Western University Scholarship@Western

Electronic Thesis and Dissertation Repository

11-14-2013 12:00 AM

Spectrum Sensing and Multiple Access Schemes for Cognitive Radio Networks

Oscar G. Filio Rodriguez The University of Western Ontario

Supervisor Prof. Serguei Primak The University of Western Ontario

Graduate Program in Electrical and Computer Engineering A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy © Oscar G. Filio Rodriguez 2013

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Systems and Communications Commons

Recommended Citation

Filio Rodriguez, Oscar G., "Spectrum Sensing and Multiple Access Schemes for Cognitive Radio Networks" (2013). *Electronic Thesis and Dissertation Repository*. 1712. https://ir.lib.uwo.ca/etd/1712

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlswadmin@uwo.ca.

Spectrum Sensing and Multiple Access Schemes for Cognitive Radio Networks

(Thesis format: Monograph)

by

Oscar G. Filio Rodríguez

Graduate Program in Electrical and Computer Engineering Dept. of Electrical and Computer Engineering

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

School of Graduate and Postdoctoral Studies The University of Western Ontario London, Ontario, Canada

© Oscar Filio Rodríguez

Abstract

Increasing demands on the radio spectrum have driven wireless engineers to rethink approaches by which devices should access this natural, and arguably scarce, resource. *Cognitive Radio* (CR) has arisen as a new wireless communication paradigm aimed at solving the spectrum underutilization problem. In this thesis, we explore a novel variety of techniques aimed at spectrum sensing which serves as a fundamental mechanism to find unused portions of the electromagnetic spectrum.

We present several spectrum sensing methods based on multiple antennas and evaluate their receiving operating characteristics. We study a cyclostationary feature detection technique by means of multiple cyclic frequencies. We make use of a spectrum sensing method called sequential analysis that allows us to significantly decrease the time needed for detecting the presence of a licensed user. We extend this scheme allowing each CR user to perform the sequential analysis algorithm and send their local decision to a fusion centre. This enables for an average faster and more accurate detection.

We present an original technique for accounting for spatial and temporal correlation influence in spectrum sensing. This reflects on the impact of the scattering environment on detection methods using multiple antennas. The approach is based on the scattering geometry and resulting correlation properties of the received signal at each CR device.

Finally, the problem of spectrum sharing for CR networks is addressed in order to take advantage of the detected unused frequency bands. We proposed a new multiple access scheme based on the Game Theory. We examine the scenario where a random number of CR users (considered as players) compete to access the radio spectrum. We calculate the optimal probability of transmission which maximizes the CR throughput along with the minimum harm caused to the licensed users' performance.

Keywords: Cognitive Radio, Dynamic Spectrum Access, Primary User, Secondary User, Spectrum Sensing, Sequential Probability Ratio Test, Sequential Analysis, Multi antenna Detectors, Game Theory, Nash Equilibrium.

Acknowledgements

To my brilliant father and best friend Professor Ernesto Filio López who has always been my role model and the most important voice inside me. You taught me how to be a man.

To my lovely mother Carmen Rodríguez Pedroza for being my muse and my inspiration. Your infinite love gave me the courage and strength to overcome anything.

To my intelligent and pretty sister Gina Filio Rodríguez for being the best sibling ever. Your smile and talent are the two most awesome gifts to the world and science respectively.

To my beautiful and talented fiancée Yadira Tejeda for being the love of my life, my pillar and ultimate reason for finishing this.

To the rest of my beloved family that supported and loved me during this whole academic journey. You made me who I am; Ernesto "Cachorro" Filio, Alfredo "El Zayas" Tirado, Clarita Tejeda and Layla Filio.

To Professor Serguei Primak for accepting me as his PhD student at the University of Western Ontario and providing me with his vast knowledge and guidance.

To Professor Valeri Kontorovich who has been my supervisor, my friend and most important, my academic father.

To my awesome friends whom stood by me and never let me down; Dan "Bro" Dechene, Daniel Wallace, Elena "Sis" Uchiteleva, Marco "Fratello" Luccini, Maysam Mirahmadi, Salvador "Krnalin" Escobedo.

To the ECE department staff Sandra Vilovksi, Chris Marriot for being so patience and helpful with me, and professors Abdallah Shami, Xianbin Wang and Lyudmil Marinov whom provided me with priceless academic discussions, advices and wisdom. To all my friends who provided me with laughs and beer when I needed them the most (which probably made this PhD last way longer than what it supposed to); Abe Escoto, Aidin Reyhani, Arturo Araujo, Bibianita Valencia, Betty Reyes, Brigitte Kok Madsen, Duane Jacques, Elham Dolatabi, Estefania Ruiz, Ferdinando Crapulli, Fernando Guiza, Fernando Ramos-Alarcón, Gejie Liu, Gustavo Ramirez, Jamie Kamula, Jay Nadeu, Jessica Dechene, Laura Ríos, Lise Doucette, Mariel Rivas, Mohamad Kalil, Mohamed Abusharkh, Nazi Torabi, Pierangelo Gobbo, Rainor "Corazón" Bode, Reyna Gómez, Sanaz Golriz, Siamack Ghadimi.

Finally, to the National Council of Science and Technology (CONACYT), México for providing me with all the necessary funding to complete with this degree.

Thank you and live long and prosper.

Oscar G. Filio Rodríguez

A la dama más bella de este mundo, mi Madre. Sin tu amor, comprensión y sabiduría, esta aventura no habría sido posible.

Table of Contents

\mathbf{A}	bstra	ct	$\ldots \ldots $
A	cknov	vledge	ments
D	edica	tion .	iv
Ta	able c	of Cont	$tents \ldots \ldots \ldots v$
Li	st of	Tables	;
Li	st of	Figure	es x
A	crony	rms .	xiii
1	Intr	oducti	on to Cognitive Radio Networks
	1.1	Introd	uction
	1.2	Funda	mentals $\ldots \ldots 3$
		1.2.1	Cognitive Radio Description
		1.2.2	Interference Temperature
		1.2.3	Cognitive Radio Tasks
		1.2.4	Network Architecture
	1.3	Spectr	um Sensing and Analysis
		1.3.1	Transmitter detection
		1.3.2	Cooperative detection
		1.3.3	Other techniques
	1.4	Dynan	nic Spectrum Allocation and Sharing
		1.4.1	Medium Access Control in CR Networks
	1.5	Thesis	Outline
	1.6	Contri	butions \ldots \ldots \ldots \ldots 16
		1.6.1	Contributions of Chapter 2 16
		1.6.2	Contributions of Chapter 3 16
		1.6.3	Contributions of Chapter 4 17
		1.6.4	Contributions of Chapter 5

2	Spe	ctrum	Sensing for Cognitive Radio Networks	18
	2.1	Introd	uction	18
	2.2	Multi-	antenna based Spectrum Sensing	19
		2.2.1	Signal Model	19
		2.2.2	Detection Algorithms	21
		2.2.3	Receiver Operating Characteristics	27
		2.2.4	Performance comparison	28
	2.3	Cyclic	Frequency Detection, Diversity Combining Approach	31
		2.3.1	Generalized Gaussian (Klovsky-Middelton) Channel Model	32
		2.3.2	Single User Multiple Cyclic Frequency Detection	34
		2.3.3	Statiscally Independent Virtual Branches with Flat Generalized	
			Gaussian Fading	37
		2.3.4	Fully Correlated Virtual Branches (Flat Fading) in GG Channel	39
		2.3.5	Special Case of Covariance Matrix for Correlated Branches at	
			the Quadratic Incoherent Addition Algorithm	41
		2.3.6	Suboptimal algorithms and their Noise Immunity	44
	2.4	Couple	ed Dynamical Systems Approach	49
		2.4.1	Coupled Dynamical System Model	50
		2.4.2	Fading Generating Algorithm	52
		2.4.3	Autoregressive Model	52
		2.4.4	Distributed Estimation Results	54
		2.4.5	Networks with Complex Topology	56
		2.4.6	Networks dynamics using a graph Laplacian presentation	58
		2.4.7	Synchronization Mode	61
		2.4.8	Simulation Results	61
3	Sea	uential	Analysis Detection in Cognitive Badio Networks	63
U	31	Introd	uction	64
	3.2	Seque	ntial Analysis of A Wald	65
	0.2	3 2 1	Wald Test for Complex Bandom Variables	68
		322	Average Number of Samples	70
	33	Seque	ntial Probability Batio Test for Partially Coherent Channels	73
	0.0	3 3 1	Average Likelihood Batio	74
	34	Cumu	lant Analysis of the PDF of the Bandom Time of Sequential	11
	0.1	Analys	sis	76
	35	Ontim	al Fusion Bule for distributed Wald detectors	81
	3.6	Dual	Sequential Spectrum Batio Test	87
	0.0	361	Optimization of Decision Time	91
		3.6.2	Numerical Results	93
	37	Conch	usions	96
	0.1	Conten		50

4	Imp	impact of Scattering Environment on Spectrum Sensing for Multi		
	Ant	cenna Detectors		
	4.1	Introduction		
	4.2	Signal Model		
	4.3	Pilot assisted detection in SIMO configuration		
		4.3.1 Constant Independent Channels		
		4.3.2 Performance of the Estimator-Correlator detector in flat block		
		fading (Constant Spatially Correlated Channel)		
		4.3.3 Independent SIMO Channels with Temporal Correlation 106		
		4.3.4 SIMO Channel with Separable Spatial and Temporal Correlation108		
	4.4	Examples and Simulation 109		
		4.4.1 Correlation models		
	4.5	Space-Time Processing Trade-Off 114		
	4.6	Conclusion		
5	Мл	Itiple Access Games for Cognitive Radio Networks 116		
9	5 1	Noncooperative games and Nash equilibrium		
	0.1	5.1.1 Nash Equilibrium 110		
		5.1.2 Mixed Strategies 120		
		5.1.2 Winted bitallegies $1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.$		
		5.1.6 Existence of Rash Equilibrium selection 122		
		5.1.5 The Prisoners' Dilemma 124		
		5.1.6 Application of the Prisoners' Dilemma: The DCF Game 125		
	5.2	Random Access Games 128		
	5.2	Poisson Games 130		
	0.0	5.3.1 Nash Equilibrium in Poisson Games 132		
		5.3.2 Examples and Motivation 132		
	5.4	System Model and Corresponding Game		
	0.1	5.4.1 Poisson Game. Single Type of Players		
		5.4.2 A Poisson Game with Two Types of Players		
	5.5	Game Model with Primary User Activity		
	5.6	Conclusion		
c	The	105		
0	$I n \epsilon$	Thesis Summary and Future Work		
	0.1 6 9	Thesis Summary 105 Future Work 169		
	0.2	Future WOIK 108 6.2.1 Conditional Came for Spectrum Sensing 160		
		6.2.2 Down Allocation of a Non Construction Course		
		6.2.2 Fower Allocation as a Non-Cooperative Game		
		6.2.4 Dadia Decourse Compatition Decoder Starbartic Learning Compatition 170		
		0.2.4 Radio Resource Competition Based on Stochastic Learning Games 170		

Appendices

A	Collaborative Spectrum Sensing with Censoring	187
В	Performance Derivation of Data Fusion Rule	195
\mathbf{C}	Curriculum Vitae	198

List of Tables

3.1	Energy savings considering the optimum time for decision (τ_o)	96
$5.1 \\ 5.2$	Example of Game Theory components in CR networks Categories of noncooperative games and corresponding equilibria	118 119

List of Figures

1.1	Shared Spectrum Company measurements, Dublin, April 2007. (Reproduced from <i>http://www.sharedspectrum.com</i>)
1.2	Spectrum hole concept. \ldots \ldots \ldots \ldots \ldots 3
1.3	Interference Temperature Concept
1.4	Cognitive Cvcle.
1.5	Network architecture of the dynamic spectrum sharing in CR 7
1.6	Classification of spectrum sensing techniques
1.7	Energy detector scheme
1.8	Transmitter detection problem: (a) Receiver uncertainty and (b) shad-
	owing uncertainty
2.1	Energy detector performance $(N = 25)$
2.2	Threshold levels λ for different false alarm probabilities (N = 25) 23
2.3	Receiver Operating Characteristics for a DC level in AWGN $(d^2 = 1)$. 29
2.4	ROC of the presented detection methods
2.5	Probability of miss detection, $eq.(2.43)$ is shown in solid line while
	eq.(2.48) is shown in the dotted line $(\beta_i^2 = 0.1, P_{fa} = 0.001, q_i^2 = 2)$. 40
2.6	Probability of miss detection, eq. (2.43) is shown in solid line while
~ -	eq.(2.48) is shown in the dotted line $(\beta_i^2 = 0.5, P_{fa} = 0.1, q_i^2 = 2)$ 40
2.7	Probability of miss detection, eq. (2.43) is shown in solid line while
0.0	eq.(2.48) is shown in the dotted line $(\beta_i^2 = 0.1, P_{fa} = 0.001, q_i^2 = 4)$. 41
2.8	Probability of miss detection comparison for the optimum (eq. (2.45)),
2.0	and quasi optimum (eq. (2.71)) cases. \ldots \ldots 48
2.9	Dynamical System Node Architecture
2.10	Fading Channel Simulator 53
2.11	Bessel Autocorrelation
2.12	Exponential Autocorrelation
2.13	Evolution of frequencies (local decisions) in time $(K = 0.3, 0.5, 0.6)$. 55 Evolution of frequencies (local decisions) in time $(K = 0.3, 0.5, 0.6)$.
2.14	Evolution of frequencies (local decisions) in time $(K = 0.9, 1.0)$ 56
2.15	Evolution of frequencies (local decisions) in time ($K = 4.0, 5.0$) 57
2.10	Rayleigh magnitude $\ldots \ldots \ldots$
2.17	Evolution of frequencies (local decisions) in time ($K = 1.0, 1.2$)
2.18	NICIAII IIIagiintude 60 Dartially connected network 60
2.19	$ \begin{array}{c} ratually conflected fletwork \dots \dots$
$\angle.20$	Dependence of bounds Λ_s and Λ_{ns} on node density 02

3.1	Sequential Detection	67
3.2	Comparison of Neyman-Pearson Test and Sequential Probability Ratio	
	Test $(P_{FA} = 0.1, P_D = 0.9)$	69
3.3	Average Number of Samples for Detection in Sequential Analysis (P_D =	
	$0.9). \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots $	71
3.4	Approximation of decision time using Wald Distribution $(P_D = 0.9)$.	72
3.5	Impact of coherency on the average number of samples	76
3.6	Variance values for the time decision pdf in sequential analysis	8
3.7	System model of data fusion system	82
3.8	Data fusion scheme considering sequential analysis decision from each sensor $(P_{M,l} = 0.3, P_{F,l} = 0.1)$.	84
3.9	Data fusion scheme considering sequential analysis decision from each sensor $(P_{M,l} = 0.1, P_{F,l} = 0.3)$.	8
3.10	Total Error Criteria	80
3.11	System Model	88
3.12	System Model for Decision.	90
3.13	Approximation of $p(Y_k; \mathcal{H}_1)$ for $L = 5$, $SNR = 0dB$, $P_{md} = 0.2$, $P_{fa} =$	
	0.1	92
3.14	Optimum τ_0 for different number of sensors (SNR = -25dB)	94
3.15	Average Number of Samples (ANS) to make a decision at the FC	9.
3.16	Optimum τ vs SNRdB	9
3.17	Energy used a DSPRT scheme. The 'o' mark shows the optimum time	
	for the FC to make a decision while the 'x' mark shows the time and	-
	energy used when the FC waits for all sensors to send their decision	96
11	System Model	100
4.1	Filtered Observations	10
4.2 // 3	Theoretical BOC and its simulation of the constant independent chan-	102
т.0	nel under AWGN Solid lines - theory x-lines - simulation	10
44	BOC approximation vs. simulation results ($\alpha = 0.8$). Solid lines -	10,
1.1	theory x-lines - simulation	10'
45	Comparison of Isotropic vs Nonisotropic scattering	11'
4.6	Eigenvalues behavior of B_c temporal correlation matrix for nonisotropic	111
1.0	scattering $(N = 10 \ \mu = 0 \text{ and } f_A = 50 Hz)$	111
$4 \ 7$	Effect of correlation between antennas in the probability of detection	11
4.8	BOC approximation for the estimator correlator considering the tem-	110
1.0	noral correlation for isotropic and nonisotropic scattering ($\kappa = 0$ and	
	$\kappa = 10$ respectively)	11.
	$n_{\nu} = 10$ respectively). \dots	TT,
5.1	Strategic form Representation of the Paper-Scissors-Rock Game	12
5.2	The Prisoners' Dilemma	12
5.3	The DCF game with two nodes	12°
5.4	Average utility for Poisson Game in Example 1 ($\lambda = 15$)	13

5.5	Pareto Optimality Utilities for Example 1	135
5.6	Average utility ($\lambda = 15, P_2 = 0.6, K_{\max_1} = 7, K_{\max_2} = 5$ and $r_1 = 0.3$).	137
5.7	Average utility ($\lambda = 15, K_{\max_1} = 7, K_{\max_2} = 5 \text{ and } r_1 = 0.3$)	137
5.8	Strategic form of multiple access game $(N = 2)$.	138
5.9	Penalization coefficient α .	141
5.10	Approximation of optimal probability P_{opt} considering $N = 20$ and	
	different values of K_{max} .	144
5.11	Approximation of optimal probability P_{opt} with $\lambda = 20$ and different	
	values of K_{\max} .	148
5.12	Probability of non-collision comparison. x - axis indicates N or λ	
	depending whether the number of players N is fixed and known or it	
	is a random Poisson variable with average number of players λ	149
5.13	Probability of non-collision comparison between the fixed number of	
	users scheme and random number of users for different values of K_{max} .	150
5.14	Throughput comparison between binary exponential backoff scheme	
	(solid line), game modelled with complete information (dashed line).	
	and game modeled as a Poisson game (dotted line) for different values	
	of K_{max} .	151
5.15	Approximation of the Pareto frontier for a complete game with two	-
0.10	type of players $(N_1 = 15, N_2 = 10, K_{max_1} = 5, K_{max_2} = 3)$,,	153
5.16	Pareto Frontier $(N_1 = 15, N_2 = 10, K_{max_1} = 5, K_{max_2} = 3)$	154
5.17	Pareto Frontier $(\lambda = 30 \ r_1 = 0.3 \ K_{max_1} = 8 \ K_{max_2} = 5)$	156
5.18	Primary User Activity	157
5 19	Restriction frontier approximation for $N_1 = 25$ $N_2 = 20$	159
5 20	Restriction Frontier and Pareto Frontier for two types of players with	100
0.20	full information $(N_1 - 25, N_2 - 20, \gamma - 0.9)$	162
5 91	Restriction Frontier and Pareto Frontier for Poisson games ($\lambda = 25, \gamma = -25, \gamma = -$	102
0.21	(x = 25, y = 0.0)	163
	0.5).	100
A.1	Probability of miss detection for the WMA. Eqs. (2.43) and (2.48) are	
	shown with a solid line whereas eq. (A.10) is shown in a dotted line.	192
A.2	Probability of miss detection for WMA with different parameters values.	192

A.3 Probability of miss detection for WMA with different parameters values.193

Acronyms

ADC	Analog to Digital Converter	
AGM	Arithmetic to Geometric Mean	
$\mathbf{A}\mathbf{M}$.M Autoregressive Model	
AoA	Angle of Arrival	
AWGN	Additive White Gaussian Noise	
BEB	Binary Exponential Backoff	
BER	Bit Error Rate	
\mathbf{CDF}	Cumulative Distribution Function	
CDI	Channel Distribution Information	
CR	Cognitive Radio	
CSCG	Circularly Symmetric Complex Gaussian	
CSI	Channel State Information	
CSMA	Carrier Sense Multiple Access	
CTS	Clear to Send	
CUSUM	Cumulative Sum Algorithm	
DCF	Distributed Coordination Function	
DIFS	Distributed Inter Frame Space	
DSA	Dynamic Spectrum Access	
EDCA	Enhanced Distributed Channel Access	
\mathbf{FC}	Fusion Center	
FDMA	Frequency Division Multiple Access	
FIFO	First-In, First-Out	
FSMC	Finite State Markov Chain	
GLRT	Generalized Likelihood Ratio Test	
$\mathbf{G}\mathbf{M}$	Geometric Mean	
IEEE	Institute of Electrical and Electronics Engineers	

LLF	Log-likelihood function
LP	Linear Programming
LTE	Long Term Evolution
MAC	Media/Medium Access Control
MDP	Markov Decision Process
MIMO	Multiple-Input, Multiple Output
MLR	Maximum Likelihood Ratio
MPR	Multipacket Reception
NE	Nash Equilibrium
NLP	Non-Linear Programming
NP	Neyman-Pearson
OC	Operation Characteristic
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PDF	Probability Density Function
PHY	Physical
PU	Primary User
\mathbf{QoS}	Quality of Service
RA	Rate Adaptive
RF	Radio Frequeny
RKRL	Radio Knowledge Representation Language
RLC	Radio Link Control
ROC	Receiver Operating Characteristic
RTS	Request to Send
RV	Random Variable
SBS	Secondary Base Station
SC-FDMA	Single Carrier Frequency Division Multiple Access
SDoF	Stochastic Degrees of Freedom
SIMO	Single-Input, Multiple-Output
SINR	Signal to Interference and Noise Ratio
SISO	Single-Input, Single-Output
SNR	Signal to Noise Ratio

A cronyms

SPRT	Sequential Probability Ratio Test
\mathbf{SU}	Secondary User
SVD	Singular Value Decomposition
TDD	Time Division Duplexing

Chapter 1 Introduction to Cognitive Radio Networks

1.1 Introduction

Nowadays, depending on the purpose, the geographical region, the particular carrier, and many other factors, the electromagnetic spectrum assigned to wireless networks is controlled by governmental agencies such as Industry Canada, the Federal Communications Commission (FCC) in the United States¹ or the Federal Commission of Telecommunications in Mexico. Responsibilities of these agencies include, allocating frequencies and call signs, managing the broadcast spectrum, and regulating other technical issues such as interference with electronics equipment. They assign the spectrum to licensed holders, also known as *Primary Users* (PUs) on a long-term basis. There have been several measurements and observations about the current usage of the radio spectrum performed by private and federal organizations [1,2]. In Figure 1.1 [2], we show an example of such measurements where can observe that some frequency bands are largely wasted, as they are unoccupied most of the time; some frequency bands are only partially occupied or are used in a sporadic manner. Finally, the rest of the frequency bands are heavily used. Particularly, it can be seen that frequency bands in the 1500 MHz to 1520 MHz range show significant unoccupied spectrum during the two time intervals when these measurements were obtained. Also, the recent increase in access to the limited spectrum for mobile services, has made it necessary to change the way in which devices are allowed to use the spectrum. The limited available spectrum and the inefficiency in spectrum usage

^{1.} Contrary to popular belief, the Canadian Radio-television and Telecommunications Commission (CRTC) is not completely equivalent to the FCC in the United States. The FCC has additional responsibilities and jurisdictions over technical matters which concerns broadcasting and other aspects of communications.



Figure 1.1: Shared Spectrum Company measurements, Dublin, April 2007. (Reproduced from *http://www.sharedspectrum.com*).

necessitate a new communication paradigm in order to exploit the existing wireless spectrum. The new paradigm aimed to solve these problems is called *Cognitive Radio* (CR) or *Dynamic Spectrum Access* (DSA) [3,4]. CR networks focus on providing high bandwidth to mobile users via heterogeneous wireless architectures and dynamic spectrum techniques. The main idea behind CR networks is to allow users with no spectrum license rights to use temporarily unused licensed spectrum. These users, from now on referred to as Secondary Users (SUs), are capable of changing their transmitter parameters according to the interactions with the operating environment. CR devices differ from conventional radio devices in that CR provides SUs with *cognitive* capability and reconfigurability. Cognitive capability is defined as the ability of the device to sense, understand, and be aware of the conditions related to the surrounding environment, e.g., presence of the PU, information about transmission frequency, bandwidth, transmission power and modulation among others [3,5]. Reconfigurability is the SUs' capacity to make decisions and rapidly adapt their operation parameters accordingly. Because of more flexible and intelligent use of the spectrum in CR, new and novel spectrum management techniques must be developed to address the



Figure 1.2: Spectrum hole concept.

new challenges, specifically those related to spectrum sensing and dynamic spectrum sharing. CR systems do allow coexistence between PUs and SUs; however, PUs will always have priority in using the spectrum. It is the responsibility of the SUs to sense their surroundings in real time and to know whether a PU is transmitting or not. Based on this information, the SUs can decide between transmitting with low power at the same time as the PU, or wait until the PU stops transmitting before using the channel.

1.2 Fundamentals

1.2.1 Cognitive Radio Description

The cognitive radio concept was first introduced in [6], where the main focus was on the Radio Knowledge Representation Language (RKRL) [7]. A few formal definitions of Cognitive Radio exist; the two most complete are given by Haykin and Thomas in [4,8] respectively:

• "Cognitive radio is an intelligent wireless communication system that is aware of its surrounding environment (i.e., outside world), and uses the methodology of understanding-by-building to learn from the environment and adapt its



Figure 1.3: Interference Temperature Concept.

internal states to statistical variations in the incoming RF stimuli by making corresponding changes in certain operating parameters (e.g., transmit-power, carrier-frequency, and modulation strategy) in real-time, with two primary objectives in mind:

-highly reliable communications whenever and wherever needed; -efficient utilization of the radio spectrum."

• "A Cognitive Radio is a radio that can change its transmitter parameters based on interaction with the environment in which it operates."

The ultimate objective of CR is to obtain the best available spectrum band through cognitive capability and reconfigurability. In order to take advantages of CR techniques we must find the unused portions of the spectrum also known as *spectrum holes* or *white spaces* [9]. If these bands are later used by a PU, the CR device has the choice of either moving to another spectrum hole or staying in the same band but adapting its transmission power or modulation scheme in order to avoid the interference to PUs. Figure 1.2 shows the spectrum hole concept. In this figure we can observe the detection of the aforementioned white spaces by real time sensing the the wideband channel followed by the selection of the more suitable frequency bands. Finally, the multiple spectrum access coordination with other SUs who finally vacate the channel when a PU needs to transmit.

1.2.2 Interference Temperature

The determination of the available portion of the spectrum can be made according to different metrics. The traditional approach is to limit the transmitter power of interfering devices, *i.e.*, the transmitted power should be no more than a prescribed noise floor with respect to a certain distance from the transmitter. Nevertheless, due to the increased mobility and variability of radio frequency (RF) emitters, constraining the transmitter power becomes more challenging, since unpredictable new sources of interference may appear. In order to address this issue, the FCC Spectrum Task Force [10] proposed a new metric to assess the interference called *interference temperature*, which enforces an interference limit perceived by receivers. The interference temperature is a measure of the RF power available at a receiving antenna to be delivered to a receiver, reflecting the power generated by other emitters and noise sources. This is depicted in detail in Figure 1.3. The interference temperature is defined as the temperature equivalent to the RF power available at a receiving antenna per unit bandwidth, *i.e.*,

$$T_I(f_c, B) = \frac{P_I(f_c, B)}{kB},$$
(1.1)

where $P_I(f_c, B)$ is the average interference power in Watts centred at the carrier frequency f_c , B is the bandwidth measured in Hertz, and k represents the Boltzmann's constant equal to 1.38×10^{-23} Joules per degree Kelvin. The FCC also defined the concept of *interference temperature limit* as the maximum tolerable interference for a given frequency band at a particular location. Any unlicensed secondary transmitter using this band must guarantee that their transmission plus the existing noise and interference will not exceed the interference temperature limit at a PU. If a regulatory body sets an interference temperature limit T_L for a particular frequency band with bandwidth B, the SUs must keep the average interference below kBT_L . Therefore, the interference temperature serves as a cap on potential RF energy that could appear on a specific band.



Figure 1.4: Cognitive Cycle.

1.2.3 Cognitive Radio Tasks

Figure 1.4, shows a typical CR duty cycle, that presents the major functions that relate to cognitive capability and reconfigurability. The cognitive cycle consists of the following tasks:

- 1. *Spectrum Sensing*: Detects unused spectrum and shares the spectrum without negative interfering with other users.
- 2. *Spectrum Analysis*: Captures the best available spectrum to meet user communication requirements.
- 3. Spectrum Management and Handoff: Enables SUs to choose the best frequency band and hop among multiple bands according to the time varying channel characteristics to meet the different Quality of Service (QoS) requirements.
- 4. *Spectrum Allocation and Sharing*: Provides a fair spectrum scheduling method between coexisting SUs and PUs.

In general, the dynamic use of the spectrum has a negative impact on the performance of conventional communication protocols that were designed for fixed frequency bands. It is important to consider this type of impact when designing CR systems.



Figure 1.5: Network architecture of the dynamic spectrum sharing in CR.

1.2.4 Network Architecture

With the inclusion of SUs in the framework of a wireless communication network, it is logical to assume that they will change the way in which the networks are formed. Moreover, since SUs use the temporarily unused licensed bands owned by the PU, the network architecture of the CR includes both a primary network and a secondary network, as seen in Figure 1.5. A secondary network is composed of a set of decentralized or centralized SUs, *i.e.*, with or without a secondary base station. However, the DSA of SUs is usually controlled and coordinated by a secondary base station (SBS). Both the SUs and the SBS must feature the four tasks of the cognitive cycle. In Figure 1.5, we can see the central network entity called the *spectrum broker*, which coordinates the spectrum usage between two or more secondary networks. It allocates the network resources by collecting the operation information from each secondary network is formed by the PUs and one or more primary base stations. Within the context of CR networks, the presence of the secondary network to the primary



Figure 1.6: Classification of spectrum sensing techniques

network should be imperceptible, *i.e.*, the PUs' transmission should be seamlessly regardless of the SUs. If a secondary network shares the licensed spectrum band with a primary network, in addition to detecting the spectrum white spaces and choosing the best available spectrum band, it must also detect the reappearance of the PUs and direct the secondary transmission to another available band or decrease the transmission power in order to avoid interfering with the primary transmissions.

1.3 Spectrum Sensing and Analysis

A major requirement of CR networks is the ability to detect the spectrum holes. Therefore, spectrum sensing and analysis are the first critical steps toward dynamic spectrum management. The spectrum sensing function enables the cognitive radio to adapt to its environment by detecting such holes. The most efficient known method of detecting spectrum holes is to detect the PUs that are transmitting data within the communication range of a CR user [11]. Generally, the spectrum sensing techniques can be classified as transmitter detection, cooperative detection, and interference-based detection, as shown in Figure 1.6 [12, 13]. Interference-based detection is out of the scope of this dissertation.

1.3.1 Transmitter detection

Transmitter detection approach is based on the detection of signals from a primary transmitter through the local observations by cognitive users. The hypothesis problem can be defined as

$$y(t) = \begin{cases} \mathcal{H}_0 : & n(t) \\ \mathcal{H}_1 : & hs(t) + n(t), \end{cases}$$
(1.2)

where y(t) is the signal received by the cognitive user, s(t) is the transmitted signal from the PU, n(t) is the Additive White Gaussian Noise (AWGN), and h is the amplitude gain of the channel. In eq.(1.2), \mathcal{H}_0 is defined as the null hypothesis, which states that there is no licensed user signal in the analyzed spectrum band. \mathcal{H}_1 is the alternative hypothesis, which indicates that there exists a PU signal. Using three different schemes, it is possible to implement transmitter detection according to the hypotheses model. These schemes are *a*) matched filter detection, *b*) energy detection, and *c*) cyclostationary feature detection [14].

1.3.1.1 Matched filter detection

When the information about the PU signal is known to the cognitive user, the optimal detector in stationary Gaussian noise is the matched filter, because it maximizes the received Signal-to-Noise Ratio (SNR) [15, 16]. Hence, the matched filter can be thought as an upper bound in detection performance. Although the main advantage of the matched filter is that it requires less time to achieve high processing gain due to coherency, it requires a *priori* knowledge of the PU signal such as the modulation type and order, the pulse shape, and the packet format. If this information is not accurate, the matched filter will perform poorly. Most wireless network systems have pilot signals, preambles, synchronization words or spreading codes which can be used for coherent detection.

1.3.1.2 Energy detection

When the receiver cannot gather sufficient information about the PU signal, for instance, if the power of the random Gaussian noise is only known at the receiver, the



Figure 1.7: Energy detector scheme

optimal detector is the energy detector [15,16]. In order to measure the energy of the received signal s(t), the output of bandpass filter with bandwidth W is squared and integrated over the observation interval T. Finally, the output of the integrator, Y, is compared with a threshold λ to decide whether a licensed user is present or not. The scheme is summarized in Figure 1.7. However, the performance of energy detector is susceptible to uncertainty in noise power. To solve this problem, a pilot tone from the primary transmitter can be used to help improve the accuracy of the energy detector. Another shortcoming is that the energy detector cannot differentiate signal types but can only determine the presence of the signal. Thus, the energy detector is prone to false detection triggered by unintended signals.

1.3.1.3 Cyclostationary feature detection

An alternative detection method is the cyclostationary feature detection [17–19]. Modulated signals are in general coupled with sine wave carriers, pulse trains, repeating spreading, hopping sequences or cyclic prefixes, which result in built-in periodicity. These modulated signals are characterized as cyclostationarity because their mean and autocorrelation exhibit periodicity. These features are detected by analyzing a spectral correlation function, which is able to differentiate the noise energy from modulated signal energy. This occurs because the noise is a wide-sense stationary signal with no correlation, while modulated signals are cyclostationary with spectral correlation due to the embedded redundancy of signal periodicity. Therefore, a cyclostationary feature detector can perform better than the energy detector in discriminating against noise due to its resilience to the uncertainty in noise power [19]. However, it is computationally complex and requires significantly long observation



Figure 1.8: Transmitter detection problem: (a) Receiver uncertainty and (b) shadowing uncertainty.

times. For more efficient and reliable performance, the enhanced feature detection scheme combines cyclic spectral analysis with pattern recognition based on neural networks [20]. Distinct features of the received signal are extracted using cyclic spectral analysis and represented by both spectral coherent function and spectral correlation density function. The neural network, then, classifies signals into different modulation types.

1.3.2 Cooperative detection

In primary transmitter detection, it could be assumed that the locations of the primary receivers are unknown due to the absence of signalling between PUs and SUs. Therefore, CR should rely on only weak primary transmitter signals based on the local observation of the SU. However, in most cases, a cognitive radio network is physically separated from the primary network so there is no interaction. Thus, with the transmitter detection, the cognitive radio user cannot avoid the interference due to the lack of the primary receiver's information, as depicted in Figure 1.8-a. Additionally, the transmitter detection model cannot prevent the hidden terminal problem [21]. A CR transmitter can have a good line-of-sight to a receiver, but may not be able to detect the transmitter due to the shadowing, as shown in Figure 1.8-b. Consequently, sensing information from other users is required for a more accurate detection [22]. In

cooperative detection spectrum sensing methods incorporate information from multiple SUs for PU detection [23]. Cooperative detection can be implemented either in a centralized or in a distributed manner. In the centralized method, the CR basestation plays the role of gathering all sensing information from the CR users and detecting the spectrum holes [24]. Distributed solutions require exchange of observations among CR users. Cooperative detection among unlicensed users is theoretically more accurate because the uncertainty introduced by a single user's detection can be minimized [25, 26]. Moreover, the multi-path fading and shadowing effect are the main factors that, in general, degrade the performance of PU detection methods. Cooperative detection schemes mitigate the multi-path fading and shadowing effects, which improve the detection probability in heavily shadowed environments. While cooperative approaches provide more accurate sensing performance, they also cause adverse effects on resource-constrained networks due to the additional operations and overhead traffic. Furthermore, the primary receiver uncertainty problem caused by the lack of the primary receiver location knowledge is still unresolved with cooperative sensing.

1.3.3 Other techniques

Although the previous techniques are the most classical in detection theory, new and novel techniques and variations have been proposed in recent literature.

1.3.3.1 Statistical Covariance-Based Sensing

Generally, the statistical covariance matrices of the received signal and noise are different. It is possible to distinguish the desired signal component from the background noise. The eigenvalues of the covariance matrix of the received signal can also be used for primary detection [27, 28]. This is done by quantizing the ratio of the maximum eigenvalue to the minimum eigenvalue and forming a detection threshold between them. This technique is particularly useful when detecting TV signals since the methods based on statistical covariances are shown to be more robust to noise uncertainty while requiring no *a priori* information of the signal, the channel, and the noise power.

1.3.3.2 Fast Sensing

In the theory of quickest detection based primarily on sequential analysis, a statistical test is performed to detect the change in the distribution of spectrum usage observations as quickly as possible, allowing for agile and robust spectrum sensing [29–31]. The unknown parameters after a PU appears can be estimated using successive refinement, which combines both generalized likelihood ratio and parallel cumulative sum tests. In [32] we proposed cumulant analysis for Dual Sequential Ratio Testing in CR networks, which uses the basis of fast sensing in a cooperative manner, improving significantly the detection performance and minimizing the energy used for detection.

1.3.3.3 Coupled Dynamical-Based Sensing

In [26], Nefedov proposed a novel technique to detect the presence of PUs based on the concept of self-organization of coupled dynamical systems. In this method, a global estimate (or decision) is obtained in a distributed fashion for complex networks without a fusion or centralized control centres. The suggested approach is based on local exchange of information among the nearby nodes within a connected (wireless) network that allows, under certain conditions, a global decision to be reached based on locally available decisions or measurements. The author considers network nodes as local dynamical systems with impulse-like coupling to establish time synchronization among the transmitted packets together with phase-coupling during packet durations to achieve distributed estimations.

1.4 Dynamic Spectrum Allocation and Sharing

Once the spectrum holes are found, the SUs are aware of the spectrum bands available for them to use; nevertheless, the quality and availability of a specific spectrum band may change rapidly due to the PUs' dynamic activity and competition from other SUs. It is therefore important to design new spectrum allocation and sharing policies to address this issue. *Open spectrum sharing* is referred to spectrum sharing among the SUs accessing the unlicensed spectrum band (*e.g.*, the open spectrum sharing in the unlicensed industrial, scientific, and medical (ISM) band). All SUs have the same rights to use the unlicensed band since no user own spectrum licenses. The *hierarchical access model* or licensed spectrum sharing can be divided in two categories: *Spectrum underlay* and *Spectrum overlay* [21, 33].

Spectrum Underlay

Allows SUs to access the channel when PUs are also transmitting as long as the interference at a PUs' receiver lies within the interference-temperature limit. Usually, due to the constrains on transmission power, only shortrange communications are achievable to the SU. Moreover, if PUs transmit all the time, spectrum underlay does not require SUs to perform spectrum sensing.

Spectrum Overlay

In this policy, SUs will use the licensed spectrum only when PUs are not transmitting *i.e.*, SUs need to sense the licensed band and detect the spectrum holes to keep from interfering with PUs.

1.4.1 Medium Access Control in CR Networks

One of the most important aspects to be analyzed in dynamic spectrum allocation and sharing is the concept of Medium Access Control (MAC), which refers to the techniques that control the way the SUs should access the primary or licensed channel. There exist several classical approaches to solve the MAC problem in the literature such as the carrier sense multiple access (CSMA) scheme or the slotted ALOHA [34]. However, since there is no concept of PUs in traditional networks, classic MAC protocols are not concerned with interference caused to PUs. Therefore, one of the main challenges in MAC for CR is to protect the PUs. The MAC protocols for CR networks should support the following two features:

- Collision avoidance amongst SUs: Since different SUs can coexist, collisions can occur if they simultaneously decide to use the same spectrum band, according to their spectrum sensing results. Thus, the MAC protocol should coordinate the spectrum access for different SUs in order to avoid the collisions.
- Interference control and avoidance for PUs: This is the ultimate goal of spectrum sharing in CR networks. There exist two modes for spectrum sharing between SUs and PUs: spectrum overlay and spectrum underlay which were explained before.

In addition to the above essential functions, the MAC layer acts as a bridge between the physical layer and the network layer in CR networks. It can exploit the spectrum sensing results from the physical layer, characterize the channels, and determine the specific channel and instant to access. It can also help the CR network layer decide on the routing path by reporting the characteristic information and listing the available channels. Moreover, the network layer can inform the MAC layer to choose a suitable channel based on the QoS requirement. By designing appropriate access probabilities for the SUs, a good tradeoff can be achieved between spectrum efficiency and fairness. In [35] we proposed a technique to calculate such probabilities based on game theory concepts. We also addressed the problem of a multiple access CR system where the number of users and their types are unknown. The framework is modelled as a non-cooperative Poisson game in which all the players (or SUs) are unaware of the total number of devices participating. In our scheme, failed attempts to transmit (collisions) are penalized, and we calculate the optimum penalization in mixed strategies. We show that this scheme conveys to a Nash equilibrium where a maximum in the possible throughput can be achieved.

1.5 Thesis Outline

This thesis consists of six chapters, including this introduction. The remaining six chapters are organized as follows: Chapter 2 presents three novel methods for spectrum sensing in CR networks using multi antenna systems; Chapter 4 assesses the impact of the scattering environment on the detection performance of multi antenna receivers; Chapter 3 introduces the concept of Sequential Analysis and, a novel detection method is proposed; Chapter 5 provides a comprehensive review of Game Theory concepts applied to CR Networks and introduces a new interpretation of the MAC problem based on game theory for CR networks; and finally, Chapter 6 is devoted to concluding remarks and suggests future directions for extending the current research topic.

1.6 Contributions

The main contributions of each chapter in the thesis are listed below.

1.6.1 Contributions of Chapter 2

- A multi antenna based spectrum sensing approach is analyzed using the Generalized Likelihood Ratio Test (GLRT).
- The concept of optimal incoherent diversity combining of virtual diversity branches is presented in cyclostationary spectrum sensing for single user detection.
- A new detection method based on coupled dynamical systems is presented within the context of complex networks.

1.6.2 Contributions of Chapter 3

- The concepts of Sequential Analysis and Sequential Probability Ratio Test (SPRT) are presented as an alternative to fixed number of samples detection methods such as the Neymann-Pearson test.
- A cumulant analysis for the probability density function of the random time sequential analysis is obtained.
- An optimal fusion rule for distributed detectors using SPRT is obtained.
- A new detection method called Dual SPRT is presented which minimizes the energy used for the detection of the PU.

1.6.3 Contributions of Chapter 4

- An approach for accounting for both spatial and temporal correlation in CR devices equipped with multiple antennas is presented.
- The impact of spatial and/or temporal correlation between antennas in Single Input Multiple Output (SIMO) spectrum sensing is assessed.
- An equivalent number of independent samples is derived based on the scattering geometry and resulting correlation properties of the received signal.

