Department of Communications and Networking

Towards Robust Spectrum Sensing in Cognitive Radio Networks

Lu Wei

Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 145/2013

Towards Robust Spectrum Sensing in Cognitive Radio Networks

Lu Wei

A doctoral dissertation completed for the degree of Doctor of Science in Technology to be defended, with the permission of the Aalto University School of Electrical Engineering, at a public examination held at the lecture hall S1 of the school on 1 November 2013 at 12 noon.

Aalto University
School of Electrical Engineering
Department of Communications and Networking

Supervising professor
Prof. Olav Tirkkonen

Thesis advisor
Prof. Olav Tirkkonen

Preliminary examiners
Prof. Ying-Chang Liang, Institute for Infocomm Research, Singapore
Prof. Markku Renfors, Tampere University of Technology, Finland

Opponents
Prof. Markku Renfors, Tampere University of Technology, Finland
Prof. Giorgio Taricco, Politecnico di Torino, Italy

Aalto University publication series
DOCTORAL DISSERTATIONS 145/2013

ISBN 978-952-60-5334-9 (printed)
ISBN 978-952-60-5335-6 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-5335-6

Unigrafia Oy
Helsinki 2013

Finland

NORDIC ECOLABEL
441 697
Printed matter

A! **Aalto University**

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Lu Wei

Name of the doctoral dissertation
Towards Robust Spectrum Sensing in Cognitive Radio Networks

Publisher School of Electrical Engineering

Unit Department of Communications and Networking

Series Aalto University publication series DOCTORAL DISSERTATIONS 145/2013

Field of research Communications Engineering

Manuscript submitted 16 May 2013 **Date of the defence** 1 November 2013

Permission to publish granted (date) 8 August 2013 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

Abstract

This thesis focuses on multi-antenna assisted energy based spectrum sensing. The studies leading to this thesis have been motivated by some practical issues with energy based detection. These include the noise uncertainty problem at the secondary receiver, the presence of multiple active primary users in cognitive cellular networks, the existence of unknown noise correlations and detection in the low signal-to-noise ratio regime.

In this thesis, the aim is to incorporating these practical concerns into the design of spectrum sensing algorithms. To this end, we propose the use of various detectors that are suitable for different scenarios. We consider detectors derived from decision-theoretical criterions as well as heuristic detectors. We analyze the performance of the proposed detectors by deriving their false alarm probability, detection probability and receiver operating characteristic. The main contribution of this thesis consists of the derived closed-form performance metrics. These results are obtained by utilizing tools from multivariate analysis, moment based approximations, Mellin transforms, and random matrix theory.

Numerical results show that the proposed detectors have indeed resolved the concerns raised by the above practical issues. Some detectors could meet the needs of one of the practical challenges, while others are shown to be robust when several practical issues are taken into account. The use of detectors constructed with decision-theoretical considerations over the heuristically proposed ones is justified as well.

Keywords cognitive radio; multivariate analysis; robust statistics; spectrum sensing

ISBN (printed) 978-952-60-5334-9 **ISBN (pdf)** 978-952-60-5335-6

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2013

Pages 144 **urn** http://urn.fi/URN:ISBN:978-952-60-5335-6

To Xin and Luke

Preface

The work leading to this thesis has been carried out in the Communication Theory group of the Department of Communications and Networking (Comnet), Aalto university, during the years 2008-2012. The work has been funded by the Graduate School of Electrical and Communications Engineering, the Academy of Finland, the Nokia Foundation, and the China Scholarship Council.

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Olav Tirkkonen for his continuous support and guidance. Prof. Tirkkonen's keen insights have always led to substantial improvement and new directions to our research work. It has been a great pleasure and an honer to work with him during the past years.

I would also like to extend my sincere gratitude to my coauthors: Prof. Matthew McKay from Hong Kong University of Science and Technology, Dr. Prathapasinghe Dharmawansa from Stanford University and Mr. Zhong Zheng from Aalto University. Thanks for Prof. McKay's hospitality during my visit to his group in the year 2010. Thanks for Dr. Dharmawansa's sharp comments and technical help, which make Publication V possible. Thanks for Mr. Zheng's efficiency in carrying on our ideas into publications. I am deeply indebted to all of you.

I would like to thank Mr. Shang Li from Columbia University, Dr. Boaz Nadler from Weizmann Institute of Science, and Prof. Dietrich von Rosen from Swedish University of Agricultural Sciences for the interesting and insightful discussions. In particular, Prof. von Rosen's encouragement had led me to the idea of Publication IV.

I would like to thank the current and former Comnet members: Dr. Chia-Hao Yu, Mr. Renaud-Alexander Pitaval, Dr. Konstantinos Koufos, and Dr. Kalle Ruttik for the small talk on academic and non-academic matters. The Comnet secretaries and supporting team are also acknowledged for making our everyday life easier.

I wish to thank Prof. Jukka Corander from University of Helsinki for giving me the opportunity to continuously pursue my current research direction.

Special thanks to Prof. Gernot Akemann from University of Bielefeld for introducing me the fruitful direction of products of random matrices.

Thanks also go to my Chinese friends in Finland for making our life colorful. Last but not least, I would like to thank my parents and parents-in-law for their love and support. This thesis is dedicated to my wife Xin and our little son Luke.

Espoo, September 12, 2013,

Lu Wei

Contents

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I Lu Wei and Olav Tirkkonen. Cooperative spectrum sensing of OFDM signals using largest eigenvalue distributions. In *Proceedings of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 2295-2299, Sept. 2009.

II Lu Wei and Olav Tirkkonen. Analysis of scaled largest eigenvalue based detection for spectrum sensing. In *Proceedings of IEEE International Conference on Communications*, pp. 1-5, June 2011.

III Lu Wei, Olav Tirkkonen, Prathapasinghe Dharmawansa, and Matthew McKay. On the exact distribution of the scaled largest eigenvalue. In *Proceedings of IEEE International Conference on Communications*, pp. 2422-2426, June 2012.

IV Lu Wei and Olav Tirkkonen. Spectrum sensing in the presence of multiple primary users. *IEEE Transactions on Communications*, vol. 60, no. 5, pp. 1268-1277, May 2012.

V Lu Wei, Prathapasinghe Dharmawansa, and Olav Tirkkonen. Multiple primary user spectrum sensing in the low SNR regime. *IEEE Transactions on Communications*, vol. 61, no. 5, pp. 1720-1731, May 2013.

VI Lu Wei and Olav Tirkkonen. Approximate condition number distribution of complex non-central correlated Wishart matrices. In *Proceedings of IEEE International Conference on Communications*, pp. 1-5, June 2011.

VII Lu Wei, Matthew McKay, and Olav Tirkkonen. Exact Demmel condition number distribution of complex Wishart matrices via the Mellin transform. *IEEE Communications Letters*, vol. 15, no. 2, pp. 175-177, Feb. 2011.

VIII Lu Wei and Olav Tirkkonen. Multiple primary user spectrum sensing for unknown noise statistics. In *Proceedings of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept. 2013.

Author's Contribution

Publication I: "Cooperative spectrum sensing of OFDM signals using largest eigenvalue distributions"

The author of this thesis had the main responsibility in formulating the idea, deriving the results, and writing of this article.

Publication II: "Analysis of scaled largest eigenvalue based detection for spectrum sensing"

The author of this thesis had the main responsibility in formulating the idea, deriving the results, and writing of this article.

Publication III: "On the exact distribution of the scaled largest eigenvalue"

The author of this thesis had a partial responsibility in formulating the idea and deriving the results, and main responsibility in writing of this article. The line of derivation based on Meijer's G-function was initiated by the third author.

Publication IV: "Spectrum sensing in the presence of multiple primary users"

The author of this thesis had the main responsibility in formulating the idea, deriving the results, and writing of this article.

Publication V: “Multiple primary user spectrum sensing in the low SNR regime”

The author of this thesis had the main responsibility in formulating the idea, deriving the results, and writing of this article. The proof of Lemma 1 is provided by the second author.

Publication VI: “Approximate condition number distribution of complex non-central correlated Wishart matrices”

The author of this thesis had the main responsibility in formulating the idea, deriving the results, and writing of this article.

Publication VII: “Exact Demmel condition number distribution of complex Wishart matrices via the Mellin transform”

The author of this thesis had the main responsibility in formulating the idea, deriving the results, and writing of this article.

Publication VIII: “Multiple primary user spectrum sensing for unknown noise statistics”

The author of this thesis had the main responsibility in formulating the idea, deriving the results, and writing of this article.

List of Abbreviations

DCN	Demmel Condition Number based
DECT	Digital Enhanced Cordless Telecommunications
DVB	Digital Video Broadcasting
ED	Energy Detector
ER	Eigenvalue Ratio based
GLR	Generalized Likelihood Ratio
i.i.d	independent and identically distributed
LBI	Locally Best Invariant
LE	Largest Eigenvalue based
RMT	Random Matrix Theory
ROC	Receiver Operating Characteristic
SLE	Scaled Largest Eigenvalue based
SNR	Signal-to-Noise Ratio
ST	Spherical Test based
TV	Television
w.r.t.	with respect to

List of Symbols

$(\cdot)^\dagger$	conjugate-transpose operation
$\|\| \cdot \|\|$	Euclidean norm
$\| \cdot \|$	matrix determinant operation
$\frac{\partial(\cdot)}{\partial(\cdot)}$	partial derivative operator
$\succ$	positive definite symbol
$A(\cdot,\cdot)$	Airy kernel
$A_\mathrm{i}(\cdot)$	Airy function
$\mathrm{adj}(\cdot)$	matrix adjoint operator
$B(\cdot,\cdot)$	Beta function
$B_{(\cdot)}(\cdot,\cdot)$	incomplete Beta function
$\mathbb{C}^K$	complex space of dimension K
$\mathbf{E}$	noise-only sample covariance matrix
$\mathbb{E}[\cdot]$	expectation
$f_n(x)$	monic orthogonal polynomial of degree n
${}_0F_1(\cdot;\cdot)$	Bessel type hypergeometric function
${}_2F_1(\cdot,\cdot;\cdot;\cdot)$	Gaussian type hypergeometric function
$F_{\mathrm{DCN}}(\cdot)$	distribution of T_{DCN} under $\mathcal{H}_0$
$F_{\mathrm{J}}(\cdot)$	distribution of T_{J} under $\mathcal{H}_0$
$F_{\mathrm{LE}}(\cdot)$	distribution of T_{LE} under $\mathcal{H}_0$
$F_{\mathrm{SLE}}(\cdot)$	distribution of T_{SLE} under $\mathcal{H}_0$
$F_{\mathrm{ST}}(\cdot)$	distribution of T_{ST} under $\mathcal{H}_0$
$F_{\mathrm{TW2}}(\cdot)$	Tracy-Widom distribution of order two
$F_{\mathrm{W}}(\cdot)$	distribution of T_{W} under $\mathcal{H}_0$
$G_{\mathrm{ER}}(\cdot)$	distribution of T_{ER} under $\mathcal{H}_1$
$G_{\mathrm{J}}(\cdot)$	distribution of T_{J} under $\mathcal{H}_1$
$G_{\mathrm{LE}}(\cdot)$	distribution of T_{LE} under $\mathcal{H}_1$
$G_{\mathrm{ST}}(\cdot)$	distribution of T_{ST} under $\mathcal{H}_1$
$h(\cdot)$	Heaviside step function

$\mathbf{h}_i$	channel vector of primary user i
$\mathbf{H}$	channel matrix
$\mathcal{H}_0$	hypothesis 0: absence of primary users
$\mathcal{H}_1$	hypothesis 1: presence of primary users
$\mathbf{I}_K$	identity matrix of dimension K
K	number of sensors
M	number of noise-only observations
m_i	i-th moment of weight function
$M_s[\cdot]$	Mellin transform operator
$\mathbf{n}$	noise vector
N	number of samples
P	number of primary users
P_{d}	detection probability
P_{fa}	false alarm probability
$\mathbb{P}(\cdot)$	probability of an event
$\mathbf{R}$	sample covariance matrix
$\mathbf{s}$	signal vector of primary users
s_i	signal of primary user i
$\mathbf{tr}(\cdot)$	matrix trace operation
T_{DCN}	Demmel condition number based detector
T_{ED}	energy detector
T_{ER}	eigenvalue radio based detector
T_{LE}	largest eigenvalue based detector
T_{J}	John's detector
T_{R}	Roy's detector
T_{SLE}	scaled largest eigenvalue based detector
T_{ST}	spherical test based detector
$w(x)$	weight function
$\mathcal{W}_{(\cdot)}\left(\cdot,\cdot\right)$	Wishart distribution
$\mathcal{W}_{(\cdot)}\left(\cdot,\cdot,\cdot\right)$	non-central Wishart distribution
$\mathbf{x}$	received data vector
$\mathbf{X}$	received data matrix
$\mathbf{Z}$	noise-only observation matrix
γ_i	transmission power of primary user i
$\gamma(\cdot,\cdot)$	lower incomplete Gamma function
$\Gamma(\cdot)$	Gamma function
$\Gamma(\cdot,\cdot)$	upper incomplete Gamma function
$\Gamma_{(\cdot)}(\cdot)$	multivariate Gamma function

Δ	monic orthogonal polynomial normalization
ζ	decision threshold
$\mathbf{\Theta}$	parameter matrix of Wishart moment generating function
θ_i	i-th eigenvalue of $\mathbf{E}^{-1}\mathbf{R}$
ϑ	degree of noise uncertainty in dB
λ_i	i-th eigenvalue of $\mathbf{R}$
$\mathbf{\Xi}$	non-central parameter matrix
ξ_i	i-th eigenvalue of $\mathbf{\Xi}$
ρ	degree of noise correlation
σ^2	noise power
$\mathbf{\Sigma}$	population covariance matrix
σ_i	i-th eigenvalue of $\mathbf{\Sigma}$ under $\mathcal{H}_1$
$\Phi(\cdot)$	Gaussian distribution function
$\mathbf{\Psi}$	noise covariance matrix
$\psi(x)$	initial approximation

1. Introduction

1.1 Motivation

The electromagnetic radio spectrum is a precious natural resource. Recent measurements have shown that the current static spectrum allocation regulations lead to severe underutilization of the spectrum. Given the ever-increasing need for higher data rate services, technologies that could offer new ways of exploiting the available spectrum are called for. Cognitive radio, firstly proposed in [1], arises to be a promising solution to the spectrum underutilization issue by introducing the concept of dynamic spectrum access. Contrary to the fixed spectrum access schemes, in dynamic spectrum access secondary (unlicensed) users are allowed to opportunistically use the frequency bands that are not heavily occupied by primary (licensed) users. The terminology primary users refers to the users who have legacy rights on the usage of a specific part of the spectrum. The secondary users, on the other hand, have lower priority, and could only exploit this spectrum in such a way that they do not cause intolerable interference to the primary users. To achieve this dynamic spectrum access scenario, the secondary users need to be able to infer reliably whether a frequency band of interest is being used by the primary users [2]. Therefore, the ability to be aware of the spectrum usage in a geographical area becomes the most important component for the establishment of cognitive radio.

