Shen, D. He, W.-h. Li, and Y.-p. Lin, "An improved cyclostationary feature detection based on the selection of optimal parameter in cognitive radios," *Journal of Shanghai Jiaotong University (Science)*, vol. 17, pp. 1–7, 2012.

[46] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data.* Cambridge University Press, 2010.

[47] J. Lunden, V. Koivunen, A. Huttunen, and H. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4182–4195, Nov. 2009.

[48] P. Rostaing, T. Pitarque, and E. Thierry, "Performance analysis of a statistical test for presence of cyclostationarity in a noisy observation," in *Proc. IEEE International Conference on the Acoustics, Speech, and Signal Processing*, Atlanta, GA, May 1996, pp. 2932–2935.

[49] Y. Sun, A. Baricz, and S. Zhou, "On the monotonicity, log-concavity, and tight bounds of the generalized marcum and nuttall Q-functions," *IEEE Transactions on Information Theory*, vol. 56, pp. 1166–1186, 2010.

[50] G. Huang and J. Tugnait, "On cyclostationarity based spectrum sensing under uncertain gaussian noise," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2042–2054, April 2013.

[51] J.-C. Shen and E. Alsusa, "An efficient multiple lags selection method for cyclostationary feature based spectrum sensing," *IEEE Signal Processing Letters*, vol. 20, no. 2, pp. 133–136, Feb. 2013.

[52] ——, "Joint cycle frequencies and lags utilization in cyclostationary feature spectrum sensing," accepted by IEEE Transactions on Signal Processing in July 2013.

[53] R. J. Serfling, *Approximation Theorems of Mathematical Statistics.* Hoboken, NJ: John Wiley & Sons, 1980.

[54] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, pp. 814–823, Feb. 2010.

[55] P. Wang, J. Fang, N. Han, and H. Li, "Multiantenna-assisted spectrum sensing for cognitive radio," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1791–1800, 2010.

[56] R. Zhang, T. J. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A GLRT approach," *IEEE Transactions on Communications*, vol. 58, no. 1, pp. 84–88, Jan. 2010.

[57] E. Axell and E. Larsson, "Eigenvalue-based spectrum sensing of orthogonal space-time block coded signals," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6724–6728, 2012.

[58] Y. Zeng and Y.-C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *IEEE Transactions on Communications*, vol. 57, no. 6, pp. 1784–1793, 2009.

[59] M. Jin, Y. Li, and H.-G. Ryu, "On the performance of covariance based spectrum sensing for cognitive radio," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3670–3682, 2012.

[60] Y. Zeng, Y.-C. Liang, and R. Zhang, "Blindly combined energy detection for spectrum sensing in cognitive radio," *IEEE Signal Processing Letters*, vol. 15, pp. 649–652, 2008.

[61] S. Kim, J. Lee, H. Wang, and D. Hong, "Sensing performance of energy detector with correlated multiple antennas," *IEEE Signal Processing Letters*, vol. 16, no. 8, pp. 671–674, Aug. 2009.

[62] J. Sala-Alvarez, G. Vazquez-Vilar, and R. Lopez-Valcarce, "Multiantenna GLR detection of rank-one signals with known power spectrum in white noise with unknown spatial correlation," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 3065–3078, 2012.

[63] B. Agee, S. Schell, and W. Gardner, "Spectral self-coherence restoral: a new approach to blind adaptive signal extraction using antenna arrays," *Proceedings of the IEEE*, vol. 78, no. 4, pp. 753–767, 1990.

[64] B. Allen and M. Ghavami, *Adaptive array systems: fundamentals and applications.* Wiley, 2006.

[65] J.-H. Lee and C.-C. Huang, "Blind adaptive beamforming for cyclostationary signals: A subspace projection approach," *IEEE Antennas and Wireless Propagation Letters*, vol. 8, pp. 1406–1409, 2009.

[66] K.-L. Du and W.-H. Mow, "Affordable cyclostationarity-based spectrum sensing for cognitive radio with smart antennas," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1877–1886, 2010.

[67] Q. Wu and K. M. Wong, "Blind adaptive beamforming for cyclostationary signals," *IEEE Transactions on Signal Processing*, vol. 44, no. 11, pp. 2757–2767, 1996.

[68] K. Jitvanichphaibool, Y.-C. Liang, and Y. Zeng, "Spectrum sensing using multiple antennas for spatially and temporally correlated noise environments," in *IEEE Symposium on New Frontiers in Dynamic Spectrum*, 2010, pp. 1–7.