1.6.4 Contributions of Chapter 5

- Some useful tools based on Game Theory are presented to model wireless communications problems.
- The concepts of noncooperative games and Nash Equilibrium are presented and the Prisoners' Dilemma example is used to model the MAC problem in IEEE 802.11e wireless networks.
- The problem of games with population uncertainty is modelled with the use of Poisson games approach.
- A new interpretation for the MAC problem in CR networks is presented for two different types of SUs competing for access to the wireless channel.
- An accurate analytical approximation of the Pareto frontier of the probability of transmission of SUs is provided regardless of their type.
- The impact of the PU based on its activity is considered and the optimal mixed strategies are calculated accordingly.

Chapter 2 Spectrum Sensing for Cognitive Radio Networks

2.1 Introduction

As explained in Chapter 1, an essential component of overlay Cognitive Radio Networks is the ability to sense the spectrum holes (*i.e.*, the absence of a PU transmitting in a specific band at a specific moment). The spectrum sensing function enables the cognitive radio to adapt to its environment and transmit or receive accordingly by detecting such holes. Currently, the most efficient way to detect spectrum holes, is to detect the PUs that are receiving data within the communication range of a SU user. In this chapter we study, analyze and review a few novel spectrum sensing techniques. The chapter is divided in three parts: Firstly, a muti-antenna based spectrum sensing is analyzed using the well-known Generalized Likelihood Ratio Test (GLRT) principle to approach the problem of spectrum sensing in a cognitive radio network. We show that under mild assumptions on the primary signal, the eigenvalued based algorithm performs better than the classical energy detector, although the system suffers from an increase in complexity. Secondly, we show that a cyclostationary spectrum sensing for CR networks, applying multiple cyclic frequencies for a single user detection, can be interpreted (under a few assumptions) in terms of optimal incoherent diversity combining of *virtual diversity branches* or a SIMO radar. Such an approach allows us to propose (analogously to diversity combining) some sub-optimal algorithms which can provide near-optimal characteristics for the NP test in the single user detection scenario. Finally, a detection method based on a coupled dynamical system approach is presented in complex networks where a global decision about the presence of the PU is obtained in a distributed fashion or centralized control centres. The latter is based on the local exchange of information among close nodes in a wireless network which allows the system, under certain conditions, to reach a global decision based on locally-available measurements or decisions.

2.2 Multi-antenna based Spectrum Sensing

In popular works on spectrum sensing for cognitive radio [40], it is usually assumed that a full or partial knowledge on the PU signal characteristics is available such as the channel from the PU to the SU, and/or the noise power level at the CR receiver. These assumptions may limit the applicability of these algorithms in realistic CR scenarios. Blind signal detection for multi-antenna sensors can be used with either no knowledge about the signal to be detected, or in cases where the sensor noise level has been studied based on information-theoretic criteria rather than in the wellknown GLRT principle. In this first section we present a few very interesting and applicable methods which avoid the requirement of prior knowledge about the PU signals or the channels from the PU to the CR [23]. We study and analyze multiantenna based spectrum sensing methods for cognitive radio networks using the GLRT approach. In attempting to sense the presence of a PU, this approach utilizes the eigenvalues calculated from the sample covariance matrix of the received signal vector from multiple antennas at the SU. It is therefore possible to take advantage of the fact that in practice, the PU signals to be detected will either occupy a subspace of dimension strictly smaller than the dimension of the observation space, or have a non-white spatial spectrum. By making some assumptions on the availability of the white noise power at the SU receiver, it is possible to implement two algorithms which improve the standard energy detector performance.

2.2.1 Signal Model

Let us consider a CR terminal sensing the presence of a PU based on a set of N discrete-time vector observations $\mathbf{x}[n]$, n = 0, ..., N - 1. The *i*-th component of
$\mathbf{x}[n]$, denoted by $x_i[n]$, i = 0, ..., M - 1, is the output of the *i*-th antenna, where M is the number of antennas at the CR terminal. For convenience and without loss of generality we define $\mathbf{X} = {\mathbf{x}[0], ..., \mathbf{x}[N-1]}$. Using this let us formulate the hypotheses testing problem as

$$\mathcal{H}_0: \quad \mathbf{x}[n] = \mathbf{w}[n], \qquad n = 0, \dots, N-1$$

$$\mathcal{H}_1: \quad \mathbf{x}[n] = \mathbf{s}[n] + \mathbf{w}[n], \quad n = 0, \dots, N-1$$
(2.1)

where $\mathbf{w}[n]$ is the additive noise at the cognitive radio receiver, modeled as an independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian (CSCG) vector with zero mean and covariance matrix $\sigma^2 \mathbf{I}$, with \mathbf{I} denoting the identity matrix, σ^2 is the noise variance, and $\mathbf{s}[n]$ is the received primary signal to be detected. In the absence of any prior knowledge of the form of $\mathbf{s}[n]$, or any attempt to estimate it, the signal $\mathbf{s}[n]$ is assumed to be an i.i.d. CSCG random vector with zero mean and the covariance matrix denoted by $\mathbf{R}_s = E[\mathbf{s}[n]\mathbf{s}^H[n]]$, where $(\cdot)^H$ denotes the Hermitian transpose. Analogously, we can also define $\mathbf{R}_x = E[\mathbf{x}[n]\mathbf{x}^H[n]]$. As we do not have always the knowledge of \mathbf{R}_x it is possible to use an empirical or sample covariance matrix denoted by

$$\hat{\mathbf{R}}_x = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{x}[n] \mathbf{x}^H[n] = \frac{1}{N} \mathbf{X} \mathbf{X}^H.$$
(2.2)

Assume that it is possible to make an eigenvalue decomposition of eq. (2.2) in such a way that the unitary eigenvector matrix \mathbf{U}_x and diagonal eigenvalue matrix Λ_x in $\hat{\mathbf{R}}_x = \mathbf{U}_x \Lambda_x \mathbf{U}_x^H$ are known at each block of N observations. The algorithm proposed in [27] relies on one of the following conditions to hold:

- \mathbf{R}_s is rank-deficient: In other words, $rank(\mathbf{R}_s) = N_s < M$, the dimension of the received signal space. In this case, the smallest $M - N_s$ eigenvalues of $\hat{\mathbf{R}}_x$ will be approximately equal to the noise variance of σ^2 , while the N_s largest eigenvalues of $\hat{\mathbf{R}}_x$ will be approximately the sum of an eigenvalue of \mathbf{R}_s and σ^2 . These approximations become exact in the limit $N \to \infty$.
- \mathbf{R}_s is *full-rank* but $\mathbf{R}_s \neq \alpha \mathbf{I}$, where α is an arbitrary positive constant: In this

case, $rank(\mathbf{R}_s) = M$, and each eigenvalue of $\hat{\mathbf{R}}_x$ will be approximately the sum of an eigenvalue of \mathbf{R}_s and σ^2 . Under this condition the eigenvalues of \mathbf{R}_s are unequal and hence so are those of $\hat{\mathbf{R}}_x$.

It is important to note that when the primary signal is not present, $\hat{\mathbf{R}}_x \to \sigma^2 \mathbf{I}$ as $N \to \infty$, *i.e.*, $\hat{\mathbf{R}}_x$ is a full-rank diagonal matrix with equal eigenvalues, which is different from $\hat{\mathbf{R}}_x$ when the primary signal is present, provided that one of the above two conditions is satisfied. Consequently, the existence of a primary signal can be directly obtained from $\hat{\mathbf{R}}_x$.

2.2.2 Detection Algorithms

There exists two very well-known methods in detection literature: the energy detector (ED) and the more general estimator-correlator (EC) detector. When the PU signal covariance matrix, \mathbf{R}_s , and the noise variance σ^2 are both known, the Neyman-Pearson scheme, makes the EC detector the optimal in the sense of maximizing the probability of detection P_D given a probability of false alarm (P_{FA}). Now, for the case when $\mathbf{R}_s = \alpha \mathbf{I}$, the optimal solution is the ED. Such detectors are used as a departure point in order to explain the derivations and advantages of the two new algorithms.

2.2.2.1 Energy Detector

This detector computes the energy in the received data and compares it to a predetermined threshold γ . Intuitively, if the signal is present, the energy of the received data increases. This is the most basic type of detectors [15,41]. Equation (2.3) shows the expression of the test statistic for the energy detector in the Neyman-Pearson (NP) criteria.

$$\mathcal{T}_{ED}(\mathbf{X}) = \sum_{n=0}^{N-1} ||\mathbf{x}[n]||^2 \stackrel{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \gamma$$
(2.3)



Figure 2.1: Energy detector performance (N = 25).

where the threshold γ can be calculated in terms of the probability of false alarm (P_{FA}) or the probability of detection (P_D) as [15]

$$\gamma = \sigma^2 Q_{\chi_N^2}^{-1} (P_{FA}),$$

$$\gamma = (1 + \sigma^2) Q_{\chi_N^2}^{-1} (P_D),$$
(2.4)

where $Q_{\chi_N^2}^{-1}(\cdot)$ is the inverse of the right-tail probability of a chi square $Q_{\chi_N^2}$ random variable with N degrees of freedom which can be calculated numerically¹. The performance characteristic of the energy detector is given in Figure 2.1. It can be seen that the performance increases monotonically with the SNR defined as σ_s^2/σ^2 . In Figure 2.2, the threshold λ is depicted for different values of P_{FA} . It can be seen that as P_{FA} gets smaller, the threshold has to increase in order to detect an absent PU.

^{1.} In fact, the equivalent test statistic $\mathcal{T}'(\mathbf{X}) = \frac{1}{N} \sum_{n=0}^{N-1} ||\mathbf{x}[n]||^2$ can be thought of as an estimator of the variance. Comparing this to a threshold, it recognizes that the variance under \mathcal{H}_0 is σ^2 but under \mathcal{H}_1 it increases to $\sigma_s^2 + \sigma^2$.



Figure 2.2: Threshold levels λ for different false alarm probabilities (N = 25).

2.2.2.2 Estimator-Correlator Detector

If the primary signal covariance matrix, \mathbf{R}_s , and the noise variance, σ^2 , are both known, the Neyman-Pearson approach leads to the following *estimator-correlator* detector that is optimal (in the sense of maximizing P_D , at a given P_{FA}) for the hypotheses testing problem in equation (2.1). The test statistic can be expressed as

$$\mathcal{T}_{EC}(\mathbf{X}) = \sum_{n=0}^{N-1} \mathbf{x}^{H}[n] \mathbf{R}_{s} \left(\mathbf{R}_{s} + \sigma^{2} \mathbf{I} \right)^{-1} \mathbf{x}[n] \begin{array}{c} \mathcal{H}_{1} \\ \gtrless \\ \mathcal{H}_{0} \end{array} \gamma,$$
(2.5)

where $(\cdot)^{-1}$ denotes the matrix inverse, while γ is set to provide a desired target P_{FA} . It is worth noting that the energy detector is a special case of the estimatorcorrelator when $\mathbf{R}_s = \alpha \mathbf{I}$. Consequently, the energy detector and the estimatorcorrelator detector can be considered extreme opposites in terms of information about the PU signal. In the ED, no prior information is estimated while performing the decision; however, in the EC, all information is known, and an optimal detection is therefore achieved. In the following we present cases where some information is known and some is somehow estimated.

2.2.2.3 Arithmetic-to-Geometric Mean Detector

Now we present the case when we do not know either the covariance matrix, \mathbf{R}_s , or the noise variance, σ^2 , and we try to estimate them in order to approach the performance given by the EC. If the secondary sensor does not know \mathbf{R}_s and/or σ^2 prior to spectrum sensing, the detection problem becomes a hypotheses testing problem in the presence of uncertain parameters, which is generally known as a composite hypothesis testing. One useful solution for this type of problems is the GLRT, which first obtains the maximum likelihood estimate (MLE) of the unknown parameters under \mathcal{H}_0 and \mathcal{H}_1 as

$$\hat{\theta}_0 = \arg \max_{\theta_0} p(\mathbf{X}|\mathcal{H}_0, \theta_0),$$
$$\hat{\theta}_1 = \arg \max_{\theta_1} p(\mathbf{X}|\mathcal{H}_1, \theta_1),$$

where θ_0 and θ_1 are the set of parameters unknown under \mathcal{H}_0 and \mathcal{H}_1 respectively. Accordingly, the GLRT test statistic can be formed as

$$L_G(\mathbf{X}) = \frac{p(\mathbf{X}|\hat{\theta}_1, \mathcal{H}_1)}{p(\mathbf{X}|\hat{\theta}_0, \mathcal{H}_0)} \stackrel{\mathcal{H}_1}{\gtrless} \gamma.$$
(2.6)

If we do not know either \mathbf{R}_s or σ^2 , the log-likelihood function (LLF) under \mathcal{H}_0 of the unknown parameter σ^2 can be expressed as

$$\ln p(\mathbf{X}|\hat{\theta}_0, \mathcal{H}_0) = -\frac{MN}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{n=0}^{N-1}||\mathbf{x}[n]||^2.$$
(2.7)

The MLE of σ^2 under \mathcal{H}_0 minimizes eq. (2.7), and is given by

$$\hat{\sigma_0}^2 = \frac{1}{MN} \sum_{n=0}^{N-1} ||\mathbf{x}[n]||^2,$$
(2.8)

which upon substitution into eq. (2.7) yields

$$\ln p(\mathbf{X}|\hat{\theta}_0, \mathcal{H}_0) = -\frac{MN}{2} \left[\ln \left(\frac{2\pi}{MN} \sum_{n=0}^{N-1} ||\mathbf{x}[n]||^2 \right) + 1 \right].$$
 (2.9)

Similarly, the LLF under \mathcal{H}_1 for both unknown \mathbf{R}_s and σ^2 , can be expressed as

$$\ln p(\mathbf{X}|\hat{\theta}_1, \mathcal{H}_0) = -\frac{MN}{2}\ln(2\pi) - \frac{N}{2}\ln(\det(\mathbf{R}_x)) - \frac{1}{2}\sum_{n=0}^{N-1} \mathbf{x}^H[n]\mathbf{R}_x^{-1}\mathbf{x}[n], \quad (2.10)$$

where det(·) denotes the matrix determinant. The MLE of \mathbf{R}_x under \mathcal{H}_1 can be derived as follows. First, by defining $\mathbf{A} = \mathbf{R}_x^{-1}$, eq. (2.10) can be written as

$$f(\mathbf{A}) = -\frac{MN}{2}\ln(2\pi) + \frac{N}{2}\ln(\det(\mathbf{A})) - \frac{1}{2}\sum_{n=0}^{N-1} \mathbf{x}^{H}[n]\mathbf{A}\mathbf{x}[n].$$
(2.11)

Since $\mathbf{R}_x \succeq 0$ i.e., \mathbf{R}_x is positive definite, so is \mathbf{A} . It is then easy to verify that $f(\mathbf{A})$ is a concave function of \mathbf{A} . By setting the first derivative of eq. (2.11) with respect to \mathbf{A} equal to the all-zero matrix, the optimal \mathbf{A} that maximizes $f(\mathbf{A})$ can be obtained. Equivalently, the MLE of \mathbf{R}_x that maximizes $p(\mathbf{X}|\mathcal{H}_1, \mathbf{R}_x)$ is obtained as

$$\hat{\mathbf{R}}_{x} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{x}[n] \mathbf{x}^{H}[n].$$
(2.12)

Using eq. (2.12), we obtain

$$\ln p(\mathbf{X}|\mathcal{H}_1, \hat{\mathbf{R}}_x) = -\frac{MN}{2}\ln(2\pi) - \frac{N}{2}\ln\left[\det(\hat{\mathbf{R}}_x)\right] - \frac{MN}{2}.$$
 (2.13)

Let the *M* eigenvalues of $\hat{\mathbf{R}}_x$ be denoted by $\lambda_x = [\lambda_{1,x}, \ldots, \lambda_{M,x}]$. It is then possible to obtain from eq. (2.13) the following

$$\ln L_G(\mathbf{X}) = \frac{MN}{2} \left\{ \ln \left(\frac{1}{M} \sum_{m=1}^M \lambda_{1,x} \right) - \frac{1}{M} \ln \left(\prod_{m=1}^M \lambda_{1,x} \right) \right\}.$$
 (2.14)

Finally, removing constant terms and using the monotonicity of the logarithm function, the GLTR statistic test is calculated as

$$\mathcal{T}_{AGM}(\lambda) = \frac{\frac{1}{M} \sum_{m=1}^{M} \lambda_{m,x}}{\left(\prod_{m=1}^{M} \lambda_{m,x}\right)^{\frac{1}{M}}} \stackrel{\mathcal{H}_{1}}{\gtrless} \gamma.$$
(2.15)

Notice that the above test statistic depends only on the eigenvalues of the sample covariance matrix, λ_x [42]. This test statistic is the ratio of the arithmetic mean (AM) to the geometric mean (GM) of the eigenvalues. Hence, this detection algorithm shall be called the Arithmetic to Geometric Mean (AGM) method [27].

2.2.2.4 Signal-Subspace Eigenvalues

Finally, we consider the case of a detector when it is assumed that σ^2 is known but \mathbf{R}_s is unknown, and, thus has to be estimated with the GLTR algorithm. The LLF under \mathcal{H}_0 is given in eq. (2.7) where σ^2 is known, but \mathbf{R}_s is now treated as an unknown parameter, and that under \mathcal{H}_1 , conditioned on the unknown parameter \mathbf{R}_s , is

$$\ln p(\mathbf{X}|\mathcal{H}_{1}, \mathbf{R}_{s}) = -\frac{MN}{2}\ln(2\pi) - \frac{N}{2}(\det(\mathbf{R}_{s} + \sigma^{2}\mathbf{I})) - \frac{1}{2}\sum_{n=0}^{N-1}\mathbf{x}^{H}[n](\mathbf{R}_{s} + \sigma^{2}\mathbf{I})^{-1}\mathbf{x}[n].$$
(2.16)

The MLE of \mathbf{R}_s under \mathcal{H}_1 can be obtained as follows. First, as in the previous scheme, we can introduce $\mathbf{A} = \mathbf{R}_x^{-1} = (\mathbf{R}_s + \sigma^2 \mathbf{I})^{-1}$ so that $\ln p(\mathbf{X}|\mathcal{H}_1, \mathbf{R}_s)$ can be rewritten as $f(\mathbf{A})$. Since $\mathbf{R}_s \succeq 0$, it follows that $\mathbf{R}_x \succeq \sigma^2 \mathbf{I}$ and hence $\mathbf{A} \preceq \frac{1}{\sigma^2} \mathbf{I}$. The MLE of \mathbf{R}_s can be obtained from the MLE of \mathbf{A} by solving the following constrained optimization problem over \mathbf{A} :

Maximize
$$f(\mathbf{A})$$

Subject to $\mathbf{A} \succeq 0$ (2.17)
 $\mathbf{A} \preceq \frac{1}{\sigma^2} \mathbf{I}.$

Because $f(\mathbf{A})$ is a concave function of \mathbf{A} , and because the previous constraints specify a convex set of \mathbf{A} , it follows that the above optimization problem is convex [43]. In [27] it is shown that by applying the Karush-Kuhn-Tucker (KKT) optimal conditions, the optimal \mathbf{A} for the above problem can be obtained as

$$\mathbf{A}^* = \mathbf{U}_x \left[\operatorname{Diag} \left(\min \left(\frac{1}{\lambda_{1,x}}, \frac{1}{\sigma^2} \right) \dots \min \left(\frac{1}{\lambda_{M,x}}, \frac{1}{\sigma^2} \right) \right) \right] \mathbf{U}_x^H, \tag{2.18}$$

where $\text{Diag}(\mathbf{x})$ denotes a diagonal matrix with the diagonal elements expressed in \mathbf{x} , while $\lambda_{m,x}$ and \mathbf{U}_x are obtained from the eigen-decomposition of $\hat{\mathbf{R}}_x$. Without any loss of generality, we also assume from this point that the eigenvalues are ordered from largest to smallest, *i.e.*, $\lambda_{1,x} \geq \lambda_{2,x} \geq \cdots \geq \lambda_{M,x}$. The MLE of \mathbf{R}_s can be then obtained as

$$\hat{\mathbf{R}}_s = \mathbf{U}_x \left[\text{Diag} \left((\lambda_{1,x} - \sigma^2)^+, \dots, (\lambda_{M,x} - \sigma^2)^+ \right) \right] \mathbf{U}_x^H,$$
(2.19)

where $(x)^+ = \max(x, 0)$. Using eq. (2.19) we get the log-GLRT statistic test as

$$\mathcal{T}_{SSE}(\lambda_{\mathbf{x}}) = \frac{Nm'}{2} \left[\frac{AM(\lambda_{\mathbf{x}}^{\mathbf{s}})}{\sigma^2} - \ln\left(\frac{GM(\lambda_x^s)}{\sigma^2}\right) - 1 \right] \begin{array}{c} \mathcal{H}_1 \\ \gtrless \\ \mathcal{H}_0 \end{array} \gamma, \quad (2.20)$$

where m' corresponds to the largest m such that $\lambda_{m,x} > \sigma^2$, $\lambda_{\mathbf{x}}^{\mathbf{s}}$ denotes the vector of signal subspace eigenvalues of $\hat{\mathbf{R}}_x$, and AM and GM denote the arithmetic mean and the geometric mean over the elements in a vector x, respectively. This algorithm is called the SSE (signal-subspace eigenvalues) method.

2.2.3 Receiver Operating Characteristics

A very common way to represent the detection performance of a NP detector is to plot P_d versus P_{fa} . For example, in [15], the authors present a DC level detection in AWGN. In eqs. (3.6), (3.7), and (3.8) from [15] we see that

$$P_{fa} = Q\left(\frac{\gamma'}{\sqrt{\sigma^2/N}}\right),$$

$$P_d = Q\left(\frac{\gamma' - A}{\sqrt{\sigma^2/N}}\right),$$
(2.21)

and

$$P_d = Q \left(Q^{-1}(P_{fa}) - \sqrt{d^2} \right),$$
 (2.22)

where $d^2 = NA^2/\sigma^2$. The plot corresponding to eq. (2.22) is shown in Figure 2.3 for $d^2 = 1$. Each point on the curve corresponds to the value of (P_{fa}, P_d) for a given threshold γ' . By adjusting γ' , any point on the curve may be obtained. As expected, as γ' increases, P_{fa} and P_d decrease; the inverse is also true. This type of performance description is called the *Receiver Operating Characteristic* (ROC). The ROC should always be above the $P_d = P_{fa}$ line, because the 45° ROC can be attained by a detector that bases its decision on flipping a coin, ignoring all data. Consider the detector that decides \mathcal{H}_1 if a flipped coin comes up a head, where $Prob\{head\} = p$. For a tail outcome we decide \mathcal{H}_0 . Then,

$$P_{fa} = Prob\{head; \mathcal{H}_0\},$$

$$P_d = Prob\{head; \mathcal{H}_1\}.$$
(2.23)

But the probability of obtaining a head does not depend upon which hypothesis is true and so $P_{fa} = P_d = p$. This detector then generates the point (p, p) on the ROC. To generate the other points on the 45° line we need only to use points with different values of p.

2.2.4 Performance comparison

In order to compare the performance of the previous detection schemes, we utilized the same conditions used by the authors in [27]. We considered the ideal case where a cognitive radio sensor with M = 8 receiving antennas is to detect Q = 3 single-



Figure 2.3: Receiver Operating Characteristics for a DC level in AWGN $(d^2 = 1)$.

antenna primary signal sources, each carrying an equal-power and independent data stream. For each data stream, the transmitted primary signals are i.i.d. CSCG random variables. A Rayleigh flat fading channel between each transmit-receive antenna pair is assumed. We performed 1,000 Monte Carlo simulations, each one consisting of $N = 10^4$ independent observations samples. The received signal-to-noise ratio (SNR) per antenna is fixed at -20dB. It is important to stress that the authors in [27] found the threshold levels for the simulation purposes using computational resources, i.e., they did not derive a closed formula as in [15]. With this in mind, we made use of some validation simulation in order to verify the correct result of our simulations. In Figure 2.2 we showed the theoretical values of the threshold γ for different probabilities of false alarm. Using these values, we reproduced the results shown in [15], which can be seen in Figure 2.1. The author of [15] does not indicate which covariance matrix to use for the simulations; we used the Jakes model, explained in detail in [44]. Consequently, Figure 2.4 shows the ROC curves, each of which constitutes all the



Figure 2.4: ROC of the presented detection methods.

achievable probability pairs of P_D and P_{FA} for each sensing algorithm. We can observe that, as expected, the estimator-correlator performs the best due to its perfect knowledge of the received primary signal covariance and the noise variance. When only the noise variance is perfectly known, the SSE performs better in comparison to the energy detector. On the other hand if we have no knowledge of the covariance matrix and the noise variance but we try to estimate it using this scheme, the detection performance of the AGM method is still better than the energy detector, which makes no effort in trying to estimate the unknown variables whatsoever². It is important to stress that the proposed schemes suffer from an increasing complexity on the system, due to the inherent need to estimate the signal covariance matrix, and perform an eigen-decomposition on such estimation.

^{2.} The results of this simulation agree with the conclusions of [27]; however the authors' plots are mislabeled.

2.3 Cyclic Frequency Detection, Diversity Combining Approach

In this section, we study a cyclostationary feature detection technique applying multiple cyclic frequencies in order to detect the presence of a single PU in CR networks. Cyclostationary detection was first proposed in [45] and it was later generalized for multiple cyclic frequencies in [14]. Cyclostationarity is not a recent development (see for example [17]), but effective tests for indication of second order cyclostationarity using a NP test statistic were proposed only in the last decade [14]. A natural generalization for multiple cyclic frequencies was recently proposed for PU detection in CR networks $[30]^3$. In the following, it will be shown that the single user detection algorithms (in the form of expected value estimation of the cyclic autocorrelation) can be interpreted as a specific case of the mixed frequency-delay incoherent diversity com*bining* block. The number of *virtual branches* is equal to the product of the number of cyclic frequencies and the time delays. This diversity technique can also be called a SIMO radar. Based on real-life scenarios, it is possible to assume that such branches suffer from fading, which, in the general case, can be modelled using generalized Gaussian statistics or the Klovski-Middleton model [46]. Moreover, depending on the frequency and the delay diversity parameters, the fading in these branches or antennas can be divided as non-homogeneous, homogeneous, and totally correlated or statistically independent. The concept of diversity approach for multiple cyclic detection is useful not only for effective development of quasi-optimal approaches as mentioned above, but also because it allows us to consider the necessary tradeoff between the number of delays and cyclic frequencies for the detection procedure and the statistical dependence on the corresponding diversity branches in order to fulfill a specific ROC requirement. Finally, it will be shown that the *diversity* concept for spectrum sensing is rather constructive for the analysis of collaborative sensing as well. For instance, a set of SUs, collaborating between themselves or operating through a Fusion Center (FC), can be interpreted as virtual branches (antennas) of the distributed detection

^{3.} This work was also presented in poster form at the 2010 School of Information Theory in addition to the journal publication in [30].

system, which can apply NP detection techniques or sequential analysis methods (see Chapter 4).

2.3.1 Generalized Gaussian (Klovsky-Middelton) Channel Model

Most of the existing fading channel models are based on the concept of the envelope and phase of the random vector with Gaussian Probability Density Functions (PDF). This PDF can be expressed, for orthogonal statistically independent quadrature components x and y, as [46,47]

$$W(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}\right\},$$
(2.24)

where σ_x^2, σ_y^2 and μ_x, μ_y are the variances and expectations of the x and y quadrature components respectively. Thus, by defining the module $z = \sqrt{x^2 + y^2}$ and the phase of the random vector $\varphi = \arctan \frac{y}{x}$, we obtain the following

$$W(z) = \frac{z}{2\pi\sigma_x\sigma_y} \int_0^{2\pi} \exp\left\{-\frac{(z\cos\varphi - \mu_x)^2}{2\sigma_x^2} - \frac{(z\sin\varphi - \mu_y)^2}{2\sigma_y^2}\right\} d\varphi.$$
(2.25)

From eq. (2.25) it is possible to obtain several representations for W(z), which depend on the four parameters $\{\mu_x, \mu_y\}$ and $\{\sigma_x^2, \sigma_y^2\}$. For this reason, the term *four parametric distribution* is adopted. Hereafter the following two equivalent forms for the four parametric distribution W(z) are used:⁴

$$W(z) = \frac{z}{\sigma_x \sigma_y} \exp\left\{-\frac{\mu_y^2 + z^2}{2\sigma_y^2} - \frac{\mu_x^2}{2\sigma_x^2}\right\} \sum_{k=0}^{\infty} \frac{H_{2k}(\alpha)}{(2k)!!2^k} \sigma^k \left(\frac{\sigma_y^2}{\mu_y}\right)^k \left(\frac{1}{\sigma_x^2} - \frac{1}{\sigma_y^2}\right)^k I_k \left(\frac{z\mu_y}{\sigma_y^2}\right)$$
(2.26)

^{4.} Considering that in the rest of this chapter the incoherent diversity combining is applied, the PDF of the phase φ is no longer needed.

$$W(z) = \sum_{k=0}^{\infty} \frac{R^k}{k!} \sigma^{2k} \frac{\partial^{2k}}{\partial_{\mu_I^k} \partial_{II}^k} \left\{ \frac{z}{\sigma^2} \exp\left(-\frac{z^2 + \mu_I^2 + \mu_{II}^2}{2\sigma^2}\right) I_0\left(\frac{z}{\sigma^2} \sqrt{\mu_I^2 + \mu_{II}^2}\right) \right\},\tag{2.27}$$

where $I_0(\cdot)$ is the modified Bessel function of order zero, $H_n(\cdot)$ stands for the Hermitian polynomials, and the following auxiliary functions are defined as

$$\mu_{I} = \frac{\mu_{x} + \mu_{y}}{\sqrt{2}}, \quad \mu_{II} = \frac{\mu_{x} - \mu_{y}}{\sqrt{2}}, \quad \sigma^{2} = \frac{\sigma_{x}^{2} + \sigma_{y}^{2}}{2},$$
$$R = \frac{\sigma_{y}^{2} - \sigma_{x}^{2}}{\sigma_{x}^{2} + \sigma_{y}^{2}}, \quad \sigma = \frac{\mu_{x}}{\sqrt{2\left(\frac{1}{\sigma_{x}^{2}} - \frac{1}{\sigma_{y}^{2}}\right)\sigma_{y}^{2}}} \quad (\sigma_{x}^{2} \le \sigma_{y}^{2}).$$

Eqs. (2.26) or (2.27) are called the Generalized Gaussian (GG) model because the Beckman, Hoyt, Rice, Rayleigh, and truncated Gaussian distributions can all be directly obtained from it. To demonstrate this, the next new parameters are defined as

$$q^{2} = \frac{\mu_{x}^{2} + \mu_{y}^{2}}{\sigma_{x}^{2} + \sigma_{y}^{2}}, \quad \beta^{2} = \frac{\sigma_{x}^{2}}{\sigma_{y}^{2}}, \quad z_{0}^{2} = \mu_{x}^{2} + \mu_{y}^{2},$$
$$z^{2} = z_{0}^{2} + \sigma_{x}^{2} + \sigma_{y}^{2}, \quad \varphi_{0} = \arctan\frac{\mu_{y}}{\mu_{x}}.$$

From these, the Beckman distribution can be obtained by making $\mu_y = 0, z_0 = |\mu_x|$, while the Hoyt PDF appears when $\sigma_x^2 \neq \sigma_y^2, z_0 = \mu_x = \mu_y = 0$. The Rayleigh distribution follows from $z_0 = 0, \mu_x = \mu_y = 0$, and finally the truncated Gaussian distribution occurs when, additionally to previous conditions, $\sigma_x^2 \to 0$. One can also obtain the *m* parameter for the equivalent Nakagami distribution as

$$m = \frac{(1+\beta^2)^2(1+q^2)^2}{2\left[1+\beta^4+2q^2(1+\beta^2)(\beta^2\cos^2\varphi_0+\sin^2\varphi_0)\right]}.$$
 (2.28)

The Nakagami distribution is only an approximation for the four-parameter case, but it does accurately represents the dynamics of the variations of the four-parameter PDF functional form.

2.3.2 Single User Multiple Cyclic Frequency Detection

The cyclostationary properties of wireless communication signals are very-well known and widely investigated (see for example [14,27,45]). When the PU signal shapes are known a priori, their cyclic frequencies of interest are also known. Let us introduce the set $A = \{\alpha_n\}_1^P$ to denote the cyclic frequencies and let $N = \sum_{n=1}^P N_n$ be the number of integers of the time delays τ for the auto covariance function at each cyclic frequency [14], with P denoting the number of cyclic frequencies. Thus, the equation of the auto covariance function is

$$\hat{R}_{xx^*}(\alpha,\tau) = \frac{1}{M} \sum_{l=1}^{M} x(l) x^*(l+\tau) \exp\left(-j2\pi\alpha l\tau\right),$$
(2.29)

where the integer time delay τ and the cyclic frequency are both fixed, M is the number of observations, and x(l) is the complex sample input with $x^*(l)$ being its complex conjugate. Now, by representing the complex exponent in eq. (2.29) in a trigonometric form and assuming that x(l) is a sample of the ergodic stochastic process, one can easily see that when M >> 1 or the time analysis T >> 1, the \hat{R}_{xx^*} estimations are simply the estimations of the complex Fourier coefficients for fixed α and τ (see also [17]). Let \hat{r}_{xx^*} be a complex vector of estimations (eq. (2.29)) of the Fourier coefficients (*F*-coefficients) for different α and τ . Hence, the GLRT for its estimation is well known and denoted by

$$\hat{r}_{xx^*}\hat{\varepsilon}^{-1}\hat{r}_{xx^*}^T \ge \Lambda_{0,} \tag{2.30}$$

where ε is a $2N \times 2N$ covariance matrix of \hat{r}_{xx^*} (in the non-asymptotic case, generally such coefficients are correlated) and

$$N = \sum_{n=1}^{P} N_n. \tag{2.31}$$

Let us define an estimation of the j-th complex F-coefficient as

$$\hat{V}_j = V_j + \tilde{V}_j, \qquad (2.32)$$

where V_j and \tilde{V}_j are the real and imaginary parts of \hat{V}_j . The real *F*-coefficients are not correlated; nevertheless, their estimations for a finite *M* corrupted by the noise correlated except in the asymptote when M >> 1 and/or T >> 1. With this assumption, eq. (2.30) can be significantly simplified by taking into account the total Gaussianity of the terms in eq. (2.32) as [14]

$$\sum_{\tau=1}^{N} \sum_{\alpha=1}^{P} \varepsilon_{\tau,\alpha}^{-1} \hat{V}^{2}(\tau,\alpha) = \sum_{i=1}^{Q} \left| \hat{V}_{i}^{2} \right| \varepsilon_{i}^{-1}, \qquad Q = NP,$$
(2.33)

where $\varepsilon_i^{-1} = \text{Diag}\left\{\frac{1}{2\gamma_1^2}, \dots, \frac{1}{2\gamma_Q^2}\right\}$. Thus the left side in eq. (2.30) can be represented as the following

$$\sum_{i=0}^{Q} \frac{V_i^2 + \tilde{V}_i^2}{2\gamma_i^2},\tag{2.34}$$

where $\gamma_i^2 = 2E_i/N_0$, $|\hat{V}_i|^2 = V_i^2 + \tilde{V}_i^2$, $E_i = P_iT$, and P_i is the average power of each *F*-coefficient (fading is not considered here). Equation (2.32) represents the optimum incoherent quadratic diversity combining algorithm having *Q* total virtual branches, where $\frac{1}{2\gamma_i^2}$, are weighting coefficients for each branch, generally related to homogenous conditions for combining. The quadratic combining in order to obtain the Neyman-Pearson Test (NPT) can be represented as

$$\sum_{i=1}^{Q} \frac{\left|\hat{V}_{i}\right|^{2}}{2\gamma_{i}^{2}} \ge \Lambda_{0},$$
(2.35)

where Λ_0 is a detection threshold. Equation (2.35) is not only a formal analogy to diversity addition or SIMO radar testing, but an essential reflection of the analogy between the auto covariance estimation and diversity communing of statistically independent data [48]. In the absence of fading, all branches are asymptotically statistically independent. In presence of fading, \hat{V}_i^2 can be statistically independent as well, but also might be correlated in scenarios of flat fading, both in the frequency and time domains. In terms of the signal hypotheses, the NPT can be formulated as follows [49]

$$\mathcal{H}_{0}: \quad \sum_{i=1}^{Q} \frac{\left|\hat{V}_{i}\right|^{2}}{2\gamma_{i}^{2}} = n(t) \\
 \mathcal{H}_{1}: \quad \sum_{i=1}^{Q} \frac{|\nu_{i}|^{2}}{2\gamma_{i}^{2}} + n(t)$$
(2.36)

where ν_i^2 is the *i*-th real *F*-coefficient and n(t) is the white Gaussian noise with power N_0 . For simplicity, let us suppose that all $2\gamma_i^2$ are the same and homogenous features of the virtual branches will be addressed to different $z_i^2 = z_0^2 + \sigma_{x_i}^2 + \sigma_{y_i}^2$. This means that in eq. (2.36) one has to consider only the routine form for quadratic combining

$$\xi = \sum_{i=1}^{Q} \left| \hat{V}_i \right|^2.$$
(2.37)

The NPT is characterized by P_{fa} and P_{md} which are respectively the probability of false alarm and the probability of miss detection [49]. In the absence of fading, ξ is formed by the squares of the normally distributed components and its PDF for different hypothesis can be defined as

$$\mathcal{H}_{0}: \qquad \chi_{2Q}^{2}(\xi) \qquad \text{Central chi-square PDF} \\ \mathcal{H}_{1}: \qquad \chi_{2Q}^{2}\left(\xi, \sum_{i=1}^{Q} \left\langle \hat{V}_{i} \right\rangle^{2}\right) \qquad \text{Non central chi-square PDF} \quad (2.38)$$

where $\sum_{i=1}^{Q} \left\langle \hat{V}_{i} \right\rangle^{2}$ is the expectation of the sum of \hat{V}_{i}^{2} and is a parameter of the non central chi-square distribution [50]. In the presence of fading, the functional forms for these distributions will differ depending on the scenarios for the GG channel model, and will be considered in the following section.

2.3.3 Statiscally Independent Virtual Branches with Flat Generalized Gaussian Fading

Let us assume that \hat{V}_j^2 can be represented as $\hat{V}_j^2 = \hat{x}_i^2 + \hat{y}_i^2$, where $\hat{x}_i = x_i V_i$, $\hat{y}_i = y_i \tilde{V}_i$ and $\{x_i, y_i\}$ are the quadrature Gaussian components of the GG fading model. For both hypotheses, each of the quadrature components in \tilde{V}_i^2 are Gaussian as before, but their means are not equal and their variances are arbitrary. Now, if

$$\xi = \sum_{i=1}^{Q} \left\{ V_i^2 + \tilde{V}_i^2 \right\}, \qquad (2.39)$$

the routine procedure for calculus of the noise immunity can be applied. For hypothesis \mathcal{H}_0 :

$$M\{V_i\} = \mu_{x_i} \sqrt{\frac{2E_i}{N_0}} \frac{1}{2\gamma_{x_i}^2} \sqrt{\frac{2\gamma_{x_i}^2}{1+2\gamma_{x_i}^2}},$$

$$M\{\tilde{V}_i\} = \mu_{y_i} \sqrt{\frac{2E_i}{N_0}} \frac{1}{2\gamma_{y_i}^2} \sqrt{\frac{2\gamma_{y_i}^2}{1+2\gamma_{y_i}^2}},$$

$$D\{V_i\} = \frac{2\gamma_{x_i}^2}{1+2\gamma_{x_i}^2}, \qquad D\{\tilde{V}_i\} = \frac{2\gamma_{y_i}^2}{1+2\gamma_{y_i}^2},$$
(2.40)

whereas for hypothesis \mathcal{H}_1 :

$$M\{V_{i}\} = \mu_{x_{i}} \sqrt{\frac{2E_{i}}{N_{0}}} \sqrt{\frac{2\gamma_{x_{i}}^{2}}{1 + 2\gamma_{x_{i}}^{2}}},$$

$$M\{\tilde{V}_{i}\} = \mu_{y_{i}} \sqrt{\frac{2E_{i}}{N_{0}}} \sqrt{\frac{2\gamma_{y_{i}}^{2}}{1 + 2\gamma_{y_{i}}^{2}}},$$

$$D\{V_{i}\} = 2\gamma_{x_{i}}^{2}, \qquad D\{\tilde{V}_{i}\} = 2\gamma_{y_{i}}^{2},$$
(2.41)

where $M\{\cdot\}$ and $D\{\cdot\}$ represent the mean and variance respectively,

$$\gamma_{x_i}^2 = \sigma_{x_i}^2 \frac{E_i}{N_0 P^2}$$

and

$$\gamma_{y_i}^2 = \sigma_{y_i}^2 \frac{E_i}{N_0 P^2}.$$

In the case of frequency diversity there are P out of Q virtual branches, with the transmitted power divided between them [51]. From eqs. (2.40) and (2.41) it follows that for both hypotheses, the PDF of $W(\xi)$ is a non-central chi-square distribution. For analytical evaluation, the special cases of high reliability detection $\gamma_{x_i}^2, \gamma_{y_i}^2 > 1$ are considered. For such conditions, it follows from eq. (2.40) that for \mathcal{H}_0 , $M\{V_i\}$ and $M\{\tilde{V}_i\}$ are close to zero while the variances are close to one. Thus, $W(\xi)$ under \mathcal{H}_0 tends to χ_{2Q}^2 . For this case P_{fa} can be obtained as

$$P_{fa} \sim \frac{1}{(Q-1)!} \Gamma(\Lambda_0, Q) = \exp(-\Lambda_0) \sum_{q=0}^{Q-1} \frac{(\Lambda_0)^q}{q!}.$$
 (2.42)

By fixing P_{fa} , one can find Λ_0 ; applying the same conditions $\gamma_{x_i}^2, \gamma_{y_i}^2 > 1$ from eq. (2.41), it follows that for the hypothesis \mathcal{H}_1 , the variances $D\{V_i\}$ and $D\{\tilde{V}_i\}$ will be extremely large. In this case P_{md} is calculated as

$$P_{md} \sim \frac{\Lambda_0^Q}{Q!} \prod_{i=1}^Q \frac{(1+\beta_i^2)(1+q_i^2)}{2\bar{\gamma}_i^2\beta_i} \exp\left\{-\frac{q_i^2(1+\beta_i^2)}{2\beta_i^2}(\cos^2\varphi_{0_i}+\beta_i^2\sin^2\varphi_{0_i})\right\}, \quad (2.43)$$

where $\bar{\gamma}_i^2 = \frac{2E_i}{N_0 P^2} \left(\sigma_{x_i}^2 + \sigma_{y_i}^2 + \mu_{x_i}^2 + \mu_{y_i}^2 \right)$. For the case of the one-sided Gaussian distribution, eq. (2.43) becomes

$$P_{md} = \frac{\Gamma(Q+1)\Lambda_0^Q}{2\bar{\gamma}^{2Q}Q!\Gamma\left(\frac{Q}{2}\right)\sqrt{\pi}}.$$
(2.44)

For the case of Nakagami fading channels, by assuming non-correlated homogenous conditions for the fading in all virtual branches,

$$P_{md} = \frac{\Lambda_0^Q}{Q!} \left(\frac{2m}{2m + \bar{\gamma}^2}\right)^{Qm}, \qquad (2.45)$$

where Λ_0 and m can be calculated using eqs. (2.42) and (2.28) respectively.