Spectrum sensing is the task of acquiring the awareness of spectrum usage, which can be obtained by using geolocation and database, by using beacons, or by local spectrum sensing at cognitive radios [3]. The focus of this thesis is on the local spectrum sensing i.e. detection based spectrum access, which is hereinafter referred to as spectrum sensing. The present literature on spectrum sensing is far from being fully developed, however most of the existing sensing algorithms can be either categorized into feature based detection or energy based detection. For comprehensive survey of spectrum sensing algorithms, we refer the readers to [3, 4] and references therein for the state-of-the-art in this direction. Feature based detection typically requires signaling information or/and statistical properties of the primary systems such as preambles, pilot patterns, spreading sequence, carrier frequency, signal power, and so on. Feature based detection includes, among others, matched filter detection [5], cyclostationary detection [6], and autocorrelation detection [7]. On the other hand, energy based detection does not require any a priori signaling information of the primary users, and thus is often called blind detection. The conventional energy detector [8], also known as radiometry, is the simplest example of energy based detection. Recently, multi-antenna spectrum sensing emerges as a promising candidate in energy based detection. The studies leading to this thesis are motivated by some of the practical issues with the multi-antenna assisted energy based detection. Specifically, in designing realistic multi-antenna spectrum sensing algorithms the following facts shall be taken into account:

Fact 1 *Perfect knowledge of noise power may not be available due to interference, noise estimation errors or non-linearity of the components.*

Fact 2 *The existence of more than one active primary users would be the prevailing condition in forthcoming CR networks.*

Fact 3 *Noise correlation at the secondary receiver is inevitable due to e.g. antenna coupling in practical systems, and the degree of correlation is usually unknown due to its time-varying nature.*

The phenomenons described in **Fact 1** give rise to the concept of noise uncertainty [9]. For systems in practice, modeling of noise uncertainty is unavoidable. Noise uncertainty may further lead to the phenomenon of Signal-to-Noise Ratio (SNR) wall [10], where the detection of primary users is impossible even if the sample size goes to infinity. **Fact 1** motivates the design of test statistics that are not functions of the noise power. The resulting detectors are blind to noise power uncertainty. **Fact 2** holds since the assumption of single primary user fails to reflect the situation in forthcoming cognitive radio networks, where the primary systems may include a cellular network. Moreover, in unlicensed bands, several unlicensed systems, such as Wi-Fi, Bluetooth, and Digital Enhanced Cordless

Telecommunications (DECT), may share the same band without coordination, resulting in multiple primary user scenarios [11]. Note that the assumption of a single primary user is made as the investigations in the literature have mainly focussed on cognitive radio networks, where the primary users are Television (TV) or Digital Video Broadcasting (DVB) systems. In these systems the single active primary user assumption is, to some extent, justifiable. Here the terminology "single primary user" refers to the scenario where the maximum number of simultaneously transmitting primary users equals one. In contrast, the term "multiple primary users" refers to an existence of more than one active primary users, but the exact number is irrelevant. **Fact 2** motivates us to design test statistics that could work efficiently in the scenarios of multiple primary users. **Fact 3** reflects realistic spectrum sensing scenarios where the assumption of a perfectly known noise covariance matrix is not realistic due to the time-varying nature of the noise statistics. The true noise covariance matrix needs to be estimated by a separate data set consisting of periodically updated noise-only observations. **Fact 3** motivates us to search for detectors that are able to work irrespectively of the actual noise covariance matrix.

1.2 Scope

Essentially, spectrum sensing is about obtaining the spectrum usage information across multiple dimensions such as time, space, and frequency. It also involves determining the types of signals that are occupying the spectrum, e.g. the modulation, waveform, bandwidth, carrier frequency, and so on [3]. However, in this thesis we chose to formulate the spectrum sensing problem as a binary hypothesis test. Namely, we are only interested in inferring the presence of absence of primary users without examining further detailed information on the spectrum usage or the primary users' signal.

This thesis is limited to the study of energy based detection of multi-antenna type. We progressively take into account the practical facts discussed in the previous section, leading to a variety of detectors that are suitable for different scenarios. The first objective is to find detection algorithms that are noise uncertainty free and to analyze the detection performance of the resulting detectors. This goal has been accomplished in Publication II and Publication III where **Fact 1** was taken into account.

The second objective is to design and analyze detectors that could incorporate **Fact 2** in addition to **Fact 1**, i.e. detectors that are noise uncertainty free and, at the same time, could work in multiple primary user scenarios. This aim has been achieved in in Publication IV, Publication V, Publication VI and Publication VII, where several detectors were proposed and studied. The last objective is to design and examine detection algorithms that will take into account all the three **facts**. This leads to the most robust detector under the framework developed in this thesis, and this final goal has been accomplished in Publication VIII. Note that the detector studied in Publication I has not taken any of the practical issues into account.

As far as the technical and system modeling aspects are concerned, the scope of this thesis is limited by the following considerations and assumptions. First of all, this thesis relies on the framework of statistical hypothesis testing theory based on multivariate complex Gaussian distributions. Specifically, the signals of the primary users as well as the noise at the secondary receiver are assumed to follow the multivariate Gaussian models. Secondly, among the vast number of performance metrics of spectrum sensing algorithms, this thesis is limited only to the analytical derivations of the false alarm probability, the detection probability, and the receiver operating characteristic. Furthermore, we focus on deriving closed-form performance metrics for a given channel realizations i.e. deterministic channels. Analytically characterizing the average detection probability or the average receiver operating characteristic over channel statistics is beyond the scope of this thesis. Finally, we mention that other practical issues such as the non-whiteness of the receiver primary user's spectrum, the multipath channel and over-sampling at the receiver are ignored in the framework of the thesis.

1.3 Contributions

The contributions of this thesis are two-fold: the technical contributions and the contributions to the literature of spectrum sensing. Technically, we contributed to the field of mathematical statistics by deriving the distributions of various test statistics built on the Wishart-Laguerre ensemble as well as the Jacobi ensemble. For some test statistics, we were able to deduce the exact distribution functions, while for others we proposed simple approximative distributions. The derived approximative distributions are crucial in understanding the behaviors of test statistics in certain asymptotic regimes whereas the exact distributions are useful for the calculations of the distributions in small dimensions.

Besides the distributional results, which consist the main contributions of this thesis, we also proposed a unified framework for the construction of robust multi-antenna detection algorithms. This is the second contribution of this thesis. Specifically, we systematically categorize detection algorithms according to the parameters they are blind to. Detection without assuming any knowledge of a certain parameter is called blind detection. The choice of these system parameters is motivated by practical issues and the development of cognitive radio concepts. For example, when secondary usage is expanded from TV secondary systems to cellular secondary networks, the concept of spectrum sensing in the presence of multiple primary users emerges. Among the conceptual contributions, our proposal of the multi-antenna based multiple primary user detection is, to the best of the author's knowledge, the first in the spectrum sensing literature. Note that the technical contributions of this thesis are applicable to other cognitive radio concepts such as spectrum awareness and spectrum management. The derived analytical results are also useful in performance analysis of certain communications systems employing multiple antennas.

1.4 Summary of publications

This thesis consists of an introductory part and eight original publications. The content of each publication is summarized as follows.

In Publication I we study the performance of the largest eigenvalue based detection. The exact false alarm and detection probabilities have been derived. It is shown that the largest eigenvalue based detector outperforms the cooperative energy detector in the presence of a single primary user.

In Publication II we examine the asymptotical performance of the scaled largest eigenvalue based detection when both the number of sensors and samples go to infinity. Specifically, an accurate and simple closed-form approximation to the false alarm probability has been derived. The scaled largest eigenvalue based detector does not depend on the noise power, and is thus free of the noise uncertainty. The technical results of this publication improve the accuracy of the false alarm probability estimation compared to the existing results in literature.

In Publication III the scaled largest eigenvalue is considered again, and its exact false alarm probability is calculated. The derived result involves only a finite sum of polynomials and its evaluation is affordable in practice. The result of this publication fills the gap in understanding the exact false alarm probability of the scaled largest eigenvalue based detection.

In Publication IV we proposed the idea of spectrum sensing in the presence of multiple but unknown number of primary users. The corresponding detector derived under the generalized likelihood ratio criterion turns out to be based on the spherical test. This detector is not only blind to the number of primary users but it is also blind to the noise power. We derived accurate analytical approximations to its false alarm probability, detection probability, and receiver operating characteristic. These derived results yield almost-exact fits to the simulations. Comprehensive performance comparisons show the superior performance of the spherical test based detector in the presence of noise uncertainty and multiple primary users.

In Publication V we have considered a similar setting as in Publication IV, where there exists more that one active primary users and the secondary receiver shall be unaware of the noise power. The difference is that here we are particularly interested in low SNR detection. The resulting test statistics is the so-called John's test, which was derived under the locally best invariant criterion. Key performance metrics of John's detector, such as the false alarm probability, the detection probability as well as the receiver operating characteristic, have been derived. Characterization of the distributions of John's detector was an open problem for applications in the science of statistics. Numerical results show performance gain of John's detector over the spherical test based detector in scenarios with relatively low SNR.

In Publication VI we study the performance of a heuristic detection algorithm based on the eigenvalue ratio. We proposed an approximation framework, under which both the finite dimensional and the asymptotic detection probabilities of the eigenvalue ratio based detector have been deduced. This detector is also blind to the number of primary users and noise power. The achieved accuracy of the proposed approximations is reasonably good.

In Publication VII we consider another heuristic multiple-primary-user detector that is free of noise uncertainty. This detector is based on the

Demmel condition number, and we derived its exact false alarm probability. The distribution of the Demmel condition number was an open problem in the literature on statistics.

In Publication VIII we extend the concept of robustness in the literature of spectrum sensing to a new direction by considering detectors that are blind to the noise covariance matrix. The resulting detector derived from the generalized likelihood ratio criterion is Wilks' test. In addition to the property of being noise uncertainty free and being able to work in the presence of multiple primary users, the proposed Wilks' detector is also blind to the degree of noise correlations. We analyze the performance of Wilks' detector by deriving an accurate approximation to its false alarm probability. Performance gain of Wilks' detector over the existing detection algorithms is observed in scenarios with arbitrary but unknown noise correlation and multiple primary users.

2. Multi-antenna Spectrum Sensing

We map the multi-antenna spectrum sensing problem to a binary hypothesis testing problem in Chapter 2.1. Based on the formulated hypothesis test, in Chapter 2.2 we outline the spectrum sensing algorithms, i.e. test statistics, considered in this thesis. Specifically, we consider test statistics derived from decision-theoretic criterions such as the generalized likelihood ratio criterion and the locally best invariant criterion, as well as test statistics that were proposed heuristically. Both types of test statistics were constructed in the literature with different assumptions and different degree of knowledge of the parameters of the primary and/or the secondary systems.

2.1 Signal model

Consider a scenario of K collaborating sensors in a secondary network trying to detect the presence of primary users' transmission. The K sensors may be e.g. K receive antennas in one secondary device or K secondary devices each with a single antenna, or any combination of these. However, this collaborative sensing scenario and the subsequent formulations are more relevant when the K sensors are in one device. For distributed collaborating sensors, accurate time synchronization between devices, and communications to the fusion center become an issue given the limited capabilities of the individual sensor. Henceforth, this collaborative spectrum sensing scenario may be referred to as multi-antenna spectrum sensing.

The standard model for K-sensor cooperative detection in the presence of possible P primary users reads

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{2.1}$$

where the K dimensional complex vector $\mathbf{x} \in \mathbb{C}^K$ is the received data vector. The $K \times 1$ vector $\mathbf{n}$ is the complex Gaussian noise with zero mean and covariance matrix $\mathbf{\Psi}$. The $K \times P$ matrix $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_P]$ represents the channels between the P primary users and the K sensors. The $P \times 1$ vector $\mathbf{s} = [s_1, \ldots, s_P]'$ denotes the transmitted signals from the primary users. It follows an independent and identically distributed (i.i.d) zero mean Gaussian distribution and is uncorrelated with the noise. As noted in Chapter 1.2, the focus of this thesis is performance analysis for a given channel realizations and thus the channel matrix $\mathbf{H}$ is assumed to be constant during sensing i.e. we consider deterministic channels.

We collect N i.i.d observations from model (2.1) to a $K \times N$ $(K \leq N)$ received data matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$. By the above assumptions, the sample covariance matrix $\mathbf{R} = \mathbf{X}\mathbf{X}^\dagger$ of the received data matrix follows a complex Wishart distribution of dimension K with N degrees of freedom and a population covariance matrix $\mathbf{\Sigma}$, denoted by $\mathbf{R} \sim \mathcal{W}_K(N, \mathbf{\Sigma})$, with the density function

$$\frac{|\mathbf{\Sigma}|^{-N}}{\Gamma_K(N)} |\mathbf{R}|^{N-K} \, \mathrm{e}^{-\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{R}\right)}, \tag{2.2}$$

where $(\cdot)^\dagger$, $|\cdot|$ and $\mathrm{tr}(\cdot)$ denote the conjugate-transpose, matrix determinant and matrix trace operations, respectively. Here the function $\Gamma_K(N)$ defines the multivariate Gamma function

$$\Gamma_K(N) = \pi^{\frac{1}{2}K(K-1)} \prod_{j=0}^{K-1} \Gamma(N-j), \tag{2.3}$$

where $\Gamma(\cdot)$ is the Gamma function.

We consider a binary hypothesis test i.e. either the presence or absence of primary users. The secondary user will not infer the number of active primary users. Hypothesis $\mathcal{H}_0$ is that there are no primary users. The corresponding population covariance matrix is

$$\mathcal{H}_0: \quad \mathbf{\Sigma} := \mathbb{E}[\mathbf{X}\mathbf{X}^\dagger]/N = \mathbf{\Psi}, \tag{2.4}$$

where $\mathbb{E}[\cdot]$ denotes expectation. Hypothesis $\mathcal{H}_1$ is that primary users are present, and the population covariance matrix is

$$\mathcal{H}_1: \quad \mathbf{\Sigma} = \mathbf{\Psi} + \sum_{i=1}^{P} \gamma_i \mathbf{h}_i \mathbf{h}_i^\dagger, \tag{2.5}$$

where $\gamma_i := \mathbb{E}[s_i s_i^\dagger]$ defines the transmission power of the i-th primary user. The received SNR of primary user i across the K sensors is defined as

$$\mathrm{SNR}_i := \frac{\gamma_i ||\mathbf{h}_i||^2}{\mathrm{tr}\left(\mathbf{\Psi}\right)/K}, \tag{2.6}$$

where $||\cdot||$ is the Euclidean norm. Note that if the transmitted signals s are constant instead of Gaussian distributed, then the hypothesis $\mathcal{H}_1$, instead of (2.5), is modeled as a non-central Wishart distribution [12]

$$\mathbf{R} \sim \mathcal{W}_K(N, \mathbf{\Psi}, \mathbf{\Xi}) \tag{2.7}$$

instead of the model (2.5), where

$$\mathbf{\Xi} = \mathbf{Hss}^\dagger \mathbf{H}^\dagger \tag{2.8}$$

is the so-called non-central parameter matrix. In [PI] and [PVI] we consider this non-central Wishart model. Finally, we note that declaring wrongly $\mathcal{H}_0$ defines the false alarm probability P_{fa}, and declaring correctly $\mathcal{H}_1$ defines the detection probability P_{d}. For a given detector, the relation between its P_{fa} and P_{d} is called the Receiver Operating Characteristic (ROC), the plot of which shows its overall detection performance.

2.2 Test statistics

By exploring the differences between the population covariance matrices under $\mathcal{H}_0$ (2.4) and under $\mathcal{H}_1$ (2.5), the presence or absence of primary users can be decided. With different assumptions on or the knowledge of the system parameters e.g. the noise covariance matrix $\mathbf{\Psi}$ or the number of primary users P, various test statistics have been proposed in literature. In the following we classify these test statistics according to whether the noise covariance matrix is assumed to be known or not.

2.2.1 Known noise covariance matrix

Since the noise covariance matrix $\mathbf{\Psi}$ is known, without loss of generality [13, pp. 338], we assume that the noise of each sensor is independent and has a common noise power σ^2 i.e.

$$\mathbf{\Psi} = \sigma^2 \mathbf{I}_K, \tag{2.9}$$

where $\mathbf{I}_K$ defines a K dimensional identity matrix. In this case, the sufficient statistics is the sample covariance matrix $\mathbf{R}$ of the received data matrix [13]. We denote its ordered eigenvalues by $0 \leq \lambda_K \leq \ldots \leq \lambda_1 < \infty$. For arbitrary but *known* $\mathbf{\Psi}$, the sufficient statistics becomes the 'whitened' sample covariance matrix $\mathbf{\Psi}^{-1}\mathbf{R}$.

In the presence of *a single* primary user, $P = 1$, the hypotheses (2.4) and

(2.5) are reduced to

$$\mathcal{H}_0 \quad : \quad \mathbf{\Sigma} = \sigma^2 \mathbf{I}_K \tag{2.10a}$$

$$\mathcal{H}_1 \quad : \quad \mathbf{\Sigma} = \sigma^2 \mathbf{I}_K + \gamma_1 \mathbf{h}_1 \mathbf{h}_1^{\dagger}. \tag{2.10b}$$

Under this hypothesis test we may further assume that the noise power σ^2 is known, and without loss of generality set at $\sigma^2 = 1$. In this case, the Largest Eigenvalue based (LE) detector

$$T_{\mathrm{LE}} := \lambda_1 \in [0, \infty) \tag{2.11}$$

was derived under the Generalized Likelihood Ratio (GLR) criterion [14, Sec. III-C]. Comparing the test statistics with a predetermined threshold ζ, the presence or absence of the primary users is decided. For the LE detector the test procedure is formally written as

$$T_{\mathrm{LE}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \zeta. \tag{2.12}$$

Performance analysis of the LE detector can be found e.g. in [14, 15] and [PI]. The assumption of known noise power, besides being impractical, leads to detectors which may suffer severe performance degradation [16] [PIV] due to noise power uncertainty [10]. Assuming an unknown noise power σ^2, the test in the GLR sense is Scaled Largest Eigenvalue based (SLE) detection

$$T_{\mathrm{SLE}} := \frac{\lambda_1}{\frac{1}{K}\sum_{i=1}^{K} \lambda_i} \in [1, K], \tag{2.13}$$

with the corresponding test procedure being

$$T_{\mathrm{SLE}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \zeta. \tag{2.14}$$

The SLE detector was first proposed in the context of spectrum sensing in [17] and further analyzed in [18, 19, 20], [PII, PIII]. Detection without assuming any knowledge of a certain parameter is often called blind detection. For example the SLE detector is a blind σ^2 detection, and is more robust than the LE detector to noise power uncertainty.