[69] R. Mahapatra and M. Krusheel, "Cyclostationary detection for cognitive radio with multiple receivers," in *IEEE International Symposium on Wireless Communication Systems*, Oct. 2008, pp. 493–497.

[70] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Processing*, vol. 86, no. 4, pp. 639–697, 2006.

[71] B. Sadler and A. Dandawate, "Nonparametric estimation of the cyclic cross spectrum," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 351–358, Jan. 1998.

[72] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[73] A. Pandharipande and J.-P. Linnartz, "Performance analysis of primary user detection in a multiple antenna cognitive radio," in *IEEE International Conference on Communications*, June 2007, pp. 6482–6486.

[74] Z. Quan, S. Cui, and A. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 28–40, Feb. 2008.

[75] Z. Quan, S. Cui, A. Sayed, and H. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1128–1140, Mar. 2009.

[76] B. Shen, S. Ullah, and K. Kwak, "Deflection coefficient maximization criterion based optimal cooperative spectrum sensing," *AEU - International Journal of Electronics and Communications*, vol. 64, no. 9, pp. 819–827, 2010.

[77] Y. Chen, G. Yu, Z. Zhang, H. hwa Chen, and P. Qiu, "On cognitive radio networks with opportunistic power control strategies in fading channels," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2752–2761, July 2008.

[78] G. W. Stewart, *Matrix Algorithms Volume II: Eigensystems.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2001.

[79] D. Kressner, *Numerical Methods for General and Structured Eigenvalue Problems.* Springer, 2005.

[80] A. Dandawate and G. Giannakis, "Nonparametric polyspectral estimators for kth-order (almost) cyclostationary processes," *IEEE Transactions on Information Theory*, vol. 40, no. 1, pp. 67–84, Jan. 1994.

[81] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis.* Wiley-Interscience, 2003.

[82] J. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, no. 3, pp. 278–305, Mar. 1991.

[83] P. O. Amblard, M. Gaeta, and J. L. Lacoume, "Statistics for complex variables and signals – Part II: signals," *Signal Processing*, vol. 53, no. 1, pp. 15–25, Aug. 1996.

[84] D. R. Brillinger, *Time series: data analysis and theory.* Holt, Rinehart, and Winston, 1975.

[85] R. A. Horn and C. R. Johnson, *Matrix analysis.* Cambridge university press, 1990.

[86] F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," in *IEEE International Conference on Communications*, vol. 5, 2003, pp. 3575–3579.

[87] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 21–24, Jan. 2007.

[88] P. Sofotasios, E. Rebeiz, L. Zhang, T. Tsiftsis, D. Cabric, and S. Freear, "Energy detection based spectrum sensing over κ -μ and κ -μ extreme fading channels," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 3, pp. 1031–1040, 2013.

[89] E. Gismalla and E. Alsusa, "Performance analysis of the periodogram-based energy detector in fading channels," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3712–3721, 2011.

[90] ——, "On the performance of energy detection using bartlett's estimate for spectrum sensing in cognitive radio systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3394–3404, 2012.

[91] K. Ruttik, K. Koufos, and R. Jantti, "Detection of unknown signals in a fading environment," *IEEE Communications Letters*, vol. 13, no. 7, pp. 498–500, 2009.

[92] S. Atapattu, C. Tellambura, and H. Jiang, "Performance of an energy detector over channels with both multipath fading and shadowing," *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3662–3670, 2010.

[93] ——, "Energy detection based cooperative spectrum sensing in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 4, pp. 1232–1241, 2011.

[94] P. D. Sutton, J. Lotze, K. E. Nolan, and L. E. Doyle, "Cyclostationary signature detection in multipath rayleigh fading environments," in *2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, 2007, pp. 408–413.

[95] M. Naraghi-Pour and T. Ikuma, "Autocorrelation-based spectrum sensing for cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 2, pp. 718–733, Feb. 2010.

[96] M. K. Simon and M.-S. Alouini, *Digital communication over fading channels.* Wiley-Interscience, 2000.

[97] Y. A. Brychkov, "On some properties of the marcum q function," *Integral Transforms and Special Functions*, vol. 23, no. 3, pp. 177–182, 2012.

[98] Y. Sun and A. Baricz, "Inequalities for the generalized marcum q-function," *Applied Mathematics and Computation*, vol. 203, no. 1, pp. 134–141, 2008.