2.3.4 Fully Correlated Virtual Branches (Flat Fading) in GG Channel

For the case of totally correlated fading processes at the virtual branches, it is known that the resulting SNR after the combining process is:

$$\bar{\gamma}^2 = \frac{1}{P^2} \sum_{i=1}^{Q} \bar{\gamma}_i^2.$$

Thus, the problem can be transferred to the quadrature addition algorithm for one equivalent branch, *i.e.*, without diversity but with the GG model of flat fading

$$V_0 + \dot{V}_0 \ge \Lambda_0, \tag{2.46}$$

where $V_0^2 = \sum_{i=1}^{Q} V_i^2$ and $\tilde{V}_0^2 = \sum_{i=1}^{Q} \tilde{V}_i^2$. Here, eqs. (2.40) and (2.41) are still valid, but for single channel conditions. Consequently, $Q_{eq} = 1$, and from eq. (2.42)

$$P_{fa} = \exp(\Lambda_0), \qquad \Lambda_0 = \ln \frac{1}{P_{fa}}, \qquad (2.47)$$

and

$$P_{md} = \frac{\ln \frac{1}{P_{fa}} (1+\beta^2)(1+q^2)}{2\beta \bar{\gamma}_i^2 \exp\left[\frac{q_i^2 (1+\beta^2)}{2\beta^2} \left(\cos^2 \varphi_0 + \beta^2 \sin^2 \varphi_0\right)\right]}.$$
(2.48)

The ROCs resulting from previous eqs. (2.43) and (2.48) are shown in Figures 2.5-2.7 for difference sets of parameters. By comparing eqs. (2.43) and (2.48) one can see that:

- For the same value of $\hat{\gamma}_i^2$ and P_{fa} fixed, the probability of miss detection P_{md} from eq. (2.48) is greater than P_{md} from eq. (2.43). This can be explained by the diversity effect in eq. (2.48) (see also [47, 52]).
- It is in fact reasonable to choose a small set of delays and multiple frequencies $(Q \leq 5 \ [52, 53])$ in order to provide (as long as the channel condition allows it) statistically independent fading in those virtual branches, *i.e.*, it is reasonable to "sacrifice" the value of P by greater intervals between α and τ so as to artificially



Figure 2.5: Probability of miss detection, eq.(2.43) is shown in solid line while eq.(2.48) is shown in the dotted line $(\beta_i^2 = 0.1, P_{fa} = 0.001, q_i^2 = 2)$.



Figure 2.6: Probability of miss detection, eq.(2.43) is shown in solid line while eq.(2.48) is shown in the dotted line $(\beta_i^2 = 0.5, P_{fa} = 0.1, q_i^2 = 2)$.



Figure 2.7: Probability of miss detection, eq.(2.43) is shown in solid line while eq.(2.48) is shown in the dotted line $(\beta_i^2 = 0.1, P_{fa} = 0.001, q_i^2 = 4)$.

create independent fading in the frequency and deal domains, leading to better noise immunity after *diversity combining*. Therefore, an appropriate choice of cyclostationary features Q = NP of the desired PU signals can significantly improve their ROC properties.

2.3.5 Special Case of Covariance Matrix for Correlated Branches at the Quadratic Incoherent Addition Algorithm

Let us consider in the following special case of the covariance matrix for quadrature components $x = \{x_l\}_1^Q$ and $y = \{y_l\}_1^Q$, and assume that across all the branches x_l or y_l the Gaussian components are correlated with coefficients R_x or R_y and there is no cross-correlation at all between x_l and y_l Gaussian components. This assumption restricts (in general) the type of the covariance matrix of the GG channel model but might be useful for the first step examination of the influence of the covariance between virtual branches at the noise immunity characteristics of the SU. It is well known that for each pair of x or y Gaussian variables, it is possible to obtain a new set of statistically independent Gaussian variables by rotating the coordinate system (linear transformation) by the angle

$$\Psi = \arctan\left\{2R\frac{\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2}\right\},\tag{2.49}$$

where R is the correlation coefficient and σ_1^2, σ_2^2 are the variances of two correlated Gaussian quadrature components, while new Gaussian variables are statistically independent [50,54]. In order to provide tractable analytical results, only the case Q = 2from eq. (2.37) will be considered. The variances for hypotheses \mathcal{H}_0 and \mathcal{H}_1 are calculated as

$$\sigma_{I,II}^{2} = \frac{2\sigma_{1}^{2}\sigma_{2}^{2}(1-R)}{(\sigma_{1}^{2} - \sigma_{2}^{2})\left[1 \pm \sqrt{1 - (1-R^{2})\frac{4\sigma_{1}^{2}\sigma_{2}^{2}}{\sigma_{1}^{2} + \sigma_{2}^{2}}}\right]},$$
(2.50)

where $\sigma_{I,II}^2$ are the new variances of the quadrature components after the angle rotation (for each two branches). Moreover, assuming for simplicity that $R_x = R_y = R$, one obtains

$$\mu_I = (\mu_1 \cos \Psi - \mu_2 \sin \Psi),$$

$$\mu_{II} = (\mu_1 \sin \Psi - \mu_2 \cos \Psi),$$
(2.51)

where μ_I and μ_{II} are the new means after the angle rotation. Consequently, this set of parameters can be considered as new parameters of the GG model with the statistically independent branches. The ROC can be calculated in the same way as eqs. (2.42) and (2.43); however, this calculation is in general rather cumbersome because the new parameters of the GG model come from rather complex expressions. Therefore, assuming that $P_{fa} \ll 1$ and $P_{md} \ll 1$, it is possible to apply an asymptotic calculation. Particularly, for the \mathcal{H}_0 hypothesis, all the means will be close to zero and all the variances will be equal to $(1 - R^2)$. Thus, P_{fa} can be calculated using eq. (2.42) but using a new threshold obtained with

$$\Lambda' = \frac{\Lambda_0}{1 - R^2}.$$

Hence, for the case $P_{md} \ll 1$ one obtains

$$P_{md} \approx \frac{\Lambda'^2}{2} \exp\left[-\frac{\mu_{I_x}^2}{2(1+\sigma_{I_x}^2)} - \frac{\mu_{II_x}^2}{2(1+\sigma_{II_x}^2)} - \frac{\mu_{Iy}^2}{2(1+\sigma_{Iy}^2)} - \frac{\mu_{Iy}^2}{2(1+\sigma_{II_y}^2)}\right] \times \left[1 + \frac{\sigma_{I_x}^2}{2(1+\sigma_{Ix}^2)} \left(1 + \frac{\mu_{Ix}^2}{\sigma_{I_x}^2(1+\sigma_{II_x}^2)}\right) + \frac{\sigma_{II_x}^2}{2(1+\sigma_{II_x}^2)} \left(1 + \frac{\mu_{II_x}^2}{\sigma_{II_x}^2(1+\sigma_{II_x}^2)}\right)\right] (2.52) + \frac{\sigma_{Iy}^2}{2(1+\sigma_{Iy}^2)} \left(1 + \frac{\mu_{Iy}^2}{\sigma_{II_y}^2(1+\sigma_{II_y}^2)}\right) + \frac{\sigma_{IIy}^2}{2(1+\sigma_{IIy}^2)} \left(1 + \frac{\mu_{II_y}^2}{\sigma_{II_y}^2(1+\sigma_{II_y}^2)}\right)\right].$$

Let us consider the special case when $\sigma_{I_x}^2 = \sigma_{II_x}^2 = \sigma_{I_y}^2 = \sigma_{II_y}^2 = \sigma^2$; by also introducing $z_{0_I}^2 = \mu_{I_x}^2 + \mu_{I_y}^2$ and $z_{0_{II}}^2 = \mu_{II_x}^2 + \mu_{II_y}^2$, we get

$$P_{md} \approx \frac{\Lambda'^2}{2\sigma^4} \left(3 + \frac{z_{0_I}^2}{2\sigma^4} + \frac{z_{0_I}^2}{\sigma^4} \right) \exp\left\{ -\frac{z_{0_I}^2}{2\sigma^2} - \frac{z_{0_{II}}^2}{2\sigma^2} \right\}.$$
 (2.53)

For this case σ^2 can be calculated as

$$\sigma^{2} = \frac{2\gamma_{0_{1}}^{2}\gamma_{0_{2}}^{2}(1-R^{2})}{(\gamma_{0_{1}}^{2}+\gamma_{0_{2}}^{2})\left[1\pm\sqrt{1-(1-R^{2})\frac{4\gamma_{0_{1}}^{2}\gamma_{0_{2}}^{2}}{(\gamma_{0_{1}}^{2}+\gamma_{0_{2}}^{2})^{2}}\right]},$$
(2.54)

and for the case of $\frac{z_{0_I}^2}{2\sigma^2}, \frac{z_{0_{II}}^2}{2\sigma^2} << 1$ (strong fading), we obtain

$$P_{md} \approx \frac{3\Lambda'^2}{2\sigma^4} = \frac{3}{2} (\Lambda')^2 \frac{1}{\gamma_{0_1}^2 \gamma_{0_2}^2 (1-R^2)}.$$
(2.55)

Equation (2.55) shows that losses related to the correlation between diversity branches depend mainly on $\frac{1}{1-R^2}$ (see [46,47]). For the case of a fading following the truncated

Gaussian PDF,

$$P_{md} \approx \frac{\Lambda'^2}{2} \frac{1}{\sqrt{(1 + \sigma_{I_y}^2)(1 + \sigma_{II_y}^2)}} \left(\frac{3}{2} + \frac{\sigma_{I_y}^2}{2(1 + \sigma_{I_y}^2)} + \frac{\sigma_{II_y}^2}{2(1 + \sigma_{II_y}^2)}\right), \qquad (2.56)$$

and for $P_{md} \ll 1$,

$$P_{md} \approx \frac{\Lambda'^2}{2} \frac{5}{8\gamma_{0_1}^2 \gamma_{0_2}^2 (1 - R^2)}$$
(2.57)

Here, the losses also depend mainly on $\frac{1}{(1-R^2)}$. Changes of the threshold, which depends on R, influence the character of the dependence of ROC on the correlation properties of the GG model in a nonlinear way. From diversity combining theory, it is well known [46, 47] that correlation between branches has an influence mainly in the ROC when the resulting SNR is rather high, *i.e.*, the correlation is much smaller than one.

2.3.6 Suboptimal algorithms and their Noise Immunity

The first suboptimal algorithm considered here will be an energy receiver where the desired signal is represented by

$$x(t) = \sum_{i=1}^{B} \alpha_i \varphi_i(t), \qquad (2.58)$$

where B is the total number of orthonormal functions $\{\varphi_i(t)\}_1^B$ applied for the expansion of the desired signal x(t). Consequently, the corresponding signal given the hypothesis can be represented as

$$\mathcal{H}_{0}:z(t|\mathcal{H}_{0}) = \sum_{j=1}^{B} \left[\int_{0}^{T} z(t)\varphi_{i}(t)dt \right]^{2},$$

$$\mathcal{H}_{1}:z(t|\mathcal{H}_{1}) = \sum_{j=1}^{B} \left[\int_{0}^{T} x(t)\varphi_{i}(t)dt \right]^{2} + \sum_{j=1}^{B} \left[\int_{0}^{T} z(t)\varphi_{i}(t)dt \right]^{2},$$
(2.59)

being

$$n(t) = \sum_{j=1}^{B} \left[\int\limits_{0}^{T} z(t) \varphi_{i}(t) dt \right]^{2}$$

and

$$x(t) = \sum_{j=1}^{B} \left[\int_{0}^{T} x(t)\varphi_{i}(t)dt \right]^{2},$$

where the T is the time of analysis. Now, for the representation of x(t) and n(t) let us apply the F-basis as in [53] (see also [48] and references therein). We now have

$$x(t) = \sum_{k=0}^{B} a_k \cos(k\omega_0 t) + b_k \sin(k\omega_0 t),$$

$$n(t) = \sum_{k=0}^{B} \alpha_k \cos(k\omega_0 t) + \beta_k \sin(k\omega_0 t),$$
(2.60)

where $\omega_0 = \frac{2\pi}{T}$, B = 2FT, $F = \frac{k_2 - k_1}{F} + 1$ - frequency bandwidth, and k_2 , k_1 are the upper and lower indexes taken into account for the *F*-series expansion. Then, the signal under hypothesis becomes,

$$\mathcal{H}_{0}: \qquad \frac{1}{2} \sum_{j=1}^{B} \left(\alpha_{k}^{2} + \beta_{k}^{2} \right), \\
 \mathcal{H}_{1}: \quad \frac{1}{2} \sum_{j=1}^{B} \left(\alpha_{k} + a_{k} \right)^{2} + \left(\beta_{k} + b_{k} \right)^{2}.$$
(2.61)

All of a_k , b_k , α_k and β_k are Gaussian distributed coefficients; the left side in eq. (2.61) therefore has central and non-central χ^2_{2B} distributions respectively. By defining the left sides in (2.61) as λ_1 and λ_2 we can define the following equations:

$$W(\lambda_1) = \frac{1}{D2^{B/2}\Gamma(B/2)} \left(\frac{\lambda_1}{D}\right)^{B/2-1} \exp\left(-\frac{\lambda_1}{2D}\right),$$

$$W(\lambda_2) = \frac{1}{2D} \left(\frac{\lambda_2}{\Delta}\right)^{\frac{B/2-1}{2}} \exp\left(-\frac{\lambda_2 + \Delta}{2D}\right) I_{B/2-1} \left(\frac{\sqrt{\lambda_2 \Delta}}{D}\right),$$
(2.62)

where $\Delta = \sum_{j=1}^{B} (\alpha_k^2 + \beta_k^2) = 2P_x$ is the average power of x(t) and the parameter $D = \frac{N_0 T}{2}$. The miss detection probability P_{md} is then

$$P_{md} = \frac{1}{2} \exp\left(-\frac{\gamma^2}{2}\right) F\left(\frac{\Lambda_0}{D}, \gamma^2, B\right) \le \frac{1}{2} \exp\left(-\frac{\gamma^2}{2}\right), \qquad (2.63)$$

where $0 \leq F\left(\frac{\Lambda_0}{D}, \gamma^2, B\right) \leq 1$ is the Cumulative Distribution Function (CDF) of the non-central χ^2_{2B} PDF. The upper bound of P_{md} for the GG channel model with flat fading is known from [47] as

$$P_{md} \le \frac{1}{C_1} \exp\left\{-\frac{q^2 \gamma^2}{2(1+q^2)} \left[\frac{(1+\beta^2)(1+q^2)\cos^2\varphi_0}{(1+\beta^2\bar{\gamma}^2)} + \left[\frac{(1+\beta^2)(1+q^2)\sin^2\varphi_0}{(1+\bar{\gamma}^2)}\right]\right]\right\},\tag{2.64}$$

where $\bar{\gamma}^2 = z^2 \frac{2E}{P^2 N_0}$ and

$$C_1 = 2\sqrt{1 + \frac{\beta^2 \bar{\gamma}^2}{(1+\beta^2)(1+q^2)}}$$

An exact tractable analytical expression of P_{md} for the GG model is not available. In the absence of fading, it is possible to obtain an analytical result. First, it is possible to represent the Bessel function as in [55] as

$$I_{B/2-1}\left(\sqrt{\gamma^2 y}\right) = \sum_{k=0}^{\infty} \frac{(\gamma^2 y)^{\frac{B/2-1}{2}+k}}{k! \Gamma(B/2+k)}.$$
 (2.65)

 P_{md} from eq. (2.63) is then

$$P_{md} = \frac{1}{2} \exp\left(-\frac{\gamma^2}{2}\right) \sum_{k=0}^{\infty} \frac{\gamma \left(B/2 + k, \frac{\Lambda_0}{2D}\right) \left(\bar{\gamma}^2\right)^{\frac{B}{4} - \frac{1}{2} + k}}{k! \Gamma(B/2 + k) 2^{k-1}},$$
(2.66)

where $B = 2F\tau$, and $\gamma(\alpha, x)$ is the lower incomplete gamma function. The analysis of eq. (2.66) shows that the influence of B can be significant and it can be shown that for fixed P_{fa} , while B increases, P_{md} increases as well. To the best of our knowledge, the influence of B and $\bar{\gamma}^2$ on the ROC of the energetic (auto covariance) receiver was first addressed in [48]. For the multiple cyclic frequency case, when the number of frequencies P is rather large while T is fixed, F is large as well and P_{md} increases. Therefore the energetic detector is not a good candidate for spectrum sensing in this scenario, as its P_{md} is much worse than for the optimum detector. Another option for suboptimal detection is to take advantage of the analogy between multiple cyclic frequency detection and quadratic diversity combining and apply a suboptimal variant of incoherent diversity addition. A selection (switching) combining method was chosen, assuming that fading has a Nakagami PDF. See for example eq. (2.28) to adjust parameters of the Nakagami PDF and four-parameter distribution. There are several different approaches for switching combining but in the following we analyze only the algorithm of selection of the *virtual branch* with

$$z' = \max_{z} z_i,$$

for $i = 1, ..., Q_0$. Let us assume here, for simplicity, homogeneous fading conditions. Then, the distribution of the maximum value of the identically distributed random values is given by [11] as

$$W(z') = QW(z = z') \left[\int_{0}^{z'} W(y) dy \right]^{Q-1}.$$
 (2.67)

If W(z) is given by [56] as

$$W(z) = \frac{2m^m y^{2m-1}}{\Gamma(m)\Omega^{2m}} \exp\left(-m\frac{y^2}{\Omega^2}\right),$$
(2.68)

 P_{md} must be averaged for one virtual branch without fading through eq. (2.67) with the help of eq. (2.69), while P_{fa} is

$$P_{fa} = \exp(-\Lambda), \tag{2.69}$$



Figure 2.8: Probability of miss detection comparison for the optimum (eq. (2.45)), and quasi optimum (eq. (2.71)) cases.

where Λ is the threshold. The P_{md} of the channel without fading is then

$$P_{md} = Q\left(\sqrt{\gamma}, \Lambda\right). \tag{2.70}$$

For the case of Q = 2 and $\gamma > 1$, one can get from eqs. (2.67)-(2.70) the approximate formula

$$P_{md} = \frac{\Lambda^2}{\left(1 + \frac{\gamma^2}{2m}\right)^m} \left\{ 1 - \sum_{i=0}^{m-1} \frac{(m-1+i)! \left(1 + \frac{\gamma^2}{2m}\right)^m}{i!(m-1)! \left(2 + \frac{\gamma^2}{2m}\right)^{m+i}} \right\},$$
 (2.71)

for *m*- integers. One can compare this method of switching combining (with fixed $\bar{\gamma}^2$ and Q = 2) with the optimum approach. Notice that in fading channel conditions when the number of virtual branches is growing, one encounters the so-called hardening effect, *i.e.* while Q is increasing, the increment of the ROC might be low. Therefore, with Q = 2, there is reason to compare the effectiveness of the selection combining method with the optimum one. In Figure 2.8, the ROC for this method is presented. One can see that the energetic losses for $P_{md} = 10^{-4}$ are rather small

and for m = 1 are negligible. In the same manner as above, the well known set of sub-optimum combining algorithms can be applied: other methods of switching combining, linear (weighted and non-weighted) addition, etc. Their application is rather straightforward. Both algorithms (eq. (2.37)) and (eq. (2.40)) rely on quadratic addition of the F-coefficients, but their noise immunity is quite different, particularly with the GG channel fading. The reason for this is rather straightforward: in eq. (2.37) the object of the quadratic addition is the F-coefficients, but from the auto covariance function of the output of the multiple cyclic frequency optimum detector, *i.e.*, after optimum processing of the quadrature components of the input signals. It is also possible to provide statistically independent fading of the virtual branches for incoherent addition by properly choosing the cyclic frequencies and delays. which drastically increase the noise immunity (through the diversity effect). In contrast, the energetic receiver, as it is in (2.59)-(2.61), does not apply specific properties of the cyclic frequencies and only extracts the total energy of the aggregate input signal. It is unlikely in this case for the F-coefficients of the input signal to exhibit statistical independency in fading conditions. Moreover, for the energetic receiver (eq. (2.61)), the noise immunity (even in the case of a constant channel without fading) decreases when the bandwidth F grows (B = 2FT, with T fixed) as the noise power grows. Therefore the energetic receiver for multiple cyclic frequency signals might not be practical when FT >> 1. For a quasi-optimum alternative for optimum quadratic combining it is possible to consider all the set of switching combining algorithms, as well as a whole set of quasi-optimum algorithms of non-coherent diversity combining,

such as a set of linear combining methods with rather low power losses for fixed P_{fa} and P_{md} .

2.4 Coupled Dynamical Systems Approach

In this section, self-organization of coupled dynamical systems are used in to establish a decentralized synchronization and to be detect the presence of an external agent (a PU for instance) in a distributed fashion with the use of only local measurements. In CR systems, such local measurements may present interference temperature at certain frequency bands at different locations to facilitate dynamic spectrum access and solve the hidden terminal problem. This method can also be implemented in the analog domain without the need of any analog-digital converters (ADC) and extensive digital signal processing.

2.4.1 Coupled Dynamical System Model

Let us assume a network with N nodes, all of them with access to a common interaction media which measures their environment and obtains a local decision about interference, temperature interference, frequencies, etc. Each node is initialized using this local decision. These decisions are shared among all nodes; through this, they attempt to achieve a globally stable behaviour, such as synchronizing the dynamics of states for all nodes (a synchronization mode). Within this context, each node is considered as a single dynamical system depicted in Figure 2.9. The node architecture shown in Figure 2.9 is divided into:

- A local decision block which performs the local decision on a variable $y_n(t_k)$ at time t_k .
- A processing block that calculates a function of the measurement, $g_n(y_n(t_k))$, whose presents an initial state of *n*-th node $x_n(t_0)$.
- A front-end that senses the environment and obtains local decisions from other nodes.
- A dynamical system, characterized by a state $x_n(t)$ which dynamics depends on a local decision $g_n(y_n(t_k))$ and decisions obtained from other states $x_m(t)$.
- An interface block to map a local state $x_n(t)$ on some physical carrier and to broadcast it to its neighbors.
- A synchronization block (dashed lines in Figure 2.9).

The system dynamics may be described by the motion equations in continuous time

$$\dot{x}_n = g_n(y_n(t)) + \frac{K}{c_n} \sum_{m=1}^N a_{nm} \psi(x_m(t - \tau_{nm}) - x_n(t)) + \eta_n(t), \qquad n = 1 \dots N$$
(2.72)



Figure 2.9: Dynamical System Node Architecture

where $\psi(\cdot)$ is a coupling function, K is a global coupling gain, and c_n are the local positive coefficients. These coefficients are related to the reliability or SNR of the local measurements, and a_{nm} and τ_{nm} are the coupling strength and propagation delay between nodes n and m respectively. The radio-wave propagation loss in wireless communications, the distance between nodes, and the effect of fading can be taken into account by the a_{nm} coefficient:

$$a_{nm}^2 = \frac{p_m |h_{nm}|^2}{d_{nm}^2},\tag{2.73}$$

where p_m is the power transmitted by the *m*-th node, h_{nm} is the amplitude fading coefficient, and d_{nm} is the distance between nodes *n* and *m*. For the local decision mapping to the model, we consider a mapping of local measurements on radio frequencies (RF) of local nodes, *i.e.* $g_n = \omega_n$. In this way, we can describe the dynamics of a local state as the dynamics of an oscillator and the coupling effect may be seen as a mean field acting on a selected oscillator. By using $\psi(\cdot) = \sin(\cdot)$ as a coupling function in equation (2.72) and neglecting the delays, the dynamics of the fully connected network can be described by the Kuramoto model of weakly phase-coupled oscillators

$$\dot{\theta}_n(t) = \omega_n + \frac{K}{N} \sum_{m=1}^N \sin(\theta_m(t) - \theta_n(t)), \quad n = 1 \dots N$$
(2.74)

where $\omega_n = g_n(y_n)$ are the local frequencies and $\theta_n(t) = x_n(t)$ are the initial phases. Hence, it is possible to define a complex mean field for N globally coupled oscillators with equal unit amplitude as follows

$$R(t) = \frac{1}{N} \sum_{n=1}^{N} e^{i\theta_n(t)} = r e^{i\phi(t)}, \qquad (2.75)$$

where r and ϕ are the mean-field amplitude and the phase, respectively. Using eq. (2.75), we can rewrite eq. (2.74) as

$$\dot{x}_n(t) = \dot{\theta}_n(t) = \omega_n + a_{nm} Kr \sin(\phi - \theta_n + \alpha_{nm}).$$
(2.76)

Notice that, unlike [25] and [26], we have introduced the fading coefficient $a_{nm} = h_{nm}$ (assuming $p_n/d_{nm}^2 = 1$) and the influence of the fading in the phase denoted by α_{nm} . This allow us to consider a more realistic physical scenario.

2.4.2 Fading Generating Algorithm

In order to generate the fading coefficients, two white random processes h_R and h_I with variances σ_R and σ_I , respectively, are used. These are later passed through a linear time invariant second order filter to produce two coloured random processes according to some target autocorrelation function (*e.g.* Bessel or exponential). We then add the corresponding means μ_R and μ_I to each process. Finally we describe the sum of all scattered components as the complex Gaussian random process [57]

$$h = h_R + jh_I. (2.77)$$

This process is described in detail in Figure 2.10.

2.4.3 Autoregressive Model

There are several methods of implementing linear time invariant filters to provide some autocorrelation function at the output. We use the well known Autoregressive



Figure 2.10: Fading Channel Simulator

Model (AR) [58]. The AR-model of a random process y(n) is defined by the following expression:

$$y(n) = \sum_{i=1}^{m} a_i y(n-i) + \sqrt{K} \xi(n), \qquad (2.78)$$

where a_i are the coefficients of the recursive filter, m is the order of the model, and ξ is a white noise sample. We determined a second order model (m = 2) provides sufficient accuracy, so eq. (2.78) is therefore

$$y(n) = a_1 y(n-1) + a_2 y(n-2) + \sqrt{K} \xi(n), \qquad (2.79)$$

where a_1 and a_2 can be calculated using the very well known Yule-Walker equations as

$$a_1 = \frac{\rho_1 - \rho_1^* \rho_2}{1 - |\rho_1|^2} \tag{2.80}$$

$$a_2 = \frac{\rho_2 - \rho_1^2}{1 - |\rho_1|^2} \tag{2.81}$$

and ρ_i is the *i*-th autocorrelation coefficient. Finally, K can be calculated as

$$K = \frac{a_1\rho_1 + a_2\rho_2 - 1}{\sigma_n^2},$$
(2.82)



Figure 2.11: Bessel Autocorrelation

where σ_n is the variance of the white noise. In Figures 2.11 and 2.12 we show the performance of the second order filter in comparison to the theoretical Bessel and Exponential curves.

2.4.4 Distributed Estimation Results

Provided that global coupling strength K in equation (2.74) is large enough compared to some frequency variations, the systems can evolve from quasi-chaotic to partial frequency synchronization where nodes with close frequencies are locked. This behaviour is analyzed in Figure 2.13 where no fading is considered. Partial synchronization is achieved when K = 0.5 and the frequency locking is achieved when K = 0.6. In Figure 2.14 we present the synchronization considering the influence of a Rayleigh fading amplitude without taking the phase into account. In Figure 2.15 we consider the influence of both amplitude and phase fading having a distribution shown in Figure 2.16. The amplitude is Rayleigh distributed while the phase is uniform dis-



Figure 2.12: Exponential Autocorrelation



Figure 2.13: Evolution of frequencies (local decisions) in time (K = 0.3, 0.5, 0.6).


Figure 2.14: Evolution of frequencies (local decisions) in time (K = 0.9, 1.0).

tributed. These were generated considering $\sigma_R = \sigma_I = 1$ and $\mu_R = \mu_I = 0$. In Figure 2.17 we plot the synchronization considering a different distribution for the amplitude and phase fading which are shown in Figure 2.18. These were generated considering $\sigma_R = \sigma_I = 1$, $\mu_R = 0$, and $\mu_I = 2$. The amplitude is Rician distributed.

2.4.5 Networks with Complex Topology

So far, we have considered only the case when the network is fully connected and a fast convergence to the sync mode can be achieved. This in general is not true for real cases, where the networks are only locally connected (seen in Figure 2.19). In general, wireless networks may be seen as locally coupled systems where the interaction strength a_{nm} depends on the distance d_{nm} between n and m nodes. As seen before, we can describe the system as

$$\dot{\theta}(t) = \omega_n + \frac{K}{c_n} \sum_{m=1}^N a_{nm} \sin\left[\theta_m(t) - \theta_n(t)\right],$$
 (2.83)



Figure 2.15: Evolution of frequencies (local decisions) in time (K = 4.0, 5.0).

where

$$a_{nm} = \frac{p_m}{(1+d_{nm})^{\alpha}} \quad \text{for } \in \text{con}_n, \qquad (2.84)$$
$$a_{nm} = 0 \quad \text{otherwise}$$

and con_n is the set of all nodes connected to node n (all the nodes inside a circle or radius r_{con} (Figure 2.19). As in [26], we have set $p_n = 1$ without loss of generality, and $a_{nm} = a_{mn}$. If each node is capable of making an estimation of the local signalto-interference ratio (SIR) we can define c_n as

$$c_n = \frac{p_n}{\sum_{m \in \operatorname{con}_2} \frac{p_m}{(1+d_{nm})^{\alpha}}}.$$
(2.85)



Figure 2.16: Rayleigh magnitude

2.4.6 Networks dynamics using a graph Laplacian presentation

The connectivity of the network can be modelled as a graph $G = \{v_n, e_n; n = 1, 2, ..., N\}$ with |v| = N nodes and |e| = E edges. This graph can be described by an oriented incidence matrix $(N \times E)$, *i.e.*, in the column of edge *e* there is a +1 in the row corresponding to one vertex of *e* and a -1 in the row corresponding to the other vertex of *e*, and all other rows have 0. Using this, it is possible to rewrite eq. (2.83) as

$$\dot{\Theta}(t) = \Omega - K D_C^{-1} B D_A \sin\left[B^T \Theta(t)\right], \qquad (2.86)$$

where the vectors and matrices are now defined as follows:

$$\Theta(t) = [\Theta_1(t), \dots, \Theta_N(t)]^T,$$
$$\Omega(t) = [\omega_1(t), \dots, \omega_N(t)]^T,$$



Figure 2.17: Evolution of frequencies (local decisions) in time (K = 1.0, 1.2).

$$D_C = \operatorname{diag} \{c_1, \dots, c_N\},\$$
$$D_A = \operatorname{diag} \{a_1, \dots, a_E\},\$$

where a_1, \ldots, a_E are weights a_{nm} indexed from 1 to E. In order to take into account the radio-wave path loss in wireless networks, the weighted Laplacian is introduced which is defined as

$$L_A = BD_A B^T. (2.87)$$

By multiplying eq. (2.86) by the row vector $c^T = \mathbf{1}_N^T D_C$ we obtain

$$c^T \dot{\Theta}(t) = c^T \Omega - K \mathbf{1}_N^T D_C D_C^{-1} B D_A \sin\left[B^T \Theta(t)\right].$$
(2.88)



Figure 2.19: Partially connected network

If the system in eq. (2.86) is in frequency sync then the derivative of the state function is constant and converges to

$$\dot{\theta}^{*}(t) = \omega^{*} = \frac{c^{T}\Omega}{1_{N}^{T}c} = \frac{\sum_{i=1}^{N} c_{i}\omega_{i}}{\sum_{i=1}^{N} c_{i}}.$$
(2.89)

2.4.7 Synchronization Mode

If we assume that a network is connected, there exists then a convergence to a stable state for a sufficiently high K. Consequently, we use the upper and lower bounds to evaluate the coupling strength when the distributed estimation converges to the sync state. The authors in [59] conclude that

• The synchronized state exists for all $K > K_s$ where

$$K_s \le \frac{2||D_C \Delta \Omega||_2}{\lambda_2(L_A)}, \text{where} \Delta \Omega = \Omega - \omega^* \mathbf{1}_N$$
 (2.90)

• No synchronized state exists for all $K < K_{ns}$ where

$$K_{ns} \ge \frac{||D_C \Delta \Omega||_{\infty}}{d_{\max}} \tag{2.91}$$

where d_{\max} denotes the maximum degree of the weighted graph, $d_{\max} = \max_i \sum_{j=1}^N a_{ij}$.

2.4.8 Simulation Results

We simulated a wireless network with N nodes randomly distributed within a unit square. By means of the distance r_{con} we derived a connectivity matrix in such a way that a connectivity exists within a circle of radius r_{con} around each node. We modelled the local measurements as random values taken from a Gaussian distribution with variance σ_{ω}^2 . In Figure 2.20, we present the behaviour of bounds K_s and K_{ns} calculated using eqs. (2.90) and (2.91) respectively as a function of the density of nodes per unit area averaged over 100 randomly generated topologies. The local estimates ω_n are simulated as normally distributed random variables with mean $\mu = 0$ and variance $\sigma_{\omega} = 0.02$. It has been shown that the influence of fading directly impacts the convergence of the distributed estimation. In Figure 2.20, the values of K_s and K_{ns} decrease when the connectivity is increased. By setting $r_{con} = 1.4$, the networks always form fully connected graphs; this case presents the lowest value of K_s , as can be seen in Figure 2.13. We can also see that if one of the local measurements has a high reliability (high SNR) and differs significantly from others nodes, this



Figure 2.20: Dependance of bounds K_s and K_{ns} on node density.

frequency will not be locked and the local mean field amplitude r will have a lower value and some significant variations in amplitude. The change in the mean field rbelow a pre-established threshold may be used in CR systems to indicate that one of the nodes senses a strong signal in a certain frequency band which is not visible for the other nodes. This can be very useful in solving the hidden terminal problem in CR networks. In the analysis of the convergence properties we do not consider the influence of fading, because the $\sin(\cdot)$ term disappears when multiplying by the row vector $c^T = 1_N^T D_C$ in equation (2.86). However we are investigating this for further studies.

Chapter 3 Sequential Analysis Detection in Cognitive Radio Networks

Most of the existing spectrum sensing schemes found in the literature are based on fixed sample size detectors, which have a preset and fixed sensing time. In this chapter, we present some novel results based on the work of Abraham Wald [29], who showed that a detector based on a sequential detection requires less average sensing time than a fixed size detector. We show that, in general, it is possible to achieve the same performance as other fixed sample based techniques using as few as half of the samples on average for low SNR scenarios. We then assess the impact of non-coherent detection with signals detected using sequential analysis, and we use the Wald test as a new of cooperative approach for sensing. This is addressed as an optimal fusion rule for distributed Wald detectors, and its performance is assessed. Later on, we present a novel methodology to evaluate the cumulants of the sample random distribution in sequential analysis. We use this to present a modification of the Dual Sequential Ratio Test algorithm used for Primary User (PU) detection in Cognitive Radio (CR) networks. In the considered scenario, the Secondary Users (SUs) utilize the sequential ratio test to sense the wireless channel looking for the presence of a transmitting PU. A Fusion Centre (FC) gathers the decisions from the SUs in order to perform a sequential ratio test and achieve a final verdict. Collection instance by the FC is optimized such that the total time to make a decision is minimized. This allows for better energy usage from the SUs along with reliable and fast detection of the PU.

3.1 Introduction

There exist several spectrum sensing algorithms aimed to solve the hypotheses testing problem explained in the previous chapters. In CR networks it is also of major importance to be able to detect the presence of PU as fast as possible, since the time of decision has great impact on the overall throughput of the system [60]. It is shown in [29,31] that Wald's Sequential Probability Ratio Test (SPRT) results in a savings of about fifty percent in the average number of observations in comparison to other well-known techniques such, as the Neyman-Pearson (NP) test. In order to improve performance and avoid hidden terminal problems, cooperative sensing can also be used [11]. There exist several papers in the literature focused on reducing the power consumption in distributed CR sensing scenarios. A distributed scheme is proposed in [61] which groups SUs into clusters and assigns one specific user as a cluster head, which gathers the spectrum sensing results from the other users in the cluster and forwards the local result to a FC. The energy savings comes from some SUs now sending their decision to their cluster head and not to the FC. In [62], a sleeping and censoring scheme for distributed networks is proposed in order to minimize energy consumption. In this case, each radio stops sensing while in sleep mode. The results from each user are sent to the FC only when they are within a reliable energy region defined by two thresholds which are to be optimized. While an energy savings occurs in [61] and [62], nothing is said about how fast their detection scheme performs. In [63], the authors use a Bayesian formulation to detect abrupt changes in multiple on-off processes. While they use a modification of the cumulative sum algorithm (CUSUM) for multiple channels, they do not address with the distribution approach. Moreover, a priori information about PU activity must be assumed. In [64], a random access-based reporting order control scheme is proposed for cooperative sensing. In this scenario, the local test statistics are reported to the FC in descending order of magnitude. They gain in time of detection by reducing the reporting time to the FC. In [65], the authors propose a Dual SPRT to perform collaborative spectrum sensing. Here, the SPRT is used at both the SU front-end and the FC in charge of taking the final decision about the presence of the PU. The FC is required to gather information about the decisions made by SUs at a particular instance. However because of the nature of sequential analysis, those decisions are not always available at the same time. If the FC waits to gather information until all SUs have made a decision, the accuracy of the sensing improves but the overall throughput decreases. On the other hand, if the FC makes a decision with very little information, the reliability of its final decision is compromised. For this reason, we suggest that decisions from SUs are only transmitted after an optimal time τ_0 in order to decrease power usage and provide sufficient number of decisions to the FC. In [66], a threshold broadcast scheme is proposed for collaborative quickest spectrum sensing. The authors obtain a reduction in the detection delay in comparison to schemes using random broadcast. Our approach differs from theirs, as we use a FC to make a final decision and do not consider limited communication slots.

3.2 Sequential Analysis of A. Wald

The sequential analysis and the sequential probability ratio test (SPRT) were introduced by A. Wald in 1943 [29] and have proven to be highly effective in taking decisions between two known hypotheses $(\mathcal{H}_0, \mathcal{H}_1)$. While most of the efforts in the analysis of detection are focused on the NP detectors [16], it is well known that, on average, sequential detection provides a substantially faster operation. As shown in [29], the SPRT frequently results in a savings of about fifty percent in the average number of observations in comparison to other well-known detection techniques, such as the aforementioned NP decision test which is based on fixed number of observations. In the NP detection test, the logarithm of the Maximum Likelihood Ratio (MLR) is compared to a single threshold Λ_0 at a predefined and fixed observation interval T. In contrast, the sequential test compares the MLR to two thresholds Λ_1 and Λ_2 until a certain condition is satisfied. The parameters of the test are calculated based on the required P_{fa} and P_{md} as described in [29, 49]. Unlike the NP test where the duration of testing is chosen based on the desired probabilities of errors, the decision time in the Wald test is allowed to fluctuate. As a result, a faster decision time can be achieved on average. The thresholds in question can be calculated (upper bounded)

as

$$\Lambda_1 = \ln \frac{1 - P_{md}}{P_{fa}} = \ln D,$$

$$\Lambda_2 = \ln \frac{P_{md}}{1 - P_{fa}} = \ln C.$$
(3.1)

The test procedure consists of sequentially accumulate m samples and calculate the cumulative sum of the m-th log-LR as

$$\Lambda_1 < \sum_{i=1}^m \mathcal{R}_i < \Lambda_2, \tag{3.2}$$

where \mathcal{R}_i is a single log-likelihood ratio sample. If eq. (3.2) is satisfied, the experiment is continued by taking an additional sample and increasing m by 1. However, if

$$\sum_{i=1}^{m} \mathcal{R}_i \ge \Lambda_2,\tag{3.3}$$

the process is terminated with the acceptance of \mathcal{H}_1 . Similarly,

$$\sum_{i=1}^{m} \mathcal{R}_i \le \Lambda_1 \tag{3.4}$$

leads to termination with the acceptance of \mathcal{H}_0 . An illustrative example of Wald test can be seen in Figure 3.1, where a decision is not made as long as the MLR remains between the thresholds Λ_1 and Λ_2 . The decision in favour of the hypothesis \mathcal{H}_0 (absence of PU) is made if the MLR becomes smaller than Λ_1 , whereas the hypothesis \mathcal{H}_1 (presence of PU) is admitted if the MLR exceeds Λ_2 . Therefore, in contrary to the NP test, decision-making instance for sequential analysis is not fixed, and is generally a random variable. It is thus possible to obtain an average number of samples (mean decision time) for accepting either one of the two hypotheses, depending on which of the two (\mathcal{H}_0 or \mathcal{H}_1) is correct, as:

$$E\{\nu|\mathcal{H}_0\} = \frac{1}{A} \left[(1 - P_{fa}) \ln C + P_{fa} \ln D \right], \qquad (3.5)$$



Figure 3.1: Sequential Detection

$$E\{\nu|\mathcal{H}_1\} = \frac{1}{B} \left[P_{md} \ln C + (1 - P_{md}) \ln D \right], \qquad (3.6)$$

where

$$A = E \{ \mathcal{R} | \mathcal{H}_0 \},$$

$$B = E \{ \mathcal{R} | \mathcal{H}_1 \},$$
(3.7)

and \mathcal{R} is the log-likelihood ratio after ν steps. It follows that the decision time is given by $T = \nu T_s$ where T_s is the sampling interval and ν is the number of steps necessary to make a decision. The decision time T is a random variable, and can be described by its PDF $p_{\nu}(t)$. An exact solution for such PDF is still unknown, although some approximations are suggested by Wald [29] for the asymptotic cases when D = const, $C \to 0$ or $C = \text{const}, D \to \infty$. This approximated solution is known as the Wald's PDF. The estimates of the variance σ_{ν}^2 are also known under the assumption that $P_{fa} \ll P_{md}$. The derivations for the parameters of the distribution $p_{\nu}(t)$ are based on the notion of the so-called *Operation Characteristic* (OC) L(a). This is defined for the low SNR case as:

$$L(a) = \frac{D^{h(a)-1}}{D^{h(a)} - C^{h(a)}},$$
(3.8)

where a is the parameter of the hypothesis testing. In other words, $a = a_0$ corresponds to the hypothesis \mathcal{H}_0 while $a = a_1$ corresponds to the hypothesis \mathcal{H}_1 . For detection problems in fading conditions, a is simply proportional to the SNR of the PU, so that $a_0 = 0$ if the PU is absent, and equal to the average link SNR (a_1) in the case of \mathcal{H}_1 . Furthermore, h(a) is a unique non-zero root of the equation:

$$\int_{-\infty}^{\infty} \left[\frac{p_{a_1}(x)}{p_{a_0}(x)} \right]^{h(a)} p_a(x) dx = 1,$$
(3.9)

where $p_a(x)$ is the PDF of the observation based on the parameter a, and $p_{a_1}(x)/p_{a_0}(x)$ is the likelihood ratio for two hypotheses \mathcal{H}_0 and \mathcal{H}_1 . It is shown in [29] that $h(a_1) = -1$ and $h(a_0) = 1$.