In the presence of *multiple* primary users, both the LE and SLE detectors are expected to suffer performance loss. To formulate a hypothesis test in the setting of multiple primary users, one needs to consider the fact that for a secondary user the most critical information is whether or not there are active primary users. The knowledge of the number of active primary users may not be relevant from the secondary user's perspective. It is worth noting that the matrix $\sum_{i=1}^{P} \gamma_i \mathbf{h}_i \mathbf{h}_i^{\dagger}$ in (2.5) is positive definite, i.e. $\sum_{i=1}^{P} \gamma_i \mathbf{h}_i \mathbf{h}_i^\dagger \succ \mathbf{0}$, if there is at least one primary user. A blind P detection for the presence of possibly multiple primary users would thus be a hypothesis test expressed as

$$\mathcal{H}_0 \quad : \quad \mathbf{\Sigma} = \sigma^2 \mathbf{I}_K \tag{2.15a}$$

$$\mathcal{H}_1 \quad : \quad \mathbf{\Sigma} \succ \sigma^2 \mathbf{I}_K. \tag{2.15b}$$

Under this hypothesis test, the corresponding detector derived from GLR criterion is the Spherical Test based (ST) detector

$$T_{\mathrm{ST}} := \frac{|\mathbf{R}|}{\left(\frac{1}{K}\mathrm{tr}(\mathbf{R})\right)^K} = \frac{\prod_{i=1}^{K} \lambda_i}{\left(\frac{1}{K}\sum_{i=1}^{K} \lambda_i\right)^K} \in [0, 1], \tag{2.16}$$

which also is a blind σ^2 detection [21]. The test procedure of the ST detector is

$$T_{\mathrm{ST}} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \zeta. \tag{2.17}$$

The spherical test was formulated in [21] as a spectrum sensing algorithm and the detection performance has been analytically addressed in [PIV]. Although in general the ST detector achieves good performance, it is not the best one in the low SNR regime. A test statistics that is optimal in detecting small deviations from $\mathcal{H}_0$ is John's detector [22, 23]

$$T_{\mathrm{J}} := \frac{\mathrm{tr}(\mathbf{R}^2)}{\left(\mathrm{tr}(\mathbf{R})\right)^2} = \frac{\sum_{i=1}^{K} \lambda_i^2}{\left(\sum_{i=1}^{K} \lambda_i\right)^2} \in \left[\frac{1}{K}, 1\right], \tag{2.18}$$

with the test procedure being

$$T_{\mathrm{J}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \zeta. \tag{2.19}$$

John's test was studied in the context of spectrum sensing in [PV]. The criterion under which John's detector is derived is known as the Locally Best Invariant (LBI) criterion. Unlike the GLR criterion, the LBI criterion often leads to detectors that perform particularly well in the low SNR regime.

Besides the ST and John's detectors, other existing blind P detectors that can be considered in the setting of multiple primary users (2.15) include the Eigenvalue Ratio based (ER) detector [24, 16, 25, 26], [PVI],

$$T_{\mathrm{ER}} := \frac{\lambda_1}{\lambda_K} \in [1, \infty), \tag{2.20}$$

as well as the Demmel Condition Number based (DCN) detector [27, 28], [PVII],

$$T_{\mathrm{DCN}} := \frac{\sum_{i=1}^{K} \lambda_i}{\lambda_K} \in [K, \infty], \tag{2.21}$$

with the test procedures being

$$T_{\text{ER}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \zeta, \tag{2.22}$$

and

$$T_{\text{DCN}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \zeta, \tag{2.23}$$

respectively. Neither the ER nor DCN detector is constructed from decision-theoretic consideration, such as the GLR or LBI criterion. It turns out that they achieve substantially worse performance than the ST and John's detectors in both single and multiple primary user scenarios [PIV]. Recently, the ER and DCN detectors have been considered in frequency selective channels [29]. Note that all considered blind P detectors, the ST, John's, the ER, and the DCN detectors, also are blind σ^2 detectors, which are robust not only to the number of primary users but also the noise power uncertainty.

To complete the story, we note that the cooperative Energy Detector derived from the LBI criterion [22]

$$T_{\text{ED}} := \mathbf{tr}(\mathbf{R}) = \sum_{i=1}^{K} \lambda_i \in [0, \infty), \tag{2.24}$$

which assumes σ^2 to be known, is also a blind P detector. The cooperative energy detector is often used as a benchmark detector for performance comparisons [18, 16], [PI, PIV], whose performance is considerably degraded by a relatively small noise power uncertainty. Finally, we note that for an arbitrary but known P, the corresponding GLR detectors have been derived in [30].

We emphasize that for the case of arbitrary but *known* $\mathbf{\Psi}$, the forms of the test statistics (2.11), (2.13), (2.16), (2.18), (2.20), (2.21), and (2.24) considered in this sub-chapter remain the same and are directly applicable. The only difference is that these test statistics are now functions of $\mathbf{\Psi}^{-1}\mathbf{R}$ instead of $\mathbf{R}$.

2.2.2 Arbitrary but unknown noise covariance matrix

The blindness of detection can be extended to a new dimension by assuming that the noise population covariance matrix $\mathbf{\Psi}$ is arbitrary and unknown. The resulting blind $\mathbf{\Psi}$ detectors are robust to any modeling assumptions on $\mathbf{\Psi}$. The concept of blindness here is different from blindness to σ^2 or P, since the knowledge of $\mathbf{\Psi}$ will be found in noise-only samples. In the context of this thesis, blind $\mathbf{\Psi}$ refers to the fact that no artificial structure is imposed on $\mathbf{\Psi}$ in contrast to, for instance, the one parameter model $\mathbf{\Psi} = \sigma^2 \mathbf{I}_K$. This extension is partially motivated by the existence of usually unknown noise correlation due to e.g. antenna coupling in practical systems. Moreover, the existing non-blind $\mathbf{\Psi}$ detectors often suffer severe performance loss even for a low degree of noise correlation as will be shown in Chapter 4. Instead of a perfectly known $\mathbf{\Psi}$, here we assume to have, in addition to the received data matrix $\mathbf{X}$, another independent noise-only observation matrix $\mathbf{Z}$ consisting of M samples from the K sensors. This noise-only observation matrix $\mathbf{Z}$ can be obtained e.g. when absence of the primary users is declared from an initial coarse sensing period. Moreover, when the signals of interest are narrow-band and located in a known frequency band, such as the case of TV primary systems, the noise-only samples collected at a frequency just outside this band can be justified as having the same noise covariance characteristics. The time-varying nature of the noise correlation is coped with by periodically updating the measurement $\mathbf{Z}$. The true but unknown noise population covariance matrix $\mathbf{\Psi}$ can be estimated via the noise-only sample covariance matrix $\mathbf{E} = \mathbf{Z}\mathbf{Z}^\dagger$, which, by the assumptions in Chapter 2.1, follows a complex Wishart distribution $\mathbf{E} \sim \mathcal{W}_K(M, \mathbf{\Psi})$. In this setting, the sufficient statistics is the sample covariance matrix of the form $\mathbf{E}^{-1}\mathbf{R}$ [13], and its ordered eigenvalues are denoted by $0 \le \theta_K \le \ldots \le \theta_1 < \infty$.

For single primary user detection the corresponding hypothesis test is

$$\mathcal{H}_0 \quad : \quad \mathbf{\Sigma} = \mathbf{\Psi} \tag{2.25a}$$

$$\mathcal{H}_1 \quad : \quad \mathbf{\Sigma} = \mathbf{\Psi} + \gamma_1 \mathbf{h}_1 \mathbf{h}_1^\dagger. \tag{2.25b}$$

Essentially we are testing the equality of population covariance matrices $\mathbf{\Sigma}$ and $\mathbf{\Psi}$ against a rank-1 perturbation alternative (2.25b) based on the received data and noise-only observation matrices $\mathbf{X}$ and $\mathbf{Z}$. Under this hypothesis test, a reasonable test statistics to choose is Roy's largest eigenvalue based detector [31]

$$T_{\mathrm{R}} := \theta_1 \in [0, \infty). \tag{2.26}$$

Nadler et al. [32, 33] were among the first to consider Roy's detector in a spectrum sensing application, and derived novel analytical expressions for the detection probability.

Although Roy's detector is a blind $\mathbf{\Psi}$ detector, it is not a blind P detector. Namely, when the actual number of primary users is more than one, Roy's detector will suffer performance loss. We now try to extend the blindness of the detection to the practical scenario of multiple primary users. Following the same line of reasoning as in (2.15), with an arbitrary and unknown noise covariance matrix the hypothesis test in the presence of multiple primary users, $P \geq 2$ but not known a priori, is

$$\mathcal{H}_0 \quad : \quad \mathbf{\Sigma} = \mathbf{\Psi} \tag{2.27a}$$

$$\mathcal{H}_1 \quad : \quad \mathbf{\Sigma} \succ \mathbf{\Psi}. \tag{2.27b}$$

For this hypothesis test, it turns out that the detector derived from the GLR criterion is

$$T_{\mathrm{W}} := \frac{|\mathbf{E}|}{|\mathbf{R}+\mathbf{E}|} = \prod_{i=1}^{K} \frac{1}{1+\theta_i} \in [0,1], \tag{2.28}$$

with test procedure being

$$T_{\mathrm{W}} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \zeta. \tag{2.29}$$

This test statistics was derived by S. S. Wilks [34] and its performance analysis in spectrum sensing application can be found in [PVIII]. Wilks' detector is blind to both $\mathbf{\Psi}$ and P i.e. its performance is robust to the noise power, the degree of noise correlation as well as to the number of primary users, which renders it the most robust detector under the framework developed in this thesis. Simulations performed in Chapter 4 show that Wilks' detector indeed outperforms other known detectors in realistic scenarios with multiple primary users and arbitrary but unknown noise correlation.

2.3 Summary: towards a unified look

The test statistics discussed in this chapter can be categorized in various ways according to different criterions. For example, the criterion can be based on the knowledge of the number of primary users or the noise covariance matrix. It is also possible to divide these tests according to whether it is constructed from decision-theoretic criterion or proposed heuristically. The development of this thesis is towards robust spectrum sensing in realistic cognitive radio scenarios. Here, robustness refers to the property that the detection performance is insensitive to deviations from the presumed modeling assumptions. The emphasis on detection algorithms derived from well-established decision-theoretic criterions such as the GLR or LBI is due to the fact that these detectors are more likely to outperform the heuristic ones.

Table 2.1. Summary of Multi-antenna Spectrum Sensing Algorithms (the performance of test statistics in blue color have been studied in this thesis.)

assumptions	single primary user	multiple primary users
$\mathbf{\Psi} = \sigma^2 \mathbf{I}_K$ with known σ^2	T_{LE}	T_{ED}
$\mathbf{\Psi} = \sigma^2 \mathbf{I}_K$ with unknown σ^2	T_{SLE}	T_{ST}, T_{J}, T_{ER}, T_{DCN}
arbitrary and unknown $\mathbf{\Psi}$	T_{R}	T_{W}

By removing the assumptions of known noise power, presence of single active primary user, and known noise covariance matrix, the resulting Wilks' detector derived from the GLR criterion turns out to be the most robust detector under the framework of this thesis. However, this holds only if the listed three facts in Chapter 1.1 are true simultaneously. If one of these facts is violated, some other detector may perform better than Wilks' detector. For example, consider a scenario of a secondary user trying to detect the presence of a possible TV transmission in the bandwidth of a certain TV channel at a specific location. In this scenario the assumption of a single active primary user is justifiable, and as a result Roy's test is preferable to Wilks'.

For convenience, the detectors considered in this chapter are summarized in Table 2.1 according to the modeling assumptions. Some remarks and discussions regarding this table are in the following:

- T_{LE} is constructed without taking into account any of the facts in Chapter 1.1; T_{SLE} is obtained by considering only **Fact 1**; T_{ED} is obtained by considering only **Fact 2**; T_{ST}, T_{J}, T_{ER} and T_{DCN} are obtained by considering both **Fact 1** and **Fact 2**; T_{R} is obtained by considering both **Fact 1** and **Fact 3**; T_{R} is derived by considering all the three facts.

- The performance of the test statistics in blue color have been analyzed, and is an essential contribution of this thesis. Specifically, Publication I is devoted to P_{fa} and P_{d} of T_{LE}; Publication II and Publication III cover P_{fa} of T_{SLE}; Publication IV is devoted to P_{fa} and P_{d} of T_{ST}; Publication V is devoted to P_{fa} and P_{d} of T_{J}; Publication VI is devoted to P_{d} of T_{ER}; Publication VII is devoted to P_{fa} and of T_{DCN}; Publication VIII is devoted to P_{fa} of T_{W}.

- Classifying the considered test statistics by the methods of construction, we have

  - GLR criterion: T_{LE}, T_{SLE}, T_{ST}, T_{W}.

  - LBI criterion: T_{ED}, T_{J}.

  - heuristic: T_{ER}, T_{DCN}, T_{R}.

  For heuristic detectors, their positions in Table 2.1 can be arbitrary to some extent. In fact, the test statistics T_{ER}, T_{DCN} were initially proposed for spectrum sensing application in the context of single-primary-user detection [24, 27].

- Implementation of the considered detectors may require numerical calculations of eigenvalue decomposition, matrix determinant and trace. As far as the computational complexity is concerned, the considered detectors can be divided into two categories according to whether eigenvalue decomposition is needed. This is because eigenvalue decomposition needs to be performed in an iterative manner and may have stability issues.

  - Detectors that require eigenvalue decomposition:
    T_{LE}, T_{SLE}, T_{ER}, T_{DCN}, T_{R}.

  - Detectors that does not require eigenvalue decomposition:
    T_{ED}, T_{ST}, T_{J}, T_{W}.

  In the first category, the detectors T_{LE}, T_{SLE} and T_{DCN} have similar computational complexity whereas T_{DCN} requires roughly twice the computing power as one needs to calculate both the smallest and the largest eigenvalues. Besides eigenvalue decomposition, Roy's detector T_{R} involves numerically inverting the noise sample covariance matrix, which becomes unstable for strongly correlated noise. As such T_{R} requires the highest computing power.

3. Distributions of Test Statistics

In order to analytically characterize the performance of the detectors discussed in Chapter 2, distributions of their test statistics under the hypotheses are needed. Specifically, closed-form expressions of test statistic distributions under hypothesis $\mathcal{H}_0$ and hypothesis $\mathcal{H}_1$ lead to closed-form false alarm probability and detection probability, respectively. With expressions for both false alarm and detection probabilities of a given detector, its receiver operating characteristic can be found. In this chapter we focus on presenting the main contribution of this thesis – the results for the distributions of the test statistics of interest, in a uniform manner. In Chapter 3.1, we outline the mathematical tools utilized to obtain these results. In Chapter 3.2, closed-form expressions for the test statistics distributions are presented for each considered detector, with sketches of proofs and the set of used tools specified.

3.1 Main mathematical tools

The order of the presentation of the following mathematical tools is determined by the usefulness in the context of this thesis. Other methodologies that are useful only in some specific problem settings will be briefly mentioned. We only focus on the knowledge that is most relevant of this thesis, the actual content of each of the following subjects is much more broader.

3.1.1 Multivariate analysis

Multivariate analysis mainly deals with the analysis of distributions and tests based on multivariate Gaussian distributions. Classical text books in multivariate analysis include [13, 35], where real matrices were dealt with. For complex matrices, A. T. James' seminal paper [36] contains a comprehensive survey. These analyses are primarily concerned with sample covariance matrices of finite dimensions. This subject dates back to as early as 1928, where the exact (finite size) distribution of the sample covariance matrix was derived as what we know today the Wishart distribution [37]. One of the key techniques in deriving the exact distributions at finite size is the calculation of Jacobians of matrix transforms. For systematical knowledge in this direction, we refer to [38]. In the context of this thesis, techniques from multivariate analysis are used to obtain test statistics distributions as well as moments expressions of finite sizes.