[99] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 7th ed. Academic Press, 2007.

[100] M. Simon and M.-S. Alouini, "Exponential-type bounds on the generalized marcum q-function with application to error probability analysis over fading channels," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 359–366, Mar. 2000.

[101] J. Filho and M. Yacoub, "Simple precise approximations to weibull sums," *IEEE Communications Letters*, vol. 10, no. 8, pp. 614–616, Aug. 2006.

[102] J.-C. Shen, E. Alsusa, and D. K. So, "Performance bounds on cyclostationary feature detection over fading channels," in *IEEE Wireless Communications and Networking Conference*, 2013, pp. 2971–2975.

[103] J.-C. Shen and E. Alsusa, "Post-combining based cyclostationary feature detection for cognitive radio over fading channels," in *IEEE Global Communications Conference*, 2013.

[104] N. Shankar, C. Cordeiro, and K. Challapali, "Spectrum agile radios: utilization and sensing architectures," in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2005, pp. 160–169.

[105] J. Bater, H. Tan, K. Brown, and L. Doyle, "Modelling interference temperature constraints for spectrum access in cognitive radio networks," in *IEEE International Conference on Communications*, 2007, pp. 6493–6498.

[106] J. Peha, "Approaches to spectrum sharing," *IEEE Communications Magazine*, vol. 43, no. 2, pp. 10–12, Feb. 2005.

[107] Y. Chen, G. Yu, Z. Zhang, H. hwa Chen, and P. Qiu, "On cognitive radio networks with opportunistic power control strategies in fading channels," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2752–2761, July 2008.

[108] H. Suraweera, P. Smith, and M. Shafi, "Capacity limits and performance analysis of cognitive radio with imperfect channel knowledge," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1811–1822, May 2010.

[109] S. Lim, H. Wang, H. Kim, and D. Hong, "Mean value-based power allocation without instantaneous csi feedback in spectrum sharing systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 874–879, March 2012.

[110] J. Vicario, A. Bel, J. Lopez-Salcedo, and G. Seco, "Opportunistic relay selection with outdated csi: outage probability and diversity analysis," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 2872–2876, June 2009.

[111] K. Phan, S. Vorobyov, N. Sidiropoulos, and C. Tellambura, "Spectrum sharing in wireless networks via qos-aware secondary multicast beamforming," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2323–2335, June 2009.

[112] N. Benvenuto and F. Santucci, “A least squares path-loss estimation approach to handover algorithms,” *IEEE Transactions on Vehicular Technology*, vol. 48, no. 2, pp. 437–447, Mar. 1999.

[113] X. Zhao, L. Razouniov, and L. Greenstein, “Path loss estimation algorithms and results for rf sensor networks,” in *IEEE 60th Vehicular Technology Conference*, vol. 7, Sept. 2004, pp. 4593–4596.

[114] T. S. Rappaport, *Wireless communications : principles and practice*, 2nd ed. Upper Saddle River, N.J.: Prentice Hall, 2002.

[115] O. Durowoju, K. Arshad, and K. Moessner, “Distributed power control for cognitive radio networks, based on incumbent outage information,” in *IEEE International Conference on Communications*, June 2011, pp. 1–5.

[116] M. Gudmundson, “Correlation model for shadow fading in mobile radio systems,” *Electronics Letters*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.

[117] N. Patwari, I. Hero, A.O., M. Perkins, N. Correal, and R. O’Dea, “Relative location estimation in wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, Aug. 2003.

[118] K. W. Sung, M. Tercero, and J. Zander, “Aggregate interference in secondary access with interference protection,” *IEEE Communications Letters*, vol. 15, no. 6, pp. 629–631, June 2011.

[119] R. M. Gray and L. D. Davisson, *An introduction to statistical signal processing.* Cambridge University Press, 2004.

[120] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge, 1985.

[121] A. Algans, K. Pedersen, and P. Mogensen, “Experimental analysis of the joint statistical properties of azimuth spread, delay spread, and shadow fading,” *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 3, pp. 523–531, Apr. 2002.

[122] J.-C. Shen and E. Alsusa, “Cooperative estimation of path loss in interference channels without primary-user csi feedback,” *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 273–276, 2013.

[123] H. Liu, Y. Tang, and H. H. Zhang, “A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables,” *Computational Statistics & Data Analysis*, vol. 53, no. 4, pp. 853–856, 2009.