3.2.1 Wald Test for Complex Random Variables

Let us consider the testing zero mean hypothesis in complex Additive White Gaussian Noise (AWGN) described as

$$\mathcal{H}_0: z_i = x_i + jy_i = w_i$$

$$\mathcal{H}_1: z_i = m + w_i$$
(3.10)

where $m = m_I + jm_Q = \mu \exp(j\phi_m) \neq 0$ is the complex non-zero mean, and w_i is the i.i.d. complex zero-mean Gaussian process of variance σ^2 . A single sample log-likelihood ratio \mathcal{R}_i is given by

$$\mathcal{R}_{i} = \ln \frac{p_{1}(z_{i}; \mathcal{H}_{1})}{p_{0}(z_{i}; \mathcal{H}_{0})} = \ln \frac{C \exp \left[-(x_{i} - \mu \cos \phi_{m})^{2} / \sigma^{2} - (y_{i} - \mu \sin \phi_{m})^{2} / \sigma^{2}\right]}{C \exp \left[-x_{i}^{2} / \sigma^{2} - y_{i}^{2} / \sigma^{2}\right]} = \frac{2\mu \left(x_{i} \cos \phi_{m} + y_{i} \sin \phi_{m}\right) - \mu^{2}}{\sigma^{2}}.$$
(3.11)

After N steps of the sequential test the cumulative log-likelihood \mathcal{R} ratio becomes

$$\mathcal{R} = \sum_{n=1}^{N} \mathcal{R}_i = \frac{2\mu}{\sigma^2} \mathcal{T}_N - \frac{N\mu^2}{\sigma^2},$$
(3.12)



Figure 3.2: Comparison of Neyman-Pearson Test and Sequential Probability Ratio Test $(P_{FA} = 0.1, P_D = 0.9)$.

where

$$\mathcal{T}_N = \cos \phi_m \sum_{n=1}^N x_n + \sin \phi_m \sum_{n=1}^N y_n.$$
 (3.13)

The rest of the test follows the procedure outlined in Section 3.2. Figure 3.2 shows the performance comparison in number of samples needed between the Wald Test and the NP Test. Notice that the NP test needs in general almost twice the number of samples in order to detect the presence of the signal. It follows from eq. (3.13) that the sufficient statistic in the case of complex observations is given by

$$\mathcal{T} = \sum_{n=1}^{N} \Re\{x \exp(-j\phi_m)\}.$$
(3.14)

The processing of the received signal is implemented in two stages: first, the data is unitary rotated by the angle ϕ_m in order to align the mean along the real axis; then, the real part of the data is analyzed using the same procedure as with purely real data.

3.2.2 Average Number of Samples

Following the sequential test procedure defined in eq.(3.2) we can rewrite the loglikelihood ratio as

$$\mathcal{R} = \sum_{i=1}^{m} \mathcal{R}_i = \ln \frac{p(z_1, \dots, z_m | \mathcal{H}_1)}{p(z_1, \dots, z_m | \mathcal{H}_0)},$$
(3.15)

where the random variable m represents the required number of samples needed to terminate the test. As stated in [29], it is possible to neglect the excess on threshold Λ_1 and Λ_2 , and the random variable can therefore have four possible combinations of terminations and hypotheses:

$$\mathcal{R} = \begin{cases} P_{FA}\Lambda_1 & \text{if } \mathcal{H}_0 \text{ is true} \\ P_D\Lambda_1 & \text{if } \mathcal{H}_1 \text{ is true} \\ (1 - P_{FA})\Lambda_2 & \text{if } \mathcal{H}_0 \text{ is true} \\ P_M\Lambda_2 & \text{if } \mathcal{H}_1 \text{ is true} \end{cases}$$
(3.16)

Following the same reasoning, we can calculate the conditional expectation for the random variable \mathcal{R} as

$$\bar{\mathcal{R}} = \begin{cases} P_{FA}\Lambda_1 + (1 - P_{FA})\Lambda_2 & \text{if } \mathcal{H}_0 \text{ is true} \\ P_D\Lambda_1 + P_M\Lambda_2 & \text{if } \mathcal{H}_1 \text{ is true} \end{cases}$$
(3.17)

It is now possible, to obtain the average number of samples (decision time) for accepting one of the two hypothesis as:

$$\bar{n}(\mathcal{H}_0) = \frac{P_{FA}\Lambda_1 + (1 - P_{FA}\Lambda_2)}{\bar{\mathcal{R}}(\mathcal{H}_0)},$$

$$\bar{n}(\mathcal{H}_1) = \frac{P_D\Lambda_1 + (1 - P_D\Lambda_2)}{\bar{\mathcal{R}}(\mathcal{H}_1)},$$
(3.18)



Figure 3.3: Average Number of Samples for Detection in Sequential Analysis $(P_D = 0.9).$

where the term $\bar{\mathcal{R}}(\mathcal{H}_0)$ can be calculated as

$$\bar{\mathcal{R}}(\mathcal{H}_0) = \frac{\sum_{i=1}^N \mathcal{R}_i}{N},\tag{3.19}$$

if no signal is present. The term $\overline{\mathcal{R}}(\mathcal{H}_1)$ can be calculated analogously assuming there is a signal present as follows:

$$\bar{\mathcal{R}}(\mathcal{H}_1) = \frac{\sum_{i=1}^N \mathcal{R}_i}{N}.$$
(3.20)

Figure 3.3 shows the average number of samples needed to achieve $P_D = 0.9$ for different SNR. The deviation at high SNRs occurs in practice for very high SNRs, when just one sample is more than enough to detect the presence of a PU.



Figure 3.4: Approximation of decision time using Wald Distribution $(P_D = 0.9)$.

3.2.2.1 Decision Time Distribution

The decision time when using sequential analysis for detection can be modelled as a random variable, which means that it can be completely described by its PDF. Although an exact shape or closed analytical solution for such a PDF is not generally known, a very good approximation is available (especially in the the low SNR region) called the Wald distribution or the inverse Gaussian distribution, defined as

$$f(x) = \frac{\lambda}{2\pi x^3} \exp \frac{-\lambda (x-\mu)^2}{2\mu^2 x} \qquad x > 0,$$
 (3.21)

where μ stands for the mean and $\lambda > 0$ is the shape parameter. Figure 3.4 shows Wald's distribution in order to approximate the decision time for $P_D = 0.9$.

3.3 Sequential Probability Ratio Test for Partially Coherent Channels

Let us consider the detection of a signal in a channel a with partially known phase. Hence, the received signal can be modelled as

$$z_i = m \exp(j\Delta) + w_i, \tag{3.22}$$

where $m = m_I + jm_Q = \mu \exp(j\phi_m)$ is a deterministic and known complex constant, and w_i is a complex zero mean Gaussian noise sample with variance σ^2 . The random variable Δ represents the uncertainty in the measurement of the phase of the carrier, and its distribution can be described by the PDF $p_{\Delta}(\Delta)$. In the following analysis, it is assumed that the phase uncertainty is described by the Von Mises (or Tikhonov) PDF defined as [67]:

$$p_{\Delta}(\delta) = \frac{\exp\left[\kappa \cos(\Delta - \Delta_0)\right]}{2\pi I_0(\kappa)}.$$
(3.23)

The parameter Δ_0 represents the bias in the determination of the carrier's phase, while κ represents the quality of the measurements. A few particular cases can be obtained from eq. (3.23) using the proper choice of parameters. We summarize three major cases as:

- 1. Perfect phase recovery (coherent detection): $\kappa = \infty$, $\Delta_0 = 0$, and, thus, $p_{\Delta}(\Delta) = \delta(\Delta)$.
- 2. No phase recovery (non-coherent detection): $\kappa = 0$ and, $p_{\Delta}(\Delta) = \frac{1}{2\pi}$.
- 3. Constant bias: $\kappa = \infty$, $\Delta_0 \neq 0$, $p_{\Delta}(\Delta) = \delta(\Delta \Delta_0)$.

We will derive the general expression first, and then investigate particular cases to isolate effects of the parameters on the performance of SPRT.

3.3.1 Average Likelihood Ratio

For a single observation z_i , the probability densities $p_1(z_i)$ and $p_0(z_i)$ corresponding to each of the hypotheses \mathcal{H}_1 and \mathcal{H}_0 are given by [31]

$$p_1(z_i) = C \exp\left[-\frac{(x_i - \mu\cos(\phi_m + \Delta))^2}{\sigma^2}\right] \exp\left[-\frac{(y_i - \mu\sin(\phi_m + \Delta))^2}{\sigma^2}\right], \quad (3.24)$$

and

$$p_0(z_i) = C \exp\left[-\frac{x_i^2 + y_i^2}{\sigma^2}\right].$$
 (3.25)

For a given Δ , the likelihood ratio L_i can be calculated to be

$$L_{i} = \frac{p_{1}(z_{i})}{p_{0}(z_{i})} = \exp\left[\frac{2\mu\left(x_{i}\cos(\phi_{m} + \Delta) + y_{i}\sin(\phi_{m} + \Delta)\right) - \mu^{2}}{\sigma^{2}}\right].$$
 (3.26)

The conditional (on Δ) likelihood ratio $L(N|\Delta)$ considered over N observation is then, the product of the likelihoods of individual observations, therefore

$$L(N|\Delta) = \prod_{n=1}^{N} \frac{p_1(z_n)}{p_0(z_n)} = \exp\left[\frac{2\mu \sum_{n=1}^{N} \left(x_n \cos(\phi_m + \Delta) + y_n \sin(\phi_m + \Delta)\right) - N\mu^2}{\sigma^2}\right]$$
$$= \exp\left[\frac{2\mu \mathcal{T}(N, \Delta)}{\sigma^2}\right] \exp\left[-\frac{N\mu^2}{\sigma^2}\right],$$
(3.27)

where

$$\mathcal{T}(N,\Delta) = \cos(\phi_m + \Delta) \sum_{n=1}^{N} x_n + \sin(\phi_m + \Delta) \sum_{n=1}^{N} y_n.$$
(3.28)

Let us introduce the new variables X(N), Y(N), Z(N) and $\Psi(N)$, defined by

$$X(N) = Z(N)\cos\Psi(N) = \sum_{n=1}^{N} x_n$$
 (3.29)

$$Y(N) = Z(N)\sin\Psi(N) = \sum_{n=1}^{N} y_n$$
 (3.30)

Using this notation, eq. (3.28) can now be rewritten as

$$\mathcal{T}(N,\Delta) = Z(N)\cos\left[\phi_m + \Delta - \Psi(N)\right].$$
(3.31)

The average likelihood [49] L(N) can now be obtained by averaging eq. (3.27) over the distribution of $p_{\Delta}(\Delta)$ to produce

$$\bar{L}(N) = \exp\left[-\frac{N\mu^2}{\sigma^2}\right] \int_{-\pi}^{\pi} \exp\left[\frac{2\mu Z(N)\cos\left[\phi_m + \Delta - \Psi(N)\right]}{\sigma^2}\right] p_{\Delta}(\Delta) d\Delta.$$
(3.32)

In turn, this expression can be further specialized if $p_{\Delta}(\Delta)$ is given by eq. (3.23) as

$$\bar{L}(N) = \exp\left[-\frac{N\mu^2}{\sigma^2}\right] \frac{1}{I_0(\kappa)} I_0 \left[\sqrt{\frac{4\mu^2 Z^2(N)}{\sigma^4} + \frac{4\mu Z(N)\kappa}{\sigma^2}} \cos\left[\phi_m - \Psi(N) - \Delta_0\right] + \kappa^2\right]$$
(3.33)

Equation (3.33) is reduced to eq. (3.13) if $\Delta_0 = 0$ and $\kappa = \infty$. Furthermore, the deterministic phase bias Δ_0 can be eliminated from consideration by taking $\tilde{z}_i = z_i \exp[-j(\phi_m + \Delta_0)]$ instead of z_i . Therefore, eq. (3.33) can be simplified to

$$\bar{L}(N) = \exp\left[-\frac{N\mu^2}{\sigma^2}\right] \frac{1}{I_0(\kappa)} I_0 \left[\frac{1}{\sigma^2} \sqrt{4\mu^2 Y^2(N) + \left[2\mu X(N) + \kappa\sigma^2\right]^2}\right].$$
 (3.34)

In the case of non-coherent detection, *i.e.* $\kappa = 0$, eq. (3.34) assumes the very well known form

$$\bar{L}(N) = \exp\left[-\frac{N\mu^2}{\sigma^2}\right] I_0\left[\frac{2\mu Z(N)}{\sigma^2}\right].$$
(3.35)

The construction of the likelihood ratio can be considered a two-step process. In the first step, the inphase and quadrature components are independently accumulated in order to lessen the effect of AWGN. In the second step, the values of X(N) and Y(N) must be combined, depending on the available information. For the case of coherent reception, it is know a priori that the quadrature component Y(N) contains only noise and it is ignored in the likelihood ratio. However, for a non-coherent reception scenario, one cannot distinguish between the in-phase and quadrature components, and their powers are equally combined to form Z(N). In the intermediate case, both



Figure 3.5: Impact of coherency on the average number of samples.

of the components are combined according to (3.34) with more and more emphasis put on the in-phase component X(N) as coherency increases along with κ . In Figure 3.5, we present the impact of the non-coherent detection in the number of samples needed in order to detect a signal with respect to a P_D target. Notice that the main repercussion of the non-coherence detection, is the increase of samples to nearly twice that of the coherent detector. The non-coherent Wald sequential test procedure can, therefore be thought as having the same efficiency (in terms of number of samples) as the coherent NP test.

3.4 Cumulant Analysis of the PDF of the Random Time of Sequential Analysis

As shown in Section 3.2.2, the expression for the average detection time in sequential detection can be easily obtained for the case of low SNR [29]. However, obtaining

other statistical characteristics, such as a complete probability density of the decision time, is a rather difficult task. Approximations by the Wald PDF or by the first few cumulants have been suggested only in cases of significantly different thresholds of detection. It is assumed that the probability of false alarm (P_{fa}) is significantly smaller than the probability of miss detection (P_{md}) (*i.e.*, $P_{fa} << P_{md}$), but this is not always applicable for CR networks. Since the analysis time T and the number of samples ν needed to make a decision are related through $T = \nu T_s$, we will focus on determining the parameters of the distribution of ν . The average time (eqs. (3.5),(3.6)) can be expressed in terms of the OC as:

$$E_a\{\nu\} = \frac{L(a)\ln C + 1 - L(a)\ln D}{E_a\{\mathcal{R}\}},$$
(3.36)

where C and D are the thresholds defined in eq. (3.1) and

$$E_a\{\mathcal{R}\} = \int_{-\infty}^{\infty} \ln \frac{p_{a_1}(x)}{p_{a_0}(x)} p_a(x) dx.$$
 (3.37)

Here

$$\mathcal{R}(x) = \ln \frac{p_{a_1}(x)}{p_{a_0}(x)} \tag{3.38}$$

is the log-likelihood ratio for the two hypotheses \mathcal{H}_0 and \mathcal{H}_1 . The cumulant generating function defined as

$$\Psi_{\nu}(\vartheta) = \ln \Theta_{\nu}(\vartheta), \qquad (3.39)$$

where

$$\Theta_{\nu}(\vartheta) = \int_{-\infty}^{\infty} p_{\nu}(x) \exp\left(j\vartheta x\right) dx, \qquad (3.40)$$

can be represented by the following series:

$$\Psi_{\nu}(\vartheta) = \sum_{i=1}^{\infty} \frac{\kappa_i}{i!} (j\vartheta)^i.$$
(3.41)

The coefficients κ_i in eq.(3.41) are the cumulants of $p_{\nu}(t)$ [50]. Given $\Theta_{\nu}(\vartheta)$ or $\Psi_{\nu}(\vartheta)$, the cumulants in question can be calculated by

$$\kappa_i = j^{-i} \frac{d^i}{d\vartheta^i} \ln \Theta_{\nu}(\vartheta) \bigg|_{\vartheta=0} = j^{-i} \Psi^{(i)}(0).$$

Particularly,

$$\kappa_1 = -j\Psi'_{\nu}(0) = E_a\{\mathcal{R}\}, \kappa_2 = -\Psi''_{\nu}(0) = \sigma_a^2.$$
(3.42)

If only two cumulants κ_1 and κ_2 are taken into account, a Gaussian approximation of the real distribution $W_{\nu}(t) \approx N(\kappa_1, \kappa_2)$ is obtained; if the first four cumulants κ_1 - κ_4 are taken into account, we obtain the so-called curtosis approximation. The curtosis approximation is a more accurate approximation for values near the mean, however it is poor in approximating the tails of the distribution. Given the complexity of calculating higher-order cumulants, we focus here only on the Gaussian approximation. It is shown in [29] that eq. (3.9) leads to the following expression for the characteristic function $\Theta_T(j\vartheta)$:

$$\theta_{\nu}(j\vartheta) = \frac{D^{t_2(\vartheta)} - D^{t_1(\vartheta)} + C^{t_1(\vartheta)} - C^{t_2(\vartheta)}}{C^{t_1(\vartheta)}D^{t_2(\vartheta)} - D^{t_1(\vartheta)}C^{t_2(\vartheta)}},\tag{3.43}$$

where $t(\vartheta)$ are the roots of the equation:

$$g(t) = \int_{-\infty}^{\infty} \left[\frac{p_{a_1}(x)}{p_{a_0}(x)} \right]^{t(\vartheta)} p_a(x) dx = e^{-j\vartheta}, \qquad (3.44)$$

which satisfy the following conditions:

$$\lim_{\vartheta \to 0} t_1(\vartheta) = 0,$$

$$\lim_{\vartheta \to 0} t_2(\vartheta) = h(a).$$
(3.45)

Using eq. (3.44) and expanding the exponent under the integral sign into Taylor series, one obtains the following relationship between moments $M_N(\mathcal{R})$ of the log-likelihood \mathcal{R} and values of $t_1(\vartheta)$ and $t_2(\vartheta)$:

$$g(t) = \int_{-\infty}^{\infty} e^{t(\vartheta)\mathcal{R}(x)} p_a(x) dx, \qquad (3.46)$$

and

$$\ln g(t) = E_a \{\mathcal{R}\} t(\vartheta) + \frac{\sigma_a^2}{2} t^2(\vartheta) + \dots = -j\vartheta.$$
(3.47)

Consequently, the algorithm for evaluating any cumulant κ_i can be accomplished using the following steps:

- 1. Differentiate the logarithm in eq. (3.43) *i*-times and evaluate it for $\vartheta \to 0$.
- 2. Differentiate eq.(3.44) or eq. (3.47) *i*-times and evaluate it for $\vartheta \to 0$ taking into account that $t_1(0) = 0$, $t_2(0) = h(a)$.
- 3. Obtain recurrent expressions for $t_k^{(i)}(0)$ from $t_k^{(i-1)}(0), t_k^{(i-2)}(0), \dots$ (k = 1, 2).
- 4. Use 3. to define $\psi^{(i)}(0)$ and, therefore, κ_i

After some simple but lengthy calculations one obtains the following values of the derivatives of t_1 and t_2 :

For the case of κ_1 , this immediately results in the expressions given by eqs. (3.5) and (3.6) for the mean values (or first cumulant). The variance $\sigma_a^2 = \kappa_2$ can then be

found from eqs. (3.43) and (3.48) to produce:

$$\begin{split} \sigma_{a_0}^2 &= \left[\frac{(D^2 - D + C - DC)\ln(C)}{E_{a_0}^3\{\mathcal{R}\}} + \frac{(C^2 - C + D - DC)\ln(D)}{E_{a_0}^3\{\mathcal{R}\}} \right] \ln(D) + \\ & \left[\frac{2C - 2DC}{E_{a_1}\{\mathcal{R}\}E_{a_0}\{\mathcal{R}\}} + \frac{DC + C - D - 1}{E_{a_1}^2\{\mathcal{R}\}} \right] \ln^2(C) + \\ & \left[\frac{4DC - 2C - 2D}{E_{a_1}\{\mathcal{R}\}E_{a_0}\{\mathcal{R}\}} + \frac{2 - 2DC}{E_{a_1}^2\{\mathcal{R}\}} \right] \ln(D)\ln(C) + \\ & \left[\frac{2D - 2DC}{E_{a_1}\{\mathcal{R}\}E_{a_0}\{\mathcal{R}\}} + \frac{DC - D + C - 1}{E_{a_0}^2\{\mathcal{R}\}} \right] \ln^2(D), \end{split}$$

$$\sigma_{a_1}^2 &= \left[\frac{(CD^2 - C^2D + C^2 - CD)\ln(C)}{E_{a_1}^3\{\mathcal{R}\}} + \frac{(D^2 - CD^2 + C^2D - CD)\ln(D)}{E_{a_1}^3\{\mathcal{R}\}} \right] \ln^2(D) + \\ & \left[\frac{2C - D^2 - 2DC}{E_{a_1}\{\mathcal{R}\}E_{a_0}\{\mathcal{R}\}} + \frac{C^2D - D^2C^2 + CD - CD^2}{E_{a_1}^2\{\mathcal{R}\}} \right] \ln^2(C) + \\ & \left[\frac{4DC - 2CD^2 - 2CD}{E_{a_1}\{\mathcal{R}\}M_{a_0}\{\mathcal{R}\}} + \frac{2C^2D^2 - 2DC}{E_{a_1}^2\{\mathcal{R}\}} \right] \ln(D)\ln(C) + \\ \end{aligned}$$

$$\left[\frac{4DC - 2CD^2 - 2CD}{E_{a_1}\{\mathcal{R}\}M_{a_0}\{\mathcal{R}\}} + \frac{2C^2D^2 - 2DC}{E_{a_1}^2\{\mathcal{R}\}}\right]\ln(D)\ln(C) + \left[\frac{2C^2D - 2CD}{E_{a_1}\{\mathcal{R}\}E_{a_0}\{\mathcal{R}\}} + \frac{DC - C^2D + CD^2 - D^2C^2}{E_{a_1}^2\{\mathcal{R}\}}\right]\ln^2(D).$$
(3.50)

Equations (3.49) and (3.50) are valid for any relations between P_{fa} and P_{md} [68]. In Figure 3.6 we show the value of the variances $\sigma_{a_0}^2$ and $\sigma_{a_1}^2$ obtained with eqs. (3.49) and (3.50) respectively for different SNRs. The second cumulant κ_2 for the fixed thresholds D and C depends on $E_a\{\mathcal{R}\}$ for different hypothesis \mathcal{H}_0 , \mathcal{H}_1 . The probability that the duration of the Wald test exceeds some preset value ν_{max} is calculated by

$$\int_{\nu_{\text{max}}}^{\infty} p(\nu) d\nu \approx Q\left(\frac{\nu_{\text{max}} - M_1}{\sigma}\right).$$
(3.51)

It is often more practical to use ν_{max} instead of the average $E_a\{\nu\}$ since ν_{max} provides a realistic figure for comparison with the NP test. Moreover, ν_{max} can be used to compare characteristics of the classical sequential analysis considered here with new



Figure 3.6: Variance values for the time decision pdf in sequential analysis. proposals of sequential analysis (e.g., [69]).

3.5 Optimal Fusion Rule for distributed Wald detectors

This section generalizes the results in [70] to the case of distributed detection using Wald sequential analysis test as explained in Section 3.2. We assume there exist Msensors making individual detections according to the SPRT algorithm. Once a decision is made at an individual sensor, it is sent in binary form to the FC for further combining with other decisions. We assume that the value u = -1 is assigned if the hypothesis \mathcal{H}_0 is accepted, u = 1 if the hypothesis \mathcal{H}_1 is accepted, and u = 0 if no decision has been made yet. Only $u = \pm 1$ are communicated to the FC. Since each node uses the SPRT detection, the decision is made at a random moments of time. Therefore, at any given moment of time t, there is a random number $L(t) \leq M$ of decisions which are available at FC, as can be seen in Figure 3.7. The probability distribution of making a decision can be approximated either by the two parametric Wald distribution [29], or by the three parametric generalized inverse Gaussian



Figure 3.7: System model of data fusion system

distribution [71] as seen in Figure 3.4. Parameters of such distributions could be found through moment/cumulant fitting, using expressions derived in [29, 68]. Following [70], we treat this problem as a two-hypothesis detection problem with an individual detector decision being the observation. For a given number L = L(t) of decisions made by the time t, the optimum decision rule is equivalent to the following likelihood ratio test

Here $P(L|\mathcal{H}_0)$ is the probability of making exactly L decisions assuming that \mathcal{H}_1 is true and u_l is the decision made by l-th sensor. Furthermore, assuming the minimum probability of error criteria, (*i.e.* by setting $C_{00} = C_{11} = 0$ and $C_{01} = C_{10} = 1$), introducing the following notation $\mathbf{u}_L = \{u_1, u_2, \cdots, u_L\}$, and using the Bayes rule, one can recast equation (3.52) as

$$\frac{P(\mathcal{H}_1|\mathbf{u}_L, \mathbf{l})P(\mathbf{l}|\mathcal{H}_1)}{P(\mathcal{H}_0|\mathbf{u}_L, \mathbf{l})P(\mathbf{l}|\mathcal{H}_0)} \stackrel{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrsim}} 1,$$
(3.53)

or, after taking the natural logarithm on both sides,

$$\ln \frac{P(\mathcal{H}_1 | \mathbf{u}_L, \mathbf{l})}{P(\mathcal{H}_0 | \mathbf{u}_L, \mathbf{l})} + \ln \frac{P(\mathbf{l} | \mathcal{H}_1)}{P(\mathbf{l} | \mathcal{H}_0)} \overset{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrsim}} 0, \qquad (3.54)$$

where \mathbf{l} is the vector representing which sensors have made their decisions. Once again, following [70], one can calculate probabilities $P(\mathcal{H}_1|\mathbf{u}_L, \mathbf{l})$ and $P(\mathcal{H}_0|\mathbf{u}_L, \mathbf{l})$ as follows. In the case of the hypothesis \mathcal{H}_1 one can write

$$P(\mathcal{H}_{1}|\mathbf{u}_{L},\mathbf{l}) = \frac{P(\mathcal{H}_{1},\mathbf{u}_{L}|\mathbf{l})}{P(\mathbf{u}_{L}|\mathbf{l})} = \frac{P_{1}}{P(\mathbf{u}_{L}|\mathbf{l})} \prod_{S_{+}} P(u_{l} = +1|\mathcal{H}_{1}) \prod_{S_{-}} P(u_{l} = -1|\mathcal{H}_{1}) = \frac{P_{1}}{P(\mathbf{u}_{L}|\mathbf{l})} \prod_{S_{+}} (1 - P_{M,l}) \prod_{S_{-}} P_{M,l},$$
(3.55)

where S_+ is the set of all *i* such that $u_i = +1$ and S_- is the set of all *i* such that $u_i = -1$. Analogously, for the case of the hypothesis \mathcal{H}_0 one obtains

$$P(\mathcal{H}_0|\mathbf{u}_L,\mathbf{l}) = \frac{P(\mathcal{H}_0,\mathbf{u}_L|\mathbf{l})}{P(\mathbf{u}_L|\mathbf{l})} = \frac{P_0}{P(\mathbf{u}_L|\mathbf{l})} \prod_{S_-} (1 - P_{F,l}) \prod_{S_+} P_{F,l}.$$
 (3.56)

Finally, using equations (3.55) and (3.56) one obtains the following expression for the conditional log-likelihood

$$\ln \frac{P(\mathcal{H}_1 | \mathbf{u}_L, \mathbf{l})}{P(\mathcal{H}_0 | \mathbf{u}_L, \mathbf{l})} = \ln \frac{P_1}{P_0} + \sum_{S_+} \ln \frac{1 - P_{M,l}}{P_{F,l}} + \sum_{S_-} \ln \frac{P_{M,l}}{1 - P_{F,l}}.$$
 (3.57)

In order to evaluate the second term in the sum in eq. (3.54), let us first consider an arbitrary node $1 \le k \le N$. The distribution $p_{T,k}(\tau)$ of the decision time in such a node is assumed to be known. Therefore, the probability $P_{D,k}(t|\mathcal{H}_i)$ that the decision is made by the time $t = mT_s$ given that a specific hypothesis is true is given by

$$P_{D,k}(t|\mathcal{H}_i) = \int_{0}^{mT_S} p_k(\tau|\mathcal{H}_i) d\tau.$$
(3.58)

The probability that no decision has been made by the time t is then simply $1 - P_{D,k}(t|\mathcal{H}_i)$. As previously noted, the parameters of this distribution can be defined



Figure 3.8: Data fusion scheme considering sequential analysis decision from each sensor $(P_{M,l}=0.3,P_{F,l}=0.1).$

as in [68]. Therefore, the second term in the equation (3.54) is given by

$$\ln \frac{P(\mathbf{l}|\mathcal{H}_1)}{P(\mathbf{l}|\mathcal{H}_0)} = \sum_{l=1}^{L} \ln \frac{P_{D,l}(t|\mathcal{H}_1)}{P_{D,l}(t|\mathcal{H}_0)} + \sum_{l=L+1}^{M} \ln \frac{1 - P_{D,l}(t|\mathcal{H}_1)}{1 - P_{D,l}(t|\mathcal{H}_0)}.$$
 (3.59)

Finally, the fusion rule in the case of nodes making a decision according to Wald's criteria can be written as

$$f(\mathbf{u}) = \begin{cases} 1 & \text{if } a_0 + \sum_{l=1}^L a_l u_l > 0\\ -1 & \text{otherwise} \end{cases}$$
(3.60)



Figure 3.9: Data fusion scheme considering sequential analysis decision from each sensor $(P_{M,l}=0.1,P_{F,l}=0.3).$

where

$$a_0 = \ln \frac{P_1}{P_0} + \sum_{l=1}^{L} \ln \frac{P_{D,l}(t|\mathcal{H}_1)}{P_{D,l}(t|\mathcal{H}_0)} + \sum_{l=L+1}^{M} \ln \frac{1 - P_{D,l}(t|\mathcal{H}_1)}{1 - P_{D,l}(t|\mathcal{H}_0)}, \quad (3.61)$$

$$a_l = \ln \frac{1 - P_{M,l}}{P_{F,l}}$$
 if $u_l = 1$, (3.62)

$$a_l = \ln \frac{1 - P_{F,l}}{P_{M,l}}$$
 if $u_l = -1.$ (3.63)

Thus, the combining rule is similar to that suggested in [70], but with some significant differences in the term a_0 . In Figures 3.8 and 3.9, we show the performance of the data fusion scheme considering that each one of the sensors takes a decision based on the sequential detection criteria. In these figures, we plot the probability of miss detection (P_{MD}) and the probability of false alarm (P_{FA}) versus the moment when



Figure 3.10: Total Error Criteria

the FC gathers the decisions of the local observers. Notice that for $t \to \infty$ all graphs converge to the data fusion rule of [70]. It is clear that for small values of time, the FC has less information (since not all the detectors might have achieved a decision by then) and the final decision taken is much less accurate than for large values of time. Nevertheless, in some practical systems it would be impossible to wait that long for getting the decision from the FC, so we can use these results as a trade-off between the performance on the detection and the time of decision [72]. We also observe the impact that $P_{D,l}$ and $P_{F,l}$ have on the performance of the data fusion detector. For very small values of false alarm probability, $a_l \approx -\ln P_{M,l}$ if u = -1 in eq. (3.61) which means that the hypothesis \mathcal{H}_0 is always less weighted in eq. (3.60). In other words, the FC "trusts" more in those sensors who decide that \mathcal{H}_1 is true. Similarly, for very small values of miss detection probability, the hypothesis \mathcal{H}_0 is more weighted in the final sum in equation (3.60). A special case occurs when $P_{D,l} = P_{F,l}$, $P_0 = P_1$, and $t \to \infty$. In this situation, the scheme converts into the more simple majority decision approach seen in Chapter 2, which sums all u_l and compares with zero. Although it is simpler, the maximum likelihood approach performs better than the majority decision scheme in the minimum probability of error criteria, as can be seen in Figure 3.10 [30]. The perceptive reader may have noticed by now that there might be some confusion at the FC when there exists an even number of sensors and there is a tie in the decision. This can be settled by considering the *a priori* probabilities P_0 and P_1 which are inherent to the system.

3.6 Dual Sequential Spectrum Ratio Test

We consider a CR Network consisting of a single PU and L SUs with perfect channel state information (CSI) cooperating with a FC. Each of the SUs makes an individual decision using the SPRT and transmits it to the FC for a final decision, as can be seen in Figure 3.11. Each SU decision is communicated to the FC by sending a constant signal b_i if \mathcal{H}_i is accepted. As discussed in previous sections, since each SU node uses sequential detection, the individual decisions are made at random moments in time. At any given instance τ there is a random number $M(\tau) \leq L$ of sensors which have made their decision (see Figure 3.12). The FC can decide when to start performing its sequential analysis. SUs transmitting before the FC starts the decision procedure is a waste of power; however, it is important that a sufficient number of SUs have made their decision in order to improve the reliability from the FC's final verdict. Therefore, it is important to identify an optimal instance τ_o for the FC to start the sequential analysis in such a way that there is a significant savings in the SUs' power and as well as an accurate sensing. Considering the SPRT algorithm explained in Section 3.2, the signal sent to the FC by the *l*-th sensor at the moment $k \ge \tau$ can be expressed as

$$Y_{k,l} = b_1 \mathbf{1}_{\{\mathcal{R}_{k,l} \ge \Lambda_{2,l}\}} + b_0 \mathbf{1}_{\{\mathcal{R}_{k,l} \le \Lambda_{1,l}\}} + 0 \cdot \mathbf{1}_{\{\Lambda_{1,l} < \mathcal{R}_{k,l} < \Lambda_{2,l}\}},$$
(3.64)

where $\Lambda_{1,l}$ and $\Lambda_{2,l}$ are the upper and lower thresholds respectively used in the sequential test by the *l*-th sensor, and $\mathcal{R}_{k,l}$ is the log likelihood ratio [73]. The thresholds, as previously mentioned, are calculated in terms of the probabilities of false alarm



Figure 3.11: System Model

and miss detection of the l-th user as:

$$\Lambda_{1,l} \approx \ln\left(\frac{1 - P_{MD_l}}{P_{FA_l}}\right),$$

$$\Lambda_{2,l} \approx \ln\left(\frac{P_{MD_l}}{1 - P_{FA_l}}\right).$$
(3.65)

The PDF of the sent signal (eq.(3.64)), under each of the hypothesis \mathcal{H}_1 and \mathcal{H}_0 , can be expressed as follows:

$$p(k|\tau; \mathcal{H}_0) = (1 - P_{FA}) P_D(\tau | \mathcal{H}_0) \delta(k - b_0) + P_{FA} P_D(\tau | \mathcal{H}_0) \delta(k - b_1) + (1 - P_D(\tau | \mathcal{H}_0)) \delta(k),$$
(3.66)

and

$$p(k|\tau; \mathcal{H}_1) = P_{MD} P_D(\tau|\mathcal{H}_1) \delta(k - b_0) + (1 - P_{MD}) P_D(\tau|\mathcal{H}_1) \delta(k - b_1)$$

+ $(1 - P_D(\tau|\mathcal{H}_1)) \delta(k).$ (3.67)

Here, $P_D(\tau | \mathcal{H}_i)$ stands for the probability that the decision is made by the time $t = \tau$ given that the hypothesis \mathcal{H}_i is true and can be calculated as [73]:

$$P_D(\tau|\mathcal{H}_i) = \Phi\left(\sqrt{\frac{\lambda}{\tau}} \left(\frac{\tau}{\mu} - 1\right)\right) + e^{\frac{2\lambda}{\mu}} \Phi\left(-\sqrt{\frac{\lambda}{\tau}} \left(\frac{\tau}{\mu} + 1\right)\right), \qquad (3.68)$$

where $\Phi(\cdot)$ is the standard Gaussian distribution cumulative distribution function, μ is the average time of decision in the sequential test, and λ is related to the variance of the sequential test as

$$\lambda = \frac{\mu^3}{\sigma_\nu^2}.\tag{3.69}$$

The expressions to calculate μ and σ_{ν}^2 are shown in eqs. (3.5) and (3.6), and eqs. (3.49) and (3.50), respectively. In the following it is assumed, for simplicity, that $P_{MD} = P_{MD_l}$ and $P_{FA} = P_{FA_l}$. At the antenna on the FC, the signals sent by individual SUs are coherently combined, and the receiver's noise is added as

$$Y_k = \sum_{l=1}^L \sqrt{\gamma_l} Y_{k,l} + W_k, \qquad (3.70)$$

where W_k is an i.i.d. AWGN realization with variance σ_N^2 and channel coefficient γ_l , which is Rayleigh distributed according to

$$p(\gamma) = \frac{\gamma}{\sigma_h^2} e^{-\frac{\gamma^2}{2\sigma\gamma^2}}.$$
(3.71)

In order to make a decision on the status of the PU, the FC also performs the SPRT according to the following:

$$F_{k} = F_{k-1} + \ln \frac{p(Y_{k}; \mathcal{H}_{1})}{p(Y_{k}; \mathcal{H}_{0})} + \ln \frac{P(L|\mathcal{H}_{1})}{P(L|\mathcal{H}_{0})}.$$
(3.72)

In eq. (3.72), $F_0 = 0$ and $P(L|\mathcal{H}_i)$ is the probability of having exactly L sensors with decisions made given that \mathcal{H}_i is true. The process is then repeated until either $F_k > \Lambda_{2_{FC}}$ or $F_k < \Lambda_{1_{FC}}$, where $\Lambda_{1_{FC}}$ and $\Lambda_{2_{FC}}$ are the two thresholds defined analogously as in eq. (3.65) for the FC. If the first situation is true, \mathcal{H}_1 is claimed to



Figure 3.12: System Model for Decision.

be correct; if the latter is true, we say that \mathcal{H}_0 is correct. The log-likelihood ratio F_k used to perform the Wald test in the FC as required by eq. (3.72) can be evaluated in terms of the PDF $p(Y_k; \mathcal{H}_i)$. In order to do so, one needs to perform L+1 convolution operations, which can be an overwhelming task. Nevertheless, it is possible to get a very accurate approximation of the PDF $p(Y_k; \mathcal{H}_i)$ for low SNR utilizing a small deviation of the Gaussian distribution as follows [74]:

$$p(Y_k; \mathcal{H}_i) \approx p_0^{(i)}(Y_k) \left[1 + \frac{\kappa_3^{(i)}}{3! \cdot \kappa_2^{3/2(i)}} H_3\left(\frac{Y_k - \mu_i}{\sigma_i}\right) + \frac{\kappa_4^{(i)}}{4! \cdot \kappa_2^{2(i)}} H_4\left(\frac{Y_k - \mu_i}{\sigma_i}\right) \right],$$

for $i = 0, 1$ (3.73)

where $p_0^{(i)}(\cdot)$ is a Gaussian PDF with mean value μ_i and variance σ_i^2 , $H_n(\cdot)$ stands for the Hermitian polynomial of the *n*-th order [75], and κ_n is the *n*-th cumulant of the distribution $p(Y_k; \mathcal{H}_i)$. In order to obtain such an approximation, one needs to estimate the first four moments $m_k^{(i)}$ (or cumulants) of the desired PDF $p(Y_k; \mathcal{H}_i)$. Given the properties of cumulants [74], we can show that:

$$\begin{split} \kappa_{1}^{(i)} &= Lm_{1}^{(i)}, \\ \kappa_{2}^{(i)} &= L[m_{2}^{(i)} - m_{1}^{2(i)}], \\ \kappa_{3}^{(i)} &= L[m_{3}^{(i)} - 3m_{2}^{(i)}m_{1}^{(i)} + 2m_{1}^{3(i)}], \\ \kappa_{4}^{(i)} &= L[m_{4}^{(i)} - 4m_{3}^{(i)}m_{1}^{(i)} + 12m_{2}^{(i)}m_{1}^{2(i)} - 3m_{2}^{2(i)} - 6m_{1}^{4(i)}], \\ \sigma_{i}^{2} &= \kappa_{2}^{(i)} + \sigma_{N}^{2}. \end{split}$$

$$(3.74)$$

The moments $m_k^{(i)}$ can be found directly from eqs. (3.66) and (3.67) as

$$\begin{split} m_{1}^{(0)} &= \sigma_{h} \sqrt{\frac{\pi}{2}} \left(\left[(1 - P_{FA}) P_{D}(\tau | \mathcal{H}_{0}) \right] b_{0} + \left[P_{FA} \cdot P_{D}(\tau | \mathcal{H}_{0}) \right] b_{1} \right), \\ m_{2}^{(0)} &= 2\sigma_{h}^{2} \left(\left[(1 - P_{FA}) P_{D}(\tau | \mathcal{H}_{0}) \right] b_{0}^{2} + \left[P_{FA} \cdot P_{D}(\tau | \mathcal{H}_{0}) \right] b_{1}^{2} \right), \\ m_{3}^{(0)} &= 3\sigma_{h}^{3} \sqrt{\frac{\pi}{2}} \left(\left[(1 - P_{FA}) P_{D}(\tau | \mathcal{H}_{0}) \right] b_{0}^{3} + \left[P_{FA} \cdot P_{D}(\tau | \mathcal{H}_{0}) \right] b_{1}^{3} \right), \\ m_{4}^{(0)} &= 8\sigma_{h}^{2} \left(\left[(1 - P_{FA}) P_{D}(\tau | \mathcal{H}_{0}) \right] b_{0}^{4} + \left[P_{FA} \cdot P_{D}(\tau | \mathcal{H}_{0}) \right] b_{1}^{4} \right), \\ m_{1}^{(1)} &= \sigma_{h} \sqrt{\frac{\pi}{2}} \left(\left[P_{MD} \cdot P_{D}(\tau | \mathcal{H}_{1}) \right] b_{0} + \left[(1 - P_{MD}) P_{D}(\tau | \mathcal{H}_{1}) \right] b_{1}^{2} \right), \\ m_{2}^{(1)} &= 2\sigma_{h}^{2} \left(\left[P_{MD} \cdot P_{D}(\tau | \mathcal{H}_{1}) \right] b_{0}^{2} + \left[(1 - P_{MD}) P_{D}(\tau | \mathcal{H}_{1}) \right] b_{1}^{2} \right), \\ m_{3}^{(1)} &= 3\sigma_{h}^{3} \sqrt{\frac{\pi}{2}} \left(\left[P_{MD} \cdot P_{D}(\tau | \mathcal{H}_{1}) \right] b_{0}^{3} + \left[(1 - P_{MD}) P_{D}(\tau | \mathcal{H}_{1}) \right] b_{1}^{3} \right), \\ m_{4}^{(1)} &= 8\sigma_{h}^{2} \left(\left[P_{MD} \cdot P_{D}(\tau | \mathcal{H}_{1}) \right] b_{0}^{4} + \left[(1 - P_{MD}) P_{D}(\tau | \mathcal{H}_{1}) \right] b_{1}^{4} \right). \end{split}$$

In order to show the accuracy of this approximation, numerical examples are shown in Figure 3.13. The approximation performs quite well for values of SNR up to about 0 dB. For greater values, the deltas resulting from the sum of random variables in eqs. (3.66) and (3.67) become more evident due to the lack of noise power.