As an example, we show in the following the derivation of the complex Wishart distribution, which is the central model throughout this thesis. For a $K \times N$ complex Gaussian matrix $\mathbf{X}$ with entries follow an i.i.d standard complex Gaussian distribution, the joint density is given by

$$p_{\mathbf{X}}(\mathbf{X}) = \frac{1}{\pi^{KN}} \mathrm{e}^{-\mathrm{tr}(\mathbf{X}\mathbf{X}^{\dagger})}. \tag{3.1}$$

The corresponding complex Wishart matrix equals $\mathbf{R} = \mathbf{X}\mathbf{X}^{\dagger}$. Consider the transformations

$$\mathbf{R} = \mathbf{T}\mathbf{T}^{\dagger}, \qquad \mathbf{X} = \mathbf{T}\mathbf{U}, \tag{3.2}$$

where $\mathbf{T}$ is a $K \times K$ lower triangular matrix with real positive diagonal elements t_{ii}, and $\mathbf{U}$ is a $K \times N$ semiunitary matrix, $\mathbf{U}\mathbf{U}^{\dagger} = \mathbf{I}_K$. The Jacobians of these transforms are given in [38, Th. 3.7] and [38, Th. 4.5] as

$$\mathrm{d}\mathbf{R} = 2^K \left(\prod_{i=1}^{K} t_{ii}^{2(K-i)+1} \right) \mathrm{d}\mathbf{T} \tag{3.3}$$

and

$$\mathrm{d}\mathbf{X} = \left(\prod_{i=1}^{K} t_{ii}^{2(N-i)+1} \right) \mathrm{d}\mathbf{T}\mathrm{d}\mathbf{G} \tag{3.4}$$

respectively. In (3.4), $\mathrm{d}\mathbf{G}$ denotes the Haar measure, the total volume of the space of semiunitary matrices $O_{K,N}$ equals [38, Coroll. 4.5.2]

$$\int_{O_{K,N}} \mathrm{d}\mathbf{G} = \frac{2^K \pi^{KN}}{\Gamma_K(N)}, \tag{3.5}$$

where $\Gamma_K(N)$ has been defined in (2.3). As a result, we have

$$\int_{O_{K,N}} \mathrm{d}\mathbf{X} = \left(\frac{\pi^{KN}}{\Gamma_K(N)} \prod_{i=1}^{K} t_{ii}^{2(N-K)} \right) \mathrm{d}\mathbf{R}, \tag{3.6}$$

and using the fact that

$$|\mathbf{R}| = |\mathbf{T}|^2 = \prod_{i=1}^{K} t_{ii}^2, \tag{3.7}$$

the density of $\mathbf{R}$ is obtained as

$$p_{\mathbf{R}}(\mathbf{R}) = \left(\frac{\pi^{KN}}{\Gamma_K(N)} |\mathbf{R}|^{N-K} \right) p_{\mathbf{X}}(\mathbf{X}) \bigg|_{\mathbf{X}\mathbf{X}^\dagger = \mathbf{R}} \tag{3.8}$$

$$= \frac{1}{\Gamma_K(N)} |\mathbf{R}|^{N-K} \mathrm{e}^{-\mathbf{tr}(\mathbf{R})}. \tag{3.9}$$

We mention in the following two auxiliary tools that were only utilized in [PV] for calculations of moments of the test statistics under $\mathcal{H}_1$. The first one is the combinatorial structure of moments of complex Wishart distribution. For a complex Wishart matrix $\mathbf{R} \sim \mathcal{W}_K(N, \mathbf{\Sigma})$ as defined in (2.2), the n-th moment of $\mathbf{R}$, by definition, equals

$$\mathbb{E}\left[\prod_{k=1}^{n} \mathbf{R}_{i_k, j_k} \right] = \frac{\partial \Omega(\mathbf{\Theta})}{\partial \theta_{j_1,i_1} \partial \theta_{j_2,i_2} \dots \partial \theta_{j_n,i_n}} \bigg|_{\mathbf{\Theta}=\mathbf{0}}, \tag{3.10}$$

where $\Omega(\mathbf{\Theta}) := \mathbb{E}\left[\mathrm{e}^{\mathbf{tr}(\mathbf{R}\mathbf{\Theta})}\right] = |\mathbf{\Sigma}|^{-N} |\mathbf{\Sigma}^{-1} - \mathbf{\Theta}|^{-N}$ denotes the moment generating function of $\mathbf{R}$ with the i, j-th entry of the Hermitian parameter matrix $\mathbf{\Theta}$ being $\theta_{i,j}$. It was proven in [39] via representation theory for symmetric group that (3.10) admits the following structure

$$\frac{\partial \Omega(\mathbf{\Theta})}{\partial \theta_{j_1,i_1} \partial \theta_{j_2,i_2} \dots \partial \theta_{j_n,i_n}} \bigg|_{\mathbf{\Theta}=\mathbf{0}} = \sum_{i=1}^{n} N^{n-i+1} B_{i-1}, \tag{3.11}$$

where the scalar B_i is a sum of *distinct* terms of the form $\prod_{k=1}^{n} \mathbf{\Sigma}_{i_k, j_{\pi(k)}}$ with an index distance i, where π defines a permutation of integers $1, \dots, n$. Here, the index distance is defined as the minimum index permutations (restricted to row-to-row or column-to-column permutations) required such that a term $\prod_{k=1}^{n} \mathbf{\Sigma}_{i_k, j_{\pi(k)}}$ is permutated to the canonical form $\prod_{k=1}^{n} \mathbf{\Sigma}_{i_k, j_k}$. For example, the index distance of the term $\mathbf{\Sigma}_{i_1,j_3}\mathbf{\Sigma}_{i_2,j_4}\mathbf{\Sigma}_{i_3,j_2}\mathbf{\Sigma}_{i_4,j_1}$ is 3. This combinatorial structure (3.11) was first observed in [40], where up to the 4-th moment of $\mathbf{R}$ were calculated.

The other tool we would like to mention is the so-called delta-method [41], which is a standard technique for estimation of statistics of functions of random variables, especially the mean and variance. The setting relevant to this thesis is the following. Consider random variables X and Y that in general are correlated. The bi-variate Taylor series expansion of a function $z(x, y)$ of X and Y about their mean values μ_x and μ_y is

$$\begin{aligned} z(x,y) = {} & z(\mu_x, \mu_y) + \left(\frac{\partial z}{\partial x} \bigg|_{x=\mu_x, y=\mu_y} \right) (x - \mu_x) + \left(\frac{\partial z}{\partial y} \bigg|_{x=\mu_x, y=\mu_y} \right) (y - \mu_y) + \\ & \frac{1}{2!} \left(\left(\frac{\partial^2 z}{\partial x^2} \bigg|_{x=\mu_x, y=\mu_y} \right) (x - \mu_x)^2 + \left(\frac{\partial^2 z}{\partial y^2} \bigg|_{x=\mu_x, y=\mu_y} \right) (y - \mu_y)^2 + \right. \\ & \left. \left(\frac{\partial^2 z}{\partial x \partial y} \bigg|_{x=\mu_x, y=\mu_y} \right) 2(x - \mu_x)(y - \mu_y) \right) + \cdots . \end{aligned} \tag{3.12}$$

If the moments and the joint moments of X and Y are known, an estimate of the mean of $z(x, y)$ is obtained by taking the expectation of (3.12) as

$$\mathbb{E}\left[z(x,y)\right] \approx z(\mu_x,\mu_y) + \frac{1}{2!}\left(\left(\left.\frac{\partial^2 z}{\partial x^2}\right|_{x=\mu_x,y=\mu_y}\right)\mathbb{E}\left[(x-\mu_x)^2\right] + \left(\left.\frac{\partial^2 z}{\partial y^2}\right|_{x=\mu_x,y=\mu_y}\right)\mathbb{E}\left[(y-\mu_y)^2\right] + \left(\left.\frac{\partial^2 z}{\partial x \partial y}\right|_{x=\mu_x,y=\mu_y}\right)2\mathbb{E}\left[(x-\mu_x)(y-\mu_y)\right]\right). \tag{3.13}$$

Estimates of high moments can be performed in a similar manner. From above formulations, it is clear that the delta-method relies on the assumption that the Taylor expansion around the mean values is effectively linear.

3.1.2 Moment based approximation

Moment based approximation is a set of techniques that are utilized to approximate the distribution of random variables based on their analytical moment expressions. Moment based approximation is very useful in practice since closed-form moments are often analytically more tractable than exact distributions. In the statistical literature, various types of moment based approximation have been proposed. For example, a family of Pearson curves can be used to approximate unknown distributions, where up to the 4-th moment are needed [42]. The drawback of this approach is that both the number of candidate curves and the moments are limited, leading to non-trivial approximation errors in some cases. Alternatively, the saddlepoint method [43] is also applicable, which is rather accurate especially in the tail of the distribution. However, the implementation of this method requires numerically solving the saddlepoint equations, which may become computationally intensive. A systematic approach to moment based approximations is to use orthogonal polynomials, which will be explained in some detail below.

We consider to approximate an unknown density $p(x)$ of random variable X by using up to its n-th moment. These moments are assumed to exist, and the approximating function is denoted by $p_n(x)$. Define a set of monic orthogonal polynomials of degree n,

$$f_n(x) = \sum_{k=0}^{n} d_{n,k}x^k, \quad d_{n,n} = 1, \tag{3.14}$$

which are orthogonal with respect to a weight function $w(x)$ supported in

$x \in [a, b]$ as

$$\int_a^b w(x) f_n(x) f_m(x) \mathrm{d}x = \begin{cases} 0 & n \neq m \\ h_n & n = m \end{cases}, \tag{3.15}$$

where h_n is the norm of the orthogonal polynomial. If the all the integer moments of the weight function are finite i.e.

$$m_i = \int_a^b w(x) x^i \mathrm{d}x < \infty, \tag{3.16}$$

the monic orthogonal polynomials $f_n(x)$ are uniquely determined by the following determinantal representation

$$f_n(x) = \frac{1}{\Delta} \begin{vmatrix} m_0 & m_1 & \cdots & m_n \\ m_1 & m_2 & \cdots & m_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n-1} & m_n & \cdots & m_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix} \tag{3.17}$$

with

$$\Delta = |m_{i+j}|, \quad i, j = 0, \ldots, n-1. \tag{3.18}$$

To find a polynomial approximation, first a proper weight function has to be selected, which has the same support as $p(X)$. Define an initial approximation to $p(X)$ via the weight function by $\psi(x) = cw(x)$, where $c = \left(\int_a^b w(x) \mathrm{d}x\right)^{-1}$. We assume that the initial approximation $\psi(x)$ is obtained by matching the first two moments of $p(x)$. With the above notations, the n-moment based orthogonal polynomial approximation is given by

$$p_n(x) = \psi(x) + \psi(x) \sum_{i=3}^{n} \eta_i f_i(x), \tag{3.19}$$

where η_i are determined from the equations

$$\int_a^b f_i(x) p_n(x) \mathrm{d}x = \int_a^b f_i(x) p(x) \mathrm{d}x, \quad i = 3, \ldots, n. \tag{3.20}$$

The choice of the initial approximation $\psi(x)$ can be decided from the support of the random variable of interest. When $x \in (-\infty, \infty)$, $x \in [a, \infty)$, and $x \in [a, b]$ (a, b being finite) reasonable choices for the initial approximations are Gaussian, Gamma, and Beta densities, respectively. These density functions in turn correspond to Hermite, Laguerre, and Jacobi orthogonal polynomials. An important property of orthogonal polynomial approximations is that for random variables of finite support $x \in [a, b]$, the approximation (3.19) becomes exact as the number of polynomials n goes to infinity. This result is known as the Weierstrass approximation theorem [44]. In practise, the choice of n reflects a trade-off between the approximation accuracy and the implementation complexity. We consider $n = 2$ in all the included publications of this thesis, where moment based approximation has been invoked. The general n-moment-based approximation, including the error analysis, can be easily obtained by following the procedures in [45, 46]. For recent applications of moment based approximations in wireless communications, we refer to [47, 48].

3.1.3 Mellin transform

Integral transforms plays an important role in calculating functions of random variables. Some manipulations can be more easily performed in the transform domain than on non-transformed functions. In this respect, Mellin transform is a powerful tool in characterizing the distribution of products and ratios of independent random variables. For a random variable X with the density function $p_X(x)$, the Mellin transform is defined as

$$M_s[p_X(x)] = \int_0^\infty x^{s-1} p_X(x) \mathrm{d}x, \tag{3.21}$$

where s is a complex number and the above Mellin transform exists if the integral converges. One basic property useful in the context of this thesis is presented here. We are interested in the distribution of product $Z = XY$ of independent random variables X and Y. The $(s-1)$-th moment of random variable Z equals

$$\mathbb{E}[Z^{s-1}] = \mathbb{E}[(XY)^{s-1}] = \mathbb{E}[X^{s-1}]\mathbb{E}[Y^{s-1}]. \tag{3.22}$$

Define the densities of random variables X, Y, and Z as $p_X(x)$, $p_Y(x)$, and $p_Z(x)$, respectively. As a consequence of the above formulations, the Mellin transform relation for the product of random variables is

$$M_s[p_Z(x)] = M_s[p_X(x)] M_s[p_Y(x)]. \tag{3.23}$$

By the Mellin inversion integral, the density of Z is uniquely determined by the contour integral

$$p_Z(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-s} M_s[p_X(x)] M_s[p_Y(x)] \mathrm{d}s, \tag{3.24}$$

where the integration path is a vertical line in the complex plane. For a review of Mellin transform techniques useful in engineering, see [49], and for recent applications in performance analysis of communication systems we refer to [50, 51].

3.1.4 Random matrix theory

Random Matrix Theory (RMT) deals with the statistics of large dimensional matrices with random variable entries. RMT embraces various branches of mathematics such as analysis, algebra, differential equations, and combinatorics. Contrary to multivariate analysis, in RMT the focus is on the asymptotic behaviors when the matrix size approaches infinity. In particular, the emphasize has been on the asymptotic eigenvalue behavior. Although multivariate analysis often leads to exact characterizations, these results may become computational prohibitive for large matrix dimensions. The corresponding results derived from RMT may serve as alternatives for computational purposes. This is the case for spectrum sensing applications, where the sample sizes are typically of the order hundreds. One of the key features of RMT is so-called universality. Namely, irrespective of the choices of measure for matrix entries, the behavior of certain metrics such as the correlation functions remain the same. The standard reference for RMT is the textbook by Mehta [52], where some of the most important methodologies are discussed. For applications in wireless communications using RMT we refer to [53], and for a recent survey of all major applications see [54]. As an example of random matrix theory, in the following we shall discuss in some detail an asymptotical result on the largest eigenvalue distribution of the sample covariance matrix $\mathbf{R}$ ($\mathbf{\Sigma} = \mathbf{I}_K$ in (2.2)), which will be used in the next section.

For complex Gaussian matrices defined in (3.1), the joint density of the eigenvalues $0 \leq \lambda_K \leq \ldots \leq \lambda_1 < \infty$ of the sample covariance matrix $\mathbf{R} = \mathbf{X}\mathbf{X}^\dagger$ reads [36]

$$C \prod_{1 \leq i < j \leq K} (\lambda_i - \lambda_j)^2 \prod_{i=1}^{K} \lambda_i^{N-K} \mathrm{e}^{-\lambda_i}, \tag{3.25}$$

where

$$C = \left(\prod_{i=1}^{K} \Gamma(N-i+1)\Gamma(K-i+1) \right)^{-1}. \tag{3.26}$$

Define the distribution function of the largest eigenvalue λ_1 as

$$F(x) = \mathbb{P}(\lambda_1 < x). \tag{3.27}$$

In the asymptotic regime $K \to \infty$, $N \to \infty$ with $K/N \in (0,1)$ fixed, it is shown in [55] based on the analysis of [56] that the scaled and centered largest eigenvalue distribution converges to a Tracy-Widom distribution of order two

$$\lim_{\substack{K\to\infty\\N\to\infty}} F(\mu+\nu s) = F_{\mathrm{TW2}}(s), \tag{3.28}$$

where

$$\mu = \left(\sqrt{K}+\sqrt{N}\right)^2, \quad \nu = \left(\sqrt{K}+\sqrt{N}\right)\left(\frac{1}{\sqrt{K}}+\frac{1}{\sqrt{N}}\right)^{1/3}. \tag{3.29}$$

Johnstone [57] gave an alternative proof of the above result and discussed its application in statistics. The Tracy-Widom distribution of order two can be represented as a Fredholm determinant

$$F_{\mathrm{TW2}}(s) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{[s,\infty)^k} |A(x_i,x_j)| \mathrm{d}x_1 \cdots \mathrm{d}x_k, \tag{3.30}$$

where $i,j = 1,\ldots,k$ and the Airy kernel

$$A(x,y) = \frac{A_{\mathrm{i}}(x)A_{\mathrm{i}}'(y) - A_{\mathrm{i}}'(x)A_{\mathrm{i}}(y)}{x-y}, \tag{3.31}$$

$A_{\mathrm{i}}(x)$ being the Airy function. The Airy kernel is obtained by the asymptotics of Laguerre polynomials [57]. Besides the Fredholm determinant representation, Tracy-Widom distribution of order two can be also represented as

$$F_{\mathrm{TW2}}(s) = \mathrm{e}^{-\int_s^{\infty}(x-s)q^2(x)\mathrm{d}x}, \tag{3.32}$$

where $q(x)$ is the solution to the Painlevé II differential equation

$$q''(x) = xq(x) + 2q^3(x) \tag{3.33}$$

with the boundary condition $q(x) \sim A_{\mathrm{i}}(x)$, when $x \to \infty$. For a survey of applications of Tracy-Widom distribution we refer to [58] and for accurate and fast algorithms for the numerical evaluation of $F_{\mathrm{TW2}}(s)$ we refer to [59, 60].

3.2 Key results

Using the above described mathematical tools, distributions of the considered test statistics can be derived. In the following we will present, without proof, these distributional results. The state-of-the-art on each result will be briefly discussed as well.

3.2.1 Distribution of T_{LE} under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$

In this case the exact distribution of T_{LE} under $\mathcal{H}_0$ is

$$F_{\mathrm{LE}}(y) = \frac{C}{(\sigma^2)^{KN}} |\mathbf{A}|, \tag{3.34}$$

where the i,j-th entry of the $K \times K$ matrix $\mathbf{A}$ equals

$$\mathbf{A}_{i,j} = \int_0^y x^{N-K+i+j-2} \mathrm{e}^{-\frac{x}{\sigma^2}} \, \mathrm{d}x, \tag{3.35}$$

and C was defined in (3.26). For $\sigma^2 = 1$ the corresponding result can be found in [61], and recently [62] gave a different derivation.

Based on the non-central Wishart model (2.7), the exact distribution of T_{LE} under $\mathcal{H}_1$ is

$$G_{\mathrm{LE}}(y) = C' \left|\mathbf{B}\right|, \tag{3.36}$$

where

$$C' = \frac{\left(\prod_{i=1}^{K-1} \Gamma(N-i)\Gamma(K-i)\right)^{-1}}{\Gamma(N-K+1)(\sigma^2)^{KN-2K+2} \, \mathrm{e}^{\xi_1/\sigma^2} \xi_1^{K-1}}, \tag{3.37}$$

and the first column of the $K \times K$ matrix $\mathbf{B}$ equals

$$\mathbf{B}_{i,1} = \int_0^y x^{N-i} \mathrm{e}^{-\frac{x}{\sigma^2}} \, {}_0F_1\left(N-K+1; \frac{\xi_1 x}{\sigma^4}\right) \mathrm{d}x, \tag{3.38}$$

and columns from the second to the K-th are

$$\mathbf{B}_{i,j} = \int_0^y x^{N+K-i-j} \mathrm{e}^{-\frac{x}{\sigma^2}} \, \mathrm{d}x, \quad j = 2, \ldots, N. \tag{3.39}$$

Here ξ_1 stands for the largest eigenvalue of the non-central parameter matrix $\mathbf{\Xi}$ in (2.8), and ${}_0F_1(\cdot;\cdot)$ is the hypergeometric function of Bessel type. For $\sigma^2 = 1$ the above result reduces to the result in [63]. For Gaussian distributed signals, the corresponding result can be derived based on [64].

The above results are obtained via multivariate analysis techniques, starting with the joint eigenvalue densities for Wishart and non-central Wishart matrices. By using the symmetry of these joint densities, the largest eigenvalue distributions can be deduced. By closely following the steps of the derivations in [61] and [63], the noise power σ^2 is incorporated into the final results. Under $\mathcal{H}_0$, in the asymptotic regime of $K \to \infty$, $N \to \infty$ with $K/N \in (0,1)$ fixed, T_{LE} follows Tracy-Widom distribution of order two, given in (3.28). Under $\mathcal{H}_1$ for Gaussian distributed signals, in the same asymptotic regime, T_{LE} follows a Gaussian distribution with mean and variance given by [65]

$$N\sigma_1 \left(1 + \frac{K/N}{\sigma_1 - 1}\right), \quad \sigma_1 \sqrt{N\left(1 - \frac{K/N}{(\sigma_1 - 1)^2}\right)}, \tag{3.40}$$

respectively, where σ_1 is the largest eigenvalue of the population covariance matrix $\mathbf{\Sigma}$ in (2.5).

Finally, by the test procedure (2.12), the false alarm and detection probabilities of the LE detector equal

$$P_{\text{fa}}(\zeta) = 1 - F_{\text{LE}}(\zeta) \tag{3.41}$$

and

$$P_{\text{d}}(\zeta) = 1 - G_{\text{LE}}(\zeta), \tag{3.42}$$

respectively. As a result, an analytical ROC is obtained as

$$P_{\text{d}} = 1 - G_{\text{LE}}\left(F_{\text{LE}}^{-1}(1 - P_{\text{fa}})\right). \tag{3.43}$$

3.2.2 Distribution of T_{SLE} under hypothesis $\mathcal{H}_0$

For the SLE detector we derived an asymptotic as well as an exact distributions of its test statistics under hypothesis $\mathcal{H}_0$. In the limit of $K \to \infty$, $N \to \infty$ with $K/N \in (0,1)$ being fixed, an asymptotic approximation to the distribution of T_{SLE} under $\mathcal{H}_0$ is

$$F_{\text{SLE}}(y) \approx c\left(A(y) - A(1)\right), \tag{3.44}$$

where

$$c = \frac{\Gamma(KN)(K\theta)^{-k}}{k\Gamma(KN-k)\Gamma(k)} \tag{3.45}$$

and

$$A(x) = {}_2F_1\left(k, 1+k-KN; k+1; \frac{x}{K\theta}\right) x^k \tag{3.46}$$

denotes the Gaussian type hypergeometric function. Here

$$k = \frac{(\mu+\nu a)^2}{\nu^2 b}, \quad \theta = \frac{\nu^2 b}{\mu + \nu a}, \tag{3.47}$$

where μ and ν have been defined in (3.29), and

$$a = -1.7711, \quad b = 0.8132 \tag{3.48}$$

are the mean and variance of Tracy-Widom distribution of order two. Techniques used to obtain the above result include RMT, moment based approximation, and Mellin transform. Based on the observations in [66] that when $K \ll N$ the extreme eigenvalues of $\mathbf{R}$ can be well approximated by Gamma distributions, we adopted a two-moment-based Gamma approximation to λ_1 using the first two asymptotic moments. With this approximation and the fact that the random variables T_{SLE} and $\sum_{i=1}^{K}\lambda_i$ are independent [67], the result in (3.44) is obtained via the Mellin transform. In literature, the classical chi-squared asymptotics of T_{SLE} was derived in [18]. When ignoring the dependence of λ_1 and $\sum_{i=1}^{K}\lambda_i$ the resulting T_{SLE} density was derived in [19]. Refined estimates of T_{SLE} when considering the correlation of λ_1 and $\sum_{i=1}^{K} \lambda_i$, as in our case, were derived in [68, 20].

For small K and N, the accuracy of the asymptotic approximation is rather loose. Motivated by this, we derived an exact representation of the T_{SLE} density for finite K and N as

$$F_{\text{SLE}}(y) = \frac{(KN-1)!}{K^{KN-1}} \sum_{i=1}^{K} \sum_{j=N-K}^{(N+K)i-2i^2} i^{KN-j-2} c_{i,j} \times \left(A(y) h\left(\frac{K}{i} - y\right) + A\left(\frac{K}{i}\right) h\left(y - \frac{K}{i}\right) - A(1) \right), \tag{3.49}$$

where

$$A(y) = \left(\frac{K}{i}\right)^{KN-j-2} \sum_{q=0}^{KN-j-1} \frac{(-i/K)^q (j+q+1)^{-1}}{(KN-j-2-q)!q!} y^{q+j+1}, \tag{3.50}$$

and

$$h(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \tag{3.51}$$

denotes the Heaviside step function. In (3.49) the coefficients $c_{i,j}$ are unknown constants. We have derived closed-form expressions of $c_{i,j}$ for $K \leq 4$ with arbitrary N. For $K > 4$, one has to resort to numerical techniques [62, 69] to obtain the values of $c_{i,j}$. For the derivations of this exact result (3.49), we have invoked tools from multivariate analysis and Mellin transform. Namely, by using the series representation of the density of the largest eigenvalue [62], its exact Mellin transform is determined. The exact T_{SLE} density is then solved via the Mellin transform relation for product of independent random variables (3.23). Note that an alternative representation for the T_{SLE} density can be found in [70], which depends on a large number of unknown constants. Numerical evaluation of these constants seems difficult.

By the test procedure (2.14), the false alarm probability of the SLE detector equals

$$P_{\text{fa}}(\zeta) = 1 - F_{\text{SLE}}(\zeta). \tag{3.52}$$

3.2.3 Distribution of T_{ST} under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$

For this case we have derived closed-form approximations to the T_{ST} distributions under both hypotheses using moment based approximation and multivariate analysis techniques. Specifically, the two-first-moment Beta approximation to the distribution of T_{ST} under $\mathcal{H}_0$ is

$$F_{\text{ST}}(y) \approx \frac{B_y(\alpha_0, \beta_0)}{B(\alpha_0, \beta_0)}, \tag{3.53}$$

where

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \tag{3.54}$$

$$B_y(\alpha,\beta) = \int_0^y x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x \tag{3.55}$$

denote the Beta function and the incomplete Beta function, respectively. The parameters α_0 and β_0 are given by

$$\alpha_0 = \frac{\mathcal{M}_1(\mathcal{M}_1-\mathcal{M}_2)}{\mathcal{M}_2-(\mathcal{M}_1)^2}, \quad \beta_0 = \frac{(1-\mathcal{M}_1)\,(\mathcal{M}_1-\mathcal{M}_2)}{\mathcal{M}_2-(\mathcal{M}_1)^2} \tag{3.56}$$

with

$$\mathcal{M}_n = \frac{\Gamma(KN)}{\Gamma_K(N)}\frac{K^{Kn}\Gamma_K(N+n)}{\Gamma(K(N+n))}, \tag{3.57}$$

where the multivariate Gamma function $\Gamma_K(N)$ has been defined in (2.3).

Similarly, the two-first-moment Beta approximation to the distribution of T_{ST} under $\mathcal{H}_1$ is obtained as

$$G_{\mathrm{ST}}(y) \approx \frac{B_y(\alpha_1,\beta_1)}{B(\alpha_1,\beta_1)}, \tag{3.58}$$

where the parameters α_1 and β_1 are given by

$$\alpha_1 = \frac{\mathcal{N}_1(\mathcal{N}_1-\mathcal{N}_2)}{\mathcal{N}_2-(\mathcal{N}_1)^2}, \quad \beta_1 = \frac{(1-\mathcal{N}_1)\,(\mathcal{N}_1-\mathcal{N}_2)}{\mathcal{N}_2-(\mathcal{N}_1)^2} \tag{3.59}$$

with

$$\mathcal{N}_n = \left(\frac{K}{b}\right)^{Kn}\frac{\Gamma(a-Kn)\Gamma_K(N+n)\,(\det(\mathbf{\Sigma}))^n}{\Gamma_K(N)\Gamma(a)} \tag{3.60}$$

and

$$a = (N+n)\frac{\left(\sum_{i=1}^K \sigma_i\right)^2}{\sum_{i=1}^K \sigma_i^2}, \quad b = \frac{\sum_{i=1}^K \sigma_i^2}{\sum_{i=1}^K \sigma_i}. \tag{3.61}$$

Recall that σ_i is the i-th eigenvalue of the population covariance matrix $\mathbf{\Sigma}$ under $\mathcal{H}_1$.

The steps that lead to the distributions under both hypotheses are similar. We first derived the finite size moments expressions (3.57) and (3.60) of T_{ST}, based on which the moment-based approximations are constructed. In the statistics literature, the T_{ST} was considered as early as 1940 [71]. The corresponding moments for real Wishart matrices under both hypotheses were derived in [13]. For $K=2$ and $K=3$ with any N, the exact densities of T_{ST} under $\mathcal{H}_0$ can be found in [72, Eq. (3.8)] and [73, Coroll. 2.1], respectively. For $K>3$ the exact T_{ST} distribution can be, in principle, obtained by the Mellin inversion integral [74]. The resulting expressions, although of theoretical interest, appear to be of limited usefulness due to their complicated forms. Under $\mathcal{H}_0$, as asymptotic T_{ST} distribution has been derived in [75] using chi-squared asymptotics, whose accuracy turns out to be inferior than our proposed approximation [PIV]. One of our motivations for the simple two-moment-based Beta approximation under $\mathcal{H}_1$ is due to the fact that the exact representations for the distribution of T_{ST}, see e.g. [76, Th. 4.1] and [77, Eq. (2.12)], involve infinite sum of products of a Zonal polynomial and the Meijer's G-function, and are difficult to compute.

From the test procedure (2.17), the false alarm and detection probabilities of the ST detector equal

$$P_{\mathrm{fa}}(\zeta) = F_{\mathrm{ST}}(\zeta) \tag{3.62}$$

and

$$P_{\mathrm{d}}(\zeta) = G_{\mathrm{ST}}(\zeta), \tag{3.63}$$

respectively. Thus, an analytical approximation to the ROC for the ST detector is obtained as

$$P_{\mathrm{d}} = G_{\mathrm{LE}}\left(F_{\mathrm{ST}}^{-1}(P_{\mathrm{ST}})\right). \tag{3.64}$$

3.2.4 Distribution of T_{J} under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$

For this case we have derived moment based approximations to the T_{J} distributions under both hypotheses. The essential calculations are closed-form moment expressions, which were open problems in the statistics literature, and are only recently addressed in [PV]. Under $\mathcal{H}_0$, the two-first-moment Beta approximation to the distribution of T_{J} is

$$F_{\mathrm{J}}(y) \approx 1 - \frac{B\left(\frac{K(1-y)}{K-1}; \beta_0, \alpha_0\right)}{B(\alpha_0, \beta_0)}, \tag{3.65}$$

where

$$\alpha_0 = \frac{(K\mathcal{M}_1 - 1)(K\mathcal{M}_1 - K\mathcal{M}_2 + \mathcal{M}_1 - 1)}{(K-1)K(\mathcal{M}_2 - \mathcal{M}_1^2)}, \tag{3.66a}$$

$$\beta_0 = \frac{(\mathcal{M}_1 - 1)(K\mathcal{M}_1 - K\mathcal{M}_2 + \mathcal{M}_1 - 1)}{(K-1)(\mathcal{M}_1^2 - \mathcal{M}_2)}. \tag{3.66b}$$

In (3.66)

$$\mathcal{M}_m := \frac{C\,\Gamma(KN)}{\Gamma(2m+KN)} \sum_{a_1+\cdots+a_K=m} \frac{m!}{a_1!\cdots a_K!} \times$$
$$\prod_{1\le i<j\le K} (2a_j - 2a_i + j - i) \prod_{i=1}^{K} \Gamma(2a_i + N - K + i), \tag{3.67}$$

where the sum is over all the non-negative integer solutions of $a_1 + \cdots + a_K = m$ and the constant C has been defined in (3.26).