3.6.1 Optimization of Decision Time

One of the most important features required from a CR sensor is making decisions regarding the presence of the PU as fast as possible. From this perspective the FC


Figure 3.13: Approximation of $p(Y_k; \mathcal{H}_1)$ for L = 5, SNR = 0dB, $P_{md} = 0.2$, $P_{fa} = 0.1$.

has to start to collect information from SU sensors as soon as possible. However, this may unnecessarily drain energy at some of the SUs. Therefore, it is important to find an optimal value of time $\tau = \tau_0$, when the FC starts to acquire decisions from the SUs. The total time to make a decision is:

$$f(\tau) = \tau + \mathcal{E}[N(\tau)] \Delta T_s, \qquad (3.76)$$

where $N(\tau)$ is the number of samples required to make a decision at the FC, given that the FC starts making decision at time τ . Therefore, our goal is to minimize the function $f(\tau)$. Notice that in eq. (3.76), the first term is clearly an increasing function, while the second term is a monotonically decreasing function due to the nature of the analysis. Therefore, when τ is small, a larger number of samples are needed because fewer sensors are contributing their decisions. Equation (3.76) is therefore a convex function, as shown in Figure 3.14. In order to calculate the average number of samples $E\{N(\tau); \mathcal{H}_i\}$ required at the FC in order to make a decision, we utilize eqs. (8) and (9) in [76] expressed as:

$$E\{N(\tau); \mathcal{H}_{0}\} \approx \frac{-(A_{FC} - B_{FC} - A_{FC}e^{A_{FC}})}{\left(e^{A_{FC}} - e^{B_{FC}}\right) \left[D(p(Y_{k}; \mathcal{H}_{0})||p(Y_{k}; \mathcal{H}_{1}))\right]},$$

$$E\{N(\tau); \mathcal{H}_{1}\} \approx \frac{A_{FC} - B_{FC} - A_{FC}e^{A_{FC}}}{\left(e^{-A_{FC}} - e^{-B_{FC}}\right) \left[D(p(Y_{k}; \mathcal{H}_{1})||p(Y_{k}; \mathcal{H}_{0}))\right]},$$
(3.77)

where $D(p(Y_k; \mathcal{H}_i)||p(Y_k; \mathcal{H}_j))$ is the Kullback-Leibler (KL) distance defined as

$$D(p(Y_k; \mathcal{H}_i)||p(Y_k; \mathcal{H}_j)) = \int_{\infty}^{\infty} p(Y_k; \mathcal{H}_i) \ln\left(\frac{p(Y_k; \mathcal{H}_i)}{p(Y_k; \mathcal{H}_j)}\right) dY_k, \quad i, j = 0, 1 \quad i \neq j.$$
(3.78)

From eq. (3.76), the nonlinear function needed to be optimized is thus given by

$$f(\tau) = \tau + p(\mathcal{H}_0)E\left\{N(\tau); \mathcal{H}_0\right\} + p(\mathcal{H}_1)E\left\{N(\tau); \mathcal{H}_1\right\}.$$
(3.79)

Using numerical simulation, we show results in Section 3.6.2, including average energy used by the SUs in order to assess energy savings. As explained in Section 3.6, the *l*-SU sends its result as soon as it becomes available. Nevertheless, the FC makes its final decision at time $t = \tau$. Thus the average energy spent by the SUs is

$$E\{\epsilon(\tau)\} = \sum_{l=1}^{L} [\tau - \mu_l] P_{D_l}(\tau | \mathcal{H}_i) b_l^2, \qquad (3.80)$$

where μ_l is the average decision time for user l.

3.6.2 Numerical Results

Figure 3.15 shows the average number of samples needed at the FC in order to make a decision from the moment it starts running the SPRT algorithm. Notice that as the FC acquires information from more sensors $(i.e., \tau \to \infty)$, it takes fewer samples on average to achieve the final decision. These results confirm that there exists an



Figure 3.14: Optimum τ_0 for different number of sensors (SNR = -25dB).

important trade-off between how much energy is used to send the result from sensors to the FC and how fast the FC achieves a decision. Figure 3.14 shows that an optimum value τ_0 exists such that the net decision time in eq. (3.79) can be minimized (*i.e.*, eq. (3.79) is convex). Figure 3.16 shows values of τ_0 for different values of SNR and different number of sensors. The importance of this result is that it allow us to deterministically configure the moment when the FC should collect the decision from the sensors. Notice that such a instance can be chosen even before the optimal result, due to the high stability of the system in terms of decision time. We can see in Figure 3.14 that although the optimal time for L = 3 sensors is 125 samples, it would be possible to configure the FC to start gathering decisions with approximately 110 samples without significantly sacrificing the overall performance of the detection. Similar assumption can be made for cases where L = 5 and L = 11. We present in Figure 3.17 the total energy used by the sensors to transmit their decision: the energy spent when the FC starts the SPRT algorithm in the optimum time proposed (τ_0) , and when it waits for all sensors to finish. Table 4.1 shows it is possible to get



Figure 3.15: Average Number of Samples (ANS) to make a decision at the FC.



Figure 3.16: Optimum τ vs SNRdB.



Figure 3.17: Energy used a DSPRT scheme. The 'o' mark shows the optimum time for the FC to make a decision while the 'x' mark shows the time and energy used when the FC waits for all sensors to send their decision.

a very high reduction in energy compared to a classical DSPRT algorithm while at the same time significantly reducing the samples required to make a decision.

No. of Sensors	Optimum Decision Time	Energy Saved
3	75	97.63%
5	100	89.78%
11	125	82.43%

Table 3.1: Energy savings considering the optimum time for decision (τ_o) .

3.7 Conclusions

This chapter was devoted to the application of the sequential analysis technique to achieve faster spectrum sensing in CR networks. By using the SPRT, it is possible to detect the presence of a PU almost twice as fast as other fixed sample approaches such as NP detectors. This can be achieved when dealing in low SNR cases, a frequent real-life scenario. The effect of errors in the estimation of the phase of the carrier on the duration of the sequential analysis was also investigated. It was shown that using non-coherent detection in sensing the presence of PUs using sequential analysis requires almost twice as many samples as a coherent detection approach. We derived an optimal fusion rule using detectors that use sequential analysis for taking decisions. We assessed the performance of the system in terms of the time that it takes to gather the decision from all detectors. It was shown that for faster decision, the FC does not consider the verdict of all sensors and therefore the performance is reduced. On the other hand, as we wait longer to gather the decisions, the detection performance is better but the system experiences a higher latency. We presented a novel methodology to obtain any order cumulant of the PDF of the sample distribution in sequential detection. We investigated some further considerations of the dual SPRT not previously covered in the literature. In addition, we derived a very accurate approximation for the log-likelihood ratio needed to run the SPRT algorithm at the FC in the dual SPRT. We showed that it is possible to reduce the energy usage on SU sensors by optimizing the instance τ when the FC gathers their local decision in order to run its own SPRT. This instance can be evaluated by calculating the average number of samples needed to make a decision at the FC and optimizing the value of τ in such a manner that the total average time of decision in the system is minimized.

Chapter 4 Impact of Scattering Environment on Spectrum Sensing for Multi Antenna Detectors

In Chapter 2, we showed that the amount of information the SUs have regarding the characteristics of the PU signal (noise variance, covariance matrix, cyclic frequencies, *etc.*) has a significant influence on the detection probability of the PU presence in CR networks. We also showed, that it is possible to improve the performance of the detection algorithms by using more than one antenna at each of the secondary sensors. Nevertheless, when working with multiple antenna systems, the spatial and/or temporal correlation between antennas has a great impact on the overall performance of the communication system. In this chapter, we present a novel approach for accounting for both spatial and temporal correlation in CR devices equipped with multiple antennas. We derive an equivalent number of independent samples based on the scattering geometry and resulting correlation properties of the received signal. The performance of the system is investigated in terms of the Neyman-Pearson detection criteria. The results of this theoretical analysis are verified through extensive numerical simulations.

4.1 Introduction

Improving spectrum sensing reliability while limiting sensing duration to preserve power is still a subject of intensive research. A large number of approaches have been developed, including energy detection [18,36], matched-filter detection [12], cyclostationarity based detection [12, 14, 37]. However, most of these studies are focused on investigating the performance of particular schemes in ideal environments such as independent antennas in cooperative scenarios or in uniform or isotropic scattering conditions [27]. Such considerations eliminate the impact of the real environment and its variation even though that it is shown in many publications [46,91] and realistic measurements that such environments change frequently, especially in highly built areas. Understanding how a particular radio environment affects the performance of CR sensing abilities is, therefore, an important issue to consider. Furthermore, it is well known [92] that the distribution of the angle of arrival (AoA) (itself defined by the scattering environment) affects both temporal and spatial correlation of signals in antenna arrays. In this chapter we utilize a simple but generic model of the AoA distribution, suggested in [91], to describe the impact of scattering on the statistical properties of received signals. We also show how to incorporate the concept of Stochastic Degrees of Freedom (SDoF) [28,93] in order to obtain approximate expressions for the probability of missed detection in terms of number of antennas, scattering parameters and number of observations. Finally, we investigate the tradeoff between the number of antennas and the required observation interval in correlated fading environments. This chapter is organized as follows: Section 4.2 presents the corresponding signal model, and in Section 4.3 we derive the performance of SIMO detection schemes for three different cases: constant independent channels, constant spatially correlated channels, and independent channels with temporal correlation. A few examples of correlation models are shown in Section 4.4. In Section 4.5 we evaluate the space-time processing trade-off of the framework while in Section 4.6 the conclusions of the chapter are presented.

4.2 Signal Model

Let us consider a primary transmitter which transmits a pilot signal s over L symbols in order to sound the primary channel. The CR network can sense the pilot signal using N_R receiving antennas as depicted in Figure 4.1a. The received signal matrix \mathbf{X} of size $N_R \times L$ can be written in terms of the $N_R \times L$ channel matrix $\mathbf{H} = \{h_{rl}\} \in \mathbb{C}$





(b) Signal received at the *r*-th antenna in the *l*-th pilot time slot.

Figure 4.1: System Model.

and the noise matrix ${\bf W}$ of the same size as

$$\mathbf{X} = \mathbf{H}s + \mathbf{W},\tag{4.1}$$

where **W** is a zero mean matrix of covariance $\sigma_n^2 \mathbf{I}_n$ and **H** is a zero mean Complex Gaussian matrix with covariance matrix $\mathbf{R}_{\mathbf{H}}$. The element h_{rl} is the channel transfer coefficient from the transmitter to the *r*-th receiving antenna measured at the *l*-th pilot time slot as depicted in Figure 4.1b. Using the vectorization operation, it is possible to rewrite eq. (4.1) as

$$\mathbf{x} = \mathbf{h}s + \mathbf{w},\tag{4.2}$$

where $\mathbf{x} = \operatorname{vec}(\mathbf{X})$, $\mathbf{h} = \operatorname{vec}(\mathbf{H})$ and $\mathbf{w} = \operatorname{vec}(\mathbf{W})$. The $\operatorname{vec}(\cdot)$ operator is defined as the $N_R L \times 1$ vector formed by stacking the columns of the $N_R \times L$ matrix *i.e.* $\operatorname{vec}(\mathbf{H}) = [\mathbf{h}'_1 \mathbf{h}'_2 \dots \mathbf{h}'_L]'$. The detection problem comes from distinguishing between the two hypotheses

$$\mathcal{H}_0: x[n] = w[n] \qquad n = 0, 1, \dots, N_R L - 1
\mathcal{H}_1: x[n] = h[n]s + w[n] \qquad n = 0, 1, \dots, N_R L - 1$$
(4.3)

The sufficient statistic in this case is given by [15, 16]

$$\mathcal{T} = \mathbf{x}^H \mathbf{Q} \mathbf{x} = |s|^2 \mathbf{x}^H \mathbf{R}_{\mathbf{h}} \left[|s|^2 \mathbf{R}_{\mathbf{h}} + \sigma_n^2 \mathbf{I} \right]^{-1} \mathbf{x}, \qquad (4.4)$$

where $\mathbf{R_h} = \mathcal{E} \left\{ \mathbf{h} \mathbf{h}^H \right\}$ is the correlation matrix of the channel vector \mathbf{h} . This correlation matrix reflects both the spatial correlation between different antennas and the time-varying nature of the channel. Let $\mathbf{R_h} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$ be the eigendecomposition of the correlation matrix $\mathbf{R_h}$. In this case, the test statistic \mathcal{T} can be recast in terms of the elements of the eigenvalues λ_i of the matrix $\mathbf{\Lambda}$ and the filtered observations $\mathbf{y} = \mathbf{U}^H \mathbf{x}$:

$$\mathcal{T} = \mathbf{y}^H \mathbf{\Lambda} \left[\mathbf{\Lambda} + \sigma_n^2 \mathbf{I} \right]^{-1} \mathbf{y} = \sum_{k=1}^{N_R L} \frac{\lambda_k^2}{\lambda_k^2 + \sigma_n^2} |y_k|^2,$$
(4.5)

which is analogous to equation (5.9) in [16]. The elements y_k of the vector \mathbf{y} could be considered as filtered version of the received signal \mathbf{x} with a set of orthogonal filters \mathbf{u}_k (columns of the matrix \mathbf{U}), *i.e.*, it could be considered as multitaper analysis [94] as depicted in Figure 4.2. Linear filtering preserves the Gaussian nature of the received signals, and, the distribution of \mathcal{T} can therefore be described by the generalized χ^2 distribution¹ [28]:

$$p(x) = \sum_{k=1}^{N_R L} \alpha_k \exp(-x/2\lambda_k)$$
(4.6)

^{1.} Assuming that all eigenvalues λ_k of $\mathbf{R_h}$ are different.



Figure 4.2: Filtered Observations

and

$$\alpha_k^{-1} = 2\lambda_k \prod_{l=1, l \neq k}^{N_R L} \left(1 - \frac{\lambda_l}{\lambda_k}\right).$$
(4.7)

Theoretically, equation (4.6) could be used to set up the detection threshold γ . However, it is difficult to use for analytical derivations. Consequently, we will consider a few particular cases of the channel where the structure of the correlation matrix can be greatly simplified to reveal its effect on the detection performance.

4.3 Pilot assisted detection in SIMO configuration

4.3.1 Constant Independent Channels

In this case, the full covariance matrix $\mathbf{R}_{\mathbf{h}} = \sigma_h^2 \mathbf{O}_L \otimes \mathbf{I}_{N_R}$ is modelled as the Kronecker product of the $N_R \times N_R$ identity correlation matrix \mathbf{I}_{N_R} and $\mathbf{O}_L = \mathbf{1}\mathbf{1}^H$ is a $L \times L$ matrix consisting of ones. Accordingly, there are N_R eigenvalues λ_k , $k = 1, \dots N_R$ equal to L. The k-th orthogonal filter \mathbf{u}_k is the averaging operator applied to the data collected from the k-th antenna. Thus, the decision statistic is simply

$$\mathcal{T}_{CI} = \sum_{k=1}^{N_R} \left| \sum_{l=1}^{L} x_{kl} \right|^2 = \sum_{k=1}^{N_R} P_k, \tag{4.8}$$

where

$$P_k = \left| \sum_{l=1}^L x_{kl} \right|^2. \tag{4.9}$$

In the absence of a signal, the samples x_{kll} are drawn from an i.i.d. complex Gaussian random variable with zero mean and variance σ_n^2 . Therefore, the distribution of P_k is exponential, with mean value $L\sigma_n^2$:

$$p(P) = \frac{1}{L\sigma_n^2} \exp\left(-\frac{P}{L\sigma_n^2}\right),\tag{4.10}$$

and the distribution of \mathcal{T} is just the gamma distribution

$$p_{CI}(\mathcal{T}|\mathcal{H}_0) = \frac{1}{\Gamma(N_R)} \frac{\mathcal{T}^{N_R - 1}}{(L\sigma_n^2)^{N_R}} \exp\left(-\frac{\mathcal{T}}{L\sigma_n^2}\right).$$
(4.11)

If γ_{CI} is a detection threshold for the statistic \mathcal{T}_{CI} then the probability P_{FA} of the false alarm is

$$P_{FA} = \int_{\gamma_{CI}}^{\infty} p(\mathcal{T}_{CI} | \mathcal{H}_0) d\mathcal{T}_{CI} = \frac{\Gamma\left[N_R, \gamma_{CI} / L\sigma_n^2\right]}{\Gamma(N_R)}, \tag{4.12}$$

or

$$\gamma_{CI} = L\sigma_n^2 \Gamma^{-1} \left[N_R, P_{FA} \Gamma(N_R) \right], \qquad (4.13)$$

where $\Gamma^{-1}[N_R, \Gamma(N_R, x)] = x$ and $\Gamma(\cdot), \Gamma(a, x)$ are respectively defined as

$$\label{eq:Gamma} \begin{split} \Gamma(x) &= \int\limits_{0}^{\infty} e^{-t} t^{x-1} dt, \\ \Gamma(a,x) &= \int\limits_{a}^{\infty} e^{-t} t^{x-1} dt. \end{split}$$

If the signal is present, *i.e.*, if the hypothesis \mathcal{H}_1 is correct, then the signal y_i has a zero mean with variance $\sigma^2 = L^2 |s|^2 \sigma_h^2 + L \sigma_n^2$. As a result, the distribution of the test statistic \mathcal{T}_{CI} under the hypothesis \mathcal{H}_1 is given by the central χ^2 distribution with N_R degree of freedom and the probability of the detection is simply

$$P_D = \int_{\gamma_{CI}}^{\infty} p(\mathcal{T}_{CI} | \mathcal{H}_1) = \frac{1}{\Gamma(N_R)} \Gamma\left(N_R, \frac{\gamma_{CI}}{\sigma^2}\right)$$

$$= \frac{1}{\Gamma(N_R)} \Gamma\left(N_R, \frac{1}{1 + L\bar{\mu}} \Gamma^{-1} \left[N_R, P_{FA} \Gamma(N_R)\right]\right)$$
(4.14)

where

$$\bar{\mu} = |s|^2 \frac{\sigma_h^2}{\sigma_n^2},$$

is the average SNR per symbol. Figure 4.3 shows the performance of optimal detector in SIMO constant channel for L = 100 and $N_R = 3$. It can be seen from both eqs. (4.8) and (4.14) that under the stated channel model, the improvement in performance of P_D comes either by means of the reduction of noise through accumulation in each of the antennas (*i.e.*, increase in the effective SNR) or as a consequence of exploitation of the diversity in N_R antennas. Consequently, increasing the number of antennas leads to a faster detection.

4.3.2 Performance of the Estimator-Correlator detector in flat block fading (Constant Spatially Correlated Channel)

Let us assume that the values of the channel remain constant over L symbols but that the values of the channel coefficients for different antennas are correlated. In other words, we will assume that $\mathbf{R_h} = \sigma_h^2 \mathbf{O}_L \otimes \mathbf{R}_s$ where \mathbf{R}_s is the spatial correlation matrix between antennas. Let $\mathbf{R}_s = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^H$ be the spectral decomposition of \mathbf{R}_s .



Figure 4.3: Theoretical ROC and its simulation of the constant independent channel under AWGN. Solid lines - theory, x-lines - simulation.

Then, the test statistic \mathcal{T}_{CC} can be expressed, according to equation (4.5), as

$$\mathcal{T}_{CC} = \sum_{k=1}^{N_R} \frac{|s|^2 \sigma_h^2 \lambda_k}{|s|^2 \sigma_h^2 \lambda_k + \sigma_n^2} |y_k|^2 = \sum_{k=1}^{N_R} \frac{\bar{\mu} \lambda_k}{\bar{\mu} \lambda_k + 1} |y_k|^2, \tag{4.15}$$

where σ_h^2 is the variance of the channel per antenna. The eigenvalues λ_k of \mathbf{R}_s reflect the accumulation of SNR in each "virtual branch" of the equivalent filtered value y_k . In general, all the eigenvalues are different, and one should utilize equation (4.6). While these calculations are relatively easy to implement numerically, it gives little insight into the effect of the correlation on the performance of the detector. Under certain scattering conditions [92], the eigenvalues of the matrix \mathbf{R}_s are either all close to some constant $\lambda > 1$ or close to zero. If there are $N_{eq} < N_R$ non-zero eigenvalues, their values have to be equal to $\lambda_k = N_R/N_{eq}$ to preserve trace, and the rest $N_R - N_{eq}$ are equal to zero. In this case, the test statistic \mathcal{T}_{CC} can be rewritten as

$$\mathcal{T}_{CC}(N_{eq}) = \sum_{k=1}^{N_{eq}} |y_k|^2, \qquad (4.16)$$

where the index k corresponds to non-zero eigenvalues. Hence, the problem is equivalent to the one considered in Section 4.3.1 with N_{eq} independent antennas and the expression for the threshold γ_{CC} and the probability of detection are given by

$$\alpha \gamma_{CC} = \sigma_n^2 \Gamma^{-1} \left[N_{eq}, P_{FA} \Gamma(N_{eq}) \right], \qquad (4.17)$$

where $0 < \alpha < 1$ performs as a corrector variable. In Figure 4.4 the ROC of the estimator correlator using this approximation are shown. The effect of the correlation between branches has a dual effect on the performance of the system. The number N_{eq} of equivalent independent branches is reduced, compared to the number of antennas N_R , therefore reducing diversity. However, an increase in correlation results in an additional accumulation of SNR (or, equivalently, an additional noise reduction through averaging) by a factor of $N_R/N_{eq} \geq 1$. Therefore,

$$P_D = \int_{\gamma_{CC}}^{\infty} p(\mathcal{T}_{CC} | \mathcal{H}_1) d\mathcal{T}_{CC} = \frac{1}{\Gamma(N_{eq})} \Gamma\left(N_{eq}, \frac{\alpha L \gamma_{CC}}{\sigma^2}\right)$$

$$= \frac{1}{\Gamma(N_{eq})} \Gamma\left(N_{eq}, \frac{1}{1 + LN_R \bar{\mu}/N_{eq}} \Gamma^{-1}\left[N_{eq}, P_{FA} \Gamma(N_{eq})\right]\right).$$
(4.18)

In general, the number N_{eq} does not need to be integer in calculations of the threshold and detection probability P_D . This allows us to account for some eigenvalues of intermediate values (further discussed in Section 4.4).

4.3.3 Independent SIMO Channels with Temporal Correlation

In the case of independent antennas but temporally correlated fading, the full correlation matrix can be represented as $\mathbf{R}_h = \mathbf{R}_T \otimes \mathbf{I}_L$ where $\mathbf{R}_T = \mathbf{U}_T^H \mathbf{\Lambda}_T \mathbf{U}$ is the



Figure 4.4: ROC approximation vs. simulation results ($\alpha = 0.8$). Solid lines - theory, x-lines - simulation.

temporal correlation matrix of an individual channel and its eigendecomposition. The decision statistic can now be be represented as

$$\mathcal{T}_{ICC} = \sum_{k=1}^{N_R} \mathbf{x}_k^H \mathbf{R}_T \left(\mathbf{R}_T + \frac{1}{\bar{\mu}} \mathbf{I}_L \right)^{-1} \mathbf{x}_k = \sum_{k=1}^{N_R} \mathcal{T}_{ICC_k}, \quad (4.19)$$

where \mathbf{x}_k is the $1 \times L$ time sample vector received by the k-th antenna. Therefore, each antenna signal is processed separately and the results are added afterwards. Taking advantage of the eigendecomposition of the correlation matrix \mathbf{R}_T , the calculation of the decision statistic \mathcal{T}_k can be recast as a multitaper analysis

$$\mathcal{T}_{ICC_k} = \mathbf{y}_k \mathbf{\Lambda}_k \left(\mathbf{\Lambda}_k + \frac{1}{\bar{\mu}} \mathbf{I}_L \right)^{-1} \mathbf{y}_k = \sum_{l=1}^L \frac{\lambda_l}{\lambda_l + 1/\bar{\mu}} |y_{kl}|^2.$$
(4.20)

We can utilize an approximation of the correlation matrix by one with constant or zero eigenvalues as in Section 4.3.2. In this case, there will be

$$L_{eq} = \frac{(\operatorname{tr} \mathbf{R}_T)^2}{\operatorname{tr} \mathbf{R}_T \mathbf{R}_T^H}$$
(4.21)

eigenvalues of size L/L_{eq} and the rest are zeros. Therefore, there are $N_R L_{eq}$ terms in the sum (4.19) each one contributing with

$$\frac{L/L_{eq}}{L/L_{eq} + 1/\bar{\mu}} = \frac{\bar{\mu}L + L_{eq}}{\bar{\mu}L}$$
(4.22)

into the variance of \mathcal{T}_{ICC} . The corresponding equations for choosing the threshold become

$$\gamma_{CC} = L\sigma_n^2 \Gamma^{-1} \left[N_R L_{eq}, P_{FA} \Gamma(N_R L_{eq}) \right], \qquad (4.23)$$

$$P_D = \int_{\gamma_{ICC}}^{\infty} p(\mathcal{T}_{ICC} | \mathcal{H}_1) d\mathcal{T}_{ICC} = \frac{1}{\Gamma(L_{eq})} \Gamma\left(L_{eq}, \frac{\gamma_{ICC}}{\sigma^2}\right)$$

$$= \frac{1}{\Gamma(L_{eq})} \Gamma\left(L_{eq}, \frac{1}{1 + LN_R\bar{\mu}/L_{eq}} \Gamma^{-1}\left[N_{eq}, P_{FA}\Gamma(L_{eq})\right]\right).$$
(4.24)

4.3.4 SIMO Channel with Separable Spatial and Temporal Correlation

The correlation matrix of the channel with separable temporal and spatial correlation has a correlation matrix of the form $\mathbf{R_h} = \mathbf{R}_T \otimes \mathbf{R}_s$. The correlation in both coordinates reduces the total number of degrees of freedom from $N_R L$ to $N_{eq} L_{eq} \leq N_R L$. The loss of degrees of freedom is offset by an accumulation of SNR due to the averaging over the correlated samples. The equivalent increase in the average SNR is $N_R L/N_{eq} L_{eq}$. Thus, the problem is equivalent to the detection using

$$K_{eq} = N_{eq} L_{eq} = \frac{(\text{tr}\mathbf{R}_s)^2}{||\mathbf{R}_s||_F^2} \frac{(\text{tr}\mathbf{R}_T)^2}{||\mathbf{R}_T||_F^2}$$
(4.25)

independent samples in the noise with an average SNR

$$\bar{\mu}_{eq} = \frac{N_R L}{N_{eq} L_{eq}} \bar{\mu},\tag{4.26}$$

where $|| \cdot ||_F$ stands for the Frobenius norm defined as

$$||A||_F = \sqrt{\operatorname{tr}(AA^*)}.$$

The sufficient test statistics in the case of a SSTCC channel can be easily obtained from the general eqs. (4.4) and (4.5). By using the Kronecker structure of $\mathbf{R_h}$, one obtains

$$\mathcal{T}_{SSC} = \sum_{k=1}^{K_{eq}} |y_k|^2.$$
(4.27)

4.4 Examples and Simulation

4.4.1 Correlation models

While the Jakes correlation function $J_0(2\pi f_D \tau)$ is almost universally used in standards on wireless channels, the realistic environment is much more complicated. A few other models can be found in the literature, with some chosen for their simplicity, and others based on experimental measurements. In most cases, we are able to calculate N_{eq} analytically, as shown below.

1. Sinc type correlation If a scattering environment is formed by a single remote cluster (as shown in [92]), the spatial covariance function $\mathbf{R}_s(d)$ as a function of electric distance between antennas d is given by

$$\mathbf{R}_{s}(d) = \exp\left(j2\pi d\,\sin\phi_{0}\right)\operatorname{sinc}\left(\Delta\phi d\,\cos\phi_{0}\right),\tag{4.28}$$

where ϕ_0 is the central angle of arrival, and $\Delta \phi$ is the angular spread. Moreover, this correlation matrix has approximately $\lfloor 2\Delta \phi \cos \phi_0 N + 1 \rfloor$ eigenvalues which are nearly equal; the rest will be close to zero [95]. 2. Nearest neighbour correlation By neglecting the correlation between any two non-neighbouring antennas, one obtains the following form of the correlation matrix \mathbf{R}_s

$$\mathbf{R}_{s} = \{r_{ij}\} = \begin{cases} 1 & \text{if } i = j \\ \rho & \text{if } i = j + 1 \\ \rho^{*} & \text{if } i = j - 1 \\ 0 & \text{if } |i - j| > 1 \end{cases}$$
(4.29)

where ρ is the correlation coefficient. The eigenvalues of eq. (4.29) are well know as [96]

$$\lambda_k = 1 - 2|\rho|\cos\frac{k\pi}{N+1}, \quad 1 \le k \le N.$$
 (4.30)

The square of Frobenius norm of eq. (4.29) $||\mathbf{R}_s||$ is given by

$$||\mathbf{R}_s||^2 = N + 2(N-1)|\rho|^2.$$
(4.31)

The equivalent number of independent virtual antennas is given by

$$N_{eq} = \frac{N^2}{N + 2(N-1)|\rho|^2} = \frac{N}{1 + 2|\rho|^2 (1 - 1/N)}.$$
(4.32)

3. Exponential spatial correlation For this case the correlation matrix \mathbf{R}_s is given by

$$\mathbf{R}_s = \left\{ r_{ij} \right\} = \left\{ \left| \rho \right|^{i-j} \right\}.$$
(4.33)

The eigenvalues of this matrix can be obtained as follows [96]

$$\lambda_k = \frac{1 - |\rho|^2}{1 + 2|\rho|\cos\psi_k + |\rho|^2},\tag{4.34}$$

where ψ_k are roots of the following equation

$$\frac{\sin(N+1)\psi - 2|\rho|\psi\sin N + |\rho|^2\sin(N-1)\psi}{\sin\psi} = 0.$$
 (4.35)

If $|\rho| \ll 1$, the eigenvalues of eq. (4.33) can be substituted by those of eq. (4.29), *i.e.* the equation (4.30) can be directly used. However, if $|\rho| \approx 1$ then

$$\lambda_1 = 1 + (N-1)|\rho|, \ \lambda_k = 1 - |\rho|, \ 2 \le k \le N.$$
(4.36)

4. Temporal correlation model for nonisotropic scattering Considering the extended case of the Clarke's temporal correlation model for the case of nonisotropic scattering around the user, the temporal correlation function is [91]:

$$\mathbf{R}_{s}(\tau) = \frac{I_0 \left(\sqrt{\kappa^2 - 4\pi^2 f_D^2 \tau^2 + j4\pi\kappa \cos(\mu) f_D \tau}\right)}{I_0(\kappa)}, \qquad (4.37)$$

where $\kappa \geq 0$ controls the width of angle of arrival (AoA), f_d is the Doppler shift, $\mu \in [-\pi, \pi)$ is the mean direction of AoA seen by the user and $I_0(\cdot)$ stands for the zeroth-order modified Bessel function. Figure 4.5a, shows the general scenario of non isotropic scattering ($\kappa = 5$) which corresponds to directional signal reception- the user receives the signal only from a particular direction through a narrow beamwidth. In the same figure we can also see the special case of isotropic scattering ($\kappa = 0$) where the user receives signals from all directions with equal probability. In Figure 4.5b, the temporal correlation function (eq.(4.37)) for the isotropic and non isotropic cases is shown. Notice that for non-isotropic cases, eq.(4.37) reduces to the classic Clarke's temporal correlation model $J_0(2\pi f_D \tau)$.

Figure 4.6 shows the eigenvalues behavior for different values of the κ factor. Notice that for $\kappa = 0$ (isotropic scattering), the values of the eigenvalues are spread in an almost equally and proportional fashion manner. As κ tends to infinity (extremely nonisotropic scattering), we obtain N - 1 zero eigenvalues and one eigenvalue with value N. In other words, as κ increases, the number of "significant" eigenvalues decreases and therefore so does the value of N_{eq} as shown in Figure 4.8.





(a) Examples of Isotropic and Nonisotropic scattering scenarios



Figure 4.5: Comparison of Isotropic vs. Nonisotropic scattering.



Figure 4.6: Eigenvalues behavior of R_s temporal correlation matrix for nonisotropic scattering $(N=10,\mu=0 ~{\rm and}~ f_d=50 Hz)$



Figure 4.7: Effect of correlation between antennas in the probability of detection.



Figure 4.8: ROC approximation for the estimator correlator considering the temporal correlation for isotropic and nonisotropic scattering ($\kappa = 0$ and $\kappa = 10$ respectively).

4.5 Space-Time Processing Trade-Off

It is common to assume that increasing the number of antennas improves performance of detection algorithms due to the increased degree of diversity. This assumption is correct when the number of time samples remains the same. However, in cognitive networks it is desirable to reduce decision time as much as possible, sometimes by introducing some added complexity in the form of additional number of antennas [28]. The goal of this section is to show the existing tradeoff between how fast it is possible to make a decision about the presence of the PU and the number of antennas needed at the receiver side. Equation (4.16) shows that the processing of the signal consists of two separate procedures: averaging in time along with accounting for diversity and suppressing noise in spatial diversity branches. Depending on the amount of noise (SNR) and fading, one of these two techniques provides a greater benefit to the net result. For relatively low levels of SNR, noise suppression is the dominant task, and it is therefore more advantageous to have a single antenna and as many samples of time as possible. However, if the SNR is somewhat higher, the noise is sufficiently suppressed even by a short time, and averaging and suppressing fading through diversity combining is more beneficial. Figure 4.7 shows the performance for different configurations of the receiver in such a way that the product $N_R L$ remains constant. The figure also shows the effect that correlation, and thus the scattering environment, plays on quality of reception. For very strong correlations $\rho \approx 1$ and $N_{eq} \approx 1$. Consequently, all the collected samples are used to reduce the noise. Such a scheme performs the best at low SNRs. However, when $\rho = 0$ and $N_{eq} = N_R$, the gain from diversity is higher and the scheme outperforms for higher SNR. The intermediate case allows for a smooth transition between these two regions. Only in the case of $\rho = 1$, is there an equivalent trade-off between the number of antennas and the samples of time, *i.e.*, the performance depends only on $Q = N_R L$ and not on the individual values of N_R and L. Nevertheless, lower correlation results in an unequal trade-off with gain or loss defined by the SNR and the amount of correlation.

4.6 Conclusion

In this chapter we assessed the impact that scattering environment has on the performance of a PU detection. We obtained approximate expressions for the probability of missed detection as function of the number of antennas, scattering parameters and number of observations. We have also shown that for low SNR scenarios, it is preferred to have only a single antenna and many time samples to emphasize noise suppression. On the contrary, for high SNR cases, noise is suppressed relatively quickly, and it is therefore better to have more antennas in order to mitigate the fading. We also showed that for very strong correlations, with $\rho \approx 1$ and $N_{eq} \approx 1$ - this scheme can be usefully applied to low SNR situations if enough time samples can be obtained. As $\rho \rightarrow 0$, the diversity gain is increased, making this approach more suitable for high SNR situations.

Chapter 5 Multiple Access Games for Cognitive Radio Networks

In this chapter, we introduce a very useful mathematical tool for wireless engineering named game theory. In recent years, interest in understanding and applying game theory concepts to wireless communications and networking problems has increased. Game theory analyses the dynamics of interactive decision-making between rational individuals who compete for common interests. It also provide us with tools to predict what might (and possibly what should) happen when agents with conflicting desires interact- in other words, it can recognize stable outcomes. As seen in Chapter 1, Secondary Users(SUs) in cognitive radio networks need to be aware of the changes in the dynamic spectrum environment and make decisions accordingly, such as adjusting their operating parameters. SUs have the ability to observe, learn and act to optimize their total performance, unlike conventional spectrum sharing where it is generally assumed that all users cooperate in a static environment. We can observe that an inherent characteristic of a SU is to compete for a licensed band with the ultimate goal of maximizing benefits. Within this context, game theory arises as a natural solution to the Cognitive Radio paradigm [97]. In this chapter some game theory concepts are presented along with their application in order to address wireless communications problems and (more specifically) cognitive radio problems. Later on, we present a new multiple access algorithm for cognitive radio networks based on the game theory. We address the problem of a multiple access system where the number of users and their types are random. In order to do this, the framework is modelled as a noncooperative Poisson game in which all players are unaware of the total number of devices participating (population uncertainty). We propose a scheme where failed attempts to transmit (collisions) are penalized and calculate the optimum penalization in mixed strategies. The proposed scheme converges to a Nash equilibrium where the maximum possible throughput is achieved.

5.1 Noncooperative games and Nash equilibrium

The three main components of any game are:

A set of Players

The players are the decision makers in the modelled scenario. Examples of these could be people, countries, a group of companies or, some biological species. In a wireless system, the players are most often the nodes of the network- secondary or PU nodes in CR networks. We denote this set by N.

A set of Strategies

Once the players have been defined, they will each have a set of possible strategies; these are the actions they may choose to follow. The action or actions taken by each player will determine the outcome of the overall game. This set is denoted by A_i , for each player $i \in N$.

A payoff or utility function

For every outcome product of an action taken by each player, there is some associated numerical utility or payoff. These payoffs represent the value of the outcome to the different players. This set will be denoted by $u_i : A \to \mathbb{R}$ which measures the outcome for player *i* determined by the strategies of the rest of the players, $A = A_i \times A_j \ \forall i \neq j \in N$.

In Table 5.1 two examples of the application of these components are shown in the context of CR networks. Generally games may be divided into noncooperative games and cooperative games. Cooperative games will be briefly explained at the end of the chapter although, they are out of the scope of this thesis. A Noncooperative game is a game modelled under the basis that all players make choices or play strategies considering *only* their own selfish interests- their final objective is to maximize their own total utility. Noncooperative games can be categorized accordingly to the players' moves and to the information availability. In a *static game*, players make their

	Open spectrum sharing		
Players	SUs who will compete for some unlicensed electromagnetic		
	spectrum.		
Strategies	General transmission parameters such as probability of		
	transmission, waveforms, power level of transmission, access		
	rates, time of transmission, receiver nodes, relay choice, <i>etc</i> .		
Utility	Non-decreasing function reflecting the QoS obtained by using		
	the unlicensed spectrum.		
	Licensed spectrum sharing		
Players	SUs who will compete for some unlicensed electromagnetic		
	spectrum and PUs willing to lease some licensed band to SUs.		
Strategies	PU: Choice of SUs allowed to transmit within leased band,		
	amount of leased band, price per Hert. SUs: Choice of band		
	to rent and price to pay for renting such spectrum band.		
Utility	PU: Revenue minus the cost of leasing the licensed spectrum.		
	SUs: Non-decreasing function reflecting the QoS obtained by		
	using the leased spectrum.		

Table 5.1: Example of Game Theory components in CR networks.

decisions simultaneously with no information about the decisions taken by other players¹. The most common representation of static games is in a table called *strategic* form or normal form of the game. On the other hand a dynamic game occurs when there is a strict order of turns that the players must obey, and players must strategize accordingly. In this type of game, players know what other players have done before having the opportunity to make a move. These kind of games are more easily depicted in the way of game trees, or most common, in the literature as the *exten*sive form of a game. This illustrates all possible actions that can be taken by all players and indicates the possible outcomes from the game. Games can also be classified, depending on the amount of information, into two types: complete information and incomplete information games. In complete information games all players are aware of the number of players, the strategies and the utility function of the rest,

^{1.} A game is also simultaneous when players choose their actions in isolation, with no information about what other players have done or will do, even if the choices are made at different points in time.

	Static game	Dynamic game	
Complete Informa-	Complete information	Complete information dy-	
tion game	static game.	namic game.	
	Nash equilibrium (John	Subgame perfect Nash equi-	
	Nash $[98, 99]$).	librium. (Reinhard Selten	
		[100]).	
Incomplete Infor-	Incomplete information	Incomplete information dy-	
mation game	static game.	namic game.	
	Bayesian Nash equilibrium	Perfect Bayesian Nash equi-	
	(John Harsanyi [101]).	librium. (Reinhard Selten	
	Poisson Games (Roger	[103]).	
	Myerson $[102]$).		

Table 5.2: Categories of noncooperative games and corresponding equilibria.

whereas incomplete information games, one of more of these components has to be estimated or assumed. Table 5.2 shows the four types of noncooperative games and their corresponding equilibrium concepts, along with the associated researchers.

5.1.1 Nash Equilibrium

Once the noncooperative game is defined the logical question arises: given a game with two or more players in conflict, what will be the most likely outcome of the game? Arguably, one could assume that in general all players would try to play the best strategy which leads to the best utility, considering what other players would play. In game theory context, this type of logic is called a Nash Equilibrium. A Nash equilibrium is a set of strategies, one for each player, with the property that no player can unilaterally change his/her stately and get a better payoff. This is the central concept and focal point of noncooperative game theory. That being said, the next questions that pop out are: Does a Nash equilibrium always exist?, and once defined is this unique? The answer to the first question is generally yes. The question about uniqueness is trickier and needs to be analyzed in a case by case basis, but indeed, more than one Nash equilibrium can exist in a game. When more than one equilibrium exists, it is important for the players to try to converge to the best one. Some equilibria selection criteria is therefore necessary. Although the existence

of a Nash equilibrium is guaranteed in noncooperative games, it does not provide us with any clues on how to converge to it. Furthermore, if players start from a strategy profile that is a Nash equilibrium, there is no reason to believe that any of the players will deviate, and the system will be in equilibrium provided no conditions (set of players, payoffs, *etc.*) change. However, what happens if the players start from a non equilibrium strategy profile? In CR networks, for instance, players can start from an arbitrary strategy, update their strategies following some rule and hope for the best towards the equilibrium convergence. Another important concept close related to the Nash equilibrium is the Pareto optimality, defined as such strategy profile in which no player can improve his utility without making any other player's utility worse. The Pareto optimality is very useful to compare multi-dimension payoff profiles so that equilibrium has some shortcomings, a significant one being fact that it is almost impossible to justify why players in a real game would necessarily play such an equilibrium. Let us define formally the Nash equilibrium as

Definition 1. A Nash equilibrium of any strategic game $\langle N, A_i, u_i \rangle$ is a strategy profile $a^* \in A$ of actions such that for every player $i \in N$ the following relation stands

$$u_i\left(a_i^*, a_{-i}^*\right) \ge u_i\left(a_i, a_{-i}^*\right) \qquad \forall i \in N \tag{5.1}$$

where a_i denotes the strategy of player *i* and a_{-i} denotes the strategies of all players other than player *i*.

If all players choose a strategy according to a Nash equilibrium criteria, no player can improve his payoff by unilaterally deviating from such equilibrium. From this we can define the best response function for the *i*-th player $a_i^* \in B_i(a_{-i}^*)$ as

$$B_{i}(a_{-i}) = \left\{ a_{i} \in A_{i} : u_{i}(a_{-i}, a_{i}) \ge u_{i}(a_{-i}, a_{i}') \right\} \quad \forall a_{i}' \in A_{i}.$$
(5.2)

5.1.2 Mixed Strategies

In the previous discussion, we assumed that each player chooses a single strategy in their strategy set and continues with it in every game. These types of strategies are

	Paper	Rock	Scissors
Paper	(0,0)	(1,-1)	(-1,1)
Rock	(-1,1)	(0,0)	(1,-1)
Scissors	(1,-1)	(-1,1)	(0,0)

Figure 5.1: Strategic form Representation of the Paper-Scissors-Rock Game

called *pure strategies*. Nevertheless, the situation in which players need to randomize their choices of strategy arises quite often in game theory scenarios. Figure 5.1 depictes the strategic form of the classic Paper-Scissors-Rock game. Following tradition, a rock will break scissors, scissors cut paper, and paper wraps rock. One can notice that there are no Nash equilibria in pure strategies by simply analyzing there is not such a strategy which satisfies eq. (5.4). For example, if one player decides to always play rock, the other player will decide to always play paper. The same situation would occur with the rest of the strategies. Hence, the question arises: which strategy should a player choose in order to be able to win at least a some of the time? The more logical and straightforward strategy to follow (tested in playgrounds all over the world) is to randomize the election among the three pure strategies, assigning a probability of $\frac{1}{3}$ to each. When players randomize over their strategy set, is called a mixed strategy. Thus, the mixed strategy available to player i can be denoted as σ_i , and the probability that σ_i assigns to the strategy a_i is denoted by $\sigma_i(a_i)$. Following the previous notation, the expected utility of player i under the joint mixed strategy σ is given by

$$u_i(\sigma) = \sum_{a \in A} \left(\prod_{j=1}^N \sigma_j(a_j) \right) u_i(a).$$
(5.3)

The space of player *i*'s mixed strategies is Σ_i , and a mixed strategy profile $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_N)$ and the mixed strategy space Σ can therefore be formed by the

Cartesian product as $\Sigma = \Sigma_i \times \Sigma_j \quad \forall i \neq j \in N$. The Nash equilibrium defined for strategic games, where players take pure strategies, can then be naturally extended, and a mixed strategy Nash equilibrium of a strategic game is a Nash equilibrium where players in the game adopt mixed strategies. We can then generalize the concept of Nash equilibrium in order to incorporate mixed strategies as follows:

Definition 2. A mixed strategy profile $\sigma \in \Sigma$ is a Nash equilibrium if

$$u_i(\sigma) \ge u_i(a_i, \sigma_i) \qquad \forall i \in N, \forall a_i \in A_i.$$
 (5.4)

5.1.3 Existence of Nash Equilibria

The existence of a Nash equilibrium is an important aspect to analyze in game modelling. The following theorem based on the fixed point establishes the existence of a Nash equilibrium [104].

Theorem 1 (Existence). Given a game in strategic form with $\langle N, A_i, u_i \rangle$, a Nash equilibrium will exist if the action set A_i of player *i* consists in a non-empty compact convex subset of a Euclidian space, and the payoff function u_i is continuous and quasi-concave on A_i for all *i*.

This results in the subsequent theorem as one of the most important keystones in game theory.

Theorem 2 (Nash [98]). Every finite strategic game has a Nash equilibrium in either mixed or pure strategies.

The case of pure strategies is just a particular case of the mixed strategies case when the assigned probability to each action is equal to one.

5.1.4 Equilibrium selection

More than one Nash equilibrium may exist in a game which produces different utilities to the players. Therefore, it is natural to question whether one equilibrium outperforms others. It is important to know if there are better Nash equilibriums in the



Figure 5.2: The Prisoners' Dilemma

same game or even if an optimal equilibrium exists. This is, most of the time, a non trivial problem since game theory solves multi-objective optimization situations, and it is not easy to define such an optimality. For example, when players have conflicting interests, an increase in one player's payoff might decrease others' payoffs. The most popular technique used in game theory models is called the *Pareto optimality*, which is a payoff profile in which no strategy can make at least one player better off without making any other player worse off.

5.1.4.1 Pareto optimality

The Pareto efficiency, or Pareto optimality, has been widely used in game theory, as well as economics, engineering and social sciences. If more than one equilibrium exists, usually the optimal ones in the Pareto sense are preferred. The Pareto optimality is defined as follows:

Definition 3. Let $U \subseteq \mathbb{R}^N$ be a set. Then $\mathbf{u} \in U$ is **Pareto efficient** if there is no \mathbf{u}' for which $u'_i \geq u_i$ for all $i \in N$ and $u'_i > u_i$ for some $i \in N$. Subsequently, the Pareto frontier is defined as the set of all $\mathbf{u} \in U$ that are Pareto efficient.

These concepts can be better explained in the context of a two application problems: *The Prisoners' Dilemma* and *The DCF Game*.

5.1.5 The Prisoners' Dilemma

In the Prisoners' Dilemma game, two prisoners, Homer and Bart, are arrested outside of a store. They are alleged to have to robbed the store, but the police can only prove that the suspects were trespassing. Thus, the police need one of the criminals to "rat out" the other in order to charge him for the greater crime of attempted robbery. To get them to confess, they will both be offered a deal simultaneously while they are being interrogated in separated rooms. The deal looks like this: if no one confesses to the robbery, the police can only charge the prisoners for trespassing and the punishment for that is one month in jail each. If one confesses and the other does not, the police will be lenient on the "rat" by letting him free, and will severely punish the other with 15 months in jail. Finally, if both confess, the police will punish them equally for attempted robbery which is more severe than simply trespassing. The punishment here is 5 months in jail each. The Prisoners' Dilemma is a complete information static noncooperative game between two players (Homer and Bart). Each player must decide between two strategies: Confess or Not confess. In Figure 5.2 the game is shown in a strategic form (payoff matrix). Bart will choose a row, and simultaneously Homer will choose one of the columns. The strategy combination {Confess, Confess} produces a payoff of 5 for each player whereas the combination {Not Confess, Not Confess} produces a payoff of 1. The combination {Confess, Not Confess} results in a payoff of 0 for Bart and a payoff of 15 for Homer, and finally the combination {Not Confess, Confess} ends in a payoff of 15 for Bart and 0 for Homer. It is normally assumed in game theory models that each player will choose in a rational manner- he will tend to choose the strategy that maximizes (or minimizes) his payoff according to what he prefers the most. At the same time, he will assume that his opponent will follow the same reasoning. Notice that in this example, a smaller payoff is preferred since it translates to a shorter time in prison. The strategy *Confess* is therefore the best strategy for a player regardless of the strategy chosen by his opponent. Moreover, both players are convinced that their counterpart will choose *Confess*. The strategy profile {*Confess*, *Confess*} is composed of the best strategy that each player chooses and, it therefore represents a Nash

equilibrium. This is the result that the police expects (both players being imprisoned for five months). Interestingly, from the players point of view, the Nash equilibrium outcome ({Confess, Confess}) is greatly inferior to the outcome from the strategy profile {Not Confess, Not Confess}, which results in 1 month of jail for each. We can conclude that the Nash equilibrium of this game is not Pareto optimal. The Prisoners' Dilemma has three Pareto optimal outcomes (where it is impossible to improve the payoff of any player without negatively affecting the payoff of at least other player): {Confess, Not Confess}, {Not Confess, Confess} and {Not Confess, Not Confess}. It is an interesting phenomena that in real life scenarios, individual rationality is often incompatible with collective rationality in noncooperative games. If the game is played not just once, but repeated infinitely, both players might cooperate by choosing the strategy Not Confess in early plays in hopes of arriving at the Pareto optimality ({Not Confess, Not Confess}). The players have also the option of playing what is called a grim strategy which consists in a Nash equilibrium strategy profile of the infinitely repeated game in order to ensure cooperation. In this strategy, each player will always choose the cooperative strategy Not confess in each sub game (a single Prisoners' Dilemma game) until his counterpart chooses the strategy Confess, and then he will always choose the strategy *Confess* in the following games to punish the traitor. Some form of coalition and cooperation is forced since each player is afraid of the punishment by the rest of players. Hence, individual rationality becomes consistent with collective rationality.

5.1.6 Application of the Prisoners' Dilemma: The DCF Game

In the IEEE 802.11 standard [105], the Distributed Coordination Function (DCF) is used as the medium access protocol in almost all of the testbeds and simulations for wireless ad hoc network research [106]. The DCF provides two access schemes: the basic scheme and the request to send/clear to send (RTS/CTS) scheme. In the basic scheme, transmitting and receiving nodes only exchange data frames and acknowledgement (ACK) frames. The RTS/CTS scheme adds a RTS/CTS dialog

before the data frame in order to reduce the probability of collisions on the channel, because the collision probability of a RTS frame (20 octets) is less than that of a data frame (up to 2346 octets). The RTS/CTS scheme works as follows: every time a node wants to transmit a data frame, it will first transmit an RTS frame in order to reserve the channel. The receiving node replies with a CTS frame if it is ready to receive. If the transmitting node successfully receives the CTS frame, it starts transmitting the data frame. After receiving the data frame, the destination node replies with an ACK frame to the source. If the CTS frame was not successfully received by the receiving node, the source times out from waiting. It will then perform the Binary Exponential Backoff (BEB) algorithm in order to calculate a new random back off time with a larger window opportunity in order to transmit the RTS frame and hence decrease the probability of collision. For every RTS retransmission, the back off time is chosen uniformly within the range (0, CW - 1), where CW is the size of the contention window which is function of the number of previous failed transmissions of the RTS frame. In the first retransmission attempt, CW is equal to the minimum contention window CW_{\min} . Every time an unsuccessful transmission occurs, the value of CWis doubled up to the maximum value CW_{max} . If this value is achieved, no further doubling is performed and the value stays constant. If reception is unsuccessful after seven tries, the RTS frame is dropped. In the IEEE 802.11 DCF standard there are no centralized infrastructures, and all nodes therefore transmit their data frames in a competitive manner. One node must compete with its neighbouring nodes in such a way that it can transmit as many packets as possible. In addition, the BEB algorithm presents a fairness problem among TCP flows in multi-hop ad hoc networks, as it always favours the latest successful nodes. In this case game theory can be used as a model to solve the unfairness problem. In order to do this we model the IEEE 802.11 DCF as a game. In the DCF game, each player (node) has two strategies: Transmit or Not transmit analogous to Confess and Not confess in the Prisoners' Dilemma. The game is presented in Figure 5.3 for the particular case of two players representing two competing nodes in the system. Figure, u_s represents the payoff when the respective node performs a successful transmission, u_i is the payoff when a node decides to stay idle for the particular slot and u_f is the payoff in case of a failed



Figure 5.3: The DCF game with two nodes

transmission. Chapter 5 analyzes in more detail the respective values of these payoffs; however, at this point we can see a self-evident relation among them as follows:

$$u_f < u_i < u_s. \tag{5.5}$$

This is obviously a noncooperative game with complete information in which the players aim to obtain higher payoffs. It can be seen that this game has two Nash equilibriums in pure strategies: {*Transmit*, *Not transmit*} and {*Not transmit*, *Transmit*}. The DCF realizes the two equilibrium strategies by first listening to the busy/idle state of the medium when a node wants to transmit packets. If the channel is idle for a period of time equal to a distributed inter frame space (DIFS), the node transmits. Otherwise, the node does not transmit and persists in monitoring the channel until the medium is determined to be idle without interruption for a DIFS. Moreover, the DCF game has another Nash equilibrium in mixed strategies, in which each node chooses the strategy *Transmit* with probability

$$\frac{u_s - u_i}{u_s - u_f},$$

and chooses the strategy Not transmit with probability

$$\frac{u_i - u_f}{u_s - u_f}$$
The DCF analyzes the mixed strategy as follows. When the channel is busy, the node persists in listening to the channel until it becomes idle for a DIFS; the node then waits a random back off interval. The random back off integral can be modelled by the mixed strategy. By analyzing the possible values of u_s , u_i and u_f further, it is possible to see that:

- u_i indicates the delay sensitivity of the traffic being transmitted. The smaller the value of u_i , the more delay-sensitive the traffic.
- u_s should be the increasing function of the length of the data frame. The longer the data packet transmitted successfully, the higher the channel utility ratio.
- u_f should be the decreasing function of the length of the data frame. A transmission failure of a long data frame does more harm to the network than that of a short frame, since a wireless node cannot sense the channel while it is transmitting.

The DCF does not consider how the priorities of different traffic affect the performance of a network, nor does it consider how the lengths of different data frames affect the performance of a network. However, it is possible to construct different DCF game models for traffic with different priorities and different lengths by adjusting the values of u_i, u_s , and u_f accordingly, giving different Nash equilibriums in mixed strategy and thus different random waiting intervals so that we can improve the performance of the DCF (*e.g.*, the fairness). In addition, if each node contends for the channel repeatedly and the network has multiple nodes, a very complex method to determine the values of u_i, u_s , and u_f is needed.

5.2 Random Access Games

As presented in previous section, because of the inherently competitive nature of CR networks, the game theory arises as a straightforward approach to deal with several application problems [97, 107–113]. The Secondary Users (SU) in Cognitive Radio (CR) networks can be considered players competing for some specific license band, resulting in different kinds of payoffs for them. Specifically, the Medium Access Control (MAC) problem has been analyzed using game theory tools in the literature

[114–123]. In [117] a game theory model is presented as a starting point for the Distributed Coordination Function (DCF) mechanism in IEEE 802.11. As seen in Chapter 5, in the DCF there are no base stations or access points which control access to the channel, and all nodes therefore transmit their data frames in a competitive manner. However, as noted in [117], the proposed DCF game does not consider how the different types of traffic can affect the sum throughput of the system. In [124]it is assumed that the total number of players is known in order to evaluate the performance of the DCF, whereas in [116] a game-theoretic model of multipacket slotted ALOHA with perfect information is studied. The authors show in [116], the Nash equilibrium must exist in this model, and its stability region is characterized. Furthermore, a pricing strategy based on slotted ALOHA with multipacket reception is proposed in [125] in order to enforce fairness among the players. In [126] the author calculates an optimal access probability based on slotted ALOHA, which maximizes the successful delivery probability in CR networks. The author also assumes that the number of transmitters and receivers is always known during the analysis and all users share a common access probability (*i.e.*, all users are treated equally.) In [127] a distributed MAC algorithm with one-slot memory is proposed in order to coordinate the access among the SUs and restrict interference to the PU. An optimal probability of attempting to access the channel for the SUs in order to maximize the throughput is obtained. A p-persistent protocol to control the selection of the contention window in the IEEE 802.11 backoff algorithm is described in [106]. The authors in [106] showed how to maximize the throughput of the scheme. Nevertheless, the authors in [127]and [106] do not consider either the possibility of having different types of users, nor the randomness in the number of SUs in the system. In contrast, this chapter addresses the problem similar to that approached by the Enhanced DCF included in IEEE 802.11e [128], which is a natural extension of the DCF mechanism. Specifically, we provide a novel interpretation and analysis in order to solve the problem of multiple access for a heterogenous and random population of SUs, based on Myerson's results for Poisson games [129]. This branch of game theory analysis addresses the problem of games with an uncertain population- when the number of $players^2$ is unknown and can be modelled as a random variable.

In this chapter, Section 5.3 summarizes the theoretical basis of Poisson Games, and two multiple access examples are given. Section 5.4 presents our novel Poisson game model, and we calculate the optimal mixed strategies and the optimal penalizations used in the game. Section 5.4.2 extends the aforementioned analysis to the case of two types of SUs and provides an accurate analytical approximation of the Pareto frontier. In Section 5.5, the impact on the PU based on its activity is considered and the optimal mixed strategies are calculated accordingly. Finally, some conclusions are drawn in Section 5.6.

5.3 Poisson Games

It is very well established in [129] that a Poisson Game is a special case of a more general type of games called Random Player Games. In games with population uncertainty [102,130], there is a nonempty finite set of players types \mathcal{T} which is known *a priori*. In the context of communications systems, this set could contain the different types of services offered by the network (voice, video, data, *etc.*). There is also a finite set of available choices or pure actions \mathcal{C} that a player may take. For instance, these could be all the different transmission powers that the SU may utilize [131] or, in the context of this chapter, the decision to transmit or not. The set of possible actions is the same regardless of the type of player. The main characteristic of a Poisson Game is that the total numbers of players of certain types, are modelled as random variables. We use the definition given by Myerson [130] which presents a Poisson Game Γ as the five-tuple ($\lambda, \mathcal{T}, r, \mathcal{C}, u$). Here, the parameter λ corresponds to the average number of users described by a Poisson random variable with probability mass function defined as

$$f(k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$
(5.6)

^{2.} Throughout this chapter the terms players and SUs shall be used interchangeably.

Thus, the number of players in the game is a Poisson random variable with average number of players³ $\lambda >> 1$. Each user from the complete population belongs to one of the types $t \in \mathcal{T}$. The probability of a user being of type t is given by r(t) = t $\operatorname{Prob}(type = t)$. This information is embedded in the vector $r \in \Delta(\mathcal{T})$, where $\Delta(\cdot)$ represents the set of probability distributions over \mathcal{T} . By applying the *decomposition* property [130, 132] of the Poisson distribution, we can establish that the number of players in the game of type t is also a Poisson random variable with parameter $\lambda r(t)$. We assume that the set \mathcal{C} of possible actions is common to all players, regardless of their type. Thus, the set $\Delta(\mathcal{C})$ is the set of mixed actions associated with the players. In Poisson games, the utility of a specific player depends on its type, the action he chooses, and on the number of players (not counting himself) who choose each possible action. The number of players for each possible element in \mathcal{C} is listed in a vector called the *action profile*. The last term of the tuple is the utility, defined as $u = (u_t)_{t \in \mathcal{T}}$, where $u_t(a, x)$ is the payoff that a player of type t receives when a pure action a is chosen and the number of players who choose action b is x(b), for all $b \in \mathcal{C}$. If the participants play in accordance to the strategy σ , we call $\sigma_t(a)$ the probability that a player of type t chooses the pure action a. Using the decomposition property again, we can establish that the number of players of type $t \in \mathcal{T}$ who choose the pure action a is Poisson distributed with mean $\lambda r(t)\sigma_t(a)$. Since the sum of independent Poisson random variables is also a Poisson variable with mean equal to the sum of the means, the total number of players who take the pure action a is Poisson distributed with mean $\lambda \tau(a)$, where

$$\tau(a) = \sum_{t \in \mathcal{T}} r(t) \sigma_t(a).$$

It follows that a player of type t who plays a pure action $a \in C$ while the rest of the players are expected to play using strategy σ has a expected utility of

$$U_t(a,\sigma) = \sum_{x \in Z(\mathcal{C})} P(x|\sigma) u_t(a,x),$$
(5.7)

^{3.} In the original paper by Myerson, the value of λ is chosen to be very large; however the validity of the results can be applied here as long as a relatively large λ is used so that the probability of having zero players in the Poisson game is negligible.

where

$$P(x|\sigma) = \prod_{b \in \mathcal{C}} e^{-\lambda\tau(b)} \frac{(\lambda\tau(b))^{x(b)}}{x(b)!},$$
(5.8)

while the expected utility when the player chooses action $\theta \in \Delta(\mathcal{C})$ is

$$U_t(\theta, \sigma) = \sum_{a \in \mathcal{C}} \theta(a) U_t(a, \sigma).$$
(5.9)

5.3.1 Nash Equilibrium in Poisson Games

It is very well known [107] that a Nash equilibrium is achieved when each strategy played by all players corresponds to the best response to all other strategies in such equilibrium. Consequently, no player has anything to gain by changing his own strategy unilaterally. The set of best responses for a player of type t against a strategy σ is then the set of actions that maximizes his expected utility given that the rest of the players (including those whose type is t) play as prescribed by σ . Let us define the set

$$B_t(\sigma) = \left\{ b \in \mathcal{C} : b \in \arg\max_{a \in \mathcal{C}} U_t(a, \sigma) \right\}$$
(5.10)

,

as the set of pure best responses against σ for a player of type t. Equally, the set for mixed best responses against σ is the set of actions $\Delta(B_t(\sigma))$. Therefore, the strategy σ^* is a Nash equilibrium if $\sigma_t^* \in \Delta(B_t(\sigma)) \ \forall t$.

5.3.2 Examples and Motivation

5.3.2.1 Example 1

Let Γ be a Poisson game with $\lambda = 15$, only one type of players, set of available choices $\mathcal{C} = \{ON, OFF\}$, and the utility function:

$$u(ON, x) = \begin{cases} R & \text{if } x(ON) \le K_{\max} \\ 0 & \text{otherwise} \end{cases}$$
$$u(OFF, x) = 0 \quad \forall x \in \mathcal{C},$$



Figure 5.4: Average utility for Poisson Game in Example 1 ($\lambda = 15$).

where K_{max} is the maximum number of players that can transmit at the same time beside one transmitting player without causing a collision, and R > 0 is the transmission rate payoff when the player achieves a successful transmission. This game follows the mixed strategies defined as $\sigma(ON) = p$ and $\sigma(OFF) = 1 - p$, where p is the probability of transmission by any given player. Using eq. (5.9), we can calculate the expected utility as

$$U(p) = R \sum_{n=1}^{K_{\max}+1} \sum_{i=n}^{\infty} \frac{n}{i} {i \choose n} p^n (1-p)^{i-n} \left\{ e^{-\lambda} \frac{\lambda^i}{i!} \right\}$$

= $R \sum_{n=1}^{K_{\max}+1} \frac{p^n \lambda^n \left[\Gamma(n+1) - n\Gamma(n, -\lambda(1-p)) \right]}{e^{\lambda} n! (-\lambda(1-p))^n}, \quad p < 1.$ (5.11)

The dependence of U(p) as a function of p is shown in Figure 5.4 for different values of K_{max} . The solid line represents the utility of the Poisson game and the dashed lines represent the utility of a game with complete information (*i.e.*, fixed number of players known for all). Considering that a Nash equilibrium predicts in a consistent manner the way in which a game will be played, it is evident that in this game there exists just one logical outcome. In other words, the Nash equilibrium is strict, which by definition must occur in non-degenerate strategies [104]. The equilibrium occurs when all players transmit all the time (p = 1). As seen in Figure 5.4, this converges to a zero utility for all the users. This game is designed in such a way that the players have no motivation to not transmit. The utilities obtained are very far from the Pareto optimal⁴, which could be achieved by playing a mixed strategy (p < 1). Figure 5.5 shows the achievable utilities if the players were motivated to play Pareto dominant strategies, and thus transmitting only a fraction of the time p < 1. By increasing the value of K_{max} , the probability of transmission by a player also increases, resulting in a higher utility. As K_{max} tends to the maximum number of players known in a game with complete information, the average utility converges to 100%. At the same time, in a Poisson game, because of the uncertainty of the number of players, the average utility does not achieve the maximum when K_{max} tends to λ . Thus, it is important to consider the Poisson game in detail.

5.3.2.2 Example 2

Consider the Poisson game defined by $\Gamma = \{\lambda, \mathcal{T}, r, \mathcal{C}, u\}$, with expected number of players $\lambda = 15$, set of types $\mathcal{T} = \{1, 2\}$ with probabilities r_1 and $r_2 = 1 - r_1$ set of choices $\mathcal{C} = \{ON, OFF\}$ and the utility function:

$$u_{1}(ON, x) = \begin{cases} R_{1} & \text{if } x(ON) \leq K_{\max_{1}} \\ 0 & \text{otherwise} \end{cases}, \\ u_{1}(OFF, x) = 0 \quad \forall x \in \mathcal{C}, \\ u_{2}(ON, x) = \begin{cases} R_{2} & \text{if } x(ON) \leq K_{\max_{2}} \\ 0 & \text{otherwise} \end{cases}, \\ u_{2}(OFF, x) = 0 \quad \forall x \in \mathcal{C}, \end{cases}$$
(5.12)

^{4.} As explained in Chapter 5, Pareto Optimality is defined as a specific set of strategies in which no player can change their strategy and have a greater utility without making any other player utility worse.



Figure 5.5: Pareto Optimality Utilities for Example 1.

where K_{\max_1} and K_{\max_2} are the maximum number of players who can transmit simultaneously with type 1 and type 2 players, respectively. Similarly to Example 1, R_1 and R_2 are the achievable rates the in case of a successful transmission. We define the mixed strategies $\sigma_1(ON) = p_1$ and $\sigma_2(ON) = p_2$ as the transmission probabilities by type 1 and type 2 players, respectively. The expected utilities can be calculated using of eqs. (5.7)-(5.9) as follows:

$$U_{1}(p_{1}, p_{2}) = R_{1} \sum_{n=1}^{K_{\max_{1}}+1} \sum_{k=1}^{n} \left\{ \sum_{i=k}^{\infty} \frac{k}{i} \binom{i}{k} p_{1}^{i} (1-p_{1})^{i-k} \\ \times \left[e^{-r_{1}\lambda} \frac{(r_{1}\lambda)^{i}}{i!} \right] \sum_{j=n-k}^{\infty} \binom{j}{n-k} p_{2}^{j} (1-p_{2})^{j-n+k} \\ \times \left[e^{-r_{2}\lambda} \frac{(r_{2}\lambda)^{j}}{j!} \right] \right\},$$

$$U_{2}(p_{1}, p_{2}) = R_{2} \sum_{n=1}^{K_{\max_{2}}+1} \sum_{k=1}^{n} \left\{ \sum_{i=k}^{\infty} \frac{k}{i} \binom{i}{k} p_{2}^{i} (1-p_{2})^{i-k} \\ \times \left[e^{-r_{1}\lambda} \frac{(r_{2}\lambda)^{i}}{i!} \right] \sum_{j=n-k}^{\infty} \binom{j}{n-k} p_{1}^{j} (1-p_{1})^{j-n+k} \\ \times \left[e^{-r_{1}\lambda} \frac{(r_{1}\lambda)^{j}}{j!} \right] \right\}.$$
(5.13)

Figures 5.6 and 5.7 show the expected utility for this example. Again, the players have no incentive to use mixed strategies, and the same strict Nash equilibrium therefore occurs- all players transmit all the time (*i.e.* $p_1 = p_2 = 1$), resulting in zero utility to all of them. However, unlike Example 1, there exist several Pareto dominant mixed strategies in terms of the pairs of probabilities (p_1, p_2) forming a Pareto frontier which will produce a set of optimal strategies [104, 133]. In the following section, we will reformulate the proposed Poisson games in order to achieve the Pareto optimality. Notice that the analysis of Examples 1 and 2 concentrates only on symmetric Nash equilibria since this game has also an asymmetric equilibria.

5.4 System Model and Corresponding Game

Let us assume for a moment that there is a fixed number N of SUs competing stations while the PU is in the idle state. The transmission queue for all SUs is assumed to be always nonempty, *i.e.*, each user always has a packet ready to be transmitted right after the completion of each transmission. We consider that the system works



Figure 5.6: Average utility ($\lambda = 15, P_2 = 0.6, K_{\max_1} = 7, K_{\max_2} = 5$ and $r_1 = 0.3$).



Figure 5.7: Average utility ($\lambda = 15$, $K_{\max_1} = 7$, $K_{\max_2} = 5$ and $r_1 = 0.3$).

in a slotted ALOHA fashion where each slot of the system is modelled as a one-stage game. At the beginning of each slot, the players have to choose between the two pos-



Figure 5.8: Strategic form of multiple access game (N = 2).

sible actions $C \in \{ON, OFF\}$, which represent their ability to transmit or to back off. Every time a player decides to transmit, he can either succeed, in which case he gains throughput, or fail due to a collision, resulting in some penalty (negative throughput) associated with such a failure. The system is capable of handling multi-packet reception (MPR) [134], *i.e.*, it is possible to receive several packets simultaneously⁵. We consider that the channel has no influence on the loss of any package, and that the only option for a failure transmission is therefore due to collisions with any package over the MPR limit ($K_{\text{max}} + 1$). As shown in Example 1 and Figure 5.4, it is possible to obtain a maximum utility by controlling the probability of transmission among the players. This is equivalent to players choosing to play mixed strategies instead of playing the single pure strategy. However, as discussed above, they do not have any incentive to cease transmission to avoid collisions, and the expected outcome would be all of them transmitting, resulting in zero throughput. It is shown in [135] that by introducing some penalty to the game, it is possible to get closer to the Pareto optimality. In terms of this we propose the following game models:

^{5.} This might be achievable by using certain enhancements to the physical layer such as beamforming in MIMO systems, frequency hopping, or multiuser detection.

5.4.1 Poisson Game, Single Type of Players

Let us reformulate Example 1 by adding a penalty in the case of a collision

$$u(ON, x) = \begin{cases} R & \text{if } x(ON) \le K_{\max} \\ -\alpha R & \text{otherwise} \end{cases},$$

$$u(OFF, x) = 0 \quad \forall x \in \mathcal{C},$$

(5.14)

where $\alpha \geq 0$ is a penalization constant. First, we consider a multiple access game with $N \geq 2$ transmitters (players) and $K_{\text{max}} = 0$. In the case of a collision the transmitter is penalized by some constant quantity $-\alpha R$. Figure 5.8 shows the game in strategic form for the case of N = 2. A similar game can be found in [117] as an alternative approach to the distributed coordination function (DCF) in the IEEE 802.11 standard. Each player transmits with probability p following a mixed strategy policy. Consequently, when the number of users is known, the multiple access could be cast as a game with the following utilities:

$$U_{OFF,k} = 0,$$

$$U_{ON,k} = p(1-p)^{N-1}R - \alpha Rp[1 - (1-p)^{N-1}] \qquad (5.15)$$

$$= (1-\vartheta)[\vartheta^{N-1}R - \alpha R(1-\vartheta^{N-1})],$$

for k = 1, 2, ..., N. Here we make use of the notation $\vartheta = 1 - p$. In order for this game to be in equilibrium we need to ensure, by choosing a proper penalty α , that $U_{OFF,k} = U_{ON,k} = 0, \forall k$. In other words the following should be true

$$\begin{split} \vartheta^{N-1} R &= \alpha R (1 - \vartheta^{N-1}), \\ \vartheta &= \left(\frac{\alpha}{1+\alpha}\right)^{\frac{1}{N-1}}, \end{split}$$

As a result, the mixed strategy in equilibrium given as p_{eq} is achieved by

$$p_{eq} = 1 - \left(\frac{\alpha}{1+\alpha}\right)^{\frac{1}{N-1}},\tag{5.16}$$

and it can be seen that

$$p_{eq} = \begin{cases} 1 & \text{if } \alpha = 0 \\ 0 & \text{if } \alpha \to \infty \end{cases}$$
(5.17)

Notice that for arbitrary $N \ge 2$ and $\alpha = 0$ (*i.e.*, no collision penalty), $p_{eq} = 1$. This shows that in general, a game without penalty would have no purpose considering no data can be transmitted at the equilibrium in pure strategies for $N > K_{\text{max}}$. The amount of data transmitted for a given $\alpha > 0$ is then

$$p\vartheta^{N-1}R = \left[1 - \left(\frac{\alpha}{1+\alpha}\right)^{\frac{1}{N-1}}\right] \left(\frac{\alpha}{1+\alpha}\right)R.$$
 (5.18)

The probability of having a particular player transmitting successfully is given as

$$P_{1,k} = p(1-p)^{N-1}, (5.19)$$

and, the maximum of $P_{1,k}$ can be found by taking the partial derivative of its logarithm as follows

$$\frac{\partial \ln P_{1,k}}{\partial p} = \frac{1}{p} - \frac{N-1}{1-p} = 0,$$

This results in the unique solution

$$p = \frac{1}{N},\tag{5.20}$$

which is clearly maximum (see also [34]). It follows from eq. (5.16), that

$$\frac{1}{N} = 1 - \left(\frac{\alpha}{1+\alpha}\right)^{\frac{1}{N-1}},$$

and, therefore

$$\alpha = \frac{1}{\left(\frac{N}{N-1}\right)^{N-1} - 1}.$$
(5.21)



Figure 5.9: Penalization coefficient α .

If N = 2, the corresponding value of α is

$$\alpha(2) = \frac{1}{2-1} = 1.$$

On the other extreme, when $N \to \infty$, one can calculate that

$$\alpha(\infty) = \lim_{N \to \infty} \frac{1}{\left(\frac{N}{N-1}\right)^{N-1} - 1} = \frac{1}{e-1} \approx 0.5.$$

Thus, the range of variation of α is $\frac{1}{e-1} \leq \alpha \leq 1$, as depicted in Figure 5.9. Collisions are frequent for smaller N, and they therefore require a higher penalty to prevent SU from continuous transmission. The case of MPR $K_{max} > 1$ could be treated in a

similar manner. The probability of transmitting without collision can expressed as

$$P_{nc}^{(p)} = \sum_{k=0}^{K_{\max}} p\binom{N-1}{k} p^{k} (1-p)^{N-1-k}$$

= $\sum_{k=0}^{K_{\max}} \binom{N-1}{k} p^{k+1} (1-p)^{N-1-k}$
= $I_{1-p}(N-1-K_{\max},K_{\max}+1),$ (5.22)

where $I_{1-p}(a, b)$ is the incomplete beta function (see [75], Chapter 6) and $N > K_{\text{max}}$. For moderately large N, the binomial distribution of interference could be considered as a sum of N-1 binary random variables. Its distribution can be very well approximated by a normal random variable $\xi \sim \mathcal{N}(\mu, \sigma^2)$, where

$$\mu = (N - 1)p,$$

$$\sigma^2 = (N - 1)p(1 - p) = (N - 1)pq.$$

Therefore,

$$P_{nc} \approx Prob(\xi < K) = \int_{-\infty}^{K_{max}} \frac{1}{\sqrt{2\pi(N-1)pq}} \exp\left(\frac{[x-(N-1)p]^2}{2(N-1)pq}\right) dx$$

$$\approx \Phi\left(\frac{K_{max} + 0.5 - (N-1)p}{\sqrt{(N-1)pq}}\right),$$
 (5.23)

for large N compared to K_{max}^{6}

$$K_{\max} < (N-1)p.$$
 (5.24)

^{6.} One can assume that the maximum $P_{nc}(p)$ is achieved when $p \ll 1$ (or more accurately $p \sim K_{\text{max}}/N$).

As the next step, let us maximize the last term in the expansion in eq. (5.22) with respect to p,

$$P_{nc}^{*}(p) = {\binom{N-1}{K_{\max}}} p^{K_{\max}+1} (1-p)^{N-1-K_{\max}},$$

$$\frac{\partial \ln P_{nc}^{*}(p)}{\partial p} = \frac{K_{\max}+1}{p} - \frac{N-1-K_{\max}}{1-p} = 0,$$

$$(K_{\max}+1)(1-p) = (N-1-K_{\max})p$$
(5.25)

or

$$p_{\max} = \frac{K_{\max} + 1}{N}.$$
(5.26)

This term represents the largest contribution to P_{nc} given by eq. (5.22). Furthermore,

$$\frac{\binom{N-1}{K_{\max}}p^{K_{\max}+1}(1-p)^{N-1-K_{\max}}}{\binom{N-1}{K_{\max}-1}p^{K_{\max}}(1-p)^{N-K_{\max}}} = \frac{N-K_{\max}}{K_{\max}}\frac{p}{1-p} \approx \frac{Np}{K_{\max}}\frac{1}{1-p} >> 1. \quad (5.27)$$

The optimized term provides the bulk contribution to P_{nc} . Taking one more term in eq. (5.22) and following the same reasoning, it is possible to obtain a very accurate approximation of the optimum probability of transmission as

$$P_{opt} \approx \frac{K_{\max} + 1}{K_{\max} + N}.$$
(5.28)

The accuracy of such an approximation can be seen in Figure 5.10, where P_{opt} calculated by (5.28) is shown as a special character, described in the plot legend. Using this value of P_{nc} we can reformulate eq. (5.15) as

$$U_{OFF,k} = 0$$

$$U_{ON,k} = RP_{nc}(P_{opt}) - \alpha RP_c(P_{opt}),$$
(5.29)

where $P_c = 1 - P_{nc}$ represents the probability of collision. Therefore, at the equilibrium

$$U_{OFF,k} = U_{ON,k} ,$$

$$P_{nc} = \alpha P_c ,$$



Figure 5.10: Approximation of optimal probability P_{opt} considering N = 20 and different values of K_{max} .

and

$$\alpha = \frac{P_{nc}}{1 - P_{nc}} \bigg|_{p = P_{opt} = \frac{K_{max} + 1}{K_{max} + N}} = \frac{\sum_{k=0}^{K_{max}} {\binom{N-1}{k} (1-p)^{N-1-k} p^k}}{\sum_{i=K_{max}}^{N-1} {\binom{N-1}{i} (1-p)^{N-1-i} p^i}} \bigg|_{p = P_{opt} = \frac{K_{max} + 1}{K_{max} + N}}.$$
(5.30)

It is consistently assumed in all modelled game theory problems that the players take decisions based on the most basic notion of rational play where dominated strategies can be iteratively eliminated [114]. It is in this sense that the Nash equilibrium predicts the most likely outcome of the game. The SUs in this game transmit with a certain probability. If such probability coincides with the one in equilibrium, then no player has any incentive to change it. We consider that the number of players is a random variable distributed according to some probability distribution $P_N(N)$. Conditioned on the number of players N, the pay-off can be rewritten as

$$U_{OFF|N} = 0,$$

$$U_{ON|N} = p\vartheta^{N-1}R - p(1 - \vartheta^{N-1})\alpha R,$$

$$= pR[(1 - \alpha)\vartheta^{N-1} - \alpha].$$
(5.31)

Averaging over the distribution of N (eq. 5.9), the unconditional utilities can be expressed as

$$U_{OFF} = \sum U_{OFF|N} P_N(N) = 0$$
$$U_{ON} = \sum U_{ON|N} P_N(N) =$$
$$pR\left[(1+\alpha)\sum_{N=1}^{\infty} \vartheta^{N-1} P_N(N) - \alpha \sum_{N=1}^{\infty} P_N(N)\right] = 0.$$
(5.32)

By making use of the notation

$$F_N(\vartheta) = \sum_{N=1}^{\infty} \vartheta^{N-1} P_N(N), \qquad (5.33)$$

eq. (5.32) can be rewritten as

$$(1+\alpha)F_N(\vartheta) = \alpha[1-P_N(0)],$$

or

$$\vartheta = F_N^{-1} \left[\frac{\alpha}{1+\alpha} (1 - P_N(0)) \right], \qquad (5.34)$$

where F_N^{-1} is the inverse function of $F_N(\vartheta)$. For instance, if $P_N(N) = \delta(N - N_0)$ $N_0 > 0$, *i.e.*, the game corresponding to a game with complete information (fixed and known number of players), then

$$P_N(0) = 0,$$

$$F_N(\vartheta) = \vartheta^{N_0 - 1},$$

which coincides with eq. (5.16). For the Poisson distribution (5.6)

$$P_N(N) = \frac{\lambda^N}{N!} e^{-\lambda}, \qquad (5.35)$$

one can easily obtain

$$P_N(0) = e^{-\lambda},$$

$$F_N(\vartheta) = \sum_{N=1}^{\infty} \frac{\vartheta^{N-1} \lambda^N}{N!} e^{-\lambda} = \frac{e^{-\lambda}}{\vartheta} \left[\exp(\vartheta \lambda) - 1 \right].$$
(5.36)

Notice that eq. (5.36) must be inverted numerically. The required equation for equilibrium is then

$$\frac{e^{-\lambda}}{\vartheta} \left[\exp(\vartheta \lambda) - 1 \right] = \frac{\alpha}{1+\alpha} \left(1 - e^{-\lambda} \right),$$

$$\frac{e^{\vartheta \lambda} - 1}{\vartheta} = \frac{\alpha}{1+\alpha} \left(e^{\lambda} - 1 \right).$$
(5.37)

If $\alpha = 0$,

$$\frac{e^{\vartheta\lambda}-1}{\vartheta}=0$$

does not have solution $\vartheta = 0$ since

$$\lim_{\vartheta \to 0} \frac{e^{\vartheta \lambda} - 1}{\vartheta} = \lambda \neq 0.$$

Consider the reduced Poisson distribution where N = 0 is not possible. For the case in which there is a random number of transmitters trying to access the channel, the question arises: which penalization α should be chosen in order to obtain the maximum throughput? We can substitute the optimum probability of transmission

for each player by averaging eq. (5.28) with a Poisson distribution as

$$P_{opt} \approx \sum_{N=0}^{K_{\max}} 1 \cdot \frac{\lambda^N}{N!} e^{-\lambda} + \sum_{N=K_{\max}+1}^{\infty} \frac{K_{\max}+1}{K_{\max}+N} \frac{\lambda^N}{N!} e^{-\lambda} = \frac{(K_{\max}+1) \left[\Gamma(K_{\max}+1) - K_{\max}\Gamma(K_{\max},-\lambda)\right]}{K_{\max}e^{\lambda}(-\lambda)^{K_{\max}}}.$$
(5.38)

Notice that eq. (5.38) provides an upper limit in the achievable P_{opt} . A more exact expression could be obtained by taking the expectation of the payoff and then maximize it, however for sake of simplicity, only the average of the maximum point is taking into consideration. For the case when $K_{\text{max}} \ll \lambda$, the first term in eq. (5.38) can be neglected; using the asymptotic of $\Gamma(x, a)$, one obtains the following approximation

$$P_{opt} \approx \frac{K_{\max} + 1}{\lambda + K_{\max} - 1}.$$
(5.39)

Figure 5.11 shows the accuracy of such an approximation for different values of K_{max} . Figures 5.12 and 5.13 show the gain in throughput that can be achieved by exact knowledge of the number of SUs compared to the averaged method described by eq. (5.39). The x-axis in Figure 5.12 represents the number of players N for the case of a game with complete information, or the average number of users λ for the Poisson game analysis. Analogously to eq. (5.30), we can calculate the optimal α using the following:

$$\alpha = \frac{\sum_{N=K_{\max}+1}^{\infty} \sum_{k=0}^{K_{\max}+1} \frac{\sum_{k=0}^{K_{\max}} {N-1 \choose k} (1-P_{opt})^{N-1-k} P_{opt}^{k} \frac{\lambda^{N}}{N!} e^{-\lambda}}{\sum_{N=K_{\max}+1}^{\infty} \sum_{i=K_{\max}}^{N-1} {N-1 \choose i} (1-P_{opt})^{N-1-i} P_{opt}^{i} \frac{\lambda^{N}}{N!} e^{-\lambda}}.$$
(5.40)

5.4.1.1 Throughput Analysis

Following the analysis of the MAC with exponential backoff with MPR [118, 119], we calculate the normalized throughput as follows. First, we obtain the conditional probability of having k packets transmitted successfully given that at least one player



Figure 5.11: Approximation of optimal probability P_{opt} with $\lambda = 20$ and different values of K_{max} .

transmits in any slot as

$$P_{succ}^{(k)} = \binom{N}{k} \frac{P_{opt}^k (1 - P_{opt})^{N-k}}{P_{Tx}},$$
(5.41)

where P_{Tx} is the probability of having at least one player transmitting in the slot time, which can be computed as

$$P_{Tx} = 1 - (1 - P_{opt})^N.$$

Therefore, the normalized throughput for the case of a game with complete information is

$$T = \sum_{k=1}^{K_{\text{max}}} k P_{succ}^{(k)} P_{Tx} = \sum_{k=1}^{K_{\text{max}}} k \binom{N}{k} P_{opt_N}^k (1 - P_{opt_N})^{N-k}, \quad (5.42)$$



Figure 5.12: Probability of non-collision comparison. x- axis indicates N or λ depending whether the number of players N is fixed and known or it is a random Poisson variable with average number of players λ .

where P_{opt_N} is given by eq. (5.28). For the case of Poisson game, eq. (5.42) becomes

$$T_R = \sum_{j=0}^{K_{\text{max}}-1} \frac{\lambda^j e^{-\lambda}}{(j-1)!} + \sum_{n=K_{\text{max}}}^{\infty} \sum_{k=1}^{K_{\text{max}}} k\binom{n}{k} P_{opt_\lambda}^k (1 - P_{opt_\lambda})^{n-k} \frac{\lambda^n e^{-\lambda}}{n!}, \quad (5.43)$$

where $P_{opt_{\lambda}}$ is obtained from eq. (5.39). In order to assess the performance of the proposed scheme, we compare the throughput obtained using Poisson games and the throughput obtained from a binary exponential backoff algorithm implemented in the IEEE 802.11 standard [124]. Standard contention windows $W_0 = 16$ and $W_0 = 32$ were used for comparison. The results are shown in Figure 5.14. In [118] it is shown that as the number of nodes (players) increases, the throughput converges to a nonzero constant in all cases ($\frac{1}{2} \ln 2$ for the case of the binary exponential backoff). The case for the game with complete information is also included. Notice that the throughput obtained with



Figure 5.13: Probability of non-collision comparison between the fixed number of users scheme and random number of users for different values of K_{max} .

the classic exponential backoff for the case of small number of users. This can be explained by the large size of the contention window compared with the number of users.