Under $\mathcal{H}_1$, the corresponding two-first-moment Beta approximation to the distribution of T_{J} equals

$$G_{\mathrm{J}}(y) \approx 1 - \frac{B\left(\frac{K(1-y)}{K-1}; \beta_1, \alpha_1\right)}{B(\alpha_1, \beta_1)}, \tag{3.68}$$

where

$$\alpha_1 = \frac{(1 - K\mu_z)\left((\mu_z - 1)(K\mu_z - 1) + K\nu_z\right)}{(K-1)K\mu_z}, \tag{3.69a}$$

$$\beta_1 = \frac{(\mu_z - 1)\left((\mu_z - 1)(K\mu_z - 1) + K\nu_z\right)}{(K-1)\mu_z}, \tag{3.69b}$$

and

$$\mu_z \approx \frac{\mu_x}{\mu_y^2} - \frac{2\mu_{xy}}{\mu_y^3} + \frac{3\mu_x\nu_y}{\mu_y^4}, \tag{3.70}$$

$$\nu_z \approx \frac{\nu_x}{\mu_y^4} - \frac{4\mu_x\mu_{xy}}{\mu_y^5} + \frac{4\mu_x^2\nu_y}{\mu_y^6}. \tag{3.71}$$

Here the quantities μ_x, μ_y, ν_x, ν_y and μ_{xy} are calculated, in terms of the population covariance matrix under $\mathcal{H}_1$ (2.5), as

$$\mu_x = \mathrm{tr}(\mathbf{\Sigma}^2) + \frac{1}{N}\left(\mathrm{tr}(\mathbf{\Sigma})\right)^2, \tag{3.72}$$

$$\nu_x = \frac{4}{N}\mathrm{tr}(\mathbf{\Sigma}^4) + \frac{2}{N^2}\left(4\mathrm{tr}(\mathbf{\Sigma})\mathrm{tr}(\mathbf{\Sigma}^3) + \left(\mathrm{tr}(\mathbf{\Sigma}^2)\right)^2\right) + \frac{2}{N^3}\left(2\left(\mathrm{tr}(\mathbf{\Sigma})\right)^2\mathrm{tr}(\mathbf{\Sigma}^2) + \mathrm{tr}(\mathbf{\Sigma}^4)\right), \tag{3.73}$$

$$\mu_y = \mathrm{tr}(\mathbf{\Sigma}), \tag{3.74}$$

$$\nu_y = \frac{1}{N}\mathrm{tr}(\mathbf{\Sigma}^2), \tag{3.75}$$

$$\mu_{xy} = \frac{2}{N}\mathrm{tr}(\mathbf{\Sigma}^3) + \frac{2}{N^2}\mathrm{tr}(\mathbf{\Sigma})\mathrm{tr}(\mathbf{\Sigma}^2). \tag{3.76}$$

In order to obtain the approximation to T_{J} under hypotheses $\mathcal{H}_0$, we first need to prove the independence of random variables $\left(\sum_{i=1}^{K}\lambda_i\right)^2$ and T_{J}. With this result, we are able to calculate the exact moments (3.67) under $\mathcal{H}_0$ using multivariate techniques. Under $\mathcal{H}_1$, the moments of the numerator and denominator of T_{J} as well as their joint moments are calculated first via representation theory of the symmetric group. With these closed-form moments, accurate estimates of the first two moments (3.70) and (3.71) under $\mathcal{H}_1$ are derived using the delta-method. In the statistics literature, for real Wishart matrix, up to the second, fourth and sixth moments of T_{J} under $\mathcal{H}_0$ were derived in [78], [79] and [46], respectively. Moreover, our derived moment expression (3.67) is an extension to the Selberg type integrals considered in [52, Eq. (17.6.5) and (17.8.1)]. Several asymptotical T_{J} distributions for real Wishart matrices under $\mathcal{H}_0$

were established in [23, 80, 81], which may be generalized to the complex Wishart case. However, simulations show that these approximations converge slowly with respect to (w.r.t.) N for a fixed K [23, 80] and w.r.t. both K and N [81]. Under $\mathcal{H}_1$, asymptotic T_{J} distributions for real Wishart matrices are available in [23, 82]. However, besides being slowly converging, these asymptotic results are only valid for some specific structures of the population covariance matrix $\mathbf{\Sigma}$.

Finally, by the test procedure (2.19), the false alarm and detection probabilities of John's detector equal

$$P_{\text{fa}}(\zeta) = 1 - F_{\text{J}}(\zeta) \tag{3.77}$$

and

$$P_{\text{d}}(\zeta) = 1 - G_{\text{J}}(\zeta), \tag{3.78}$$

respectively. As a result, an analytical ROC is obtained as

$$P_{\text{d}} = 1 - G_{\text{J}}\left(F_{\text{J}}^{-1}(1 - P_{\text{fa}})\right). \tag{3.79}$$

3.2.5 Distribution of T_{ER} under hypothesis $\mathcal{H}_1$

In this case, we derived finite size as well as asymptotic approximations to the distributions of the ER detector under $\mathcal{H}_1$. As for T_{LE} we consider here the non-central Wishart matrix model (2.7)

$$\mathbf{R} \sim \mathcal{W}_K\left(N, \mathbf{\Psi}, \mathbf{\Xi}\right). \tag{3.80}$$

For finite K and N, the distribution of T_{ER} can be approximated by

$$G_{\text{ER}}(y) \approx -a^2 \int_0^\infty \mathbf{tr}\left(\mathbf{adj}(\mathbf{B})\mathbf{D}\right)|\mathbf{C}|\mathrm{d}x, \tag{3.81}$$

where the constant

$$a = \frac{\left(\prod_{i=1}^{K} \upsilon_i\right)^{-N}}{\left|-\upsilon_j^{1-i}\right| \prod_{i=1}^{K}(N-i)!}, \quad i,j = 1,\ldots,K, \tag{3.82}$$

with υ_i being the i-th eigenvalue of the matrix

$$\mathbf{\Upsilon} = \mathbf{\Psi} + \mathbf{\Xi}/N. \tag{3.83}$$

Here the i,j-th entries of the $K \times K$ matrices $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are

$$b_{i,j} = \upsilon_i^{N-K+j}\Gamma(N-K+j, x/\upsilon_i), \tag{3.84}$$

$$c_{i,j} = \upsilon_i^{N-K+j}\gamma(N-K+j, xy/\upsilon_i), \tag{3.85}$$

$$d_{i,j} = -x^{N-K+j-1}\mathrm{e}^{-x/\upsilon_i}, \tag{3.86}$$

where $\Gamma(\cdot,\cdot)$ and $\gamma(\cdot,\cdot)$ are respectively the upper and lower incomplete Gamma functions. In deriving the above result, we first formulated the non-central Wishart model to an equivalent central Wishart model. We then utilized the finite size distributions of the extreme eigenvalues densities of the central Wishart derived from multivariate analysis to construct the approximation. To achieve the final result, we ignored the independence of the extreme eigenvalues. Note that under $\mathcal{H}_0$, the exact distribution of T_{ER} was derived in [25].

In order to gain insight into the behaviors of the ER detector for large K and N, we derived an asymptotical approximation using recent results in RMT. Specifically, in the limit of $K \to \infty$, $N \to \infty$ with $K/N \in (0,1)$ fixed, the distribution of T_{ER} can be approximated by a Gaussian distribution

$$G_{\mathrm{ER}}(y) \approx \Phi\left(\frac{a_K y - a_1}{\sqrt{b_K^2 y^2 + b_1^2}}\right), \tag{3.87}$$

where

$$a_i = N\upsilon_i\left(1 + \frac{K/N}{\upsilon_i - 1}\right), \tag{3.88}$$

$$b_i = \upsilon_i\sqrt{N\left(1 - \frac{K/N}{(\upsilon_i - 1)^2}\right)}, \tag{3.89}$$

and υ_i has been defined in (3.83). The derivation of this result is similar to that of the corresponding finite size result (3.81). The only difference is that here the asymptotical distributions of the extreme eigenvalue densities [65] have been utilized. Under $\mathcal{H}_0$, the asymptotic distribution of T_{ER} can be found in [16, 24].

By the test procedure (2.22), the detection probability of ER detector equals

$$P_{\mathrm{d}}(\zeta) = 1 - G_{\mathrm{ER}}(\zeta). \tag{3.90}$$

3.2.6 Distribution of $T_{\mathbf{DCN}}$ under hypothesis $\mathcal{H}_0$

In this case the exact distribution of the DCN detector under $\mathcal{H}_0$ is derived using Mellin transform, and multivariate analysis techniques, as

$$F_{\mathrm{DCN}}(y) = \Gamma(KN) \sum_{n=N-K}^{(N-K)K} \frac{c_n\big(B(n+1, KN-n-1) - B_{\frac{K}{y}}(n+1, KN-n-1)\big)}{K^{n+1}\Gamma(KN-n-1)}, \tag{3.91}$$

where $B(\cdot,\cdot)$ and $B_{(\cdot)}(\cdot,\cdot)$ are the Beta and incomplete Beta functions defined in (3.54) and (3.55), respectively. In (3.91), the constants c_n are generally unknown, and numerical calculations are needed [62, 69]. Up to $K = 4$, closed-form expressions for these constants were calculated in [83]. Under $\mathcal{H}_0$, the distribution of T_{DCN} for $K = 2$ with arbitrary N was derived in [25], and for $K = N$ the corresponding result was derived in [84].

By the test procedure (2.23), the false alarm probability of the DCN detector is

$$P_{\text{fa}}(\zeta) = 1 - F_{\text{DCN}}(\zeta). \tag{3.92}$$

3.2.7 Distribution of T_{W} under hypothesis $\mathcal{H}_0$

For this case, the idea of the derivations are rather similar to those of T_{ST} and T_{J}, where we first calculated the exact moments using multivariate analysis then the desired approximation was constructed by moment based approximation techniques. Under $\mathcal{H}_0$, the two-moment-based Beta approximation to the distribution of T_{W} is

$$F_{\text{W}}(y) \approx \frac{B(y; \alpha_0, \beta_0)}{B(\alpha_0, \beta_0)}, \tag{3.93}$$

where

$$\alpha_0 = \frac{\mathcal{M}_1(\mathcal{M}_1 - \mathcal{M}_2)}{\mathcal{M}_2 - (\mathcal{M}_1)^2}, \quad \beta_0 = \frac{(1 - \mathcal{M}_1)(\mathcal{M}_1 - \mathcal{M}_2)}{\mathcal{M}_2 - (\mathcal{M}_1)^2}, \tag{3.94}$$

and

$$\mathcal{M}_m := \frac{\Gamma_K(N + M)\Gamma_K(M + m)}{\Gamma_K(M)\Gamma_K(N + M + m)}. \tag{3.95}$$

In literature, an exact density representation for real Wishart matrices via Meijer's G-function is derived in [85]. Although of theoretical interest, it is too complicated for computational purposes. For complex Wishart matrices, exact T_{W} densities were derived for a few limited cases, i.e. $K = 2$ and $K = 3$ in [86]. An asymptotic T_{W} distribution for real Wishart matrices can be found in [87, Eq. (5.4)], which is slowly converging and rather computationally intensive.

By the test procedure (2.29), the false alarm probability of Wilks' detector is

$$P_{\text{fa}}(\zeta) = F_{\text{W}}(\zeta). \tag{3.96}$$

Evaluations of the accuracy of the proposed approximative distributions of the considered test statistics in Chapter 3.2 will not be reproduced here. For results in this direction, the reader may follow each included paper for detailed discussions.

4. Performance Comparisons

In this chapter we compare the performance of the detectors discussed in Chapter 3 by means of the receiver operating characteristic. Since a ROC curve shows the achieved detection probability as a function of the false alarm probability, it reflects the overall detection performance for a given detector. The plots of false alarm probability versus threshold will not be reproduced here, which can be found in the included individual papers. Note that all the considered detectors in this thesis, except for the largest eigenvalue based detector and the energy detector, belong to the so-called constant false alarm rate detectors. Namely, for a given false alarm probability the threshold remains unchanged irrespective of the noise level. Contrary to the instantaneous ROC plots for a given channel realization in the included papers, we consider in the following average ROCs over fading channels. This metric is more informative than the instantaneous ROC curve. For the considered detectors, the corresponding analytical ROCs over fading channels are not available in literature, we thus resort to Monte-Carlo simulations for these ROC curves. For performance comparisons conditioned on a channel realizations, including the evaluations of the derived instantaneous ROCs and the corresponding error analysis, the readers are referred to the included individual papers for details. Moreover, we focus on detection performance comparisons in relatively low SNR regime, which is a practical and challenging issue in spectrum sensing.

To see a complete picture within the described three facts in Chapter 1.1, we will examine the impact of the noise uncertainty, the number of primary users, as well as the knowledge of the noise covariance matrix. In order to identify the effect of each of the above factors on detection performance we divide this chapter into two subchapters according to the criterion we have chosen to divide detectors in Chapter 2.2. Specifically, in

Table 4.1. Summary of the Simulation Scenario Considered in Each Figure

scenarios	$P = 1$	multiple P	
known $\boldsymbol{\Psi}$	Figure 4.1	Figure 4.2	Chapter 4.1
unknown $\boldsymbol{\Psi}$	Figure 4.3	Figure 4.4	Chapter 4.2

Chapter 4.1 we focus on performance comparisons among detectors that have perfect knowledge of the noise covariance matrix. In Chapter 4.2 we then concentrate on comparing blind noise covariance detectors to non-blind ones. For a quick reference, we summarize in Table 4.1 the scenarios considered in each subsection and for each figure.

The considered values of the parameters K, N and M in this chapter reflect practical spectrum sensing scenarios. The number of samples N and the number of noise-only observations M can be as large as a couple of hundred whereas the number of sensors K is at most eight due to physical constraints of the device size. Contrary to the instantaneous ROC curves generated for a fixed channel realization $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_P]$ i.e. a fixed covariance matrix under $\mathcal{H}_1$

$$\boldsymbol{\Sigma} = \boldsymbol{\Psi} + \sum_{i=1}^{P} \gamma_i \mathbf{h}_i \mathbf{h}_i^{\dagger}, \tag{4.1}$$

each of the following curves is generated by averaging over 10^4 ROC curves obtained by 10^4 independent realizations of $\boldsymbol{\Sigma}$. For each realization, the single-tap flat fading channel matrix $\mathbf{H}$ is independently drawn from a standard complex Gaussian distribution. The channel is assumed to be constant during each sensing interval. As a result, we are considering average detection performance over Rayleigh fading channels. Each instantaneous ROC is generated by using 10^6 independent realizations of the data matrix $\mathbf{X}$ under both hypotheses. Each average ROC curve is obtained by averaging over the detection probabilities for a fixed false alarm probability.

4.1 Impact of noise uncertainty and number of primary users on the detection performance

In this subchapter the noise covariance matrix $\boldsymbol{\Psi}$ is assumed to be known and is chosen, without loss of generality, to be $\boldsymbol{\Psi} = \sigma^2 \mathbf{I}_K$. In Figure 4.1,

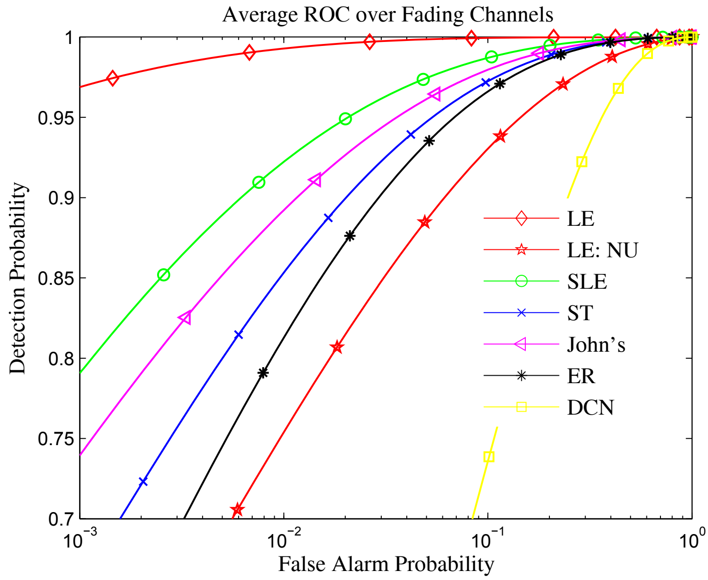

Figure 4.1. Average ROC over Rayleigh fading channels: assuming the presence of a single primary user with SNR $= -2$ dB using $K = 4$ sensors and $N = 150$ samples per sensor. The noise covariance matrix is assumed to be $\mathbf{\Psi} = \sigma^2 \mathbf{I}_K$. The degree of noise uncertainty (NU) for the LE detector is chosen to be 0.3 dB.

we first consider the scenario of a single primary user $P = 1$. We include for comparisons all the non-blind $\mathbf{\Psi}$ detectors discussed in Chapter 2.2.1, including the non-blind noise power σ^2 detector, i.e. the LE detector. The effect of noise uncertainty on the performance of the LE detector is studied as well. The LE detector is subject to noise uncertainty due to its dependence on the noise power. All the other detectors in Figure 4.1 are blind to σ^2, and thus their performance will not change irrespective of the degree of noise uncertainty. The model of noise uncertainty is as follows. If ϑ denote the degree of noise uncertainty in dB, the actual noise power thus falls in the interval $[\sigma^2/c, c\sigma^2]$, where $c = 10^{\vartheta/10}$. As in [9, 10, 88], here we consider the worst performance degradation due to noise uncertainty, where the noise power for the LE detector is $c\sigma^2$ under $\mathcal{H}_0$ and σ^2/c under $\mathcal{H}_1$. In Figure 4.1 we consider the noise uncertainty level to be 0.3 dB, which is generally realistic. For example, it was remarked in [9] that the noise uncertainty can be at least 1 to 2 dB due to limitations of devices only and in [88] the authors considered noise uncertainty levels up to 3 dB. In Figure 4.1, the number of sensors and samples per sensor equal $K = 4$ and $N = 150$, respectively. The SNR of the primary user is set at -2 dB.