5.4.2 A Poisson Game with Two Types of Players

In this section we extend the Poisson game interpretation of Multiple Access to the case of two types of SUs, defined by different Quality of Services (QoS) requirements on their rate. For example, some users might use voice services while the rest require a video streaming service. Within this framework, we consider that type 1 and type 2 players transmit with rates R_1 and R_2 , respectively. Depending on the QoS for each type of user, we assume that the maximum number of simultaneous transmissions supported by users of first type could be no more than $K_{\max_1} + 1$, and the maximum supported by the second type could be no more than $K_{\max_2} + 1$. The corresponding



Figure 5.14: Throughput comparison between binary exponential backoff scheme (solid line), game modelled with complete information (dashed line), and game modeled as a Poisson game (dotted line) for different values of K_{max} .

Poisson game can be modelled as

$$u_{1}(ON, x) = \begin{cases} R_{1} & \text{if } x(ON) \leq K_{\max_{1}} \\ -\alpha R_{1} & \text{otherwise} \end{cases},$$

$$u_{1}(OFF, x) = 0 \quad \forall x \in \mathcal{C},$$

$$u_{2}(ON, x) = \begin{cases} R_{2} & \text{if } x(ON) \leq K_{\max_{2}} \\ -\beta R_{2} & \text{otherwise} \end{cases},$$

$$u_{2}(OFF, x) = 0 \quad \forall x \in \mathcal{C},$$

(5.44)

where $\alpha, \beta > 0$ are the two penalization constants in order to guarantee the convergence to a Nash equilibrium in mixed strategies⁷. Similar to eq. (5.29) for the case of a single type of player, it is possible to write the utilities functions in terms of the

^{7.} Notice that x(ON) is the sum of all transmitting players regardless of their type.

probabilities of non-collision and collision as

$$U_{OFF}^{(1)} = 0,$$

$$U_{ON}^{(1)}(p_1, p_2) = R_1 P_{nc}^{(1)}(p_1, p_2) - \alpha R_1 P_c^{(1)}(p_1, p_2),$$

$$U_{OFF}^{(2)} = 0,$$

$$U_{ON}^{(2)}(p_1, p_2) = R_2 P_{nc}^{(2)}(p_1, p_2) - \beta R_2 P_c^{(2)}(p_1, p_2).$$
(5.45)

Here p_1 and p_2 are the probabilities of transmission (mixed strategies) of type 1 and type 2 players, respectively. Let us assume, as in the case of a single type of player, that this is a game with complete information with N_1 and N_2 players of each type. The probability of non-collision for both types can be found using the following expressions:

$$P_{nc}^{(1)}(p_1, p_2) = \sum_{j=0}^{K_{\text{max}_1}} \sum_{i=1}^{j} {N_1 - 1 \choose i} {N_2 \choose j - i} p_1^{i+1} (1 - p_1)^{N_1 - i - 1} p_2^{j-i} (1 - p_2)^{N_2 - j + i},$$

$$P_{nc}^{(2)}(p_1, p_2) = \sum_{j=0}^{K_{\text{max}_2}} \sum_{i=1}^{j} {N_2 - 1 \choose i} {N_1 \choose j - i} p_2^{i+1} (1 - p_2)^{N_2 - i - 1} p_1^{j-i} (1 - p_1)^{N_1 - j + i}.$$
(5.46)

As shown in Example 2 (6.2.2.2), there is a tradeoff in the choice of p_1 and p_2 . It is clear that if p_1 is fixed and p_2 is decreased, type 1 players benefit, and vice versa. Therefore, it is desirable to assign penalties α and β in such a way that the system works in the boundaries of the Pareto frontier, *i.e.*, ensure all choices of mixed strategies be Pareto efficient. The question is how to find such a frontier. We start by noting that if $N_1 = 0$ or $N_2 = 0$, the maximum utility for each type of users is given by eq. (5.28) when

$$P_{nc}^{(1)}\left(\frac{1+K_{\max_1}}{N_1+K_{\max_1}},0\right) \text{ or } P_{nc}^{(2)}\left(0,\frac{1+K_{\max_2}}{N_2+K_{\max_2}}\right).$$



Figure 5.15: Approximation of the Pareto frontier for a complete game with two type of players $(N_1 = 15, N_2 = 10, K_{\max_1} = 5, K_{\max_2} = 3)$.

Similarly, using eq. (5.28) we can see that for the case of $p_1 = p_2$, the maximum probability of non collision for both utilities is achieved by

$$P = \frac{1 + \bar{K}_{\max}}{N_1 + N_2 + \bar{K}_{\max}},\tag{5.47}$$

where \bar{K}_{max} is the arithmetic mean of K_{max_1} and K_{max_2} . Using these three points, one can construct an approximation of the Pareto frontier by connecting them with straight lines, as shown in Figure 5.15. The analytical approximation of such a frontier follows:

$$p_{2} = \begin{cases} m_{1}p_{1} + \frac{1+K_{\max}}{N_{2}+K_{\max}} & 0 \le p_{1} \le \frac{1+\bar{K}_{\max}}{N_{1}+N_{2}+\bar{K}_{\max}} \\ m_{2}\left(p_{1} - \frac{1+K_{\max}}{N_{1}+K_{\max}}\right) & \frac{1+\bar{K}_{\max}}{N_{1}+N_{2}+\bar{K}_{\max}} \le p_{1} \le \frac{1+K_{\max}}{N_{1}+K_{\max}} \end{cases}$$
(5.48)



Figure 5.16: Pareto Frontier $(N_1 = 15, N_2 = 10, K_{\max_1} = 5, K_{\max_2} = 3).$

where m_1 and m_2 are calculated as

$$m_{1} = \frac{\left(\frac{1+\bar{K}_{\max}}{N_{1}+N_{2}+\bar{K}_{\max}} - \frac{1+K_{\max}}{N_{2}+K_{\max}}\right)\left(N_{1}+N_{2}+\bar{K}_{\max}\right)}{1+\bar{K}_{\max}},$$

$$m_{2} = -\frac{1+\bar{K}_{\max}}{\left(N_{1}+N_{2}+\bar{K}_{\max}\right)\left(\frac{1+K_{\max}}{N_{1}+K_{\max}} - \frac{1+\bar{K}_{\max}}{N_{1}+N_{2}+\bar{K}_{\max}}\right)}.$$
(5.49)

The accuracy of the approximation achieved by eq. (5.48) is shown in Figures 5.15-5.16, where we compare it with the Pareto frontier obtained by simulation using a genetic algorithm with 30 iterations in Matlab [136]. All solutions in a Pareto set are equally optimal, so it is up to the wireless designer to select a solution in that set depending on the application or the QoS goal. Furthermore, extending this to the case of a Poisson game, the probability of non collision for both types can be obtained using the equations

$$P_{nc}^{(1)} = \sum_{j=0}^{K_{\max_1}} \sum_{i=0}^{j} \mathcal{P}_1(i) \mathcal{P}_2(j-i),$$

$$P_{nc}^{(2)} = \sum_{j=0}^{K_{\max_2}} \sum_{i=0}^{j} \mathcal{P}_1(j-i) \mathcal{P}_2(i),$$
(5.50)

where

$$\mathcal{P}_{n}(x) = \sum_{s=0}^{x-1} \frac{(r_{n}\lambda)^{s}}{s!} e^{-r_{n}\lambda} + \sum_{l=x}^{\infty} \binom{l}{x} p_{n}^{x} (1-p_{n})^{l-x} \frac{(r_{n}\lambda)^{l}}{l!} e^{-r_{n}\lambda}.$$
 (5.51)

The Pareto frontier for the case of Poisson games can be calculated as

$$p_{2} = \begin{cases} m_{1}p_{1} + \frac{1+K_{\max_{2}}}{r_{2}\lambda+K_{\max_{2}}-1} & 0 \le p_{1} \le \frac{1+\bar{K}_{\max}}{\lambda+\bar{K}_{\max}} \\ m_{2}\left(p_{1} - \frac{1+K_{\max_{1}}}{r_{1}\lambda+K_{\max_{1}}-1}\right) & \frac{1+\bar{K}_{\max}}{\lambda+\bar{K}_{\max}} \le p_{1} \le \frac{1+K_{\max_{1}}}{r_{1}\lambda+K_{\max_{1}}-1}, \end{cases}$$
(5.52)

where m_1 and m_2 are calculated as

$$m_{1} = \frac{\left(\frac{1+\bar{K}_{\max}}{\lambda+\bar{K}_{\max}} - \frac{1+K_{\max}}{r_{2}\lambda+K_{\max}^{2}-1}\right)\left(\lambda+\bar{K}_{\max}\right)}{1+\bar{K}_{\max}},$$

$$m_{2} = -\frac{1+\bar{K}_{\max}}{\left(\lambda+\bar{K}_{\max}\right)\left(\frac{1+K_{\max}}{r_{1}\lambda+K_{\max}^{2}-1} - \frac{1+\bar{K}_{\max}}{\lambda+\bar{K}_{\max}}\right)},$$
(5.53)

following the same technique used in deriving eqs. (5.48) and (5.49). The Pareto Frontier for the Poisson game with two types of players is shown in Figure 5.17. The solid line represents the utility obtained by using eq. (5.52) to calculate the mixed strategies in eq. (5.50). Consequently, the necessary condition for the system to be in equilibrium is

$$U_{OFF}^{(1)} = U_{ON}^{(1)}(p_1, p_2) = 0$$

$$U_{OFF}^{(2)} = U_{ON}^{(2)}(p_1, p_2) = 0$$
(5.54)



Figure 5.17: Pareto Frontier ($\lambda = 30, r_1 = 0.3, K_{\max_1} = 8, K_{\max_2} = 5$).

and the penalization factor α and β can be obtained using

$$\alpha = \frac{P_{nc}^{(1)}(p_1, p_2)}{1 - P_{nc}^{(1)}(p_1, p_2)} \text{ and } \beta = \frac{P_{nc}^{(2)}(p_1, p_2)}{1 - P_{nc}^{(2)}(p_1, p_2)}.$$
(5.55)

Here p_1 and p_2 are obtained from the Pareto frontier using eq. (5.52).

5.5 Game Model with Primary User Activity

In this section we consider the existence of a single PU transmitting within the same channel as the SUs. We assume that the PU transmits in a slot by slot basis with probability P_T , and that the slots are synchronized between the PU and the SUs. In this sense, we consider a PU transmitting to be in an ON state. Let \bar{N}_{ON} be the average number of consecutive slots in which the PU is in an ON state, and let us observe the PU over $\nu >> \bar{N}_{ON}$ sequential time slots. Then, on average, there will



Figure 5.18: Primary User Activity

be νP_T ON states, and the average number of transitions from the OFF to ON state is therefore

$$\eta_{OFF \to ON} \approx \frac{\nu P_T}{\bar{N}_{ON}}.$$
(5.56)

We assume that all SUs have perfect detection of the PU activity; when the PU is transmitting, all SUs remain silent. Let $P_{SU,T}$ be the probability that there is a transmission from one or more SUs when the state of the PU is turned ON for the first time after being in an OFF state, creating a collision with the PU. The average number of collisions \bar{N}_{col} is given by

$$\bar{N}_{col} = P_{SU,T} \cdot \eta_{OFF \to ON} = \frac{\nu P_T P_{SU,T}}{\bar{N}_{ON}}.$$
(5.57)

Consequently, the average probability of collision between SUs and the PU is

$$P_{col,PU} = \frac{\bar{N}_{col}}{\nu P_T} = \frac{P_{SU,T}}{\bar{N}_{ON}}.$$
(5.58)

For the game with full information, the probability $P_{SU,T}$ represents the probability of having at least one SU transmitting at the same time as the PU. This can be calculated as follows

$$P_{SU,T}(p) = 1 - (1-p)^N.$$
(5.59)

As an example of eq. (5.59), let us assume that the transmissions from a single PU operate in a channel inversion with cut-off mode [137]. The transmissions follow

a Gilbert-Elliot (GE) model (particularly one which imposes a correlation ρ in the time domain) [138] [139] where the "GOOD" state occurs with probability P_T and the "BAD" state occurs with probability $1 - P_T$. In this case, the probability of transmission P_T is directly related to the fading channel as

$$P_T = \int_{\gamma_0}^{\infty} p_{\gamma}(\gamma) d\gamma, \qquad (5.60)$$

where γ_0 is an energy threshold above which a transmission is possible and $p_{\gamma}(\gamma)$ is the fading distribution of the channel. Within this context, we consider that the PU will transmit in a slot only within a "GOOD" state and otherwise remain silent. As seen in Figure 5.18, we can model the dynamic traffic from the PU with r and q defined as follows [139]:

$$q = P_T(1 - \rho),$$

$$r = (1 - P_T)(1 - \rho),$$
(5.61)

where $0 \le \rho \le 1$ is the correlation coefficient. The average duration of the PU in the ON state \bar{N}_{ON} can be calculated as

$$\bar{N}_{ON} = \sum_{i=1}^{\infty} i(1-r)^{i-1}r = \frac{1}{r} = \frac{1}{(1-P_T)(1-\rho)}.$$
(5.62)

Let P_{col}^{Th} be a pre-established tolerance threshold defined as the maximum average probability of collision $P_{col,PU}$ the PU would be able to tolerate. It follows from equation (5.58) that

$$P_{SU,T}(p) \le \bar{N}_{ON} P_{col}^{Th}.$$
(5.63)

Note that for $P_{col}^{Th} \geq 1/\bar{N}_{ON}$, any value of p satisfies eq. (5.63). This means that the SUs can transmit with any probability, and simply stop transmitting when they detect the presence of a transmitting PU. When $P_{col}^{Th} < 1/\bar{N}_{ON}$, there is a value p^* such that

$$P_{SU,T}(p^*) = \bar{N}_{ON} P_{col}^{Th}, \tag{5.64}$$



Figure 5.19: Restriction frontier approximation for $N_1 = 25, N_2 = 20$.

which can be rewritten as

$$p^*(N) = 1 - \left(1 - \bar{N}_{ON} P_{col}^{Th}\right)^{\frac{1}{N}}.$$
(5.65)

If $p^* \geq \frac{K_{\max}+1}{K_{\max}+N}$ in eq. (5.28), the impact to the PU can once again be ignored. However, if $p^* < \frac{K_{\max}+1}{K_{\max}+N}$, a different value of P_{opt} must be used in order to calculate the penalty α in eq. (5.30). Consequently, P_{opt} in eq. (5.28) will be given as

$$P_{opt} = \min\left(p^*(N), \frac{K_{\max} + 1}{K_{\max} + N}\right).$$
(5.66)

Considering the example, one can note that for a fixed tolerance P_{col}^{Th} , the SUs can transmit using the optimal strategy without significantly affecting the PU if it has very poor channel conditions and/or the channel is uncorrelated. By extending this analysis to the case of two types of players, the analogous effect of eq. (5.66) is to create a restriction frontier under which the SUs can calculate their penalization factor α and β using equation (5.55). Such a frontier can be obtained by rewriting eq. (5.59) as

$$P_{SU,T}(p_1, p_2) = 1 - (1 - p_1)^{N_1} (1 - p_2)^{N_2} \le \bar{N}_{ON} P_{col}^{Th},$$
(5.67)

and noting that the maximum $P_{SU,T}$ for each type of player that satisfies eq. (5.67) occurs when either all type 1 SUs transmit with probability $p_1^*(N_1)$ and none of the type 2 SUs transmit (*i.e.*, $P_{SU,T}^{(1)}(p_1^*(N_1), 0)$), or all type 2 SUs transmit with probability $p_2^*(N_2)$ and none of the type 1 SUs transmit (*i.e.*, $P_{SU,T}^{(2)}(0, p_2^*(N_2))$). Therefore, an accurate approximation of the restriction frontier can be formed by connecting the two points with a straight line, as shown in Figure 5.19. It can be seen that the quality of such an approximation is very good for the selected values of parameters. We omit a detailed analysis due to lack of space. Finally, in order to calculate p^* for a Poisson game with the average number of SUs denoted as λ , we have to average eq. (5.65) over the Poisson distribution analogously with eq. (5.38) as

$$p^{*}(\lambda) = \sum_{k=0}^{\infty} p^{*}(k) \frac{\lambda^{k} e^{-\lambda}}{k!}$$
$$= \sum_{k=0}^{\infty} \left\{ 1 - \left(1 - \bar{N}_{ON} P_{col}^{Th}\right)^{\frac{1}{k}} \right\} \frac{\lambda^{k} e^{-\lambda}}{k!}.$$
(5.68)

Here $p^*(k)$ can be expanded in terms of a Taylor series with respect to the number of players k around the average value λ to produce

$$p^{*}(k) = e^{\frac{\ln(1-\bar{N}_{ON}P_{col}^{Th})}{\lambda}} - e^{\frac{\ln(1-\gamma)}{\lambda}} \frac{\ln(1-\bar{N}_{ON}P_{col}^{Th})(k-\lambda)}{\lambda^{2}} + \mathcal{O}\left\{(k-\lambda)^{2}\right\}.$$
(5.69)

By taking the first two terms of eq. (5.69), we can approximate the solution of eq. (5.68) for large λ (as assumed throughout this chapter) as

$$p^*(\lambda) \approx 1 - \frac{\lambda e^{\lambda} - \lambda - \ln\left(1 - \bar{N}_{ON} P_{col}^{Th}\right)}{\lambda e^{\lambda - \ln\left(1 - \bar{N}_{ON} P_{col}^{Th}\right)}}.$$
(5.70)

Therefore P_{opt} in eq. (5.39) can be obtained simply by

$$P_{opt} = \min\left(p^*(\lambda), \frac{K_{\max} + 1}{K_{\max} + \lambda - 1}\right).$$
(5.71)

Analogously, for the case of two types of players, eq. (5.67) becomes

$$P_{SU,T}(p_1, p_2) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \left\{ 1 - (1 - p_1)^i (1 - p_2)^j \right\} \times e^{\lambda} \frac{(r_1 \lambda)^i}{i!} \frac{(r_2 \lambda)^j}{j!} \le \bar{N}_{ON} P_{col}^{Th}.$$
(5.72)

The restriction frontier can therefore be formed, simply by connecting the two points corresponding to

$$P_{SU,T}^{(1)}(p_1^*(r_1\lambda), 0) \text{ and } P_{SU,T}^{(2)}(0, p_2^*(r_2\lambda)).$$

In Figures 5.20 and 5.21, the influence of the PU activity on the strategy choice by SUs in the case of two types of players is shown for different values of K_{max} . We present three different scenarios. In Figures 5.20(a) and 5.21(a), the restriction frontier is below the Pareto frontier. This means that the SUs must limit their transmission probabilities within the boundaries of the shown area in order to avoid significantly affecting the PU. However, in Figures 5.20(b) and 5.21(b), the Pareto frontier below the restriction frontier. Here, the SUs can choose their strategies based on the Pareto frontier already calculated with the guarantee that the PU's performance will remain unaltered and at the same time, obtain a maximum throughput. In Figures 5.20(c)-(d) and 5.21(c)-(d), we show the case of an intersection between the Pareto frontier and the restriction frontier. Here, the SUs can use any transmission strategy within



Figure 5.20: Restriction Frontier and Pareto Frontier for two types of players with full information $(N_1 = 25, N_2 = 20, \gamma = 0.9)$.

the area shown in order to minimize the effect on the PU to a minimum. However, it is always better to choose strategies which lie on the Pareto frontier boundaries as opposed to the restriction frontier boundaries in order to obtain a better throughput for all players.



Figure 5.21: Restriction Frontier and Pareto Frontier for Poisson games $(\lambda = 25, \gamma = 0.9).$

5.6 Conclusion

In this chapter, we provided a comprehensive overview of game theory and its application to research on cognitive radio networks. The concept of noncooperative game was explained along with the Nash equilibrium for mixed and pure strategies. We explained the notion of Pareto optimality as a criteria for equilibrium selection when more than one Nash equilibrium are present in a game. Finally, we reviewed the well known Prisoners' Dilemma problem and we used to analogously model the multiple access problem in the IEEE 802.11 DCF standard. Also, in this chapter we present a game-theoretic perspective on the SUs multiple access problem in cognitive radio networks. We have showed how to design a game with a fixed number of homogeneous SUs with penalization. The Nash equilibrium from such a game results in an optimal throughput of the network per SU. Corresponding analytical expressions have been
obtained for the SU nodes with MPR properties. Furthermore, we have extended these results to be incorporated into a game with a random number of players (SUs). The Poisson games allow us to account for a dynamically changing number of active SUs. An optimal probability of transmission for the SUs (a mixed strategy) is calculated in order to achieve the Pareto optimality in the system. This was proven to achieve a better performance for a small number of users than other well known approaches, such as the DCF. We extended the game-theoretic analysis to the case of two different types of SUs with different QoS requirements. We showed that the Pareto frontier can be accurately approximated by means of connecting a piece-wise function based on the optimal probability of transmission obtained from each type of player in isolation. Finally, we considered the impact of the dynamic activity of the PU on the optimal strategy for the SUs. We showed that the optimal probabilities of transmission for the SUs are influenced by the PU pattern only under certain conditions. More specifically, we derived the conditions for the Pareto frontier under which the SUs' activities are limited by either the equilibrium strategy in games without a PU, by constrains on the SINR of a PU, or by both in a piece-wise manner. In particular, we showed that if the PU has low intermittency in the transmitting intervals, its effect on the SU strategy can be neglected. However, for PUs with relatively bursty activities, the strategy of SUs is limited by the SINR requirements at the PU.

Chapter 6 Thesis Summary and Future Work

In this thesis, we explored a number of topics surrounding spectrum sensing and spectrum sharing for Cognitive Radio networks. The available electromagnetic spectrum useful for wireless communication devices is of limited physical extent. Cognitive Radio (CR) has been raised as a feasible solution to the spectrum underutilization problem by dynamically accessing to the channel. A CR is an intelligent entity that is aware of its surroundings and adapts its transmission parameters accordingly. In this way, the performance of the of wireless transmissions can be optimized, and the utilization of the frequency spectrum can be enhanced. The major functionalities of a CR device include spectrum sensing, spectrum management, and spectrum mobility. Through spectrum sensing, information on the target radio spectrum is obtained. Different spectrum sensing techniques have been developed, *e.g.*, energy detection, matched filter detection, and cyclostationary detection among others This information is then used by the spectrum-management function to determine spectrum opportunities and make decisions on spectrum access. If the status of the target spectrum changes, the spectrum mobility function can change the operational parameters.

6.1 Thesis Summary

In Chapter 1, a thorough review of the concept, applications and challenges of CR networks was provided. It was discussed that the major factor that leads to an inefficient use of radio spectrum is the spectrum licensing scheme itself. Two kinds of users were defined, Primary Users (PUs) who are licensed to use the radio spectrum allocated to them, and Secondary Users (SUs), who are not licensed. To improve the

efficiency and utilization of the available spectrum, new spectrum licensing models have been introduced. The idea is to make spectrum access more flexible by allowing SUs to access the radio spectrum under certain conditions. Two major approaches for spectrum sharing in CR networks were introduced: spectrum overlay approach and spectrum underlay approach. In the spectrum underlay approach, power control is important for SUs not only for maximizing the transmission rate but for maintaining the interference below target levels so that the PUs are not harmed. For this purpose, the concept of interference temperature was introduced as a new metric on interference assessment. In the spectrum overlay approach, based on the shared user model, Medium Access Control (MAC) is important to SUs for detecting and accessing spectrum opportunities. Several classic transmitter detection techniques for spectrum sensing and analysis were described, along with a few new techniques. We explained the need for dynamic spectrum allocation and sharing methods in CR networks once the spectrum holes or spectrum opportunities are found.

In Chapter 2, a multi-antenna based spectrum sensing in CR networks is studied using the well known Generalized Likelihood Ratio Test (GLRT) approach and the Neyman-Pearson (NP) criteria. Depending on the information about the PU signal, different GLRT approaches can be used. If neither the signal covariance matrix \mathbf{R}_{s} nor the noise variance σ^{2} are known, the easiest and less complex option is to use the energy detector. Nevertheless, if we are willing to increase the complexity of the detector by trying to estimate the unknown parameters, we can significantly increase the performance of the energy detector using the arithmetic to geometric detector. On the other hand, if we assume that the noise power σ^2 is known, but \mathbf{R}_s is unknown, the optimal approach in the GLRT criteria is given by the signal-subspace eigenvalues method which attempts to estimate the covariance matrix σ^2 . Finally, if both parameters are known, the optimal detection in the GLRT sense is to apply the estimator-correlator detector. In Chapter 2, we also showed that cyclostationary spectrum sensing as well as collaborative spectrum sensing in CR networks can be interpreted as a special case of the concept of optimum or sub-optimum incoherent diversity combining approach (SIMO radar). The number of virtual branches is equal to the product of the number of cyclic frequencies and the time delays. The concrete detection algorithms using the NP test lead to the SIMO radar algorithms, and their performance was analyzed for the case of the Generalized Gaussian channel fading models. Finally, the idea of self-organization of coupled dynamical systems was presented as a decentralized synchronization approach that is able to detect the presence of a PU (acting as an external agent). One significant advantage of the method is its ability to be implemented with analog circuitry without any need of analog to digital converters and extensive digital signal processing. This approach is also very useful in addressing the hidden terminal problem in CR networks.

Chapter 4 presents a novel analysis that allows us to take into account the influence of both spatial and temporal correlation in CR receivers equipped with more than one antenna. An approximate expression for the probability of miss detection as a function of the number of antennas, scattering parameters, and number of observations is obtained. Using this expression, we are able to assess the impact that scattering has on the detection performance of the PUs. We shown that in low Signalto-Noise Ratio (SNR) scenarios, it is more convenient to include only one antenna and consider many time samples. This allows for a better noise suppression. In the high SNR regime, it is more appropriate to have more receiving antennas to mitigate the fading due to the high noise immunity.

In Chapter 3, we introduced the sequential analysis or Wald's Sequential Probability Ratio Test (SPRT) detection method for PU detection in CR networks. We show that SPRT can result in a savings of nearly fifty percent of the average number of observations (in comparison to the NP test) needed to detect the presence of the PU. We also show that non-coherent detection requires almost twice as many samples as a coherent detection approach. We derived an optimal fusion rule with detectors using sequential analysis and showed that in order to achieve a faster decision, the Fusion Center (FC) does not consider the information of sensors that have not made a decision at some point. Subsequently, we presented a novel procedure to obtain any order cumulant of the probability density function (PDF) of the average sample distribution in sequential analysis. Using this procedure, we present a new approach to the dual SPRT scheme. We showed that SU sensors can save energy if we optimize the instance when the FC gathers their local decisions and then performs itself the SPRT algorithm.

Chapter 5 reviews several important Game Theory (GT) concepts and links their application to wireless communications systems by means of some practical examples. Specifically, the topics of non-cooperative games and Nash equilibrium were covered in great detail. Using GT concepts, we showed that, especially in multichannel environments, channel selection/allocation to avoid congestion among SUs can be formulated as a non-cooperative game. We presented a GT model of the Distributed Coordination Function (DCF) in the IEEE 802.11 standard and addressed the MAC problem from a different perspective in such networks. This Chapter also presents a game theoretic treatment to the MAC problem in CR networks. We proposed a game which penalizes failed transmission attempts by the SUs. We showed that the Nash equilibrium achieved by this game results in an optimal throughput of the network per SU. A very novel result and model was obtained by introducing population uncertainty into the game, *i.e.*, a random number of players. To the best of our knowledge, this analysis has not been considered in any previous literature. We also considered two types of SUs or players in this game, and their optimal probabilities of transmission are calculated in order to achieve the Pareto optimality of the system. Finally, we examine the impact of the dynamic activity of the PU on the SUs' optimal strategies. We derived useful conditions for the Pareto frontier under analysis such that the SU transmissions are limited by either the equilibrium strategy in games with no PU or by constrains on the SINR (or interference temperature) of a PU.

6.2 Future Work

The work in this thesis investigates some of the major aspects and challenges involved in CR networks; however, we have just scratched the surface concerning this new communication paradigm. Here we present a few possible future directions that could be taken to extend this important topic even farther.

6.2.1 Coalitional Game for Spectrum Sensing

As presented in Chapters 2 and 3, cooperative spectrum sensing can lead to more accurate decisions by SUs than independent sensing, especially when the signal from a PU presents fading or shadowing. The FC processes the individual results and achieves a final decision. However, the performance of the spectrum sensing method can be improved if cooperation among SUs is allowed. The SUs form coalitions (*i.e.*, groups) and share individual sensing results. With more information from the coalition, the decisions on spectrum sensing can be made more accurate.

6.2.2 Power Allocation as a Non-Cooperative Game

In underlay DSA, SUs and PUs can access the same spectrum simultaneously. In this case, the interference to the PU can be limited by controlling the transmit power of the SUs. This is similar to the concept of the CDMA cellular system. Power/control allocation is crucial in order for the SUs to achieve the best performance, while interference to the PU is maintained below a target level. Power allocation by SUs becomes more challenging when the SUs are non-cooperative. All SUs need equilibrium strategies (*i.e.*, transmission power) to ensure not only that none of them deviates from the equilibrium, but also that the interference requirement is not violated. This problem of power allocation can be formulated as a non-cooperative game.

6.2.3 Spectrum Leasing and Cooperation

In CR networks, the PU needs an incentive to share the spectrum with SUs. This incentive could be done through pricing. Alternatively, SUs can help the PU transmit data so that the PU's transmission rate and reliability are improved. For instance, we can consider the incentive in which the cooperative diversity technique (*e.g.*, decode-and-forward) is used so that SUs can relay the transmitted data of the PU. In return, the PU could allow the SUs to access the spectrum. This approach could save a great deal of spectrum sensing processing on the SU side since they already know when the PU will cease or start transmitting. This exchange of resources can be formulated as a hierarchical game.

6.2.4 Radio Resource Competition Based on Stochastic Learning Games

In order to obtain spectrum access, a SU can bid competitively for the spectrum from a central spectrum moderator (*i.e.*, a spectrum broker) as explained in Chapter 2. To bid for the spectrum, not only the channel state but also the local state (*e.g.*, buffer occupancy) of the SU will impact the strategy selection. In a dynamic environment, a stochastic game model can be formulated to obtain a competitive strategy for spectrum bidding. Additionally, if information about the other SUs is not publicly available, each user has to learn and adapt its strategy dynamically to achieve the highest reward, or equivalently, the lowest cost.

References

- Federal Communications Commission, "Spectrum policy task force," FCC, Tech. Rep., 2002 November. 1
- [2] T. Erpek, K. Steadman, and D. Jones, "Dublin ireland spectrum occupancy measurements collected on april 16-18,2007," Shared Spectrum Company, Tech. Rep., 2007. 1
- [3] L. E. Doyle, Essentials of Cognitive Radio, ser. The Cambridge Wireless Essentials Series, W. Webb, Ed. Cambridge University Press, 2009. 2
- [4] S. Haykin, "Cognitive radio: brain-empowered wireless communications," vol. 23, no. 2, pp. 201–220, February 2005. 2, 3
- [5] K. Hamdi, W. Zhang, and K. B. Letaief, "Opportunistic spectrum sharing in cognitive mimo wireless networks," *Trans. Wireless. Comm.*, vol. 8, no. 8, pp. 4098–4109, 2009. 2
- [6] J. Mitola and J. Maguire, G.Q., "Cognitive radio: making software radios more personal," *Personal Communications, IEEE*, vol. 6, no. 4, pp. 13–18, 1999. 3
- [7] F. K. Jondral, "Software-defined radio: basics and evolution to cognitive radio," EURASIP J. Wirel. Commun. Netw., vol. 2005, no. 3, pp. 275–283, 2005.
- [8] R. Thomas, L. DaSilva, and A. MacKenzie, "Cognitive networks," in *Proceed-ings IEEE DySPAN*, November 2005, pp. 352–360. 3
- [9] R. Tandra, M. Mishra, and A. Sahai, "Extended edition: What is a spectrum hole and what does it take to recognize one?" EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2008-110, Aug 2008. 4, 17

- [10] FCC, Ed., Establishment of interference temperature metric to quantify and managa interference and to expand available unlicensed operation in certain fixed mobile and satellite frequency bands, no. 03-289, Federal Communications Commission. FCC, 2003. 5
- [11] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: Challenges and solutions," *EURASIP Journal on Advances in Signal Processing*, no. ID 381465, 2010. 8, 17, 64, 188
- [12] I. Akyildiz, W. Lee, M. Vuran, and S. Mohanty, "A Survey on Spectrum Management in Cognitive Radio Networks," *Communications Magazine*, *IEEE*, vol. 46, no. 4, pp. 40–48, 2008. 8, 17, 98
- B. Wang and K. Liu, "Advances in cognitive radio networks: A survey," Selected Topics in Signal Processing, IEEE Journal of, vol. 5, no. 1, pp. 5–23, 2011.
- [14] J. Lunden, V. Koivunen, A. Huttunen, and H. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," vol. 57, no. 11, pp. 4182–4195, November 2009. 9, 17, 31, 34, 35, 98
- S. Kay, Statistical Signal Processing: Detection Theory. Upper Saddle River, NJ: Prentice Hall PTR, 1998. 9, 10, 17, 21, 22, 27, 28, 29, 101
- [16] H. van Trees, Detection, Estimation, and Modulation Theory: Part I, 1st ed. New York: John Wiley & Sons, 2001. 9, 10, 65, 101
- [17] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Processing*, vol. 86, no. 4, pp. 639 – 697, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/ S0165168405002409 10, 31, 34
- [18] D. Cabric, S. Mishra, and R. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," vol. 1, November 2004, pp. 772 – 776 Vol.1. 10, 17, 98

- [19] S. Haykin, D. Thomson, and J. Reed, "Spectrum sensing for cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 849-877, May 2009. 10
- [20] A. Fehske, J. Gaeddert, and J. Reed, "A new approach to signal classification using spectral correlation and neural networks," in New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on, 2005, pp. 144–150. 11
- [21] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *COMPUTER NETWORKS JOURNAL (ELSEVIER*, vol. 50, pp. 2127–2159, 2006. 11, 14
- [22] Y.-C. Liang, Y. Zeng, E. Peh, and A. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, April 2008. 11
- [23] G. Ganesan and Y. Li, "Cooperative spectrum in cognitive radio networks," in Proceedings IEEE International Symposium New Frontiers Dynamic Spectrum Access Networks (DySPAN), 2005. 12, 19
- [24] Y. Zhang, J. Xiang, Q. Xin, and G. E. Øien, "Optimal sensing cooperation for spectrum sharing in cognitive radio networks," in *European Wireless 2009 (EW* 2009), 2009. 12
- [25] N. Nefedov, Decentralized Synchronization and Estimation in Wireless Networks, ser. Lecture Notes in Computer Science. Berlin: Springer Berlin / Heidelberg, August 2008, vol. Volume 5174/2008, ch. One. 12, 52
- [26] N. Nefedov, "Distributed estimation and synchronization in wireless networks: Coupled dynamical systems approach," in *Nonlinear Dynamics and Synchronization, 2009. INDS '09. 2nd International Workshop on*, July 2009, pp. 110– 115. 12, 13, 52, 57

- [27] R. Zhang, T. J. Lim, Y.-C. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: A glrt approach," *IEEE Transactions on Communications*, vol. 58, no. 1, pp. 84–88, January 2010. 12, 20, 26, 27, 28, 29, 30, 34, 99
- [28] O. Filio, S. Primak, and V. Kontorovich, "Impact of scattering environment on the performance of spectrum sensing in multiantenna cognitive radio systems," in 26th Biennial Symposium on Communications (QBSC), Kingston, Canada, 2012, pp. 38–43. 12, 99, 101, 114
- [29] A. Wald, Sequential Analysis, 6th ed. Mineola, NY: Dover Phoenix, 2004. 13, 63, 64, 65, 67, 68, 70, 76, 78, 81, 82, 187
- [30] O. Filio, V. Kontorovich, S. Primak, and F. Ramos-Alarcon, "Collaborative spectrum sensing for cognitive radio: Diversity combining approach," Wireless Sensor Network, vol. 3, no. 1, pp. 24–37, January 2010. 13, 31, 87, 188
- [31] O. Filio, S. Primak, and V. Kontorovich, "On performance of the wald test in partially coherent cognitive radio channels," in *Communications and Information Technology (ICCIT), 2011 International Conference on*, march 2011, pp. 214–217. 13, 64, 74
- [32] —, "Cumulant analysis of dual sequential ratio testing for cognitive radio spectrum sensing," Wireless Personal Communications, 2013. 13
- [33] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," Signal Processing Magazine, IEEE, vol. 24, no. 3, pp. 79–89, 2007. 14
- [34] L. Kang and L. Ni, "Revisiting ALOHA with physical layer network coding," technical report, CSE Dep. HKUST, Tech. Rep., 2009. 14, 140
- [35] O. Filio, S. Primak, V. Kontorovich, and A. Shami, "A game theory interpretation for multiple access in cognitive radio networks with random number of secondary users," *IEEE Transactions on Communications*, 2013. 15

- [36] R. Tandra and A. Sahai, "Fundamental limits on detection in low SNR under noise uncertainty," vol. 1, June 2005, pp. 464 – 469 vol.1. 17, 98
- [37] A. Ghasemi and E. Sousa, "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs," vol. 46, no. 4, pp. 32–39, 2008. 17, 98
- [38] V. Kontorovich, F. Ramos-Alarcon, O. Filio, and S. Primak, "Cyclostationary spectrum sensing for cognitive radio and multiantenna systems," in *The 2010 International Conference on Wireless Communications and Signal Processing*, Suzhou, China, October 2010. 17
- [39] X. Kang, Y.-C. Liang, H. Garg, and L. Zhang, "Sensing-based spectrum sharing in cognitive radio networks," in *Proc. IEEE GLOBECOM 2008*, New Orleans, LA, November, 30- December, 4 2008. 17
- [40] T. Yucek and Huseyin Arslam, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Comm. Surveys & Tutorials*, vol. 11, no. 1, pp. 116–130, 2009. 19, 97
- [41] H. van Trees, Detection, Estimation and Modulation Theory: Optimal Array Processing, 1st ed. New York: Wiley-Interscience, 2002. 21
- [42] Y. Zeng and Y.-C. Liang, "Eigenvalue-based spectrum sensing algorithms for cognitive radio," *Communications, IEEE Transactions on*, vol. 57, no. 6, pp. 1784–1793, 2009. 26
- [43] S. Boyd and L. Vandenberghe, *Convex Optimization*. The Edinburgh Building, Cambridge, CB2 8RU, UK: Cambridge University Press, 2009. 27
- [44] J. Salz and J. Winters, "Effect of fading correlation on adaptive arrays in digital mobile radio," vol. 43, no. 4, pp. 1049–1057, July 1994. 29
- [45] A. Dandawate and G. Giannakis, "Statistical tests for presence of cyclostationarity," Signal Processing, IEEE Transactions on, vol. 42, no. 9, pp. 2355–2369, 1994. 31, 34

- [46] M. Simon and M.-S. Alouini, Digital Communication over Fading Channels:A Unified Approach to Performance Analysis. New York: John Wiley & Sons, 2000. 31, 32, 43, 44, 99
- [47] D. Klovski, Digital Data Transmission over Radio Channels. Moscow, USSR, (in Russian): Sviaz, 1982. 32, 39, 43, 44, 46
- [48] I. Okunev and O. Okunev, Phase and Phase-Difference Modulation in Digital Communications, ser. Artech House Communications Library. Artech House, Incorporated, 1997. [Online]. Available: http://books.google.ca/books?id= gidmQgAACAAJ 35, 45, 47
- [49] Middleton, Introduction to Statistical Communications Theory, 1st ed. New York: McGraw-Hill, 1960. 36, 65, 75
- [50] A. Papoulis, Probability, Random Variables, and Stochastic Processes, 3rd ed. Boston, MA: McGraw-Hill, 1991. 36, 42, 78, 188
- [51] A. Paulraj, R. Nabar, and D. Gore, Introduction to Space-Time Wireless Communications. Cambridge, UK: Cambridge University Press, 2003. 38
- [52] V. Kontorovich and S. Primak, "Performance analysis of ostbc over generalized gaussian mimo channels," in *Proceedings of the 2009 International Conference* on Wireless Communications and Mobile Computing: Connecting the World Wirelessly, ser. IWCMC '09. New York, NY, USA: ACM, 2009, pp. 1223–1227. [Online]. Available: http://doi.acm.org/10.1145/1582379.1582647 39, 189
- [53] —, "Autocovariance receiver for dtsm: performance characteristics for a generalized gaussian mimo channel," in *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*, ser. IWCMC '09. New York, NY, USA: ACM, 2009, pp. 1228–1232. [Online]. Available: http://doi.acm.org/10.1145/1582379.1582648 39, 45

- [54] P. Beckmann and A. Spizzichino, The scattering of electromagnetic waves from rough surfaces. New York, USA: Pergamon Press, 1963. 42
- [55] I. Gradshteyn and I. Ryzhik, Tables of Integrals, Series, and Products, 6th ed. New York: Academic Press, 2000. 46
- [56] M. Nakagami, "The m-distribution, a general formula of intensity of rapid fading," in *Statistical Methods in Radio Wave Propagation: Proceedings of a Symposium held June 18-20, 1958*, W. C. Hoffman, Ed. Permagon Press, June 1960, pp. 3–36. 47, 190, 191
- [57] M. Patzold, Mobile Fading Channels. New York, NY, USA: John Wiley & Sons, Inc., 2003. 52
- [58] S. Haykin, Adaptive Filter Theory. Pearson Education, Limited, 2013.
 [Online]. Available: http://books.google.ca/books?id=J4GRKQEACAAJ 53
- [59] S. Barbarossa and G. Scutari, "Decentralized maximum-likelihood estimation for sensor networks composed of nonlineary coupled dynamical systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3456–3470, 2007. 61
- [60] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," vol. 7, no. 4, pp. 1326 –1337, April 2008. 64
- [61] J. Wei and X. Zhang, "Energy-efficient distributed spectrum sensing for wireless cognitive radio networks," in *INFOCOM IEEE Conference on Computer Communications Workshops*, 2010, march 2010, pp. 1–6. 64
- [62] S. Maleki, A. Pandharipande, and G. Leus, "Energy-efficient distributed spectrum sensing for cognitive sensor networks," *Sensors Journal, IEEE*, vol. 11, no. 3, pp. 565 –573, march 2011. 64
- [63] Q. Zhao and J. Ye, "When to quit for a new job: Quickest detection of spectrum opportunities in multiple channels," in *Military Communications Conference*, 2008. MILCOM 2008. IEEE, nov. 2008, pp. 1–6. 64

- [64] G. Noh, H. Wang, J. Jo, B.-H. Kim, and D. Hong, "Reporting order control for fast primary detection in cooperative spectrum sensing," *Vehicular Technology*, *IEEE Transactions on*, vol. 60, no. 8, pp. 4058 –4063, oct. 2011. 64
- [65] J. K. Jithin and V. Sharma, "A novel algorithm for cooperative distributed sequential spectrum sensing in cognitive radio," in WCNC'11, march 2011, pp. 1–5. 64
- [66] H. Li, H. Dai, and C. Li, "Collaborative quickest spectrum sensing via random broadcast in cognitive radio systems," Wireless Communications, IEEE Transactions on, vol. 9, no. 7, pp. 2338 –2348, july 2010. 65
- [67] R. von Mises, Mathematical theory of probability and statistics. New York: Academic Press, 1964. 73
- [68] O. Filio, V. Kontorovich, and S. Primak, "Characteristics of sequencial detection in cognitive radio networks," in *Proc. ICACT 2011*, Phoenix Park, Korea, February 13-16 2011. 80, 82, 84
- [69] M. Guerriero, V. Pozdnyakov, J. Glaz, and P. Willett, "A repeated significance test with applications to sequential detection in sensor networks," vol. 58, no. 7, pp. 3426 –3435, July 2010. 81
- [70] Z. Chair and P. Varshney, "Optimal data fusion in multiple sensor detection systems," vol. 22, no. 1, pp. 98 –101, 1986. 81, 82, 83, 85, 86, 197
- [71] B. Jørgensen, Statistical Properties of the Generalized Inverse Gaussian Distribution. Lecture Notes in Statistics. Vol. 9. New York, Berlin: Springer-Verlag, 1982. 82
- [72] N. Gosan, L. Jemin, W. Hano, K. Sungtae, C. Sooyong, and H. Daesik, "Throughput analysis and optimization of sensing-based cognitive radio systems with markovian traffic," vol. 59, no. 8, pp. 4163–4169, 2010. 86
- [73] O. Filio, S. Primak, and V. Kontorovich, Recent Advances in Wireless Communications and Networks, J.-C. Lin, Ed. InTech, 2011. 87, 89

- [74] S. Primak, V. Kontorovich, and V. Lyandres, Stochastic methods and their applications to Communications: SDE Approach. Chichester: John Wiley & Sons, 2004. 90, 91
- [75] M. Abramowitz and I. Stegun, Eds., Handbook of Mathematical Functions. New York: Dover, 1965. 90, 142
- [76] Q. Zou, S. Zheng, and A. Sayed, "Cooperative sensing via sequential detection," vol. 58, no. 12, pp. 6266 –6283, 2010. 93, 97
- [77] H. Chenhui, W. Xinbing, Y. Zichao, Z. Jianfeng, X. Youyun, and G. Xinbo, "A geometry study on the capacity of wireless networks via percolation," vol. 58, no. 10, pp. 2916 – 2925, 2010. 97
- [78] S.-C. Lin, C.-P. Lee, and H.-J. Su, "Cognitive Radio with Partial Channel State Information at the Transmitter," vol. 9, no. 11, pp. 3402–3413, 2010. 97
- [79] H. Li and Z. Han, "Dogfight in spectrum: Combating primary user emulation attacks in cognitive radio systems, part i: Known channel statistics," vol. 9, no. 11, pp. 3566 – 3577, 2010. 97
- [80] —, "Catch me if you can: An abnormality detection approach for collaborative spectrum sensing in cognitive radio networks," vol. 9, no. 11, pp. 3554 – 3565, 2010. 97
- [81] P. Mitran, L. B. Le, and C. Rosenberg, "Queue-aware resource allocation for downlink ofdma cognitive radio networks," vol. 9, no. 10, pp. 3100 – 3111, 2010. 97
- [82] Z. Rui, "On active learning and supervised transmission of spectrum sharing based cognitive radios by exploiting hidden primary radio feedback," vol. 58, no. 10, pp. 2960 – 2970, 2010. 97
- [83] X. Shengli, L. Yi, Z. Yan, and Y. Rong, "A parallel cooperative spectrum sensing in cognitive radio networks," vol. 59, no. 8, pp. 4079 – 4092, 2010. 97

- [84] M.-G. Song, D. Kim, and G.-H. Im, "Recursive channel estimation method for ofdm-based cooperative systems," vol. 14, no. 11, pp. 1029 – 1031, 2010. 97
- [85] H. Sun, D. Laurenson, and C.-X. Wang, "Computationally tractable model of energy detection performance over slow fading channels," vol. 14, no. 10, pp. 924 –926, October 2010. 97
- [86] S. Almalfouh and G. Stuber, "Interference-aware power allocation in cognitive radio networks with imperfect spectrum sensing," in *Proc. of ICC 2010*, May 2010, pp. 1–6. 97
- [87] P. K. Tang and C. Han, "On the modeling and performance of three opportunistic spectrum access schemes," vol. 59, no. 8, pp. 4070 – 4078, 2010. 97
- [88] R. Wang and T. Meixia, "Blind spectrum sensing by information theoretic criteria for cognitive radios," vol. 59, no. 8, pp. 3806 – 3817, 2010. 97
- [89] Z. M. Wenyi, "Spectrum shaping: a new perspective on cognitive radio-part i: coexistence with coded legacy transmission," vol. 58, no. 6, pp. 1857 – 1867, 2010. 97
- [90] Z. Wenyi, "Spectrum shaping: A new perspective on cognitive radioópart ii: Coexistence with uncoded legacy transmission," vol. 58, no. 10, pp. 2971 – 2983, 2010. 97
- [91] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," vol. 20, no. 3, pp. 550–560, March 2002. 99, 111
- [92] S. J. Haghighi, S. Primak, V. Kontorovich, and E. Sejdic, Mobile and Wireless Communications Physical layer development and implementatiom. InTech Publishing, 2010, ch. Wireless Communications and Multitaper Analysis: Applications to Channel Modelling and Estimation. 99, 105, 109

- [93] Y. Do-Sik and W. Stark, "Characterization of wssus channels: normalized mean square covariance and diversity combining," vol. 4, no. 4, pp. 1307 – 1310, July 2005. 99
- [94] D. Thomson, "Spectral estimation and harmonic analysis," vol. 70, no. 9, pp. 1055–1096, September 1982. 101
- [95] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V-The discrete case," *Bell System Technical Journal*, vol. 57, pp. 1371– 1430, 1978. 109
- [96] S. Kotz and J. W. Adams, "Distribution of sum of identically distributed exponentially correlated gamma-variables," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 277–283, March 1964. 110
- [97] A. Hjørungnes, Z. Han, D. Niyato, W. Saad, and T. Basar, Game Theory in Wireless and Communication Networks. Cambridge University Press, 2012. 116, 128
- [98] J. F. Nash, "Equilibrium points in n-person games," Proceedings of the National Academy of Sciences of the United States of America, vol. 36, no. 1, pp. 48–49, January 1950. [Online]. Available: http://www.jstor.org/pss/88031 119, 122
- [99] —, "Non-cooperative games," Annals of Mathematics, vol. 54, no. 2, pp. 286–295, 1951. 119
- [100] R. Selten, "Spieltheoretische behandlung eines oligopolmodells mit nachfrageträgheit," Zeitschrift für die gesamte Staatswissenschaft, vol. 121, pp. 301 – 324, 1965 1965. 119
- [101] J. C. Harsanyi, "Games with incomplete information played by bayesian players, i-iii," *Manage. Sci.*, vol. 50, no. 12 Supplement, pp. 1804–1817, dec 2004. [Online]. Available: http://dx.doi.org/10.1287/mnsc.1040.0270 119

- [102] R. B. Myerson, "Large poisson games," Journal of Economic Theory, vol. 94, no. 1, pp. 7–45, September 2000. [Online]. Available: http: //ideas.repec.org/a/eee/jetheo/v94y2000i1p7-45.html 119, 130
- [103] R. Selten, "A simple model of imperfect competition, where 4 are few and 6 are many," Bielefeld University, Center for Mathematical Economics, Working Papers 008, Feb 1973. [Online]. Available: http: //ideas.repec.org/p/bie/wpaper/008.html 119
- [104] D. Fudenberg and J. Tirole, Game Theory. Cambridge, MA: MIT Press, 1991. 122, 134, 136
- [105] "Ieee 802.11 wireless lan: Capacity analysis and protocol enhancement," 1998.125
- [106] F. Cali, M. Conti, and E. Gregori, "Dynamic tuning of the ieee 802.11 protocol to achieve a theoretical throughput limit," *Networking, IEEE/ACM Transactions on*, vol. 8, no. 6, pp. 785–799, dec 2000. 125, 129
- [107] B. Wang, Y. Wu, and K. R. Liu, "Game theory for cognitive radio networks: An overview," *Comput. Netw.*, vol. 54, no. 14, pp. 2537–2561, Oct. 2010.
 [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2010.04.004 128, 132
- [108] X. Feng, H. Wang, and X. Wang, "A game approach for cooperative spectrum sharing in cognitive radio networks," Wireless Communications and Mobile Computing, pp. n/a–n/a, 2013. [Online]. Available: http: //dx.doi.org/10.1002/wcm.2364 128
- [109] O. Gharehshiran, A. Attar, and V. Krishnamurthy, "Collaborative sub-channel allocation in cognitive lte femto-cells: A cooperative game-theoretic approach," *Communications, IEEE Transactions on*, vol. 61, no. 1, pp. 325–334, 2013. 128
- [110] Y. Zhang, D. Niyato, P. Wang, and E. Hossain, "Auction-based resource allocation in cognitive radio systems," *Communications Magazine*, *IEEE*, vol. 50, no. 11, pp. 108–120, 2012. 128

- [111] G. Kasbekar and S. Sarkar, "Spectrum pricing games with random valuations of secondary users," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 11, pp. 2262–2273, 2012. 128
- [112] H. Inaltekin and S. Wicker, "The analysis of nash equilibria of the one-shot random-access game for wireless networks and the behavior of selfish nodes," *Networking, IEEE/ACM Transactions on*, vol. 16, no. 5, pp. 1094–1107, 2008. 128
- [113] V. Srivastava, J. Neel, A. MacKenzie, R. Menon, L. Dasilva, J. Hicks, J. Reed, and R. Gilles, "Using game theory to analyze wireless ad hoc networks," *Communications Surveys Tutorials, IEEE*, vol. 7, no. 4, pp. 46–56, 2005. 128
- [114] A. MacKenzie and L. DaSilva, Game Theory for Wireless Engineers. USA: Morgan & Claypool Publishers, 2006. 129, 144
- [115] Y. Zhang and M. Guizani, Game Theory for Wireless Communications and Networking, ser. Wireless Networks and Mobile Communications. Taylor and Francis, 2011. [Online]. Available: http://books.google.ca/books?id= YKWtPwAACAAJ 129
- [116] A. MacKenzie and S. Wicker, "Stability of multipacket slotted aloha with selfish users and perfect information," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 3, march-3 april 2003, pp. 1583 – 1590 vol.3. 129
- [117] Y. Xiao, X. Shan, and Y. Ren, "Game theory models for IEEE 802.11 DCF in wireless ad hoc networks," *Communications Magazine*, *IEEE*, vol. 43, no. 3, pp. S22 – S26, march 2005. 129, 139
- [118] B.-J. Kwak, N.-O. Song, and L. Miller, "Performance analysis of exponential backoff," *Networking*, *IEEE/ACM Transactions on*, vol. 13, no. 2, pp. 343 – 355, april 2005. 129, 147, 149

- [119] P. X. Zheng, Y. J. Zhang, and S. C. Liew, "Analysis of exponential backoff with multipacket reception in wireless networks," in *Local Computer Networks*, *Proceedings 2006 31st IEEE Conference on*, nov. 2006, pp. 855–862. 129, 147
- [120] S. Ghez, S. Verdú, and S. Schwartz, "Stability properties of slotted aloha with multipacket reception capability," *Automatic Control, IEEE Transactions on*, vol. 33, no. 7, pp. 640–649, 1988. 129
- [121] M. Bloem, T. Alpcan, and T. Başar, "A stackelberg game for power control and channel allocation in cognitive radio networks," in *Proceedings* of the 2nd international conference on Performance evaluation methodologies and tools, ser. ValueTools '07. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2007, pp. 4:1–4:9. [Online]. Available: http://dl.acm.org/ citation.cfm?id=1345263.1345268 129
- [122] D. Niyato and E. Hossain, "Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 1, pp. 192–202, 2008. 129
- [123] T. Cui, L. Chen, and S. Low, "A game-theoretic framework for medium access control," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 7, pp. 1116–1127, 2008. 129
- [124] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," vol. 18, no. 3, pp. 535 –547, March 2000. 129, 149
- [125] D. Wang, C. Comaniciu, and U. Tureli, "Cooperation and fairness for slotted aloha," Wirel. Pers. Commun., vol. 43, no. 1, pp. 13–27, Oct. 2007. [Online]. Available: http://dx.doi.org/10.1007/s11277-006-9240-5 129
- [126] Y. H. Bae, "Analysis of optimal random access for broadcasting with deadline in cognitive radio networks," *Communications Letters, IEEE*, vol. 17, no. 3, pp. 573–575, 2013. 129

- [127] J. Park and M. Van Der Schaar, "Cognitive MAC protocols using memory for distributed spectrum sharing under limited spectrum sensing," vol. 59, no. 9, pp. 2627–2637, 2011. 129
- [128] S. Choi, J. del Prado, S. S. N, and S. Mangold, "IEEE 802.11 e contentionbased channel access (EDCF) performance evaluation," in *Communications*, 2003. ICC '03. IEEE International Conference on, vol. 2, may 2003, pp. 1151 -1156 vol.2. 129
- [129] I. Milchtaich, "Random-player games," Games and Economic Behavior, vol. 47, no. 2, pp. 353–388, May 2004. [Online]. Available: http: //ideas.repec.org/a/eee/gamebe/v47y2004i2p353-388.html 129, 130
- [130] R. B. Myerson, "Population uncertainty and Poisson games," International Journal of Game Theory, vol. 27, no. 3, pp. 375–392, 1998. [Online]. Available: http://ideas.repec.org/a/spr/jogath/v27y1998i3p375-392.html 130, 131
- [131] S. Koskie and Z. Gajic, "A nash game algorithm for sir-based power control in 3g wireless cdma networks," *Networking*, *IEEE/ACM Transactions on*, vol. 13, no. 5, pp. 1017 – 1026, oct. 2005. 130
- [132] F. D. Sionopoli and C. G. Pimienta, "Undominated (and) perfect equilibria in Poisson games," Universidad Carlos III, Departamento de Economia, Economics Working Papers we073117, Apr. 2007. [Online]. Available: http://ideas.repec.org/p/cte/werepe/we073117.html 131
- [133] Y. Xiao, J. Park, and M. Van der Schaar, "Repeated games with intervention: Theory and applications in communications," *Communications, IEEE Transactions on*, vol. 60, no. 10, pp. 3123–3132, 2012. 136
- [134] D. S. Chan, T. Berger, and L. Tong, "Carrier sense multiple access communications on multipacket reception channels: Theory and applications to IEEE 802.11 wireless networks," vol. 61, no. 1, pp. 266 –278, January 2013. 138
- [135] M. Xiao, N. B. Shroff, and E. K. P. Chong, "A utility-based powercontrol scheme in wireless cellular systems," *IEEE/ACM Trans. Netw.*,

vol. 11, no. 2, pp. 210–221, Apr. 2003. [Online]. Available: http: //dx.doi.org/10.1109/TNET.2003.810314 138

- [136] M. Mitchell, An Introduction to Genetic Algorithms. Cambridge, MA, USA: MIT Press, 1998. 154
- [137] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," vol. 43, no. 6, pp. 1986–1992, November 1997. 157
- [138] L.N.Kanal and A.R.K.Sastry, "Models for channels with memory and their applications to error control," *Proceedings of the IEEE*, vol. 66, no. 7, pp. 724– 744, July 1978. 158
- [139] S. Primak, "On a dynamic model of cognitive radio systems," in Communications and Information Technology (ICCIT), 2012 International Conference on, June 2012, pp. 144 –148. 158
- [140] Y. Goldstein and V. Kontorovich, "Noisy immunity of weighted discrete addition of diversity signals," *Telecommunications and Radio Engineering*, pp. 104–106, 1972. 189, 193
- [141] D. da Costa, M. Yacoub, and J. Filho, "An improved closed-form approximation to the sum of arbitrary nakagami- variates," *Vehicular Technology, IEEE Transactions on*, vol. 57, no. 6, pp. 3854–3858, 2008. 191
- [142] Z. Khan, J. Lehtomaki, K. Umebayashi, and J. Vartiainen, "On the selection of the best detection performance sensors for cognitive radio networks," vol. 17, no. 4, pp. 359 –362, April 2010. 193, 194

Appendix A Collaborative Spectrum Sensing with Censoring

It is understood that spectrum sensing is based on the cyclostationary properties of the signals from the PU and that the SUs are spatially distributed within a certain area. All the SUs can sense the whole frequency band of interest or each SU may sense just a partial band. Hereafter it will be assumed that all SU are sensing the same frequency band. In both cases of spectrum monitoring, the SUs must share the sensing information between them or it can be coordinated by a Fusion Centre (FC). It seems reasonable that, no matter what kind of exchange information is used, the local decision information must be obtained by a minimum set of observations Min eq. (2.29) having N and P fixed¹. In other words, the time of analysis must be reduced as much as possible; it is then opportunistic to apply the sequential analysis of A. Wald [29] where the ML test, in contrary to NPT, must be compared with two thresholds related to the requirements of P_{fa} and P_{md} . Let us suppose that for the latter, highly reliable final results for the test are predefined, so P_{fa} and P_{md} must be rather low. Each SU will obtain those reliable final results at different time instants. This information must be sent to other fellow SUs or to a FC in a binary way. Let us consider, several rather general but different scenarios of collaborative spectrum sensing:

• Each of the *n*-th (n = 1...K) SUs sends, after time T, the information (not

^{1.} The asymptotic conditions for M(T) are assumed to be valid here in order to preserve the uncorrelated conditions for F-coefficients.

binary) of ξ_n to the system. The quadratic combining at the FC is then

$$\Xi = \sum_{n=1}^{K} \xi_n^2. \tag{A.1}$$

 Ξ can then be analyzed by the NPT, assuming that the channel from the SUs and the FC is error free. Hence, after the final addition, the result of the quadratic diversity combining of KQ virtual branches (or of K SU) is analyzed, assuming statistically independent fading along all summations (see eqs. (2.36) and (2.37)). This scenario can be called a distributed optimum incoherent *SIMO passive radar*, and its characteristics are equal to eqs. (2.42) and (2.43) with the number of virtual branches equal to KQ.

• Each of the *n*-th SUs makes an individual decision regarding the presence of the PU and then sends the binary decision to the FC through an error free channel. Assuming that all those decisions are statistically independent, the final result at the FC can be obtained according to the majority rule with the majority not-weighted (or weighted) diversity addition method. This scenario can be also called *SIMO* radar but contrary to the first one, it is non-optimum.

If the majority principle is applied at the FC, then the decision is made by analysis of the partial decisions at each SU (here the SU acts as a *virtual* diversity branch), and the decision which takes place at the majority of the branches is favoured. This method is called *majority diversity combining* [30]. If partial solutions are binary and the number of virtual branches is odd, there cannot be any confusion in the final decisions. Let K = 2q - 1 and P_0 denote the event of the existence of the PU after qtests on the branches. If after m - 1 probes on the virtual branches, one gets q - 1results of the existence of the PU and the *m*-th probe gives the same, then for the qtest one gets the probability of this event as

$$Prob(P_0) = \binom{m-1}{q-1} P_1^q (1-P_1)^{m-q} \qquad m \ge q,$$
(A.2)

where P_1 can be P_{fa} or P_{md} . $Prob(P_0)$ is therefore a final probability of false alarm or miss detection [11, 50], depending on which hypothesis is considered. From the theory of diversity combining, it is also known that majority addition is equivalent to the optimal incoherent addition with the number of branches (here virtual ones) q, *i.e.*, to incoherent (quadratic) addition with almost half as many branches. By comparing the characteristics of majority addition with those of the optimum SIMO radar, one can see several limitation of the former:

- Optimum SIMO radar with incoherent addition actually operates with almost twice as many virtual branches and therefore provides significantly better detection characteristics (ROC's).
- The majority addition operates successfully only with odd number of virtual branches, while optimum SIMO radar operates with any number of branches.

The price one has to pay for the advantages of the optimum SIMO radar is a more complex data transmission scheme: in the majority addition, simply binary results are transmitted, whereas for SIMO radar the information of the ξ value for each SU must be transmitted to FC through error free channels. In order for the noise immunity properties of the majority addition to approach those of the optimum incoherent addition, some modifications of the former have been proposed. The so-called *weighted* majority addition is proposed in [140]. In this method, the channel gains for each partial solution are introduced as *weights* in the majority addition algorithm. In this way the channel gains for the diversity branches work as weighting coefficients in the process of majority selection. It was shown that this suboptimal method provides very close results to those of the optimum incoherent addition [140] if the communication scenario allows taking advantages of channel gains. It is not the case for one of the scenarios at the FC: the results of detection at the SU were obtained through the optimum quadratic addition by the SU itself, and the resulting fading at the SU has a very low variance when Q is rather large (a hardening effect) [52]. Therefore, it is difficult to improve the results of majority addition by introducing weighting coefficients, as all the weights might be practically equal. However it is known that if the channels are sufficiently heterogeneous, the hardening effect does not appear, or it appears very slowly when $Q \to \infty$ at the SU. Let us consider another extreme special case. Let us assume that the fading at the SUs are so heterogeneous that almost no quadrature addition algorithms work as the diversity combining algorithm,

and each SU has $Q \approx 1$ (single reception) with an *m*-distributed fading² (generally heterogenous). With this assumption, one can see that the problem is covered to the case of SIMO radar: the SUs send the binary information of their weights in order to provide the FC with weighted addition. Let us formulate here an assumption: if the final decisions are taken at the FC by applying the technique of weighted addition of partial decisions, then the SUs must transmit to the FC not only the information of partial decisions, but information of their reliability as well. The whole system (PU, SU, and FC) works as a distributed quasi-optimum SIMO radar. In the previous scenario, the decision of the PU existence in the majority of branches can be obtained by the algorithm:

$$\sum_{j=q+1}^{K} \mu_j \bigg|_{\mathcal{H}_0} > \sum_{j=1}^{q} \mu_j \bigg|_{\mathcal{H}_1}, \qquad (A.3)$$

where $\{\mu_j\}_1^K$ are magnitudes of the channel gains. From Bayes theorem, each of the summands in eq. (A.3) have the following PDFs:

$$W(\mu_j | \mathcal{H}_0) = \frac{W(\mu_j) P_{fa} \big|_{\mu_j}}{P_{fa}},$$

$$W(\mu_j | \mathcal{H}_1) = \frac{W(\mu_j) P_{md} \big|_{\mu_j}}{P_{md}}.$$
(A.4)

Let us assume that [56]

$$W(\mu_j) = \frac{2m_j^{m_j} \mu_j^{2m_j - 1}}{\Gamma(m_j)(z_i^{-2})^{m_j}} \exp\left(-m_j \frac{\mu_j^2}{z_i^{-2}}\right),\tag{A.5}$$

where m_j and the corresponding parameters for the four-parametric distribution are related by $m_j \geq \frac{1}{2}$. Then introducing the new variable

$$x_j = \mu_j \frac{\sqrt{m_j + \frac{\gamma^2}{2}}}{m_j},\tag{A.6}$$

^{2.} In relation to the fading model assumed here for simplicity, see eq. (2.28) for the definition of the parameter m through the parameters of the Generalized Gaussian model.

we obtain

$$W\left(x_{j}; \mathcal{H}_{0}\right) = \frac{2m_{j}^{m_{j}}x_{j}^{2m_{j}-1}}{\Gamma(m_{j})(z_{i}^{-2})^{m_{j}}} \exp\left(-m_{j}\frac{x_{j}^{2}}{(z_{i}^{-2})^{m_{j}}}\right),$$

$$W\left(x_{j}; \mathcal{H}_{1}\right) = \frac{2m_{j}^{m_{j}}x_{j}^{2m_{j}-1}}{\Gamma(m_{j})\left(\frac{\gamma^{2}}{2m_{j}}z_{i}^{-2}\right)^{m_{j}}} \exp\left(-m_{j}\frac{x_{j}^{2}}{\left(\frac{\gamma^{2}}{2m_{j}}z_{i}^{-2}\right)^{m_{j}}}\right),$$
(A.7)

for $P_{fa} \ll 1$ and $P_{md} \ll 1$. It is clear from eq. (A.7) that $W(x_j; \mathcal{H}_0)$ and $W(x_j; \mathcal{H}_1)$ have the Nakagami PDF form. Recalling eq. (A.3) and introducing the following variables

$$\zeta_q = \sum_{j=1}^q x_j \Big|_{\mathcal{H}_0},$$

$$\eta_{Q-q-1} = \sum_{j=q+1}^Q x_j \Big|_{\mathcal{H}_1},$$
(A.8)

one can formally calculate the error probability in eq. (A.3). In the general case of the heterogenous scenarios according to [141] it is possible to find distributions of ζ_q and η_{Q-q-1} in a Nakagami PDF form after rather complicated calculation. In the most straightforward manner, following equations (77)-(89) in [56], it is possible to provide the error analysis for the following special case when

$$\frac{z_1^{-2}}{m_1} = \frac{z_2^{-2}}{m_2} = \dots = \frac{z_n^{-2}}{m_n},$$

with *n* denoting the number of Nakagami variables at eq. (A.8). Thus, the sums in eq. (A.8) will have an equivalent Nakagami parameters $\hat{m} \cong mn$ and $\hat{z}^{-2} \cong nz^{-2}$. Notice that for the general case, the calculus of \hat{m} and \hat{z}^{-2} can be done mainly numerically. From eq. (A.3) it is possible to get the conditional error probability (with q fixed):

$$P(q) = \frac{q^{\hat{m}(K-q)}\Gamma(\hat{m}K)}{\Gamma(\hat{m}(K-q)+1)\Gamma(\hat{m}K)\left[\frac{\gamma^2(K-q)}{2\hat{m}}\right]^{\hat{m}(K-q)}}.$$
 (A.9)



Figure A.1: Probability of miss detection for the WMA. Eqs. (2.43) and (2.48) are shown with a solid line whereas eq. (A.10) is shown in a dotted line.



Figure A.2: Probability of miss detection for WMA with different parameters values.

The number of virtual branches q with errors, both for P_{md} and P_{fa} is a random variable with Bernoulli PDF when the virtual branches have statistically independent



Figure A.3: Probability of miss detection for WMA with different parameters values.

fading [140]. It follows that

$$P_{error} = \sum_{q=1}^{K-1} P_K(q) P(q) + P_K(K), \qquad (A.10)$$

where $P_K(q) = {q \choose K} P_1^q (1 - P_1)^{K-q}$ and P_1 stands for P_{fa} or P_{md} from eqs. (2.42) and (2.43) respectively, when Q = 1. Notice that at eqs. (2.42) and (2.43), the four parameters must be modified from eq. (2.28) with the value of \hat{m} and \hat{z}^{-2} at γ^{-2} . Equation (A.10) is universal in the sense that the final P_{md} and P_{fa} can be calculated through it because the inequality in the eq. (A.3) type can be applied for calculus of false alarm as well. The ROCs for the WMA is shown in Figures A.1-A.3. For $P_{md} = 10^{-4}$, the energetic losses are less than 1.5-2 dB.

Finally, let us compare the technique majority addition with some of the approaches mentioned in [142]. The *simple counting* approach is simply selecting for the FC decision only *highly weighted* SUs. This addition is *less optimum* than the approach in [140] because some of the SUs with small weights do not participate in the decision making process at the FC. Two other methods *Partial Agreement Counting* and *Collision Detection*, assume the existence of a feedback channel between SUs and FC, which can be used for comparing partial decisions at the SU and final decisions at

the FC in order to select the "true" final decision. The collision detection method is not considered in the current analysis. Indeed, the application of the feedback channel opens the possibility of improving the reliability of the final decision at the FC. Taking into account that weighted majority addition is practically optimum incoherent addition, the final characteristics might be better than those obtained in [142].

Appendix B Performance Derivation of Data Fusion Rule

Let us introduce the following notations

$$a_{i} = \begin{cases} \ln \frac{1 - P_{MD}}{P_{FA}} & \text{if } u_{i} > 0\\ \ln \frac{1 - P_{FA}}{P_{MD}} = -\ln \frac{P_{MD}}{1 - P_{FA}} & \text{if } u_{i} < 0 \end{cases},$$
(B.1)

and

$$\xi_{i} = \begin{cases} a_{i} = \ln \frac{1 - P_{MD}}{P_{FA}} & \text{if } u_{i} > 0\\ b_{i} = -a_{i} = -\ln \frac{1 - P_{FA}}{P_{MD}} & \text{if } u_{i} < 0 \end{cases}$$
(B.2)

Consider a \mathcal{T} (test statistic) given by eq. (3.60)

$$\mathcal{T} = a_0 + \sum_{k=1}^{K} a_k u_k = a_0 + \sum_{k=1}^{K} \xi_k |u_k| = a_0 + \sum_{k=1}^{K} \xi_k.$$
 (B.3)

Here ξ_k could be considered as a random variable with PDF

$$P_{\xi}(x) = P_{+}\delta(x-a) + P_{-}\delta x - b$$

= $P\delta(x-a) + (1-P)\delta(x-b),$ (B.4)

where

$$a = a_i = \ln \frac{1 - P_{MD}}{P_{FA}}$$

$$b = -a_i = -\ln \frac{1 - P_{FA}}{P_{MD}}.$$
(B.5)

and P_+ is probability of u = +1 decision, equal to

$$P_{+} = p(\mathcal{H}_{1})(1 - P_{MD}) + p(\mathcal{H}_{0})P_{FA}$$

= $p(\mathcal{H}_{1})(1 - P_{MD}) + [1 - p(\mathcal{H}_{1})]P_{FA}$, (B.6)

The corresponding characteristic function of ξ is then given by

$$\Theta_{\xi}(s) = P_{+}e^{-sa} + P_{-}e^{-sb}, \tag{B.7}$$

and the characteristic function of \mathcal{T} could be evaluated as

$$\Theta_{\mathcal{T}} = \Theta_{\xi}^{K} e^{-sa_{0}} = \left[P_{+}e^{-sa} + (1 - P_{+})e^{-sb} \right]^{K} e^{-sa_{0}}$$
$$= \sum_{k=0}^{K} {K \choose k} P_{+}^{k} (1 - P_{+})^{K-k} e^{-s[ka + (K-k)b + a_{0}]}$$
(B.8)

Equivalently, the PDF is given by

$$P_{\mathcal{T}}(x) = \sum_{k=0}^{K} {\binom{K}{k}} P_{+}^{k} (1 - P_{+})^{K-k} \delta[x - (ka + (K-k)b + a_{0})].$$
(B.9)

If k = 0 then

$$ka + (K - k)b - a_0 = Kb + a_0$$

= $-K \ln \frac{1 - P_{FA}}{P_{MD}} + \ln \frac{P(\mathcal{H}_1)}{1 - P(\mathcal{H}_1)}.$ (B.10)

If $P(\mathcal{H}_1) \approx 1$ such that

$$\frac{P(\mathcal{H}_1)}{1 - P(\mathcal{H}_1)} > \left(\frac{1 - P_{FA}}{P_{MD}}\right)^K,\tag{B.11}$$

then the FC makes only \mathcal{H}_1 decisions i.e.

$$P_{MD} = 0,$$

$$P_{FA} = P(\mathcal{H}_0) = 1 - p(\mathcal{H}_1).$$
(B.12)

If (B.11) is not satisfied then there is $k_{\text{max}} > 0$ such that

$$k_{\max}a + (K - k_{\max})b + a_0 < 0,$$
 (B.13)

and

$$(k_{\max} + 1)a + (K - k_{\max} - 1)b + a_0 > 0.$$
(B.14)

In this case the scheme suggested in [70] is equivalent to $(k_{\max} + 1)$ out of K scheme (this is assuming that are statistically equivalent). Let \mathcal{H}_1 be true. Then the target is missed if there are no more than k_{\max} positive decisions, or, equivalently, no less than $K - k_{\max}$ negative decisions. The probability of miss detection at FC is then given by

$$P_{MD_F} = \sum_{k=0}^{k_{\max}} {\binom{K}{k}} P_{MD}^k (1 - P_{MD})^{K-k}.$$
 (B.15)

To more decisions \mathcal{H}_1 there should be at least $k_{\max} + 1$ partial 1. If \mathcal{H}_0 is true, the probability of false alarm at the fusion center is then:

$$\mathcal{H}_{1} : P_{MD_{F}} = \sum_{k=0}^{k_{\max}} {\binom{K}{k}} (1 - P_{MD})^{k} P_{MD}^{K-k}$$

$$\mathcal{H}_{0} : P_{FA_{F}} = \sum_{k=k_{\max}}^{K} {\binom{K}{k}} (P_{FA})^{k} (1 - P_{FA})^{K-k}$$
(B.16)

Appendix C Curriculum Vitae

OSCAR GUILLERMO FILIO RODRÍGUEZ

Profile

- Excellent ability to relay concepts to others through extensive tutoring and teaching experience
- Strong communication skills developed through teaching and conference presentations
- A proven and productive researcher who thrives working both independently and in a team environment
- A quick learner who is able to adapt to changing research environments and directions

Education

- Ph.D., 2013, In Progress—Communications Systems Engineering: The University of Western Ontario
 - Research Focus: Cognitive Radio Networks, 4G and LTE systems, Spectrum Sensing, Power allocation of Secondary Users, Optimization Techniques.
 - Thesis Work: Primary User Detection Using Multiple Antenna Receivers and Power Allocation Schemes for Cognitive Radio Systems
- Masters of Engineering Science, 2008—Electrical & Computer Engineering: Research and Advanced Studies Centre of the National Polytechnique Institute, Mexico City, Mexico
 - Research Focus: MIMO systems, Multi User Diversity Schemes, Scheduling Algorithms
 - Thesis Work: Multiuser Diversity in MIMO systems
- Bachelor of Engineering, 2005—Electrical Engineering: Professional Interdisciplinary Unit on Engineering and Advanced Technologies, Mexico City, Mexico
 - Areas of Concentration: Signal Processing, Control Systems and Power Electronics
- Guitar and Composition Studies, 2003—: College of Music Technology & Audio, Mexico City, Mexico
 - Focused on Application of Electronic and Computer Tools as well as Digital Signal Processing in Music Composition and Guitar Performances

Skills

- Languages and Tools:
 - Fluent in C, C++, Java, MATLAB script/Simulink, Maple, Mathematica, Graphmatica, HTML, ET_EX, SQL, MS/Open Office
 - Basic Experience with Verilog
- Operating Systems (Development Experience): Linux (Ubuntu and Debian), Mac Os X 10.x, Windows 98 through 7, Android 2.x, iOS 5.x

Training Courses

- Communication in the Canadian Classroom (2009)
- Teaching Assistant Training Program (2009)
- Teaching in the Canadian Classroom (2010)
• The Language of Teaching in Engineering (2012)

Relevant Experience

- National Polytechnique Institute, Mexico City, Mexico—Full Time Professor: Jan. 2008 Aug 2009
 - Key Experience: Knowledge Transfer to Others, Confidence of Speaking in front of Big Audiences
 - Courses: Communications I (Fourier Analysis and Modulation), Communications II (Digital Signal Processing, Filter Design), Information Theory, Telematic Systems (Networking).
- The University of Western Ontario—Teaching Assistant: Sept. 2010 April 2011
 - Key Experience: Knowledge Transfer to Others
 - Responsible for supervision and teaching of lab material on Introduction to Electrical Instrumentation and and Electrical Laboratory I & II.
- Bell Centre for Information Engineering—Researcher: Sept. 2009 present
 - Key Experience: The Research Lifecycle From Concept to Implementation to Dissemination
 - Extensive experience in phenomenon modelling, parameter extraction and optimization techniques
 - Supervising and advising other graduate students and researchers
- Centre of Research and Advanced Studies, Mexico City, Mexico—Researcher: Aug. 2007 Sept. 2009
 - Key Experience: Simulation of Wireless Communications Systems
 - Simulation of MIMO Communication Systems
 - Research of Noise Modelling with Chaos Theory
 - Supervising and advising other graduate students and researchers.
- INTEL corp. Guadalajara, Mexico-Intern Verilog Developer: August 2006 February. 2007
 - Key Experience: Implementation of Communication Algorithms in Real Time Devices and Adherence to Tight Deadlines
 - Worked on development of a MIMO system using an FPGA chip.

Referred Publications

- Journal Papers
 - [J1] O. Filio, V. Kontorovich, S. Primak and F. Ramos-Alarcón, Collaborative Spectrum Sensing for Cognitive Radio: Diversity Combining Approach. Wireless Sensor Network, vol.3 No.1, pp. 24-37, 2011.
 - [J2] O. Filio, V. Kontorovich, and F. Ramos-Alarcón, *Time-Energy Limited Feedback Algorithms for Multiuser Diversity Scenarios on Nakagami Fading Channels*. IEEE Transactions on Wireless Communications, vol.60 No.4, pp.1045-1057, April 2012.
 - [J3] O. Filio, S. Primak, V. Kontorovich, Characteristics of Sequential Detection in Cognitive Radio Networks Journal of Communication and Computer, vol.9 No.5, pp. 587-595, May 2012.
- Books
 - [B1] O. Filio, V. Kontorovich, S. Primak, Recent Advances in Wireless Communications and Networks. InTech, 2011.

[B2] **Technical reviewer of the spanish version**, *Principios de Transferencia de Calor (Principles of Heat Transfer)*, Seventh Edition, CENGAGE Learning, 2012

• Conference Papers

- [C1] O. Filio, V. Kontorovich, and F. Ramos-Alarcón, "A New Approach for Selective Multi-User Diversity in Nakagami Fading Channels," in Proc of IEEE Wireless Communications and Mobile Computing Conference, 2008. IWCMC '08. International, 2008, pages 452-457.
- [C2] O. Filio, V. Kontorovich, and F. Ramos-Alarcón, "A novel Selective Multiuser Diversity algorithm for Nakagami fading channels," in Proc of 25th Biennial Symposium on Communications (QBSC), 2010, pages 106-109.
- [C3] V. Kontorovich, F. Ramos-Alarcón ,O. Filio, S. Primak, "Cyclostationary spectrum sensing for Cognitive Radio and multiantenna systems," in Proc of Wireless Communications and Signal Processing (WCSP), 2010, pages 1-6.
- [C4] O. Filio, S. Primak and V. Kontorovich, "On performance of the Wald test in partially coherent Cognitive Radio channels," in Proc of Communications and Information Technology (ICCIT), 2011, pages 214-217.
- [C5] O. Filio, S. Primak and V. Kontorovich, "Characteristics of sequential detection in Cognitive Radio Networks," in Proc of Advanced Communication Technology (ICACT), 2011, pages 307-312
- [C6] O. Filio, S. Primak and V. Kontorovich, "Power allocation for cognitive users applying OFDM under dynamic spectrum activity", in Proc of Devices, Circuits and Systems (ICCDCS), 2012, pages 1-4
- [C7] O. Filio, S. Primak and V. Kontorovich, "Impact of Scattering Environment on the Performance of Spectrum Sensing in Multiantenna Cognitive Radio Systems," in Proc of 26th Biennial Symposium on Communications (QBSC), 2012

Additional manuscripts are under review and under preparation

Awards and Scholarships

2005-2007 & 2010-2013
2011
2010 & 2011 & 2012
2013
2005 – present
2005 – present