From Figure 4.1 we can see that in the case of perfectly known σ^2 the

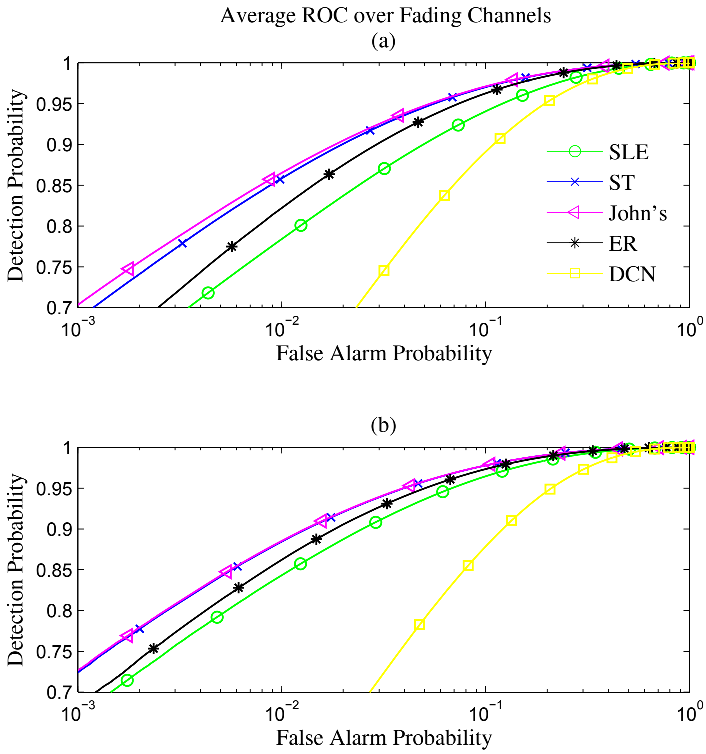

Figure 4.2. Average ROC over Rayleigh fading channels: assuming the presence of three primary users. The noise covariance matrix is assumed to be $\mathbf{\Psi} = \sigma^2\mathbf{I}_K$. In subfigure (a): $\text{SNR}_1 = -3$ dB, $\text{SNR}_2 = -4$ dB, and $\text{SNR}_3 = -5$ dB with $K = 4$ sensors and $N = 150$ samples per sensor. In subfigure (b): $\text{SNR}_1 = -16$ dB, $\text{SNR}_2 = -18$ dB, and $\text{SNR}_3 = -20$ dB with $K = 4$ sensors and $N = 5 \times 10^4$ samples per sensor.

LE detector performs best, and with only 0.3 dB noise uncertainty its performance degrades severely. The superior performance of the LE detector shown here will not likely occur in practise due to the presence of noise uncertainty in the real world. As expected, the SLE performs best among the blind-σ^2 detectors, since it is designed for (in the GLR sense) single-primary-user detection with known $\mathbf{\Psi}$ by construction. On the other hand, the blind P detectors indeed suffer performance loss in the presence of a single primary user. Among these, it is seen that the heuristically proposed ER and DCN detectors perform particularly worse.

Having seen the effect of noise uncertainty on the relative detection performance, in the next figure we only include for comparisons these noise-uncertainty-free detectors. In the presence of noise uncertainty, the behavior of the ED detector is similar to that of the LE detector. Here, our focus will be on the impact of number of primary users on the detection performance. Specifically, in Figure 4.2 we assume a scenario of three simultaneously transmitting primary users. In Figure 4.2 (a) we consider the case when $\mathrm{SNR}_1 = -3$ dB, $\mathrm{SNR}_2 = -4$ dB, and $\mathrm{SNR}_3 = -5$ dB, with $K = 4$ sensors and $N = 150$ samples per sensor. In Figure 4.2 (b) we consider detection in the very low SNR regime when $\mathrm{SNR}_1 = -16$ dB, $\mathrm{SNR}_2 = -18$ dB, and $\mathrm{SNR}_3 = -20$ dB, with $K = 4$ sensors and $N = 5 \times 10^4$ samples per sensor. This consideration is motivated by the fact that the recent Federal Communications Commission regulations require that the secondary devices must be able to detect signals with SNR as low as -18 dB [89, 90]. We observe from Figure 4.2 that in this case John's detector achieves the best detection performance. The very similar performance of John's and the ST detectors as observed in Figure 4.2 (b) is not surprising. In fact, it was proven in [91] that their asymptotic performances are the same as measured by the Pitman efficiency. The performance loss of the SLE detector is due to its non-blind P nature. Again we see that neither of the heuristic detectors, the ER or the DCN detector, perform well. Both Figure 4.1 and Figure 4.2 seem to indicate that the use of decision-theoretically constructed detectors over these heuristically proposed ones is justifiable.

4.2 Impact of knowledge of noise covariance matrix on the detection performance

We now focus on comparing performance of blind-$\boldsymbol{\Psi}$ detectors. The exponential correlation model [92]

$$(\boldsymbol{\Psi})_{i,j} = \rho^{|i-j|}, \quad \rho \in [0, 1), \tag{4.2}$$

is chosen for the noise covariance matrix, where ρ specifies the degree of noise correlation. Note that unlike the cases studied above, here the noise covariance matrix is unknown at the secondary receivers. We start by studying the scenario of a single primary user in Figure 4.3, where we set $\mathrm{SNR} = -2$ dB, $(K, N, M) = (4, 150, 100)$ and $\rho = 0.4$. For this case, we compare Roy's detector and Wilks's detector, both of which are blind-$\boldsymbol{\Psi}$ detectors by construction. As the focus is now the single-primary-user detection in addition to the blind-$\boldsymbol{\Psi}$ setting, we also include for comparison the SLE detector derived from the GLR criterion. Comparisons with other non-blind $\boldsymbol{\Psi}$ detectors in Chapter 2.2.1 are excluded. For a fair com-

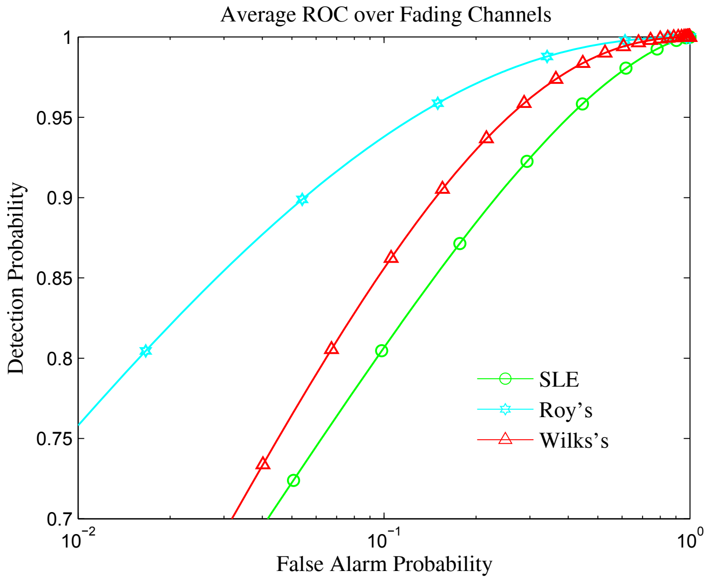

Figure 4.3. Average ROC over Rayleigh fading channels: assuming the presence of a single primary user with SNR $= -2$ dB using $K = 4$ sensors, $N = 150$ samples per sensor, and $M = 100$ noise-only observations. The noise covariance matrix is assumed to be $(\mathbf{\Psi})_{i,j} = \rho^{|i-j|}$ with $\rho = 0.4$, but is unknown at the secondary receiver.

parison, the SLE detector also needs to utilize the available noise-only observations M. To this end, we replace $\mathbf{R}$ by the 'whitened' sample covariance matrix $\mathbf{E}^{-1}\mathbf{R}$ i.e. the SLE detector now becomes

$$\frac{\theta_1}{\frac{1}{K}\sum_{i=1}^{K}\theta_i}. \tag{4.3}$$

This modification is motivated by the fact that $\mathbf{E}$ is the maximum likelihood estimate of $\mathbf{\Psi}$, and for a known $\mathbf{\Psi}$ the SLE detector becomes a function of $\mathbf{\Psi}^{-1}\mathbf{R}$ [13]. Moreover, the modified SLE detector (4.3) can be considered as a heuristic blind-$\mathbf{\Psi}$ detector. We can see from Figure 4.3 that Roy's detector performs best in this scenario, and indeed its usefulness in detecting a single primary user in the case of arbitrary and unknown $\mathbf{\Psi}$ has been justified in [32, 33]. The performance loss of Wilks' detector is due to its multiple-primary-user assumption induced from construction.

In Figure 4.4 we investigate the impact of the number of primary users on the relative performance of the blind-$\mathbf{\Psi}$ detectors. In addition to Roy's and Wilks' detectors, we also include for comparisons the non-blind $\mathbf{\Psi}$ multiple-primary-user detectors: the ST and John's detectors. Comparisons with heuristically proposed non-blind-$\mathbf{\Psi}$ detectors are excluded. For a fair comparison, both the ST and John's detectors are modified to in-

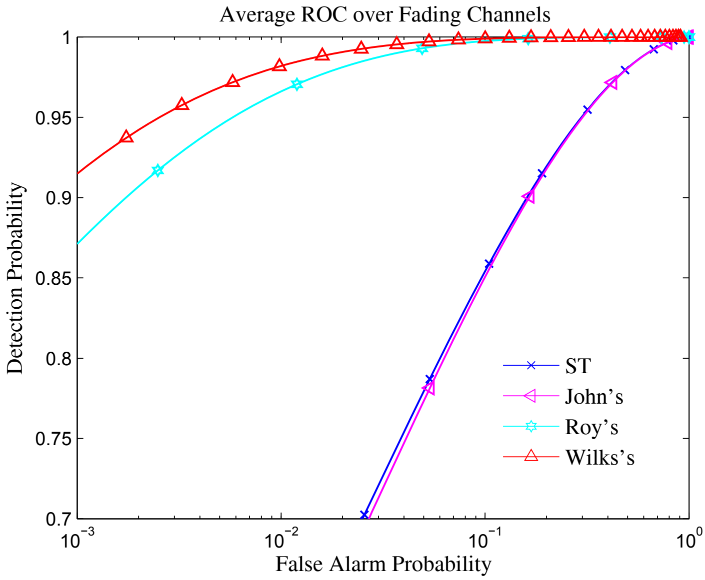

Figure 4.4. Average ROC over Rayleigh fading channels: assuming the presence of three primary users with $\mathrm{SNR}_1 = -3$ dB, $\mathrm{SNR}_2 = -4$ dB, and $\mathrm{SNR}_3 = -5$ dB using $K = 4$ sensors, $N = 150$ samples per sensor, and $M = 100$ noise-only observations. The noise covariance matrix is assumed to be $(\mathbf{\Psi})_{i,j} = \rho^{|i-j|}$ with $\rho = 0.4$, but is unknown at the secondary receiver.

corporate the available noise-only observations M in the same way as the modified SLE detector (4.3). In Figure 4.4 we consider a scenario of three active primary users with $\mathrm{SNR}_1 = -3$ dB, $\mathrm{SNR}_2 = -4$ dB, $\mathrm{SNR}_3 = -5$ dB and we set the degree of noise correlation at $\rho = 0.4$. The sensor size, the number of samples per sensor, and the number of noise-only observations equal $K = 4$, $N = 150$, and $M = 100$, respectively. From Figure 4.4 we see that Wilks' detector outperforms Roy's detector, which is expectable since the former is designed for (in the GLR sense) multiple P detection when $\mathbf{\Psi}$ is arbitrary and unknown. It is seen that the heuristically proposed detectors i.e. the modified ST and John's detectors perform substantially worse. This is intuitively clear since Wilks' detector was derived from a decision-theoretic criterion i.e. the GLR criterion whereas these modified detectors utilize the noise-only samples in a heuristic manner. Our intensive simulations show that both Roy's and Wilks' detectors perform much better than these modified non-blind $\mathbf{\Psi}$ detectors for $\rho > 0.4$. Finally we note that the observations in Figure 4.3 and Figure 4.4 regarding the relative performance of Roy's and Wilks' detectors are in line with those of [93, 94].

5. Conclusion and Future Work

5.1 Conclusion

Current static spectrum allocation schemes lead to severe underutilization of the spectrum, and thus can not meet the demands of an ever increasing number of higher data rate devices. The concept of cognitive radio has been introduced as a promising solution to the problem of spectrum underutilization. Cognitive radio provides the opportunity to autonomously exploit locally unused spectrum, giving rise to new paths to spectrum access. One of the key steps towards this dynamic spectrum access is the ability of the secondary user to reliably detect the primary users. Therefore, spectrum sensing becomes an important enabler for cognitive radio networks. In this thesis, we have focused on multi-antenna assisted energy based spectrum sensing, where the detection problem is modeled as a binary hypothesis test. The investigations leading to this thesis were motivated by relevant practical issues related to energy based detection. These issues are the noise uncertainty problem, the presence of more than one active primary users as well as the existence of unknown noise correlations. By taking the above practical issues into considerations, we proposed different detection algorithms that are suitable for different scenarios. Furthermore, we analyzed the performance of the relevant detectors by deriving closed-form expressions for the false alarm probability, the detection probability, and the receiver operating characteristic, which are the main contributions of this thesis. These results are obtained by making use of tools from multivariate analysis, moment based approximations, Mellin transforms, and random matrix theory.

Comprehensive simulations performed in Chapter 4 have shown that the proposed detectors indeed resolved the concerns raised by the above listed practical issues. Specifically, the SLE detector, studied in [PII] and [PIII], performs best in the presence of noise uncertainty and a single active primary user, see Figure 4.1. The ST detector studied in [PIV] and John's detector studied in [PV] have shown their superiority in the presence of multiple primary users, outperforming single primary detectors and heuristic detectors that are with the same assumptions, see Figure 4.2. In the most general scenario of unknown noise covariance matrix with multiple primary users, Wilks' detector studied in [PVIII] is shown to perform best, see Figure 4.4. Wilks' detector stands out to be the most robust detection algorithm under the framework developed in this thesis. In general, the choice of detection algorithm should be based on a priori knowledge on the primary and/or secondary systems. However, given the fact that a priori knowledge is hardly available in practice, Wilks' detector becomes a viable candidate in realistic spectrum sensing scenarios. Finally, we would like to remark that there are no theoretical justifications for assuming that test statistics derived from the GLR or LBI criterion are optimal tests, i.e. uniformly most powerful tests. However, strong numerical evidence suggests that it is wise to use detectors that are derived from decision-theoretic considerations over the corresponding heuristic detectors.

5.2 Future work

Based on the results and conclusions of this thesis, several viable directions for future work are listed in the following.

- The derived detection probabilities in Chapter 3.2 are conditioned on a channel realization, i.e. they represent instantaneous detection probability. It would be interesting to capture the average detection probability over fading channels, and to study the relative detection performance in different fading scenarios. To obtain analytical expressions for the average detection probability, by definition, one needs to integrate an instantaneous detection probability over the density of the corresponding channel model. For this type of analysis, the tools from multivariate analysis and from random matrix theory such as Selberg's integrals [52] may be useful.

- Under the same setting as Wilks' detector, the corresponding detector derived from the locally best invariant criterion is Lawley-Hotelling' test [13, 35]

$$T_{\mathrm{LH}} := \sum_{i=1}^{K} \theta_i \in [0, K]. \tag{5.1}$$

Similar to Wilks' detector, Lawley-Hotelling' detector is blind to the noise covariance matrix and the number of primary users. Moreover, it is likely to outperform Wilks' detector in the low SNR regime by construction. Analytically capturing the detection probability of Lawley-Hotelling' detector is still an open problem in literature. The tools for this purpose include representation theory for symmetric group, which has been invoked in deriving the detection probability of John's detector [PV].

- In situations where we do have reliable a priori information of primary and/or secondary systems, the advantages of different detectors can be dynamically exploited by switching between detectors that are most suitable for the current system state. The a priori information may be acquired algorithms estimating system parameters. How to incorporate the a priori information to capture this 'estimation-assisted detector' remains interesting future work as well.

References

[1] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Personal Commun.,* vol. 6, no. 4, pp. 13-18, Aug. 1999.

[2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.,* vol. 23. no. 2, pp. 201-220, Feb. 2005.

[3] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys and Tutorials,* vol. 11, no. 1, pp. 116-130, Mar. 2009.

[4] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: challenges and solutions," *EURASIP J. Adv. Signal Process.,* Article ID 381465, 15 pages, 2010.

[5] D. Cabric, A. Tkachenko, and R. Brodersen, "Spectrum sensing measurements of pilot, energy, and collaborative detection," in *Proc. IEEE Military Commun. Conf.,* pp. 1-7, Oct. 2006.

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

[7] P. Jallon, "An algorithm for detection of DVB-T signals based on their second-order statistics," *EURASIP J. Wireless Commun.,* Article ID 538236, 9 pages, 2008.

[8] H. Urkowitz, "Energy detection of unkown deterministic signals," *Proc. IEEE,* vol. 55, no. 4, pp. 523-531, Apr. 1967.

[9] A. Sahai and D. Cabric, "Spectrum sensing: fundamental limits and practical challenges," tutorial in *IEEE DySPAN,* Nov. 2005.

[10] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Topics in Signal Process.,* vol. 2, no. 1, pp. 4-17, Feb. 2008.

[11] B. Zhao, Y. Chen, C. He, and L. Jiang, "Performance analysis of spectrum sensing with multiple primary users," *IEEE Trans. Veh. Technol.,* vol. 61, no. 2, pp. 914-919, Feb. 2012.

[12] O. Besson, S. Kraut, and L. L. Scharf, "Detection of an unknown rank-one component in white noise," *IEEE Trans. Signal Process.,* vol. 54, no. 7, pp. 2835-2839, July 2006.

[13] R. J. Muirhead, *Aspects of Multivariate Statistical Theory.* New York: Wiley, 1982.

[14] A. Taherpour, M. N. Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Trans. Wireless Commun.,* vol. 9, no. 2, pp. 814-823, Feb. 2010.

[15] Y. Zeng, C. L. Koh, and Y.-C. Liang, "Maximum eigenvalue detection: Theory and application," in *Proc. IEEE Int. Conf. Commun.,* pp. 4160-4164, May 2008.

[16] F. Penna, R. Garello, and M. A. Spirito, "Cooperative spectrum sensing based on the limiting eigenvalue ratio distribution in Wishart matrices," *IEEE Commun. Lett.,* vol. 13, issue 7, pp. 507-509, July 2009.

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

[18] P. Wang, J. Fang, N. Han, and H. Li, "Multiantenna-assisted spectrum sensing for cognitive radio," *IEEE Trans. Veh. Technol.,* vol. 59, no. 4, pp. 1791-1800, May 2010.

[19] P. Bianchi, M. Debbah, M. Maida, and J. Najim, "Performance of statistical tests for single-source detection using random matrix theory," *IEEE Trans. Inf. Theory,* vol. 57, no. 4, pp. 2400-2419, Apr. 2011.

[20] B. Nadler, F. Penna, and R. Garello, "Performance of eigenvalue-based signal detectors with known and unknown noise power," in *Proc. IEEE Int. Conf. Commun.,* pp. 1-5, June 2011.

[21] R. Zhang, T. J. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: a GLRT approach," *IEEE Trans. Commun.,* vol. 58, no. 1, pp. 84-88, Jan. 2010.

[22] S. John, "Some optimal multivariate tests," *Biometrika*, vol. 58, no. 1, pp. 123-127, Apr. 1971.

[23] N. Sugiura, "Locally best invariant test for sphericity and the limiting distributions," *Ann. Math. Statist.,* vol. 43, no. 4, pp. 1312-1316, Aug. 1972.

[24] Y. Zeng and Y.-C. Liang, "Eigenvalue based spectrum sensing algorithms for cognitive radio," *IEEE Trans. Commun.,* vol. 57, no. 6, pp. 1784-1793, June 2009.

[25] M. Matthaiou, M. R. McKay, P. J. Smith, and J. A. Nossek, "On the condition number distribution of complex Wishart matrices," *IEEE Trans. Commun.,* vol. 58, no. 6, pp. 1705-1717, June 2010.

[26] A. Kortun, T. Ratnarajah, M. Sellathurai, C. Zhong, and C. B. Papadias, "On the performance of eigenvalue-based cooperative spectrum sensing for cognitive radio," *IEEE J. Sel. Topics in Signal Process.,* vol. 5, no. 1, pp. 49-55, Feb. 2011.

[27] IEEE Standard for Information Technology – Telecommunications and information exchange between systems Wireless Regional Area Networks (WRAN) – Specific requirements, Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and Procedures for Operation in the TV Bands, IEEE Std 802.22, 2011.

[28] C. Zhong, M. R. McKay, T. Ratnarajah, and K.-K. Wong, "Distribution of the Demmel condition number of Wishart matrices," *IEEE Trans. Commun.,* vol. 59, no. 5, pp. 1309-1320, May 2011.

[29] S. Dikmese and M. Renfors, "Performance analysis of eigenvalue based spectrum sensing under frequency selective channels," in *Proc. Int. Conf. Cognitive Radio Oriented Wireless Net. Commun.,* June 2012.

[30] D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank-P signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Trans. Signal Process.,* vol. 59, no. 8, pp. 3764-3774, Aug. 2011.

[31] S. N. Roy, "On a heuristic method of test construction and its use in multivariate analysis," *Ann. Math. Statist.,* vol. 24, no. 2, pp. 220-238, June 1953.

[32] B. Nadler and I. M. Johnstone, "Detection performance of Roy's largest root test when the noise covariance matrix is arbitrary," in *Proc. IEEE Stat. Signal Process. Conf.,* pp. 681-684, June 2011.

[33] B. Nadler and I. M. Johnstone, "On the distribution of Roy's largest root test in MANOVA and in signal detection in noise," Department of Statistics, Stanford University, Technical Report No. 2011-04, 2011.

[34] S. S. Wilks, "Certain generalizations in the analysis of variance," *Biometrika*, vol. 24, no. 3/4, pp. 471-494, 1932.

[35] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis.* New York: Wiley, 2003.

[36] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.,* vol. 35, no. 2, pp. 475-501, 1964.

[37] J. Wishart, "The generalized product moment distribution in samples from a normal multivariate population," *Biometrika,* A20, 32-52, 1928.

[38] A. M. Mathai, *Jacobians of Matrix Transformations and Functions of Matrix Arguments.* Singapore: World Scientific, 1997.

[39] P. Graczyk, G. Letac, and H. Massam, "The complex Wishart distribution and the symmetric group," *Ann. Statist.,* vol. 31, no. 1, pp. 287-309, 2003.

[40] D. Maiwald and D. Kraus, "Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices," *IEE Proc.-Radar, Sonar Navig.,* vol. 147, no. 4, pp. 162-168, Aug. 2000.

[41] H. Cramér, *Mathematical Methods of Statistics.* Princeton: Princeton University Press, 1999.

[42] H. Solomon and M. Stephens "Approximations to density functions using Pearson curves," *J. Amer. Statist. Assoc.,* vol. 73, pp. 153-160, 1978.

[43] H. E. Daniels, "Saddlepoint approximations in statistics," *Ann. Math. Statist.,* vol. 25, pp. 631-650, 1954.

[44] H. Hochstadt, *Special Functions of Mathematical Physics.* New York: Holt, Rinehart and Winston, 1961.

[45] H. T. Ha, *Advances in Moment-Based Density Approximation Methods.* Ph.D thesis, University of Western Ontario, 2006.

[46] R. J. Boik, "Algorithm AS 284: Null distribution of a statistics for testing sphericity and additivity: a Jacobi polynomial expansion," *J. Roy. Statist. Soc. Ser. C,* vol. 42, no. 3, pp. 567-576, 1993.

[47] Z. Zheng, L. Wei, J. Hämäläinen, and O. Tirkkonen, "Approximation to distribution of product of random variables using orthogonal polynomials for lognormal density," *IEEE Commun. Lett.,* vol. 16, no. 12, pp. 2028-2031, Dec. 2012.

[48] Y. Chen and M. R. McKay, "Coulumb fluid, Painleve transcendents and the information theory of MIMO system," *IEEE Trans. Inf. Theory,* vol. 58, no. 7, pp. 4594-4634, July 2012.

[49] G. Fikioris, "Integral evaluation using the Mellin transform and generalized hypergeometric functions: Tutorial and applications to antenna problems," *IEEE Trans. Antennas Propag.,* vol. 54, no. 12, pp. 3895-3907, Dec. 2006.

[50] S. Ahmed, L.-L. Yang, and L. Hanzo, "Mellin-transform-based performance analysis of FFH M-ary FSK using product combining for combatting partial-band noise jamming," *IEEE Trans. Veh. Technol.,* vol. 57, no. 5, pp. 2757-2765, Sept. 2008.

[51] M. D. Renzo, F. Graziosi, and F. Santucci, "Further results on the approximation of log-normal power sum via Pearson type IV distribution: A general formula for log-moments computation," *IEEE Trans. Commun.,* vol. 57, no. 4, pp. 893-898, Apr. 2009.

[52] M. L. Mehta, *Random Matrices.* 3rd Edition, London: Elsevier, 2004.

[53] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications.* Delft: Now Publishers, 2004.

[54] G. Akemann, J. Baik, P. Di Francesco (Eds.), *The Oxford Handbook of Random Matrix Theory.* Oxford: Oxford University Press, 2011.

[55] K. Johansson, "Shape fluctuations and random matrices," *Commun. Math. Phys.,* vol. 209, pp. 437-476, 2000.

[56] C. Tracy and H. Widom, "Level-spacing distributions and the Airy kernel," *Commun. Math. Phys.* vol. 159, pp. 151-174, 1994.

[57] I. Johnstone, "On the distribution of the largest eigenvalue in principal components analysis," *Ann. Statist.,* vol. 29, no. 2, pp. 295-327, 2001.

[58] C. Tracy and H. Widom, "Distribution functions for largest eigenvalues and their applications," in *Proc. of ICM,* vol. 1, pp. 587-596, 2002.

[59] P. O. Perry, I. M. Johnstone, Z. Ma, and M. Shahram, "RMTstat: Distributions and Statistics from Random Matrix Theory," 2009, R software package version 0.1.

[60] F. Bornemann, "On the numerical evaluation of distributions in random matrix theory: A review," *Markov Processes Relat. Fields,* vol. 16, pp. 803-866, 2010.

[61] C. G. Khatri, "Distribution of the largest or the smallest characteristic root under null hyperthesis concerning complex multivariate normal populations," *Ann. Math. Statist.,* vol. 35, pp. 1807-1810, 1964.

[62] P. A. Dighe, R. K. Mallik, and S. R. Jamuar, "Analysis of trasmit-receive diversity in Rayleigh fading," *IEEE Trans. Commun.,* vol. 51, no. 4, pp. 694-703, Apr. 2003.

[63] M. Kang and M. S. Alouini, "Largest eigenvalue of complex wishart matrices and performance analysis of MIMO MRC systems," *IEEE J. Selec. Areas Commun.,* vol. 21, no. 3, pp. 418-426, Apr. 2003.

[64] C. G. Khatri, "Non-central distribution of i-th largest characteristic roots of three matrices concerning complex multivariate normal populations," *Ann. Inst. Statist. Math.,* vol. 21 pp. 23-32, 1969.

[65] Z. D. Bai, and J. F. Yao, "Central limit theorems for eigenvalues in a spiked population model," *Ann. Inst. H. Poincaré,* vol. 44, no. 3, pp. 447-474, 2008.

[66] A. Edelman and B. D. Sutton, "Tails of condition number distributions," *SIAM J. Matrix Anal. Appl.,* vol. 27, 2005.

[67] O. Besson and L. L. Scharf, “CFAR matched direction detector,” *IEEE Trans. Signal Process.,* vol. 54, no. 7, pp. 2840-2844, July 2006.

[68] B. Nadler, “On the distribution of the ratio of the largest eigenvalue to the trace of a Wishart matrix,” *J. Multivariate Anal.,* vol. 102, pp. 363-371, Feb. 2011.

[69] A. Maaref and S. Aissa, “Closed-form expressions for the outage and ergodic Shannon capacity of MIMO MRC systems,” *IEEE Trans. Commun.,* vol. 53, no. 7, pp. 1092-1095, Jul. 2005.

[70] F. J. Schurmann, P. R. Krishnaiah, and A. K. Chattopadhyay, “On the distributions of the ratios of the extreme roots to the trace of the Wishart matrix,” *J. Multivariate Anal.,* vol. 3, pp. 445-453, 1973.

[71] J. W. Mauchly, “Significance test for sphericity of a normal n-variate distribution,” *Ann. Math. Statist.,* vol. 11, no. 2, pp. 204-209, June 1940.

[72] B. N. Nagarsenker and M. M. Das, “Exact Distribution of sphericity criterion in the complex case and its percentage points,” *Comm. Statist. Theory Methods,* 4(4), pp. 363-374, 1975.

[73] D. K. Nagar, S. K. Jain and A. K. Gupta “Distribution of LRC for testing sphericity of a complex multivariate Gaussian model,” *Internat. J. Math. & Math. Sci.,* vol. 8, no. 3, pp. 555-562, 1985.

[74] P. C. Consul, “The exact distributions of likelihood criteria for different hypotheses,” *Multivariate Analysis 2.* New York: Academic Press, 1969.

[75] D. B. Williams and D. H. Johnson, “Using the sphericity test for source detection with narrow-band passive arrays,” *IEEE Trans. Acoust. Speech Signal Process.,* vol. 38, no. 11, pp. 2008-2014, Nov. 1990.

[76] K. C. S. Pillai and B. N. Nagarsenker, “On the distribution of the sphericity test criterion in classical and complex normal populations having unknown covariance matrices,” *Ann. Math. Statist.,*, vol. 42, no. 2, pp. 764-767, Apr. 1971.

[77] C. G. Khatri and M. S. Srivastava, “On exact non-null distributions of likelihood ratio criteria for sphericity test and equality of two covariance matrices,” *Sankhyā*, vol. 33, no. 2, pp. 201-206, June 1971.

[78] S. John, "The distribution of a statistic used for testing sphericity of normal distributions," *Biometrika,* vol. 59, no. 1, pp. 169-173, Apr. 1972.

[79] S. John, "Fitting sampling distribution agreeing in support and moments and tables of critical values of sphericity criterion," *J. Multivariate Anal.,* vol. 6, pp. 601-607, 1976.

[80] H. Nagao, "On some test criteria for covariance matrix," *Ann. Statist.,* vol. 1, no. 4, pp. 700-709, 1973.

[81] O. Ledoit and M. Wolf, "Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size," *Ann. Statist.,* vol. 30, no. 4, pp. 1081-1102, 2002.

[82] H. Nagao and M. S. Srivastava, "On the distribution of some test criteria for a covariance matrix under local alternatives and bootstrap approximations," *J. Multivariate Anal.,* vol. 43, pp. 331-350, 1992.

[83] C. S. Park and K. B. Lee, "Statistical multimode transmit antenna selection for limited feedback MIMO systems," *IEEE Trans. Wireless Commun.,* vol. 7, no. 11, pp. 4432-4438, Nov. 2008.

[84] A. Edelman, "On the distribution of a scaled condition number," *Math. Comp.,* vol. 58, pp. 185-190, 1992.

[85] T. Pham-Gia, "Exact distribution of the generalized Wilks's statistic and applications," *J. Multivariate Anal.,*, vol. 99, pp. 1698-1716, 2008.

[86] A. K. Gupta, "Distribution of Wilks' likelihood-ratio criterion in the complex case," *Ann. Inst. Statist. Math.,*, vol. 23, no. 1, pp. 77-87, 1971.

[87] B. N. Nagarsenker and J. Suniaga, "Distributions of a class of statistics useful in multivariate analysis," *J. Amer. Statist. Assoc.,*, vol. 78, no. 382, pp. 472-475, 1983.

[88] A. Sonnenschein and P. M. Fishman, "Radiometric detection of spread spectrum signals in noise of uncertain power," *IEEE Trans. Aerosp. Electron. Syst.,* vol. 28, pp. 654-660, July 1992.

[89] FCC, "In the matter of unlicensed operation in the TV broadcast bands: second memorandum opinion and order, Federal Communications Commission," FCC 10-174, Sept. 2010.

[90] H.-S. Chen and W. Gao, "Spectrum sensing for TV white space in north America," *IEEE J. Sel. Areas Commun.,* vol. 29. no. 2, pp. 316-326, Feb. 2011.

[91] A. P. Grieve, "Tests of sphericity of normal distributions and the analysis of repeated measures designs," *Psychometrika,* vol. 49, no. 2, pp. 257-267, Jun. 1984.

[92] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.,* vol. 48, no. 3, pp. 502-513, Mar. 2000.

[93] K. C. S. Pillai and K. Jayachandran, "Power comparisons of tests of equality of two covariance matrices based on four criteria," *Biometrika*, vol. 55, no. 2, pp. 335-342, 1968.

[94] T. Sakata, "Likelihood ratio test for one-sided hypothesis of covariance matrices of two normal populations," *Comm. Statist. Theory Methods,* 16 (11), pp. 3157-3168, 1987.

Aalto-DD 145/2013

ISBN 978-952-60-5334-9
ISBN 978-952-60-5335-6 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934
ISSN 1799-4942 (pdf)

Aalto University
School of Electrical Engineering
Department of Communications and Networking
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS