Software Defined Radio & MIMO Signal Processing

Amor Nafkha

UNIVERSITÉ DE RENNES 1

HABILITATION À DIRIGER DES RECHERCHES

L'UNIVERSITÉ DE RENNES 1
COMUE UNIVERSITÉ BRETAGNE LOIRE

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Télécommunications-Électronique*

Titre:

Software Defined Radio & MIMO Signal Processing :

From Theoretical to Practical Results

Par:

Amor NAFKHA

HDR présentée et soutenue à Rennes, le 24/11/2021 à 10 :30
Unité de recherche : IETR UMR CNRS 6164

Rapporteurs avant soutenance :

Mme Marie-Laure BOUCHERET, Professeur des universités, INP-ENSEEIHT Toulouse, FRANCE.
M. Christophe JÉGO, Professeur des universités, INP/ENSEIRB-MATMECA Bordeaux, FRANCE.
M. Mehdi BENNIS, Full Professor, Directeur du groupe ICON, University of Oulu, FINLANDE.

Composition du Jury :

Président :	M. Guy GOGNIAT, Professeur des universités, Université de Bretagne-Sud, FRANCE.
Examinateurs :	M. Mérouane DEBBAH, Professor, Chief Researcher at TII, Émirats Arabes Unis.
	M. Christophe MOY, Professeur des universités, Université de Rennes 1, FRANCE.
	M. Yves LOUET, Professeur de CentraleSupélec, Campus de Rennes, FRANCE.

Invité(s) :
M. Faouzi Carlos Bader, Enseignant-Chercheur (HDR) de CentraleSupélec, Campus de Rennes, FRANCE.

Acknowledgement

BismiALLAH ar-Rahman ar-Raheem, In the name of God, the infinitely Compassionate and Merciful. All Praise and thanks goes to ALLAH, Lord of the Worlds, and may the peace and blessings be on his Prophets and Messenger Muhammad and on his family and all of his Companions.

I would like to thank the members of the jury: Pr. Guy Gogniat (University of Southern Brittany, Lorient), Pr. Mérouane Debbah (Chief Researcher at TII, United Arab Emirates), Pr. Christophe Moy (University of Rennes 1, Rennes), Pr. Yves Louet (CentraleSupélec, Rennes) and Pr. Faouzi Carlos Bader (CentraleSupélec, Rennes) – and the three reviewers – Pr. Marie-Laure Boucheret (INP - ENSEEIHT, Toulouse), Pr. Mehdi Bennis (University of Oulu, Finland), and Pr. Christophe Jégo (ENSEIRB-MATMECA, Bordeaux) – for kindly accepting my invitation, complying with all the HDR admin constraints and honouring me with your participation, which lead to a very interesting and rich scientific discussion.

My deep and sincere gratitude to my brothers and sisters for their continuous and unparalleled love, help and support. I am grateful for my parents (Mohammed and Daklia) whose constant love and support keep me motivated and confident. My accomplishments and success are because they believed in me.

Finally, I owe my deepest gratitude to my wife Boutheina and daugthers Sirine and Ryma. I am forever thankful for their unconditional love and support throughout the entire HDR process and every day.

Table of Contents

List of acronyms

ACO	Ant colony optimization
ADC	Analog to Digital Converter
AFE	Analog Front-End
AHSD	Average-based Hybrid spectrum Sensing Detector
ARM	Advanced RISC Machines
ASPC	Application-specific programmable circuits
ASRC	Arbitrary Sample Rate Conversion
BCH	Bose–Chaudhuri–Hocquenghem codes
BER	Bit Error Rate
CAF	Cyclic Auto-correlation Function
CFAP	Constant False Alarm Probability
CFD	Cyclostationary Feature Detector
CFR	Crest Factor Reduction
CLT	Central Limit Theorem
CMOS	Complementary Metal–Oxide–Semiconductor
CR	Cognitive Radio
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
DAC	Digital to Analog Converter
DDC	Digital Down Converter
DFE	Digital Front-End
DMA	Direct Memory Access
DPD	Digital Pre-Distortion
DPR	Dynamic Partial Reconfiguration
DSP	Digital Signal Processor
DTED	Double Thresholds Energy Detector
DUC	Digital Up Converter
EAHSD	Enhanced Average-based Hybrid spectrum Sensing Detector
EBD	Eigenvalue-Based Detector
ECC	Error-Correcting Code
ED	Energy Detector
EME	Energy with Minimum Eigenvalue
FA	Firefly Algorithm

FIR	Finite Impulse Response
FPGA	Field Programmable Gate Array
GALS	Globally Asynchronous Locally Synchronous
GDGI	Geometrical Diversification and Greedy Intensification
GEV	Generalized Extreme Value
GISD	Geometrical Intersection and Selection Detector
GPP	General Purpose Processor
GRC	GNURadio Companion
ICAP	Internal Configuration Access Port
IED	Ideal Energy Detector
IID	Independent and Identically Distributed
IoT	Internet of Things
L2E	List Exploration and Exploitation
LDPC	Low Density Parity Check
LGISD	List Geometrical Intersection and Selection Detector
LRT	Likelihood Ratio Test
MCF	Multi-Core Fiber
MIMO	Multiple-In Multiple-Out
ML	Maximum Likelihood
MME	Maximum-Minimum Eigenvalue
MMF	Multi-Mode Fiber
MMSE	Minimum Mean Square Error
OKOP	one Kernel One Peak
OMP	Orthogonal Matching Pursuit
OSHSD	Order Statistic based Hybrid spectrum Sensing Detector
PSO	Particle Swarm Optimization
PU	Primary User
RF	Radio Frequency
RO	Ring Oscillator
ROC	Receiver Operating Characteristic
SCN	Standard Condition Number
SD	Sphere Decoding
SDM	Space Division Multiplexing
SDR	Software Defined Radio
SED	Sequential Energy Detector
SINR	Signal-to-Interference-plus-Noise Ratio
SLE	Scaled Largest Eigenvalue
SM	Spatial multiplexing

SMF	Single Mode Fiber
SNR	Signal-to-Noise Ratio
SPCAF	Symmetry Property based Cyclic Autocorrelation Function
SRC	Sample Rate Conversion
TR	Tone Reservation
SU	Secondary User
UHD	USRP Hardware Driver
USRP	Universal Software Radio Peripheral
ZF	Zero-Forcing

List of Figures

List of Tables

SHORT INTRODUCTION

In the Beginning

This year marks 15 years after my PhD graduation and I feel that applying for the accreditation to supervise research (Habilitation à Diriger des Recherches, HDR) is probably the greatest achievement I could hope for. The main reasons that motivated me to apply for the HDR diploma, from University of Rennes 1, can be summarized as follows:

- In September 2020, I realized that it was time to synthesize my research findings and acquired supervisory experiences and to serve as a point of departure to my research activities in the near to mid-term future (3 - 5 years).
- After my PhD, I have worked on several research projects which have allowed me to grow and improve my theoretical and practical skills (typically in hardware design, signal processing, information theory, compressive sensing theory, random matrix theory,...), combine my passion for wireless communications and embedded systems together, and meet extraordinary researchers from all over the world. I have arrived at the point today where I have reached a degree of scientific maturity to develop my own researches in the future, and serve the scientific community as PhD/Postdoctoral full-supervisor and Phd/HDR examiner or reporter.
- As the highest university degree that can be awarded, the HDR is required to apply for the French qualification as university professor. The HDR is the cornerstone of my future academic and research careers.

Structure of the manuscript

The present HDR manuscript is divided into seven chapters as follows. As my research interest lies in the intersection between wireless communications and adaptive reconfigurable hardware, the first chapter is dedicated to give the reader a short overview over my main research activities of the last 15 years. Chapter 2 and Chapter 3 will respectively provide an extended version of Curriculum Vitae and a complete list of all my publications. Chapter 4 will be dedicated to present main results achieved in the context of MIMO channel capacity and MIMO detection techniques. It will incorporate an important theoretical result relating the ergodic capacity of MIMO Rayleigh fading channel (*i.e.* case of wireless communication) to that of MIMO Jacobi fading channel (*i.e* case of multicore and multimode fibers). The second

part of Chapter 4 will present extended studies concerning MIMO detection techniques using bio-inspired algorithms or geometrical approaches. Chapter 5 will address the spectrum sensing problem in cognitive radio systems through modeling it as a binary hypothesis testing task. Finite dimensional random matrix theory, moment matching method and compressive sensing theory are used to derive accurate detection thresholds and to propose new spectrum sensing techniques. Chapter 6 will address software defined radio platforms from three different aspects: flexible digital front-end design, virtual hardware through dynamic partial reconfiguration, and experimental validation of some spectrum sensing algorithms developed in Chapter 5 using Ettus USRP-based software defined radio platforms. In Chapter 7, I conclude this HDR thesis by discussing several (short-, mid- and long-term) perspectives and interesting open issues.

Notations

In the sequel of the present document, the sets of natural, real, and complex numbers are denoted by $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{C}$, respectively. Matrices, vectors and scalars are respectively denoted by boldface upper case symbols, bold-face lower case symbols, and italic lower case symbols. The transpose and Hermitian transpose of a vector $\mathbf{u}$ (respectively a matrix $\mathbf{U}$) are denoted by $\mathbf{u}^T$ and $\mathbf{u}^\dagger$ (respectively $\mathbf{U}^T$ and $\mathbf{U}^\dagger$). $\mathbb{E}_{\mathbf{x}}[.]$ stands for the mathematical expectation with respect to $\mathbf{x}$. Symbols $\det(.)$ and $\mathrm{Tr}(.)$ denote the determinant and the trace of a matrix, respectively. $vect[\mathbf{x}, \mathbf{y}]$ denotes the concatenation of the two vectors $\mathbf{x}$ and $\mathbf{y}$.

Chapter 1

The Big Picture of my Research Activities

Since the beginning of my PhD thesis, my research activities have mainly addressed wireless communication systems and their implementation on reconfigurable hardware devices, and it can be divided to four main topics as depicted in Fig. 1.1. Under the umbrella of other research topics, I mentioned all research studies where I have been involved but they are outside the scope of the present manuscript:

- MIMO signal processing (channel capacity and detection techniques).
- Spectrum sensing techniques for cognitive radio.
- Dynamic reconfiguration of FPGA-based SDR platforms.
- DFE signal processing: Arbitrary sample rate conversion in SDR platforms.

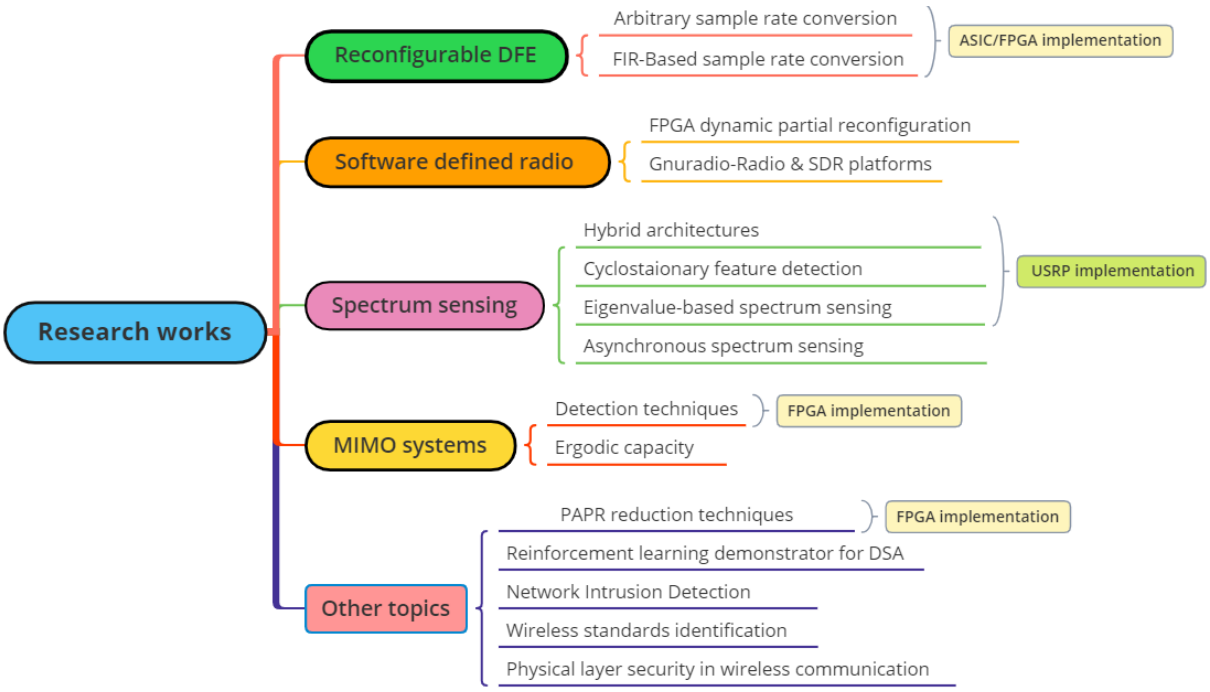


Figure 1.1 – Addressed research topics between 2003 and 2021.

In January 2003, I joined the Lab-STICC laboratory (formerly LESTER), CNRS UMR 6285, at South Brittany University (UBS) in Lorient as PhD student under the supervision of Prof. Emmanuel Boutillon. The main objective of my PhD was to propose a novel near maximum likelihood (ML) detection algorithm for the multiple-input multiple-output (MIMO) system. The ML detection is the process to find the nearest lattice point to a given vector in an N-dimensional search space. In general, the ML detection problem is NP-hard due to the discrete nature of solution space. I addressed the dual challenges of quasi-optimal ML detection performances and a low fixed computational complexity associated to parallel hardware structure. During my PhD period, I developed a new parallel low-complexity MIMO detection scheme, called geometrical diversification and greedy intensification (GDGI) technique, in order to achieve near optimal maximum likelihood detection performance [J2]. The proposed algorithm was based on diversification and intensification steps. The diversification step, also called *exploration*, is about discovering new promising initial feasible solutions for the ML problem, and ensures that the proposed algorithm explores the totality of search space. Based on the k strongest right singular eigen-vectors of the MIMO channel matrix and a geometrical interpretation of the objective function, the diversification phase provides a list of initial "smart" solutions. In contrast, the intensification step, also called *exploitation*, involves the refinement of the "smart" solutions by performing a 1-flip or 2-flip neighborhood local search algorithm. The proposed algorithm has three properties that make it very attractive for several practical wireless communication systems. First, it achieves near optimal maximum likelihood detection performance. Second, it has a constant polynomial-time computational complexity. Finally, the flexibility and parallel structure of the GDGI algorithm allow high-throughput and energy-efficient FPGA implementation.

From 2018 to 2021, research works on low-complexity MIMO detection algorithms were continued within the Phd thesis of Bastien Trotobas. He has worked on exploration-exploitation trade-off in the context of MIMO detection techniques. The main objective of Bastien's research is to investigate the relevance of adding exploratory elements inside common tree-based MIMO detection algorithms. To achieve this purpose, we have reformulated the standard firefly algorithm (FA), initially proposed by Xin-She Yang in [1] and used for large MIMO detection in [2], as a tree based search algorithm associated to an exploratory factor. We proposed a novel soft-output MIMO detector called randomized tree-search with lower complexity and better detection performance than tree-based and firefly-based frameworks. Moreover, the designed soft-output detector was adapted to high-order modulation schemes, and three parameters have been proposed to handle the exploration-exploitation trade-off of the derived detector. The corresponding results was recently published in [CI56]. The geometrical-based MIMO detector proposed in [J2] was restricted to lower order modulation schemes (such as BPSK or QPSK) whereas higher-order modulation schemes were not investigated at all. The geometrical heuristic was compared with a

tree-based MIMO detection algorithms on both computational complexity and detection performance but these two criteria were not considered simultaneously within a trade-off perspective. In order to overcome the previous drawbacks, we investigated the performance-complexity trade-off between the new detector and tree-based algorithms through Pareto efficiency. Moreover, we proposed a new exploration techniques to improve the performance-complexity trade-off and extend the geometrical-based detection technique to higher-order modulation schemes. The Pareto front allow us to pick up the optimal tuning parameters and it demonstrate that the geometrical-based detector is suitable for lower order modulation schemes [J20].

Immediate after my PhD completion at the South Brittany University, Lorient in March 2006, I joined l'école supérieure d'électricité (currently CentraleSupélec) as postdoc fellow in the IETR/SCEE team, with Prof. Jacques Palicot and Prof. Christophe Moy. During my postdoctoral period (Apr. 2006 - Dec. 2007), my research has largely focused on the topic of software defined radio (SDR). In particular, I worked on the dynamic partial reconfiguration (DPR) technique in FPGAs to facilitate the design of SDR equipment. The DPR allows reconfiguration of some part or region of the FPGA logic resources at run-time, while the rest of the device continues its normal operation. I was involved in two projects: E2R phase II (EU FP6-IST) and IDROMEL (ANR RNRT 2005). In collaboration with Prof. Pierre Leray (IETR/SCEE), we have proposed a high-throughput module, called *ICAP32DMA*, for partial FPGA bitstream data management which can significantly reduce the reconfiguration time and the FPGA logic resources overhead in comparison with Xilinx's standard solution [CI11].

In January 2008, I joined CentraleSupélec as assistant professor in electronics and communication engineering. My main research objective was to provide proof-of-concept demonstrations of the dynamic partial reconfiguration technique in context of industrial applications. In collaboration with my former postdoctoral researcher Julien Delorme, we have provided the following two hardware demonstrations: **(i)** Reconfigurable hardware implementation of convolutional encoder, constellation mapper, and 32-tap FIR filter using DPR on Sundance platform, and **(ii)** Reconfigurable network on chip based on FPGA dynamic partial reconfiguration technique (FAUST platform). I have continued to work sporadically on DPR technique during the following years. In 2012, I investigated the possibility to use this design technique to reduce leakage power consumption in FPGAs [BC4]. In 2016, I managed to accurately measure the power consumption overhead during FPGA dynamic partial reconfiguration process [CI46].

Between 2008 and 2010, I was the principal investigator of DTTv2 project supported by Media and Networks cluster (Pôle Images & Réseaux). In collaboration with my former postdoctoral researcher (Mohamad Mroue), we have evaluated the performance of different tone reservation (TR) based techniques for peak to average power ratio (PAPR) in DVB-T2 context where the number of reserved subcarriers are less than 1%. Moreover, we have proposed a new iterative TR-based technique, called one kernel one peak (OKOP), with a reduced average-

power increment. The corresponding work was published in [J4]. In 2011, the DTTv2 project was awarded the third prize of Loading the Future organized by Pôle Images & Réseaux.

I started in September 2009 the co-supervision of the PhD thesis of Ziad Khalaf titled: "*contributions on spectrum sensing in the context of cognitive radio systems*". The main objective of this dissertation is to propose and develop simple and efficient spectrum sensing techniques for cognitive radio. Several spectrum sensing techniques have been proposed in this dissertation mainly focusing on hybrid architectures or on the sparsity of the cyclic autocorrelation function (CAF). My involvement in this thesis was up February 2013, the date of its defence. From 2013 to 2016, research works on spectrum sensing were continued with the Phd thesis of Hussein Kobeissi entitled: "eigenvalue based detector in finite and asymptotic multi-antenna cognitive radio systems". The main concern of this dissertation was to investigate eigenvalue-based spectrum sensing techniques using results from finite random matrix theory. This dissertation allowed me to deepen my skills in probability theory, model reduction via moments matching technique, and random matrix theory. My involvement in this excellent thesis (5 journal and 3 conference papers) was up 2016, the date of its defence. Since 2014, I was interested in experimental development of a proof of concept of different spectrum sensing techniques using Gnuradio tools and universal software radio peripheral platforms. From 2012 to 2015, I'm also involved in the European funded network of excellence Newcom# [1], focusing my researches on spectrum sensing form theoretical to hardware implementation aspects.

In October 2014, I joined the institute of research and technology (IRT-Bcom) as a part-time member (20% equivalent to 1 day/week). I was participating in two ANR projects: UniThing (2016-2018) and IoTrust (2018-2021). UniThing was focused on the design and optimization of reconfigurable architecture for a multi-standard radio equipment. This project was mainly related to Internet of Things (IoT) devices and networks, it addressed the following research issues: **(i)** adaptable, flexible and scalable hardware architecture for analog and digital front-end processing, **(ii)** various localization and data fusion methods, and **(iii)** quality of service (network densification and collision management). At IRT-BCOM, from 2016 to 2019, I have been involved in the co-supervision of a PhD student Ali Zeineddine on the design of a generic digital front-end for the internet of things (IoT). Currently, I am involved in ANR IoTrust project, which aims to safe connected devices and networks in the internet of things (IoT).

Based on the works of Dar *et al.* [3] and Karadimitrakis *et al.* [4] related to the capacity of Jacobi MIMO channel, I have been interested in the capacity of multimode and/or multimode fibers since 2016. In 2017, with Rémi Bonnefoi, a former PhD student in IETR/SCEE team, we published two papers [CI47] and [J13], where we derived upper and lower bounds of the ergodic capacity of optical MIMO Jacobi fading channels. In collaboration with Prof. Nizar Demni from IRMAR - UMR CNRS 6625, we recently published in IEEE Access journal a study

1. Network of Excellence in Wireless Communications

on unified exact closed-form expressions for the ergodic capacity of MIMO Jacobi and Rayleigh fading channels [J18].

I'm currently involved in the PhD supervision of Jeremy Guillaume, on his thesis on "The screaming gate array: study and characterization of IP data leakages in mixed-signal FPGA SoCs". The new UltraScale+ RFSoC device integrates multiple ARM cores, programmable logic, and programmable radio frequency stage (mixed-signal end, analog front-end) modules in the same chip. Therefore, due to substrate coupling, traces of digital switching noises will reach the analogue parts of the RFSoC device through the substrate. Once in the analog domain, this sensitive noise finds its way out of the system by: **(i)** being modulated, amplified and transmitted by radio (screaming channels) or **(ii)** being picked up by an ADC (leaky noise). As a result : **(i)** screaming channels can scream sensitive information through wireless communications, and **(ii)** ADCs can pick it up as conversion noise that will no longer be statistically independent. The objectives of Jeremy's thesis are threefold: **(a)** understand how far high sensitive signals circulating in programmable logic can leak internally, **(b)** look for potential side channels putting at risk these high sensitive signals, and **(c)** understand how the internal security of IPs may be compromised by exploiting screaming channels.

Bibliography

[1] X. S. Yang, "Firefly algorithm, stochastic test functions and design optimisation," *International Journal of Bio-Inspired Computation*, vol. 2, pp. 78–84, Jun 2010.

[2] A. Datta and V. Bhatia, "A near maximum likelihood performance modified firefly algorithm for large mimo detection," *Swarm and Evolutionary Computation*, vol. 44, pp. 828–839, 2019.

[3] M. S. R. Dar, M. Feder, "The jacobi mimo channel," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2426–2441, 2013.

[4] A. Karadimitrakis, A. L. Moustakas, and P. Vivo, "Outage capacity for the optical mimo channel," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 4370–4382, 2014.

Chapter 2

Extended Curriculum Vitae

This chapter presents an extended version of my curriculum vitae which contains my PhD thesis, Postdoctoral and associate professor research works.

2.1 General Information

Name:	NAFKHA
Surname:	Amor
Date of Birth:	7^{th} March 1977
Marital status:	Married with two children
Actual position:	Associate Professor/Enseignant-chercheur
Actual institution:	CentraleSupélec, Campus of Rennes, France
Address:	Avenue de la Boulaie CS 47601, F-35576, Rennes, France
E-mails:	amor.nafkha@centralesupelec.fr
Contacts:	Tel: +33(0)2.99.84.45.56, Fax: +33(0)2.99.84.45.99
Home page:	http://www.rennes.supelec.fr/ren/perso/anafkha/Welcome.php

2.1.1 Research Interests

- **Cognitive radio**: Spectrum sensing techniques, blind wireless standards identifications.
- **Signal processing**: Sample rate conversion, variable fractional delay filter
- **MIMO communication technology**: Low-complexity and high-throughput MIMO detection, iterative joint detection and decoding, multi-(mode/core) optical fiber capacity.
- **Parallel computing**: Dynamic partial reconfiguration in FPGA, fixed-point conversion, low power design techniques, software-defined radio platforms.
- **Security**: Side-channel attacks, secure FPGA design, Wireless Physical Layer Security.

2.1.2 Education

- **2003-2006**: PhD degree in Information and Communication Technology from University of Southern Brittany (UBS), LESTER CNRS-FRE 2734, Lorient, France.

- **Dissertation titled**: "*A geometrical approach detector for solving the combinatorial optimization problem : application in wireless communication systems*", Mar. 2006.
- **Adviser**: Prof. Emmanuel Boutillon, Université Bretagne Sud, Lorient, France.
- **Examination committee**:
  - ⋆ Prof. Michel Jézéquel, Institut télécom, ENST Bretagne, (President)
  - ⋆ Prof. Jean-Francois Hélard, INSA Rennes,
  - ⋆ Prof. Marie Laure Boucheret, ENSEEIHT Toulouse,
  - ⋆ Emerit Prof. Maurice Bellanger, CNAM Paris,
  - ⋆ Prof. Emmanuel Boutillon, Université Bretagne Sud, Lorient,
  - ⋆ Ass. Prof. Christian Roland, Université Bretagne Sud, Lorient.

- **1998-2001**: Engineering degree in electronics and communication from Higher School of Communication of Tunis (Sup'Com), Ariana, Tunisia.
  - **Dissertation titled**: "*Behavioral level design of a channel turbo decoder*", Jun. 2001.
  - **Adviser**: Prof. Emmanuel Boutillon, Université Bretagne Sud, Lorient, France.
  - **Examination committee**:
    - ⋆ Prof. Adel Ghazel, Sup'Com Tunis,
    - ⋆ Hichem Ben Hamida, director of STMicroelectronics-Tunis,
    - ⋆ Prof. Neji Youssef, Sup'Com, Tunis.
- **June 1998**: Success in national exams to access to engineering schools (rank: 5).
- **June 1996**: Baccalaureate science of technology graduated with honor after succeeding in the national Tunisian baccalaureate exam.

2.1.3 Current academic and research appointments

- **Jan. 2015 - present**: Associate Professor (Enseignant-Chercheur) at Centralesupélec, Campus of Rennes, France.
  - **Research**: Signal, Communications and Embedded Electronics (SCEE) lab of the "Institut d'Électronique et des Technologies du numéRique", (IETR UMR CNRS 6164).
  - **Teaching**: CentraleSupélec campus of Rennes.
- **Sep. 2010 - present**: Part-time teacher at the University of Rennes 1 (UR1), France.
  - **Teaching**: ISTIC Computer Science and Electrical Engineering, Master M2.
- **Oct. 2014 - present**: Part-time Researcher at the Institutes of Research and Technology (IRT-BCOM), Rennes, France.
  - **Research**: Network interfaces laboratory & AI for cybersecurity team.

2.1.4 Previous academic/ research appointments

- **Jan. 2008 - Dec. 2014**: Assistant Professor at Supélec, Campus of Rennes, France.

- **Research**: Signal, Communications and Embedded Electronics (SCEE) lab of the "Institut d'Électronique et des Technologies du numéRique", (IETR UMR CNRS 6164).
- **Teaching**: Supélec campus of Rennes.

- **Sep. 2012 - Jun. 2013**: Part-time teacher at the Institut Supérieur d'électronique de Paris (ISEP), France.
  - **Teaching**: High level digital embedded system design.
- **Apr. 2006 - Dec.2007**: Postdoctoral scholar at Supélec, campus of Rennes, France.
  - **Research**: FPGA dynamic partial reconfiguration for software defined radio system design.
- **Sep. 2004 - Feb. 2006**: Part-time teacher at the institute of technology (IUT), Lorient, France.
  - **Teaching**: Labs for time series characteristics and analysis.

2.1.5 Publications

- **Journals**: 20 including 1 editorial (List in section 3.3).
- **Book Chapters**: 6 (List in section 3.4).
- **International conferences**: 56 with peer review process (List in section 3.5).
- **National conferences**: 3 with peer review process (List in section 3.6).
- **Patents**: 1 patent number 05-02664, 2005.

2.1.6 Professional experience

- **Sep. 2001 - Dec. 2002**: Electronics design engineer at ST-Microelectronics, Ariana, Tunisia.
  - Responsible of hardware and software co-verification of parametric on-chip communication IP ST-BUS.
  - Bus-cycle accurate transaction level model of ST-BUS IP using SystemC.

2.1.7 Demonstrations at scientific meeting

- **Machine Learning for Cognitive Radio**: Opportunistic spectrum access proof-of-concept based on universal software radio peripheral (USRP) platform
  - MiWEBA, Green 5G networks Workshop (2015)
  - DySPAN, International Symposium on Dynamic Spectrum Access Networks (2015)
- **Reconfigurable Network-on-Chip based on FPGA**: Dynamic partial reconfiguration (DPR) technique is used to reconfigure a part of the NOC in real-time.
  - DySPAN, International Symposium on Dynamic Spectrum Access Networks (2015)

  - Next-GWiN, International Workshop on Next Generation Green Wireless Networks (2014)
  - International Conference on Cognitive Radio Oriented Wireless Networks (2010)
  - SDR'FORUM, SDR Forum Technical Conference (2008)
- **Blind Wireless Standards Identification**: Experimental measurements using USRP platforms to evaluate performances of blind techniques.
  - Partenariat Hubert Curien (PHC) program Platon GRIC (2014)
  - International Conference on Cognitive Radio Oriented Wireless Networks (2010)
- **Spectrum Sensing Algorithms for Cognitive Radio**: Experimental measurements using USRP platforms.
  - Partenariat Hubert Curien (PHC) program RILA PUSLO Platform for Cooperative Spectrum Sensing and Primary User Localization (2015)

2.1.8 Distinctions, awards and honors

- **2016**: Best Paper Award
  - The Twelfth Advanced International Conference on Telecommunications (AICT)
  - Paper entitled: *Implementation of an Energy Detection Based Cooperative Spectrum Sensing on USRP Platforms for a Cognitive Radio Networks.*
- **2016**: The Best Booth Award
  - International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM)
  - Demonstration entitled: *Spectrum Utilization and Reconfiguration Cost Comparison of Various Decision Making Policies for Opportunistic Spectrum Access Using Real Radio Signals.*
- **2016**: Nominated IEEE Senior Member
- **2012**: Best Readings in Green Communications (IEEE ComSoc)
  - Book: *Green Communications,Theoretical Fundamentals, Algorithms and Applications*, CRC Press, Edited by Jinsong Wu, Sundeep Rangan, and Honggang Zhang.
  - Chapter: *Moving a processing element from hot to cool spots, is this an efficient method to decrease leakage power consumption in FPGAs ?*
- **2011**: Third Prize Trophies
  - Images & Networks Pole: Loading The Future
  - French regional DTTV2 project
- **2010**: Best Paper Award
  - The Sixth Advanced International Conference on Telecommunications (AICT)
  - Paper entitled: *Hybrid Spectrum Sensing Architecture for Cognitive Radio Equipment.*

2.1.9 Visiting appointments

- **Technical University of Sofia, Bulgaria**
  - **Duration**: One week in July 2015
  - **Collaborator**: Prof. Galia Marinova
  - **Publications**: 1 book chapter [BC5], 1 conference paper [CI45]
- **High School of Communications of Tunis, Tunisia**
  - **Duration**: One week in June 2011
  - **Collaborator**: Prof. Adel Ghazel, Prof. Mohamed Siala
  - **Publications**: 2 Conference papers [CI13], [CI31]
- **Nanyang Technological University, Singapore**
  - **Duration**: Two weeks in May 2010
  - **Collaborator**: Prof. Vinod A Prasad

2.1.10 Professional memberships

- **Institute for Electrical and Electronics Engineers**
  - IEEE Green ICT Community (2012 - present)
  - IEEE Computer Society (2009 - 2015)
  - IEEE Communications Society (2008 - 2015)
  - IEEE Circuits and Systems Society (2009-2015)
- **Research Networks**
  - WUN Cognitive Communications Consortium (WUN CogCom) (2010-present)
  - ICT COST IC0902 Cognitive Radio and Networking (2010-present)
  - Technical Committee on Green Communications and Computing (2014-present)
  - Technical Committee on Cognitive Networks (2014-present)
  - Groupement de Recherche en Information Signal Image et viSion (2008-present)
  - Union Radio-Scientifique Internationale, URSI-France (2008-present)
  - Institutes of Research and Technology (IRT-B<>COM) (2014-present)

2.1.11 Main scientific collaborations

- International collaborators:
  - **Pr. Honggang ZHANG** (Zhejiang University, China)
  - **Pr. Francesco Regazzoni** (ALaRI Institute, Switzerland)
  - **Pr. Mohamed Siala** (Sup'Com Tunis, Tunisia)
  - **Pr. Adel Ghazel** (Sup'Com Tunis, Tunisia)
  - **Pr. Adrian Kliks** (Poznań University of Technology, Poland)
  - **Pr. Youssef Nasser** (American University of Beirut, Lebanon)

  - **Pr. Oussama Bazzi** (Lebanese University, Beirut, Lebanon)
  - **Pr. Galia Marinova** (Technical University of Sofia, Bulgaria)
  - **Pr. Djamel Slimani** (University of Setif 1, Algeria)
  - **Pr. Jordi Pérez-Romero** (Universitat Politècnica de Catalunya, Spain)
  - **Pr. Taoufik Hmidi** (New York University, Abu Dhabi)
- National/ local collaborators:
  - **Pr. Yves Louet** (CentraleSupélec)
  - **Pr. Christophe Moy** (UR1)
  - **Pr. Emmanuel Boutillon** (UBS Lorient)
  - **Pr. Maxime Pelcat** (INSA Rennes)
  - **Pr. Nizar Demni** (I2M, Marseille)
  - **Pr. Zied Ammari** (IRMAR, Rennes)
  - **Dr. Stéphane Paquelet** (IRT-B<>COM)
  - **Ass. Pr. Fakhreddine Ghaffari** (ENSEA)
  - **Ass. Pr. Xun Zhang** (ISEP)
  - **Ass. Pr. Ruben Salvador** (CentraleSupélec)

2.1.12 Technical program Committees

- IEEE International Conference on Internet of Things and Intelligence System 2020
- IEEE Symposium on Personal, Indoor and Mobile Radio Communications 2020-2017
- IEEE International Conference on Communications 2021-2014
- IEEE Global Communications Conference 2019
- IEEE International Symposium on Wireless Communication Systems 2016-2014
- IEEE Wireless Communications and Networking Conference 2019-2018
- IARIA Advanced International Conference on Telecommunications 2020-2016
- IEEE Conference on Wireless Communications & Signal Processing 2014-2013

2.1.13 International journals reviewer

- IET Electronics Letters
- IEEE Photonics Technology Letters
- China Communications
- IEEE Transactions On Communications
- IEEE Transactions On Cognitive Communications and Networking
- Transactions on Emerging Telecommunications Technologies
- Elsevier Embedded Hardware Design (Microprocessors and Microsystems)
- Elsevier Signal Processing
- Elsevier Integration, the VLSI journal

- Journal of the Optical Society of America A
- Mathematical Problems in Engineering
- EURASIP Journal on Wireless Communications and Networking
- IEEE Transactions on Signal Processing
- IEEE Wireless Communications
- IEEE Transactions on Very Large Scale Integration Systems
- MDPI Sensors, MDPI Electronics, MDPI Applied Sciences
- Springer Annals of Telecommunications

2.2 Teaching activities

Since 2004, I teach mainly at post-graduate level in several institutions including: University of South Brittany (UBS), University of Rennes 1 (UR1), École supérieure d'électricité (Supélec), Institut supérieur d'électronique de Paris (ISEP), CentraleSupélec (CS). Moreover, since 10 years, I participate in several continuing training for engineers and industry leaders. My teaching activities belong to embedded electronics and signal processing areas. The standard teaching duty is **192** hours per year. During the academic years between 2008 and 2015, I was with Supélec and my teaching duties amounted up to an average of **283** HETD per year (*i.e.* **teaching surcharge = 47**%). Since January 2015, I have been working as an associate professor at CentraleSupélec. From academic years 2015–2016 until 2018–2019, my teaching duties were heavier with an average close to **365** HETD per year (*i.e.* **teaching surcharge = 90**%). During the last academic year, part of my teaching hours were passed over to a new associate professor. Hence, my teaching duties were reduced (*i.e.* **239** HETD $\Leftrightarrow$ **teaching surcharge = 24**%), so that I could have more time for my research activities. Table 2.1 gives the partition of my teaching activities in terms of HETD between different institutions and across different academic years between 2004 and 2020.

Table 2.2 gives a detailed summary of my teaching areas for the academic year 2019-2020 in terms of: courses titles, amount of hours (CM: Cours Magistral, TD: Travaux Dirigés, TP: Travaux Pratiques), and student levels.

2.2.1 Module responsibilities

In the last 5 years, I have been in charge of organizing the following modules at CentraleSupélec and University of Rennes 1. The main tasks included: contacting instructors, updating the content lectures and labs, updating of the course schedule, proposing the exam.

- Re-configurable Computing;
- Advanced Digital Design;
- Introduction to MIMO Systems;

Year	HETD (Supélec/CS) [1]	HETD (CT) [2]	HETD (Others) [3]	Total
2004-2008	-	-	160	160
2008-2009	≈ 250	≈ 10	8	268
2009-2010	≈ 250	≈ 10	8	268
2010-2011	≈ 270	≈ 23	8	301
2011-2012	≈ 270	≈ 23	8	301
2012-2013	≈ 270	≈ 23	20	313
2013-2014	≈ 270	≈ 23	8	301
2014-2015	≈ 270	≈ 23	8	301
2015-2016	294	50	16	360
2016-2017	294	23	16	333
2017-2018	341	21	16	378
2018-2019	379	58	16	453
2019-2020	205	34	53	292
Total (Per.)	3363 (83.6%)	321 (7.9%)	345 (8.5%)	4029 (100%)

Table 2.1 – Teaching during 2004–2020 in number of hours (Heure équivalent TD). $(.)^1$ refers to 1^{st} year, 2^{nd} year, and 3^{rd} year at Supélec/CS, $(.)^2$ refers to continuing training, and $(.)^3$ refers to UBS, UR1 and ISEP.

- Introduction to Analog and Digital Electronics;
- Embedded C programming;
- Hardware Design Tools;
- MicroC/OS-II and FreeRtos Real-Time Operating Systems.

2.2.2 Continuing training

- From software radio to cognitive radio (MR19) (New training programme)
  - De la radio logicielle à la radio intelligente
  - https://exed.centralesupelec.fr/formation/de-la-radio-logicielle-a-la-radio-intelligente/
  - Co-responsible (with Pr. Yves Louet) of this training.
- SDR Platforms & Security (MR40) (New training programme)
  - Plateformes radio logicielle: Analyse de la sécurité des objets connectés
  - https://exed.centralesupelec.fr/formation/plateformes-radio-logicielle-analyse-de-la-securite-des-objets-connectes/
  - Responsible of this training.
- Formation electronics and instrumentation (ER12) (2015 - present)
  - Électronique analogique: Méthodologie et mise en oeuvre
  - https://exed.centralesupelec.fr/formation/electronique-analogique/
  - Responsible of this training.
- Formation electronics and instrumentation (ER18) (2015 - present)

Courses	CM	TD	TP	Level
Intro. Analog Electronics	9	3	-	1A (CS)
Intro. Digital Electronics	9	3	-	1A (CS)
Statistics	-	12	-	1A (CS)
Software Radio	13.5	-	12	2A (CS)
Computer Architecture	6	-	6	2A (CS)
Intro. Information Theory	2	-	-	2A (CS)
Projects	-	5	-	2A (CS)
Reconfigurable Computing	13.5	6	-	3A (CS)
Real-Time OS	4.5	3	-	3A (CS)
Low Power Design	4.5	-	-	3A (CS)
μVision IDE - Keil	-	3	-	3A (CS)
Introduction to VHDL	4.5	6	-	3A (CS)
Design Tools	-	6	-	3A (CS)
FPGA Project	-	-	25	3A (CS)
SystemC	4.5	3	-	3A (CS)
Digital Communications	4.5	-	-	3A (CS)
Introduction to MIMO	9	-	-	3A (CS)
Project Screaming Channel	-	15	-	3A (CS)
Internship	-	8	-	3A (CS)
Advanced Digital Design	12	-	-	M2 (UR1)
Digital Design	19	-	-	M2 (UR1)
Embedded Software	12	-	-	M2 (UR1)
Projets RISCV	-	10	-	M2 (UR1)
Nios II Processor	4.5	-	6	M2 (CS)
Design VGA Controller	4.5	-	9	M2 (CS)
MicroC/OS-II	4.5	-	6	M2 (CS)
Internship	-	7	-	M2 (CS)
Total	141 (48.2%)	87 (29.8%)	64 (22%)	= 292 HETD

Table 2.2 – Summary of courses taught during the academic year (2019-2020) in number of hours (Heure équivalent TD).

  - Conception de Systémes "On Chip" (SOC)
  - https://exed.centralesupelec.fr/formation/conception-de-systemes-on-chip-soc/
  - Instructor in this training.
- Formation electronics and instrumentation (ER22) (2015 - present)
  - Comprendre et utiliser les FPGAs.
  - https://exed.centralesupelec.fr/formation/comprendre-et-utiliser-les-fpga/
  - Instructor in this training.
- Formation electronics and instrumentation (ER30) (2015 - 2019)
  - Executive Certificate Conception de systémes électroniques numériques

  - http://www.rennes.centralesupelec.fr/en/fc
  - Instructor in this training.
- Formation electronics and instrumentation (ER19) (2015 - 2018)
  - Systémes numŕiques: Architecture et Conception.
  - www.exed.centralesupelec.fr/en/trainings/er19-18-systemes-numeriques-architecture-et-conception
  - Instructor in this training.

2.2.3 Administrative responsibilities

- **2012-present:** Jury member of the Master I-MARS
  - Micro-technologies, Architecture, Réseaux & Systèmes de communication
  - Jointly accredited by: INSA of Rennes, CentraleSupélec, University of South Brittany, IMT Atlantique.
- **2017-present:** Oral examinator (Concours CentraleSupélec)
  - Competitive entrance exam and admission for engineering program.
- **2008-present:** Referee and internship tutor
  - Short-term and long-term study abroad programs.
  - Engineer internships for 1A, 2A and 3A levels.
- **2015:** Village des Sciences
  - Maths, Réseaux & Numérique.
  - Les secrets des transmissions radio.
  - Place: Diapason, campus de Beaulieu, Rennes.
- **2014:** Village des Sciences
  - Maths, Réseaux & Numérique.
  - Les Ondes: Naissance, Transformations & Mort.
  - Place: Chartres de Bretagne.

2.3 Research activities

Since January 2008, my research activities as associate professor have been performed mainly at the Institut d'Électronique et des Technologies du numéRique (**IETR**), UMR, CNRS-6164, Rennes, within the Signal & Communications (**SC**) department, and more specifically with the Signal, Communications & Embedded Electronics (**SCEE**) Team. My main research activities were focused on the **cognitive radio** and **software-defined radio** technologies as well as on the real-time FPGA-based re-configurable computing. In particular, in the following list, my former and current research activities are listed, including references to related major publications.

- **My current research topics are:**

(1) MIMO detection techniques [J20], [J16], [BC6], [J1].

(2) Arbitrary sample rate conversion [J17], [J15].

(3) Spectrum sensing techniques [J19], [J14], [J12], [J11], [J10], [J9], [J8], [J7], [J6],[J3].

- **My former research topics are:**

(1) Real-Time Dynamic Partial Reconfiguration of FPGA [CI11].

(2) Leakage Power Reduction Techniques [J5], [BC4].

(3) Peak-to-Average Power Ratio Reduction Techniques [J4].

(4) MIMO optical fiber communication [J18], [J13], [CI47].

2.3.1 Supervising activities

Below, I outline my experience in supervising Master degree projects and visiting/local PhD students. In addition to this, I have acted as advisor for researchers at post-doctoral level at the Signal, Communication, and Embedded Electronics Lab (IETR/SCEE). In Table 2.3, I included the official advising percentages.

Students	M2	Local PhD	Visiting PhD	Post-Doc	Advising (%)
Marwa Jellali	x	-	-	-	100%
Abbe Ahmed Khalifa	x	-	-	-	30%
Bastien Trotobas	x	-	-	-	100%
Gurvan Priem	x	-	-	-	100%
Narjes Karmani	x	-	-	-	100%
Ali Zeineddine	x	-	-	-	50%
Sara Bahamou	x	-	-	-	100%
Khalil Ajmi	x	-	-	-	100%
Tri Huan Hoang Dinh	x	-	-	-	100%
Jeremy Guillaume	-	x	-	-	30%
Bastien Trotobas	-	x	-	-	75%
Ali Zeineddine	-	x	-	-	25%
Hussein Kobeissi	-	x	-	-	25%
Ziad Khalaf	-	x	-	-	30%
Zdravka Chobanova	-	-	x	-	100%
Julio Vásquez	-	-	x	-	100%
Aziz Babar	-	-	-	x	50%
Mohamad Mroué	-	-	-	x	70%
Julien Delorme	-	-	-	x	70%
Total	9	5	2	3	

Table 2.3 – Summary of my experience in supervising (2008-2021).

I. Ongoing/former Master students

- **Mar. 2021 - Jul. 2021**: Marwa Jellali
  - **Subject entitled**: *Malicious user detection in asynchronous cooperative sensing environment.*
  - **From**: *Sup'Com Tunis - Tunisia.*
  - **Funding**: *IETR/SCEE grant.*
  - **Abstract**: Cooperative spectrum sensing among a few sensors has been shown to offer significant gain in the performance of the cognitive radio spectrum sensing system by countering the shadow-fading effects. During this internship, we will investigate new schemes to identify the malicious users based on outlier detection techniques for an asynchronous cooperative sensing system employing different spectrum sensing algorithms at the sensors side.
- **Mar. 2020 - Sep. 2020**: Abbe Ahmed-Khalifa
  - **Subject entitled**: *Circuit Diversification for Improved Security of FPGA Computing Architectures.*
  - **From**: *Eurecom Sophia Antipolis.*
  - **Funding**: *IETR PEPS.*
  - **Abstract**: This internship has participated to a long-term objective of understanding the relationship between configuration bitstream of a given FPGA and its resilience to side channel attacks. We have focused on studying information leakage due to power consumption, which can be exploited with power side-channel attacks. Hence, during the internship, we have worked on understanding the bitstream structure which describes the FPGA configuration and on setting up a power side-channel measurement environment using FPGA internal power sensors. Finally, we worked with the open-source toolchain: ChipWhisperer. This tool allowed us to realize power attacks, to observe their efficiency and compare them to our own configuration.
- **Apr. 2018 - Aug. 2018**: Bastien Trotobas
  - **Subject entitled**: *Hardware Implementation of Reconfigurable MIMO detector.*
  - **From**: *École Normale Supérieure de Rennes.*
  - **Funding**: *ENS Rennes grant.*
  - **Abstract**: In this internship, we propose a new soft-output MIMO detection technique based on two complementary methods: *exploration* and *exploitation*. The proposed detector, called list exploration and exploitation (L2E), achieves near-optimal performance with low and fixed computational complexity. It has a high parallelism degree, which makes it suitable for an efficient FPGA implementation. The average bit error rate performances of the L2E are compared to the state-of-the-art results. The proposed technique provides near maximum likelihood performance with signifi-

cantly reduced hardware complexity.
  - **Current position**: PhD student at SCEE/IETR team (Rennes).
- **Apr. 2017 - Aug. 2017**: Marouan Maghzaoui
  - **Subject entitled**: *FPGA implementation of a reconfigurable Massive MIMO detector.*
  - **From**: *ISTIC/UR1.*
  - **Funding**: *SCEE/IETR.*
  - **Abstract**: In this internship, In this internship, The main objective involved examining the FPGA implementation aspects of the GISD MIMO detector to achieve good understanding of the complexity of various FPGA implementation architectures. Moreover, to reduce the number of multipliers required in the design, we use $l1$-norm instead of the optimal squared $l2$-norm. The impact of the use of $l1$-norm and the floating-to-fixed-point conversion have been studied and illustrated upon the bit-error-rate performance of the detector for different MIMO system configurations.
  - **Current position**: Embedded systems engineer at SERMA INGENIERIE (Paris).
- **Apr. 2016 - Sep. 2016**: Gurvan Priem
  - **Subject entitled**: *SCN-based spectrum sensing technique under asynchronous primary user traffic.*
  - **From**: *Politecnico di Milano.*
  - **Funding**: *SCEE/IETR.*
  - **Abstract**: In this internship, we provide a theoretical formulation of the standard condition number based spectrum sensing under an asynchronous primary user traffic. We assume the case where the primary users generate asynchronous slotted traffic. The standard condition number distributions under null and alternative hypotheses are derived where the secondary user is equipped with two receive antennas. We provide some simulation results to validate the accuracy and the effectiveness of the derived mathematical expressions.
  - **Current position**: Statistic engineer at Biosency (Rennes).
- **Mar. 2016 - Jul. 2016**: Narjes Karmani
  - **Subject entitled**: *Reconfigurable Low-power MIMO Detector Based on Dynamic Partial Reconfiguration.*
  - **From**: *Tunisia Polytechnic School.*
  - **Funding**: *SCEE/IETR.*
  - **Abstract**: In this internship, we consider the problem of design a reconfigurable low-power MIMO detection using FPGA dynamic partial reconfiguration technique. The main objectives of this work are: **(i)** understand a quasi-optimal MIMO detector based on two phases are as follows: geometrical diversification and greedy intensification,

(ii) evaluate the complexity/performance tradeoffs of the MIMO detector through examination of its parameters including: the MIMO channel condition number, the number of studied dominant noise axis, and the number of best initial candidates for the intensification phase, **(iii)** Experimental results and comparisons between different architectures using dynamic partial reconfiguration and a performance analysis of the area, power consumption and maximum frequency are analyzed.

- **Current position**: Data scientist at Foton-IT (Tunisia).

- **Apr. 2016 - Aug. 2016**: Ali Zeineddine
  - **Subject entitled**: *Efficient Implementation of Sample Rate Converter.*
  - **From**: *École CentraleSupélec.*
  - **Funding**: *IRT-BCOM.*
  - **Abstract**: In this internship, we consider available options of sampling rate conversion (SRC) in digital front-ends (DFE). We start with the fundamentals of SRC from a digital signal processing perspective, and distinguish three configurations of SRC depending on its re-sampling factor: integer, rational, and arbitrary. Following this identification, the work focuses on three solutions that are be the best adapted for DFEs: the cascaded-integrator-comb (CIC) filter for SRC applications with very large factors from one side, and the polynomial based filters and the Newton structure from the other, where these last two offers an efficient solution for implementing fractional delay filters which are used for SRC applications of fine-tuned factors. Used together, the different solutions can answer to almost all SRC needs in a DFE.
  - **Current position**: Signal processing engineer at TDF (Rennes).
- **Mar. 2012 - Jul. 2012**: Sara Bahamou
  - **Subject entitled**: *Spectrum sensing techniques under noise uncertainty.*
  - **From**: *INSA Toulouse.*
  - **Funding**: *SCEE/IETR.*
  - **Abstract**: Energy detection is a simple non-coherent approach used for spectrum sensing that offers linear computational complexity and low latency. Its main shortcoming is the well-known noise uncertainty problem, which results in failed detection in the very low signal-to-noise ratio. In this internship, we give a theoretical analysis of the noise uncertainty modeling with energy detection to drive a relationship between bounded and unbounded uncertainty approximation. In order to increase the detection accuracy and reduce spectrum sensing cost, we tried to extend the formulation to cooperative spectrum sensing based on energy detector.
  - **Current position**: PhD student at IA laboratory (Morocco).
- **Mar. 2010 - Jul. 2010**: Khalil Ajmi
  - **Subject entitled**: *Energy detector technique under MIMO uncorrelated Rayleigh*

fading channels.

- **From**: *Sup'Com Tunisia.*
- **Funding**: *SCEE/IETR.*
- **Abstract**: Spectrum sensing is an essential mechanisms of cognitive radio equipment to find the unused frequency bands. The energy detector based spectrum sensing technique is known as the simplest. In fact, it has both low hardware cost and low computation complexity (compared to other sensing algorithms). It also has the advantage that it does not require any a prior knowledge about the primary user signal. The main idea of the internship is to implement energy detector with Matlab in case of single-input single-output Rayleigh fading channel, derive different mathematical expressions, and extend the previous work to the MIMO uncorrelated Rayleigh fading channels.
- **Current position**: Test automation engineer at Nokia (Belgium).

II. Ongoing/former PhD students

- **Jan. 2021 - Dec. 2023**: Jeremy Guillaume
  - **PhD. title**: *The Screaming Gate Array: Study and characterization of IP data leakages in mixed-signal FPGA SoCs.*
  - **From**: *INSA Rennes.*
  - **Funding**: *1/2 ARED-PEC grant and 1/2 IETR/SCEE.*
  - **Joint supervision with**: Professor Maxime Pelcat (INSA Rennes, 30% [4]), Ass. Prof. Ruben Salvador (CentraleSupélec, 40%) and Ass. Prof. Amor Nafkha (CentraleSupélec, 30%)
  - **Abstract**: Previous work exploiting digital to analog substrate couplings have considered only CPU-based mixed-signal systems, which feature fixed hardware. However, mixed-signal reconfigurable platforms known as RFSoCs (Radio Frequency SoCs) now combine CPU cores, FPGA, and programmable analog/RF (radio transmitters, data converters) modules in the same chip. RFSoCs are thus potential targets for malicious sensor circuits. The combination of the different leakages issues makes highly probable the presence of unexpected new leakage channels to be discovered and exploited. In this context, the thesis will focus on studying the data leakage mechanisms in mixed-signal reconfigurable devices. An objective will be to look for vulnerabilities specific to this new type of devices and linked to their reconfigurable nature, as well as to the fact that information is crossing ever more system layers (memory, CPUs, reconfigurable HW, analog RF and data converters). The rapid spread of highly integrated reconfigurable devices is the main driver for this thesis.

4. PhD. Supervising rate

- **Publications**: No related papers right now.

- **Sep. 2018 - Aug. 2021**: Bastien Trotobas
  - **PhD. title**: *Exploration-Exploitation trade-off for MIMO detection problem.*
  - **From**: *École Normale Supérieure de Rennes.*
  - **Funding**: *CDSN grant.*
  - **Joint supervision with**: Prof. Yves Louet (CentraleSupélec, 25%) and Ass. Prof. Amor Nafkha (CentraleSupélec, 75%)
  - **Abstract**: In order to approach MIMO capacity, there has been a tremendous effort to develop the coded MIMO systems, which include an iterative detection and decoding (IDD) algorithms. The main goal of an IDD algorithm is to combine an efficient soft-input soft-output (SISO) detection method and a SISO decoding technique. Due to the existence of a high number of antennas in large-scale MIMO systems, the receiver poses challenges in several design aspects such as, complexity of detection and decoding algorithms, channel estimation and hardware implementation. The main objective of this thesis can be divided into two parts: **(i)** Part I investigate a MIMO detector algorithm based on *diversification* and *intensification* process. The proposed detector scheme will be developed with respect to four axioms: Near-optimal bit-error rate performance, parallel processing structure, constant and fixed computational complexity, and finally a very high throughput performance. The soft-input soft-output version of the proposed detector will be investigated for iterative detection and decoding scheme. **(ii)** Part II investigate the hardware implementation of the proposed MIMO detector in a FPGA using dynamic partial reconfiguration design flow. Different hardware architectures will be analyzed to efficiently utilize the available FPGA resources.
  - **Publications**: The related papers and book chapters to date are [J20], [J16], [BC6], [CI52], [CI56].
- **Oct. 2016 - Sep. 2019**: Ali Zeineddine
  - **PhD. title**: *Design of a Generic Digital Front-End for the Internet of Things (IoT).*
  - **Funding**: *CIFRE TDF.*
  - **Joint supervision with**: Prof. Christophe Moy (University of Rennes 1, 25%), Ass. Prof. Amor Nafkha (CentraleSupélec, 25%), Stéphane Paquelet (IRT-BCOM, 25%), Pierre-Yves Jézéquel (TDF, 25%).
  - **Examination committee**:
    - ⋆ Prof. Christophe Jego, IMS Bordeaux, France.
    - ⋆ Prof. Markku RENFORS, Tampere University, Finland.
    - ⋆ Prof. Marie Laure Boucheret, INP-ENSEEIHT/IRIT, France (The jury president)
    - ⋆ Dr. Dominique Morche, Research Engineer at CEA-LETI, France.

- **Abstract**: The number of wireless communication technologies and standards is constantly increasing to provide communication solutions for today's technological needs. This is particularly relevant in the domain of the Internet of Things (IoT), where many standards are available, and many others are expected. To efficiently deploy the IoT network, the interoperability between the different solutions is critical in order to avoid the fragmentation of this network, which increases its installation and operation costs, and complicates its management. Interoperability on the physical level is achieved through multi-standard modems that support the largest number of standards with minimal added implementation costs. These modems are made possible through the digital front-end (DFE), that offers a flexible radio front-end able of processing a wide range of signal types. This thesis first develops a generic architecture of both transmission and reception DFEs, which can be easily adapted to support different IoT standards. These architectures highlight the main role of sample rate conversion (SRC) in the DFE, and the importance of optimizing the SRC implementation. This optimization is then achieved through an in-depth study of the SRC functions, and the development of new structures of improved efficiency in terms of implementation complexity and power consumption, while offering equivalent or improved performance. The final part of the thesis addresses the optimization of the DFE hardware implementation, which is achieved through developing an optimal quantization method that minimizes the use of hardware resources while guaranteeing a given performance constraint. The obtained results are finally highlighted through implementing and comparing different implementation strategies on both field programmable gate array (FPGA) and application specific integrated circuit (ASIC) targets.
- **Related Publications**: [J17], [J15], [CI54], [CI53], [CI49], [CI48].
- **Keywords**: Arbitrary Sample Rate Conversion, Digital Front-End, Newton Structure, ideal fractional delay filter, Newton fractional delay filter, Lagrange interpolation, reconfigurable Newton structure, uniform convergence, FPGA, power consumption, ASIC.
- **Current position**: Signal processing engineer at TDF (Rennes).

- **Nov. 2013 - Dec. 2016**: Hussein Kobeissi
  - **PhD. title**: *Eigenvalue Based Detector in Finite and Asymptotic Multi-antenna Cognitive Radio Systems.*
  - **Funding**: *1/2 ARED + 1/2 Lebanon grant.*
  - **Joint supervision with**: Prof. Yves Louet (CentraleSupélec, 25%), Ass. Prof. Amor Nafkha (CentraleSupélec, 25%), Prof. Oussama Bazzi (Lebanese University, 25%), Prof. Youssef Nasser (American University of Beirut, 25%).
  - **Examination committee**:

  - ⋆ Professor Abed Ellatif Samhat, Lebanese University (The jury president).
  - ⋆ Professor Inbar Fijalkow, ETIS/ENSEA - University Cergy-Pontoise, France.
  - ⋆ Professor Chafic MOKBEL, Balamand University, Lebanon.
  - ⋆ Ass. Prof. Maher Jridi, ISEN Brest, France.
- **Abstract**: In cognitive radio, spectrum sensing is the task of obtaining awareness about the spectrum usage. The first part of this thesis concerns the standard condition number (SCN) and the scaled largest eigenvalue (SLE) based detectors. The main focus is on the complexity of the statistical distributions of the SCN and the SLE decision metrics since this will imply a complicated mathematical expressions for the performance probabilities as well as the decision threshold if it could be derived. We derive exact expressions for the probability density function (PDF) and the cumulative distribution function (CDF) of the SCN using results from finite random matrix theory (RMT). In addition, we derived exact expressions for the moments of the SCN and we proposed a new approximation based on the generalized extreme value (GEV) distribution. Moreover,we proved that the SLE decision metric could be modeled using Gaussian function. Therefore, we derived expressions of PDF, CDF, Pfa, Pd and decision threshold. In addition, we also considered the correlation between the largest eigenvalue and the trace in the SLE study. The second part of this thesis concerns the massive MIMO technology and how to exploit the large number of antennas for spectrum sensing. Two antenna exploitation scenarios are studied: **(i)** Full antenna exploitation and **(ii)** Partial antenna exploitation in which we have two options: **(a)** Fixed use or **(b)** Dynamic use of the antennas. We considered the largest eigenvalue (LE) based detector if noise power is perfectly known and the SCN and SLE based detectors when noise uncertainty is considered. For fixed approach, we derived the optimal threshold which minimizes the error probabilities. For the dynamic approach, we derived the equation from which one can compute the minimum requirements of the system. For full exploitation, asymptotic approximation of the threshold is derived using the GEV distribution. Finally, comparisons between these scenarios and different detectors are provided in terms of system performance and minimum requirements.
- **Related Publications**: [J14], [J12], [J11], [J10], [J9], [CI43], [CI42], [CI35].
- **Keywords**: Cognitive radio, MIMO systems, eigenvalue based detector, spectrum sensing, Wishart matrix, generalized extreme value distribution, Taylor approximation, ROC curves.
- **Current position**: Assistant professor at AUB (Lebanon).

- **Sep. 2009 - Feb. 2013**: Ziad Khalaf
  - **PhD. title**: *Contributions on spectrum sensing in the context of cognitive radio sys-*

tems.

- **Funding**: *MENRT grant.*
- **Joint supervision with**: Prof. Jacques Palicot (Supélec, 70%), Ass. Prof. Amor Nafkha (Supélec, 30%).
- **Examination committee**:
  - ⋆ Prof. Jean-Marie Gorce, INSA de Lyon, France.
  - ⋆ Prof. Aawatif Hayar, UH2C, Casablanca (The jury president)
  - ⋆ Prof. Didier Le Ruyet, CNAM of Paris, France.
  - ⋆ Ass. Prof. Jean-Yves Baudais, Chercheur CNRS, France.
  - ⋆ Ass. Prof. Carlos Faouzi Bader, CTTC, Barcelona.
- **Abstract**: The wireless communications systems continue to grow and has become very essential nowadays. This growth causes an increase in the demand of spectrum resources, which have become more and more scarce. To solve this problem of spectrum scarcity, Joseph Mitola III introduced the idea of dynamic spectrum allocation. Mitola defines the term *cognitive Radio*, which is widely expected to be the next Big Bang in wireless communications. In this thesis work, we focus on the problem of spectrum sensing which is the detection of the presence of primary users in licensed spectrum, in the context of cognitive radio. The main objective of this work is to propose effective detection methods at low-complexity and/or using short observation time, using minimal a prior knowledge about primary users. The first part of the thesis deals with the problem of detecting a random signal in noise. Two main methods of detection are used: **(i)** energy detection and **(ii)** cyclostationary feature detector. We propose a hybrid architecture for detecting primary user signals, which combines the simplicity of the energy detector and the robustness of the cyclostationary feature detector. Two detection methods are proposed that are based on this hybrid architecture. Given the flexibility of the proposed architecture, the overall computation complexity of the hybrid detector decreases and tends to the ideal energy detector complexity. Moreover, the performance in terms of false alarm rate and detection probability, under low SNR and noise uncertainty environment, of the proposed architecture tends to be very close to performance of cyclostationary feature detector. In the second part of this thesis, we exploit the sparse property of the cyclic auto-correlation function (CAF) to propose a new blind estimator based on compressed sensing technique that estimates the cyclic auto-correlation vector (CAV) which is a particular presentation of the CAF for a given delay. It is shown by simulation that this new estimator gives better performances than those obtained with the classical estimator, which is non-blind, under the same conditions and using the same number of samples. Using the new estimator, we propose two blind spectrum

sensing techniques that require a smaller number of samples as compared to the conventional cyclostationary feature detector algorithm. The first proposed detector uses only the sparse property of the CAV while the second detector exploits the symmetry property of the CAV in addition to its sparse property, which allows better sensing performances.

- **Related Publications**: [J3], [CI30], [CI26], [CI25], [CI24], [CI22], [CI19], [CN3].
- **Keywords**: cognitive radio, dynamic spectrum access, spectrum sensing, compressed sensing, sparsity, OMP algorithm, cyclic frequency domain, energy detector, cyclostationary feature detector, cyclic auto-correlation function estimation, blind spectrum sensing.
- **Current position**: Network architect with EDF (Paris).

III. Visiting PhD students

- **Mar. 2017 - Aug. 2017**: Zdravka Chobanova
  - **Visiting objective**: *Experimental Evaluation of Energy detector Based Cooperative Spectrum Sensing using USRP Platforms.*
  - **From**: Technical University of Sofia, Bulgaria.
  - **Supervision**: Amor Nafkha (CS, **100%**).
  - **Related Publications**: 1 Chapter, 1 International Conference ([BC5], [CI45]).
  - **Abstract**: The main objective is to develop a centralized cooperative spectrum sensing system, implemented on Universal Software Radio Peripheral (USRP) hardware platforms driven by GnuRadio tool and its graphical user interface gnuradio-companion. Spectrum sensing is realized by energy detection and a new block of energy detector with uncertainty is developed using GNU Radio out-of-tree implementation. A centralized scheme for cooperative spectrum sensing is applied and a hard global decision is taken in a fusion center which collects the local decisions from secondary users, selects those of them which will be taken for global decision estimation and performs classical decision fusion logic, such as logic-and, logic-or, and majority rules. Based on measured data, the probabilities of detection for different signal to noise ratios are built for each secondary user and for different scenarios of cooperative spectrum sensing.
  - **Keywords**: Energy detector, Cooperative spectrum sensing, Gnu-Radio, USRP N210, Noise power estimation.
- **Mar. 2015 - Aug. 2015**: Julio César Manco Vásquez
  - **Visiting objective**: *Experimental Evaluation of Bayesian-Based Spectrum Sensing using USRP Platforms.*
  - **From**: University of Cantabria, Spain.

- **Supervision**: Amor Nafkha (CS, **100%**).
- **Related Publications**: Journal paper in progress.
- **Abstract**: The main objective is to evaluate a novel detection scheme that incorporate learning techniques for the development of intelligent radios in cognitive radio (CR) networks. Bayesian inference is applied at each sensing period and the posterior probabilities are utilized for detection. This spectrum seining scheme is implemented and tested in a SDR platform, USRP N210. The numerical results obtained in simulation-like environments are corroborated with those obtained in realistic scenarios.
- **Keywords**: Bayesian methods, Detectors, Covariance matrix, Cognitive radio, Approximation methods, Gnu Radio, USRP N210.

IV. Ongoing/former Post-Doc researchers

- **Sep. 2012 - Sep. 2013**: Babar Aziz
  - **Post-doc subject**: *Spectral Analyzer based on Software Defined Radio.*
  - **Funding Project**: PME SOFTRF.
  - **Advisors**: Prof. Jacques Palicot (Supélec , 20%), Prof. Daniel Le Guennec (Supélec , 30%), and Ass. Prof Amor Nafkha (Supélec , **50%**).
  - **Abstract**: We focus on blind identification of standards for Green Communications. The aim of Green Communications is to propose solutions to overcome the rapid increase in energy consumption in communication and networking devices. Cognitive Radio (CR) presents itself as a set of concepts which can help achieve the goals of Green Communications, since a CR device collects information from surroundings and improves its behavior accordingly. One way to achieve Green Communication is to select a standard for communication which requires less transmission power. Thus, the CR device must be capable of identifying all the standards in its surroundings in order to select one of them. We implement a bandwidth shape detector and a pilot based detector. These detectors are the building blocks of the so called Blind Standard Recognition Sensor (BSRS). The performance of the proposed method is evaluated both through simulations and experiments conducted using USRP n210 with Matlab/Simulink.
  - **Related Publications**: 2 journals ([J8], [J6]) and 6 international conferences ([CI40], [CI38], [CI37], [CI36], [CI33], [CI32]).
- **Sep. 2009 - Sep. 2010**: Mohamad Mroué
  - **Post-doc subject**: *Study and development of signal processing algorithms related to hierarchical modes and PAPR reduction for DVB-T2 system.*
  - **Funding Project**: PME DTTV2.

- **Advisors**: Prof. Jacques Palicot (Supélec, 30%) and Ass. Prof Amor Nafkha (Supélec , **70%**).
- **Abstract**: High Peak to Average Power Ratio (PAPR) is a critical issue in multi-carrier communication systems using Orthogonal Frequency Division Multiplexing (OFDM), as Second Generation Terrestrial Digital Video Broadcasting (DVB-T2) system. This problem can result in large performance degradation due to the non-linearity of the High Power Amplifier (HPA) or in its low power efficiency. We analyze and evaluate the performance of different Tone Reservation based techniques for PAPR reduction in DVB-T2 context. Also, we propose an iterative TR based technique called "One Kernel One Peak" (OKOP). It offers the advantage of controlling the power variation of the reserved subcarriers.
- **Related Publications**: 1 journal ([J4]) and 1 international conference ([CI15]).
- **Awards and Honors**: The DTTv2 project received the third prize Trophies Loading The Future, during the general assembly, Media & Networks Competitiveness Cluster, France, May 2011.

- **Sep. 2007 - Feb. 2009**: Julien Delorme
  - **Post-doc subject**: *Dynamically reconfigurable SoC/FPGA architecture for Software Defined Radio.*
  - **Funding Project**: EU E2R-Phase II, ANR IDROMEL.
  - **Advisors**: Prof. Christophe Moy (Supélec, 30%) and Ass. Prof Amor Nafkha (Supélec , **70%**).
  - **Abstract**: A cognitive radio is the final point of software-defined radio platform evolution: a fully reconfigurable radio that changes its communication modules depending on network and/or user demands. We only focus on the heterogeneous reconfigurable hardware platform for cognitive radio. software defined radio (SDR) basically refers to a set of techniques that permit the reconfiguration of a communication system without the need to change any hardware module. The goal of Software Defined Radio is to produce communication devices which can support several different services. These terminals must adapt their hardware structure in function of the wireless networks such as GSM, UMTS, IEEE 802.11a/b/g As a consequence, NoC offer good perspectives to future SoC in the way to satisfy SDR concept. Conception, validation and evaluation of solutions for NoC design is conducted through simulations. We are proposing to extend the NoC structure to a FPGA where dynamic partial reconfiguration is used to dynamically reconfigure the requested IP block of the wireless communication chain.
  - **Related Publications**: 6 international conference papers ([CI11], [CI10], [CI9], [CI8], [CI7], [CI6]).

2.3.2 Research Projects

Since 2008, I have participated/or currently participate as **principal investigator** (PI) or **project member** (PM) in over 14 research projects (internal, regional, national and international). A non-exhaustive list of projects is shown in Table 2.4.

Project	Internal	Regional	National	European	PHC	
SSR	x	-	-	-	-	PI
PHC-RILA	-	-	-	-	x	PI
SOFTRF	-	x	-	-	-	PI
DTTV2	-	x	-	-	-	PI
GREAT	-	x	-	-	-	PM
IoTrust	-	-	x	-	-	PM
UniThing	-	-	x	-	-	PM
WONG5	-	-	x	-	-	PM
SHARING	-	-	-	x	-	PM
NEWCOM#	-	-	-	x	-	PM
NEWCOM++	-	-	-	x	-	PM
Total	1	3	3	3	1	11

Table 2.4 – Summary of the projects where I am/was involved.

I. Principal Investigator

- **[2020-20XX] SSR (70 Keuros)**
  - **Project type**: AAP Mi-lourd 2019 (CentraleSupélec & Rennes Metropole)
  - **Collaborators**: Romain Bourdais (IETR) / Jean-François Lalande (IRISA)
  - **Title**: Smart and Secure Room
  - **Main focus**: The Smart and Secure Room project has two objectives: in the short term, it aims to develop a benchmarking platform in order to experimentally validate a set of works by three research teams: SCEE, AUT and IRISA. In the medium term, this experimental platform will be at the heart of emerging research topics related to multi-level detection (hardware, software and physics) of failures and attacks. Thus, the Smart and Secure Room project intends to equip a room with a set of heterogeneous systems (connected devices, programmable logic controllers, and server racks, ...) to make an experimental platform representative of the buildings of the future.
  - **Related publications**: No publications right now.
- **[2015-2017] PHC RILA (5 Keuros)**
  - **Project type**: Partenariats Hubert Curien (Campus France)
  - **Project Partners**: Technical University of Sofia / CentraleSup'elec

  - **Title**: USRP-based SDR for Cognitive Radio: Platform for cooperative spectrum Sensing and primary User LOcalization (**PSULO**)
  - **Main focus**: This proposal aims at pursuing the research activities led in collaboration between the Department of Technology and Management of Telecommunication Systems – Technical University, Sofia, Bulgaria and Signal, Communication & Embedded Electronics research (SCEE) team of the IETR. The main challenge of this project is to design spectrum sensing algorithm with a desired level of accuracy and sensing time taking into account the wireless channel hostile environment. The other requirement is that spectrum sensing has to be able to perform across wide band of frequencies which brings in additional implementation challenges. The main objective is both theoretical and practical investigation of advanced signal processing techniques used for cooperative spectrum sensing and primary user localization application in cognitive radio networks. All theoretical cooperative sensing and primary user localization algorithms will be implemented using GnuRadio tools and USRP platforms.
  - **Related publications**: [BC5] and [CI45].
- **[2012-2014] PME SOFTRF (105 Keuros)**
  - **Project type**: Media & Networks cluster funding (PME project)
  - **Project Partners**: CAPS / ENENSYS / Supélec
  - **Title**: SOFTRF
  - **Main focus**: The objective of the SOFTRF project is to specify and implement a low-cost SDR (Software-Defined Radio or Software Defined Radio) type system, allowing real-time analysis of an RF frequency band in the UHF range in order to determine: **(i)** the types of signals transmitted (Digital TV, Analogue TV, Digital Radio...), **(ii)** the main interference source affected these signals. The scientific and technological innovation of the SOFTRF project resided in the development of spectral estimation algorithms based on non-uniform sampling while taking into account the spectral occupancy rate of different received signals, the detection of free spectral resources, as well as in the identification of different existing standards inside a given frequency band.
  - **Related publications**: [J6], [CI36], [CI33], [CI32] and [CI29].
- **[2008-2010] PME DTTV2 (105 Keuros)**
  - **Project type**: Media & Networks cluster funding (PME project)
  - **Project Partners**: SIRADEL / ENENSYS / Supélec
  - **Title**: DTTv2
  - **Main focus**: The main objective of the DTTv2 project was to work on improvements of the digital terrestrial television standard and personal mobile television broad-

casting networks. The project covers three axis: **(i)** study of the state of the art and research work to define a strategy to improve the radio coverage of a DVB-T2 transmitter, **(ii)** elaboration of signal processing algorithms (hierarchical modulation, PAPR reduction, ...) to improve broadcasting, and **(iii)** development of a DVB-T2 prototype (modulator and propagation simulator) in order to make a real measurements of performances. The scientific innovation of the DTTv2 project resided in the development of new PAPR reduction techniques while taking into account the DVB-T2 standard. Thus, two main conditions have to be respected: **(i)** respect the DVB-T2 standard (*e.g.* number of reserved null sub-carriers to reduce PAPR), **(ii)** reduction of the hardware complexity (in terms of FPGA resources).

- **Related publications**: [J4] and [CI15].

II. Project Member

- **[2018-2021] IoTrust**
  - **Project type**: The French National Research Agency (ANR)
  - **Project Partners**: Orange / IRT-BCOM / Secure-IC / TDF / UR1 / CentraleSupélec
  - **Title**: Internet of Things Trust
  - **Main focus**: This project proposes to study of the cybersecurity of different IoT protocols corresponding to Low Power Wide Area Network. Those protocols can cover communication in area under 100km through radio signal with low power and are ideal for communication with huge number of IoT devices. The protocols under study are LoRaWAN, Sigfox, LTE-M, and NB-IoT. The project will provide an overview of the risk that may affect each protocol and its components and provide countermeasures that benefit to them. The objective is to give adapted protections versus exposure relatives to malicious users using internet network or radio communication by using dedicated countermeasures.
  - **Related Livrables**: 4.1 and 4.2.
  - **Collaborators**: Stéphane Paquelet (BCOM), David Bernard (Orange)
- **[2016-2018] UniThing**
  - **Project type**: The French National Research Agency (ANR)
  - **Project Partners**: Orange / IRT-BCOM / Kerlink / TDF / UR1 / CentraleSupélec
  - **Title**: Internet of Things devices integration in 5G networks
  - **Main focus**: UniThing was focused on the design and optimization of reconfigurable architecture for a multi-standard radio equipment. This project was mainly related to Internet of Things (IoT) devices and networks and it addressed the following research issues: **(i)** adaptable, flexible and scalable hardware architecture for analog and digital

front-end processing, **(ii)** various localization and data fusion methods, and **(iii)** quality of service (network densification and collision management.
  - **Related publications**: [J17], [J15], [CI55], [CI53], [CI49] and [CI48].
  - **Collaborators**: Stéphane Paquelet (BCOM), Ali Zeineddine (TDF), Prof. Christophe Moy (UR1), Pierre-Yves Jézéquel (TDF)
- **[2016-2019] WONG5**
  - **Project type**: The French National Research Agency (ANR)
  - **Project Partners**: CEA LETI / THALES / CNAM / CentraleSupélec
  - **Title**: Waveforms MOdels for Machine Types CommuNication inteGrating 5G Networks (ID: ANR-15-CE25-0005)
  - **Main focus**: The main objectives of the WONG5 project are to study and propose the most appropriate waveforms that are adapted to critical machine type communications (C-MTC). The derived waveforms will be used for the definition of a new physical layer adapted to this context. The C-MTC market is estimated at 50 billion connected machines by 2020, with the emergence of the Internet of Things. Requirements for C-MTC systems can be summarized as follows: **(i)** low latency, **(ii)** very high reliability and data integrity, **(iii)** high energy efficiency for mobile systems, and **(iv)** resistance to asynchronous users (time and frequency).
  - **Related publications**: [CI54].
  - **Collaborators**: Yves Louet (CentraleSupélec), Daniel Roviras (CNAM), Hmaied Shaiek (CNAM), Rafik Zayani (CNAM)
- **[2012-2016] SHARING**
  - **Project type**: Celtic-Plus project
  - **Project Partners**: Orange / Avea Labs / Ericsson / Mitsubishi Electric / THALES/ Magister Solutions/ European Communications Engineering / Sequans Communications / SIRADEL / IDATE Consulting / TTI Norte / University of Oulu/ Eurecom / CEA-LETI / Supélec
  - **Title**: Self-organized Heterogeneous Advanced RadIo Networks Generation (ID: C2012-1/8)
  - **Main focus**: The main objective of the SHARING project is to propose cost and power efficient high capacity broadband solutions by: **(i)** enabling a flexible interference management concept in order to trigger spectral efficiency increase in future heterogeneous networks, **(ii)** introducing smart and innovative off-loading strategies, as well as joint radio resource management (RRM) solutions across radio access technologies (RATs), and **(iii)** proposing a novel integrated architecture incorporating: seamless Inter-RAT service continuity, machine type communications, device-to-device transmissions and efficient licensed /unlicensed spectrum usage

  - **Related publications**: [CI52].
- **[2012-2015] NEWCOM#**
  - **Project type**: European Commission
  - **Project Partners**: CNRS, Eurecom, CNIT, Poznan University of Technology, CTTC, Bilkent Universitesi, Aalborg Universitet, Technische Universitaet Dresden, University of Cambridge, IASA, INON, Oulun Yliopisto, Université Catholique de Louvain, Technische Universitaet Wien, THALES, Orange, NEC, Avea, Agilent, Aeroflex, Renesas, Samsung, Telecom Italia, Telefonica
  - **Title**: Network of Excellence in Wireless Communications#
  - **Main focus**: One of the main objectives of Newcom# is to improve the dissemination of research results to European industry. The first task toward this objective is to make a survey and find out how the research topics covered by Newcom# are in line with the needs of European companies in terms of research in wireless communications. Newcom# pursues long-term, interdisciplinary research on the most advanced aspects of wireless communications like finding the ultimate limits of communication networks, opportunistic and cooperative communications, or energy-bandwidth efficient communications and networking.
  - **Related publications**: [CI41], [CI40], [CI38], [CI37] and [CI34].
  - **Collaborators**: Adrian Kliks (PUT), Krzysztof Cichon (PUT), Malek Naoues (CNRS), Aziz Babar (CNRS)
- **[2012-2014] GREAT**
  - **Project type**: UEB / CominLabs Chair
  - **Project Partners**: Supélec / UEB / CominLabs
  - **Title**: Green Cognitive Radio for Energy-Aware wireless communication Technologies evolution
  - **Main focus**: The main objective of this research program is to define a novel framework and representative functionalities of energy-efficient cognitive radio to enable green radio future wireless communications with efficient and sustainable energy utilization. Moreover, this project provides some demonstrations as proof-of-concept, in order to validate the advantage, efficiency and feasibility of the proposed green cognitive radio approach.
  - **Related publications**: [CI32], [CI30] and [CN3].
  - **Collaborators**: Ziad Khalaf (Supélec), Aziz Babar (Supélec), Pr. Honggang Zhang (UEB/CominLabs Chair Professor, Zhejiang University)
- **[2008-2011] NEWCOM++**
  - **Project type**: UEB / CominLabs Chair
  - **Project Partners**: ISMB, BILKENT / KHAS, TECHNION, IASA, CNIT, UPC,

CTTC, IST-TUL, CNRS, CEA-LETI, LNT-TUM, RWTH, UGENT, TUW, PUT, CHALMERS / KAU, AAU.

- **Title**: Network of Excellence in Wireless COMmunications ++
- **Main focus**: The fundamental premise behind Newcom++ can be summarized as the research challenges posed by the target outcome "Ubiquitous network infrastructure and architectures" are by their nature very complex and interdisciplinary, thus requiring a fully committed and large-scale collaboration between strong research groups in different disciplines, a collection that can hardly be found at a single institution/country. The main manner by which Newcom++ intends to promote solutions to the above mentioned problems and challenges is by creating a trans-European virtual research centre linking a proper number of leading European research groups in a highly integrated, carefully harmonized, cooperative fashion.
- **Related publications**: [CI11], [CI10] and [CI9].
- **Collaborators**: Julien Delorme (Supélec), Christophe Moy (Supélec)

2.3.3 Editorial Service

- **Guest Editor**
  - **Journal**: Hindawi International Journal of Antennas and Propagation
  - **Special Issue**: Antennas and RF Front Ends for Cognitive Radio
  - **Date**: April 2015.

2.3.4 Invited Talks

- **RILA Project**
  - **Talk on**: Overview on spectrum sensing techniques for cognitive radio networks.
  - **Project**: *Platform for cooperative spectrum sensing and primary user localization.*
  - **Place/Date**: Sofia, Bulgaria, Jul. 26 - Aug. 1, 2015.
- **Journée thématique de la DGA-MI**
  - **Talk on**: Détecteur aveugle associant la structure parcimonieuse et la propriété de symétrie de la fonction d'autocorrélation cyclique.
  - **Thematic day**: *Sensing appliqué à la Radio Cognitive.*
  - **Place/Date**: Bruz, France, March 21, 2013.

2.3.5 Conference Organization

- **URSI GASS 2020 Commission C**
  - **Technical session**: Multi-antenna technologies and Massive MIMO
  - **Special Issue**: Antennas and RF Front Ends for Cognitive Radio

- **Co-organizer**: Prof. M. Crussiere
- **Place/Date**: Sapienza University Campus, Rome, Italy. Aug. 29 – Sep. 5, 2020.

- **IEEE ICT 2018**
  - **Organisation committee**: Publication co-chairs
  - **Session Chair**: MIMO Systems and related Signal Processing
  - **Conference**: IEEE International Conference on Telecommunications
  - **Place/Date**: Saint-Malo, France, Jun. 26 - Jun. 28, 2018.
- **IEEE ISWCS 2016**
  - **Special Session**: "Low Power Design Techniques for Embedded Systems"
  - **Co-organizer**: Prof. Yves Louet
  - **Conference**: IEEE International Symposium on Wireless Communication Systems
  - **Place/Date**: Poznan, Poland, Sep. 20 – Sep. 23, 2016.
- **IEEE Next-GWiN 2014**
  - **Organisation committee**: Demo Chair
  - **Conference**: International Workshop on Next Generation Green Wireless Networks.
  - **Place/Date**: Rennes, France, Oct. 1 – Oct. 3, 2014.
- **IARIA AICT 2010**
  - **Special Area Chairs**: Cognitive Radio
  - **Panelists**: Challenges in Advanced Communications and Services
  - **Session chair**: Cognitive Radio III
  - **Conference**: Advanced International Conference on Telecommunications
  - **Place/Date**: Barcelona, Spain, May 09 – May 15, 2010.
- **IEEE DTIS 2010**
  - **Session chair**: SOC, SIP Design and Wireless Systems
  - **Conference**: International Conference on Design & Technology of Integrated Systems in Nanoscale Era
  - **Place/Date**: Hammamet, Tunisia, Mar. 23 – Mar. 25, 2010.

2.4 About my PhD thesis

In comparison with conventional single-input single-output (SISO) system, the multiple-input multiple-output (MIMO) system can make full use of space resources and increase the channel capacity without increasing the bandwidth since each antenna at the receiver can receive signals transmitted from all the transmitting antennas simultaneously [1, 2, 3, 4, 5, 6, 7, 8]. MIMO techniques can be categorized as:

- Diversity techniques, (*increase the link quality*), intend to receive several replicas of the same transmitted signal, while assuming that at least some of them are not severely atten-

uated by the channel. Therefore, diversity techniques are able to combat the fluctuations in the signal-to-noise ratio and obtain an improved bit error rate (BER) performance.

- Spatial multiplexing (SM) techniques, (*increase the data rate*), send independent data streams on different transmit antennas in order to improve the spectral efficiency of a wireless link and increase data rate but not necessarily reliability.
- Beam-forming intend to use multiple antennas to form beams and increase the signal-to-interference-plus-noise ratio (SINR), and thereby improve the system capacity.

The main focus of my PhD dissertation (2003-2006) is to propose a low-complexity and high-performance detection technique for spatial multiplexing (SM) MIMO systems with perfect channel state information (CSI) at the receiver side. One major implementation difficulty of the spatial multiplexing MIMO scheme is the signal separation (detection) problem at the receiver side. This detection problem can be defined as follows: Given an $t \in \mathbb{N}_{>0}$ input linear system whose transfer function is described by a matrix, $\mathbf{H} \in \mathbb{C}^{r \times t}$, having non-orthogonal columns and its $r \in \mathbb{N}_{>0}$ outputs are contaminated by additive random noise. The relationship between the inputs and the outputs of this linear system can be expressed as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{b} \tag{2.1}$$

where $\mathbf{y} \in \mathbb{C}^r$ is the received signal vector, and $\mathbf{b} \in \mathbb{C}^r$ denotes complex zero-mean additive white Gaussian noise vector with covariance $\sigma^2\mathbf{I}_{r\times r}$. $\mathbf{H}$ is the channel matrix, that characterizes the input-output relation, whose entries are independent identically distributed (i.i.d) complex Gaussian random variables with zero mean and unit variance, and $\mathbf{x} \in \mathbb{A}^t$ is a transmitted symbol vector, uncorrelated with the noise vector $\mathbf{b}$, whose elements are drawn from a symbol constellation set $\mathbb{A}$. The MIMO detection problem aims to detect the transmitted vector $\mathbf{x}$ based on the received signal $\mathbf{y}$ and the channel matrix $\mathbf{H}$. For uncoded SM–MIMO, the maximum likelihood detection problem for the system model of (2.1) can be reformulated as integer constrained least squares optimization problem of

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{A}^t} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 \tag{2.2}$$

$\mathbb{A}^t$ is the lattice whose elements represent all possible combinations of transmitted vector $\mathbf{x}$. Due to the discrete nature of $\mathbb{A}$, (2.2) is a non-deterministic polynomial-time hard (NP-hard) problem [9], and can only be exactly solved by exhaustive search over all possible transmitted vector $\mathbf{x} \in \mathbb{A}^t$. In this case, the detection's complexity grows exponentially grows exponentially with the cardinality of constellation set $\mathbb{A}$ and the antenna dimensions.

The main objective of my PhD was to propose a novel near-optimal maximum likelihood detection algorithm for the SM-MIMO system. In order to develop the novel detection scheme, I addressed the dual challenges of quasi-optimal ML detection performances and a low fixed

computational complexity associated with an inherent parallel hardware structure. My dissertation was approximately 150 pages and included seven chapters. Main technical contributions are presented in Chapter 4 on "*novel algorithm for near-optimal MIMO detection*", Chapter 5 on "*the extension of the proposed algorithm in higher-order modulation schemes*", and Chapter 6 on "*FPGA implementation of proposed algorithm for SM-MIMO system*". In Chapter 4, I developed a new parallel low-complexity MIMO detection schema, called geometrical intersection and selection detector (GISD) technique, in order to achieve near optimal maximum likelihood detection performance. The proposed algorithm was based on diversification and intensification steps. The diversification step, also called *exploration*, is about discovering new promising initial feasible solutions for the ML detection problem (2.2), and ensures that the proposed algorithm explores the totality of search space $\mathbb{A}^t$. Based on the D smallest right singular vectors of the MIMO channel matrix $\mathbf{H}$, and a geometrical interpretation of the objective function ($\|\mathbf{y}-\mathbf{H}\mathbf{x}\|_2^2$), the diversification phase provides a list of initial "smart" solutions. In contrast, the intensification step, also called *exploitation*, involves the refinement of the "smart" solutions by performing a 1-flip or 2-flip neighborhood local search algorithm. The performances of proposed scheme have been analyzed and published in [PA1], [J1], [CI2], [CI5], [CI28], and [CN1]. The main contributions in Chapter 5 are the extension of the proposed scheme to sixteen quadrature amplitude modulation (16-QAM) signal constellation and the derivation of a novel list-based soft-output MIMO detection algorithm that can be used in iterative detection and decoding (IID) for coded SM-MIMO systems. It should be noted that for 16-QAM constellation, the performance of the hard-output version of the proposed detection algorithm is not as promising as I expected after interesting results obtained in the case of BPSK and QPSK modulation schemes. A soft-output MIMO detector, called list-based geometrical intersection and selection detector (LGISD), is proposed. Unlike the existing list-based soft-output methods like list-based sphere decoding, the worst-case computational cost of the LGISD is bounded by a low-order polynomial of the number of bits to be demodulated. Moreover, simulation results suggest that the LGISD computational advantage is obtained without incurring a significant degradation in bit error rate (BER) performance. The third thesis contribution (Chapter 6) is related to FPGA-based hardware implementation of the GISD detector for conventional small-scale MIMO systems. Globally asynchronous locally synchronous (GALS) technique, which divide the chip area into several independent synchronous clusters, has been intensively used to design and analyze the hardware architecture of the GISD MIMO detector.
The main results of my PhD dissertation can be summarized as follows

- I have proposed and investigated the performance of a novel detection technique for conventional small-scale SM-MIMO systems. The proposed algorithm, called GISD (geometrical intersection and selection detector), is based on diversification and intensification strategies. The GISD detector provided an easier hardware implementation while keeping

a quasi-optimal BER performance and fixed-low computational complexity.

- In contrast to the existing MIMO detection methods, the numerical simulations shown that the proposed GISD detector offers excellent performance-versus-complexity tradeoff. The GISD has a computational complexity of the order of $\mathcal{O}(t^3)$, and typically it has a much more attractive error probability performance than the existing MIMO detectors. The proposed algorithm is highly parallelizable and seems well suited for FPGA-based hardware implementation.
- Diversification phase is motivated by the geometrical interpretation of the ML detection problem (2.2). The principal idea is that the optimal ML solution will be in the proximity of the D lines, originated from the r-dimensional point $\mathbf{x}_{zf} = \mathbf{H}^{+} \times \mathbf{y}$, and directed by the D smallest right singular vectors of the channel matrix $\mathbf{H}$. Then, for each of the D lines, a subset $\mathbb{S}_k \subset \mathbb{A}^t$ can be generated using one of the following developed methods: (**i**) hypercube intersection and selection (HIS), (**ii**) plane intersection and selection (PIS), and (**iii**) basis intersection and selection (BIS). Finally, a subset $\mathbb{S} = \cup_{k=1}^{D} \mathbb{S}_k \subset \mathbb{A}^t$, called "smart" initial feasible points, will be generated for the intensification phase.
- Intensification phase, also called greedy search, perform a local search around each "smart" feasible point $\mathbf{x}_s \in \mathbb{S}$ by evaluating a set of points belonging the neighborhood of Hamming distance one from $\mathbf{x}_s$. The greedy search process continues until convergence to local minima in neighborhood of $\mathbf{x}_s$ or until the maximum number of iterations is reached. In general, the maximum number of iterations will be fixed to 2. For a given initial "smart" feasible point, the intensification phase has a computational complexity of the order of $\mathcal{O}(t^2)$.
- Novel soft-output version of the geometrical intersection and selection detector (GISD) previously proposed for uncoded SM-MIMO systems, was also presented. Thus, the soft-GISD, also called LGISD, can be used in turbo-SM-MIMO systems to exchange extrinsic soft-information with the outer decoder. The effect of channel estimation error on the performance of the LGISD detector was investigated. Simulation results shown that the proposed hard-output (*resp.* soft-output) detector introduced only a small performance degradation compared to optimal sphere decoder [10, 11, 12, 13, 14] (*resp.* list sphere decoder [15, 16, 17, 18]).
- The GISD functionality is implemented and optimized with VHDL on RTL level. I used globally asynchronous locally synchronous method to design different GISD blocks on synchronous commercial XC2V2000-6FF896C virtex-ii FPGA. The fixed-point arithmetic is used for the GISD architecture and I demonstrated that the proposed detector running on a FPGA device can achieve a throughput of the order of a 2 Mbps which is sufficient for tasks like voice and/or video communications over 5×5 SM-MIMO system with QPSK modulations.

BIBLIOGRAPHY

[1] J. H. Winters, "On the capacity of radio communication systems with diversity in a rayleigh fading environment," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 5, pp. 871–878, 1987.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[3] E. Telatar, "Capacity of multi?antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.

[5] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[6] J. B. Andersen, "Antenna arrays in mobile communications: gain, diversity, and channel capacity," *IEEE Antennas and Propagation Magazine*, vol. 42, no. 2, pp. 12–16, 2000.

[7] G. J. Foschini, D. Chizhik, M. J. Gans, C. Papadias, and R. A. Valenzuela, "Analysis and performance of some basic space-time architectures," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 303–320, 2003.

[8] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of mimo communications - a key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–218, 2004.

[9] D. Micciancio, "The hardness of the closest vector problem with preprocessing," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1212–1215, 2001.

[10] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Mathematics of Computation*, no. 170, pp. 463–471, 1985.

[11] Z. Xie, C. K. Rushforth, and R. T. Short, "Multiuser signal detection using sequential decoding," *IEEE Transactions on Communications*, vol. 38, no. 5, pp. 578–583, 1990.

[12] C. P. Schnorr and M. Euchner, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," *Mathematical Programming*, vol. 66, no. 1, pp. 181–199, 1994.

[13] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1639–1642, 1999.

[14] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Transactions on Information Theory*, vol. 48, no. 8, pp. 2201–2214, 2002.

[15] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 389–399, 2003.

[16] J. Boutros, N. Gresset, L. Brunel, and M. Fossorier, "Soft-input soft-output lattice sphere decoder for linear channels," in *GLOBECOM '03. IEEE Global Telecommunications Conference*, vol. 3, 2003, pp. 1583–1587 vol.3.

[17] C. Studer, A. Burg, and H. Bocskei, "Soft-output sphere decoding: algorithms and vlsi implementation," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 2, pp. 290–300, 2008.

[18] C. Studer and H. Bocskei, "Soft-input soft-output single tree-search sphere decoding," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4827–4842, 2010.

Chapter 3

Scientific Production

This chapter gives the complete list of my publication: patents, book chapters, journals, and conferences. Note that the underlined authors are the Post-Doc and/or the PhD students under my advising or co-advising, and the bold authors are visiting PhD students under my supervision. [CNxx] and [CIxx] denote national and international conference paper number xx, respectively.

Publication in	**Number**
Patents [PAxx]	1
Book chapters [CHxx]	6
Journals [Jxx]	20
International conferences [CIxx]	56
National conferences [CNxx]	3
Total	**84**

Table 3.1 – Summary of all my published papers (up to Apr. 1^{st}, 2021)

3.1 Publication Indicators

- **ORCID iD**: 0000-0002-1164-7163
- **DBLP Publications list**: https://dblp.uni-trier.de/
- **CNRS HAL ID**: http://haltools.archives-ouvertes.fr/
- **IEEE Publications list**: https://ieeexplore.ieee.org/
- **Google Scholar**: https://scholar.google.com/

3.2 Patents

- [PA1] A. Nafkha, E. Boutillon, "*Procédé de décodage de codes matriciels, module de décodage associé, et applications mettant en oeuvre un tel procédé* ", Brevet Français, number: 05 02664.

3.3 Journals papers

Below, I listed all my journals papers and their corresponding impact factor according to the latest impact factor list of 2020 provided by the journal citation report (JCR).

- [J20] B. Trotobas, A. Llave, A. Nafkha and Y. Louet, *When Should We Use Geometrical-Based MIMO Detection Instead of Tree-Based Techniques? A Pareto Analysis*, in IEEE Access, 2020. **IF. 3.745**
- [J19] A. Nafkha, *Standard Condition Number Based Spectrum Sensing Under Asynchronous Primary User Activity*, in IEEE Access, 2020. **IF. 3.745**
- [J18] A. Nafkha, N. Demni, *Closed-Form Expressions of Ergodic Capacity and MMSE Achievable Sum Rate for MIMO Jacobi and Rayleigh Fading Channels*, in IEEE Access, 2020. **IF. 3.745**
- [J17] A. Zeineddine, A. Nafkha, S. Paquelet, C. Moy, P. Jezequel, *Comprehensive Survey of FIR-based Sample Rate Conversion*, VLSI signal processing systems, 2020. **IF. 1.013**
- [J16] B. Trotobas, Y. Akourim, A. Nafkha, Y. Louet, J. Weiss, *Evaluation of the complexity, performance and implementability of geometrical MIMO detectors: The example of the exploration and exploitation list detector*, in International Journal on Advances in Telecommunications. Vol.13(1&2), 2020. **IF. 0.654**
- [J15] S. Paquelet, A. Zeineddine, A. Nafkha, P. Jezequel, C. Moy, *Convergence of the Newton Structure Transfer Function to the Ideal Fractional Delay Filter*, in IEEE Signal Processing Letters, vol. 26, no. 9, pp. 1354-1358, 2019. **IF. 3.105**
- [J14] H. Kobeissi, Y. Nasser, O. Bazzi, A. Nafkha, Y. Louet, *ELASTIC-Enabling Massive-Antenna for Joint Spectrum Sensing and Sharing: How Many Antennas Do We Need?*, in IEEE Transactions on Cognitive Communications and Networking, vol. 5, no. 2, 2019. **IF. 4.574**
- [J13] A. Nafkha, R. Bonnefoi, *Upper and lower bounds for the ergodic capacity of MIMO Jacobi fading channels*, in OSA Opt. Express, vol. 25, pp. 12144-12151, 2017. **IF. 3.561**
- [J12] H. Kobeissi, Y. Nasser, A. Nafkha, O. Bazzi and Y. Louet, *Asymptotic Approximation of the Standard Condition Number Detector for Large Multi-Antenna Cognitive Radio Systems*, in EAI Endorsed Transactions on Cognitive Communications 17 (11): e1, vol. 3, 2017. **IF. n.i.**
- [J11] H. Kobeissi, A. Nafkha, Y. Nasser, O. Bazzi and Y. Louet, *On the Performance Analysis and Evaluation of Scaled Largest Eigenvalue in Spectrum Sensing: A Simple Form Approach*, in EAI Endorsed Transactions on Cognitive Communications 17 (10): e5, vol. 3, 2017. **IF. n.i.**
- [J10] H. Kobeissi, A. Nafkha, Y. Nasser, O. Bazzi, Y. Louet, *On Approximating the Standard Condition Number for Cognitive Radio Spectrum Sensing with Finite Number of Sensors*, in IET Signal Processing, vol. 11, Issue 2, pp. 145-154, 2017. **IF. 1.754**

- [J9] H. Kobeissi, Y. Nasser, A. Nafkha, O. Bazzi, Y. Louet, *On the Detection Probability of the Standard Condition Number Detector in Finite Dimensional Cognitive Radio context*, in EURASIP Journal on Wireless Communications and Networking, 2016. **IF. 1.487**
- [J8] A. Nafkha, B. Aziz, *Closed-Form Approximation for the Performance of Finite Sample-Based Energy Detection Using Correlated Receiving Antennas*, in IEEE Wireless Communications Letters, vol. 3, no. 6, pp. 577-580, 2014. **IF. 4.66**
- [J7] M. Al-Husseini, A. El-Hajj, M. Bkassiny, S. El-Khamy, A. Nafkha, *Antennas and RF Front Ends for Cognitive Radio*, (editorial) Special Issue on Antennas and RF Front Ends for Cognitive Radio, Hindawi International Journal of Antennas and Propagation, 2014. **IF. 1.207**
- [J6] L. Safatly, B. Aziz, A. Nafkha, Y. Louet, Y. Nasser, A. El-Hajj, K. Kabalan, *Blind spectrum sensing using symmetry property of cyclic auto-correlation function: from theory to practice*, in EURASIP Journal on Wireless Communications and Networking, 26, 2014. **IF. 1.487**
- [J5] M. Hentati, A. Nafkha, P. Leray, J-F. Nezan, M. Abid, *Software defined radio equipment: What's the best design approach to reduce power consumption and increase reconfigurability?*, International Journal of Computer Applications, 45 (14), pp. 26-32, 2012. **IF. 0.45**
- [J4] M. Mroue, A. Nafkha, J. Palicot, B. Gavalda, N. Dagorne, *Performance and Implementation Evaluation of TR PAPR Reduction Methods for DVB-T2*, in Hindawi Journal on Digital Multimedia Broadcasting, Article ID 797393, 2010. **IF. n.i.**
- [J3] Z. Khalaf, A. Nafkha, J. Palicot, M. Ghozzi, *Low Complexity Enhanced Hybrid Spectrum Sensing Architectures for Cognitive Radio Equipment*, in International Journal on Advances in Telecommunications, vol. 3, no. 3 and 4, pp. 215-227, 2010. **IF. 0.654**
- [J2] A. Nafkha, E. Boutillon and C. Roland, *Quasi-maximum-likelihood detector based on geometrical diversification greedy intensification*, in IEEE Transactions on Communications, vol. 57, no. 4, pp. 926-929, 2009. **IF. 5.646**
- [J1] D. Gnaedig, E. Boutillon, E. Martin, A. Nafkha, J. Tousch, M. Jézéquel, N. Brengarth, *High-level synthesis for behavioral design of the MAP algorithm for turbo decoder*, in Les Annales des Télécommunications, 2004. **IF. 1.412**

3.4 Book Chapters

- [BC6] B. Trotobas, A. Nafkha, Y. Louet, *A Review to Massive MIMO Detection Algorithms: Theory and Implementation*, Radio Frequency Antennas For 5G, IOT and Medical Applications, Intech open, ISBN 978-1-83968-345-9, DOI: 10.5772/intechopen.93089, Edited by Dr. Albert Sabban, 2020.

- [BC5] Z. Tchobanova, G. Marinova, A. Nafkha, *USRP-based Implementations of Various Scenarios for Spectrum Sensing*, Chapter 15, Advances in Networks, Security and Communications: Reviews, Volume 1, IFSA Publishing, pp. 363-388, ISBN:978-84-697-8994-0, Edited by Sergey Y. Yurish, 2018.
- [BC4] A. Nafkha, P. Leray, Y. Louet, J. Palicot, *Moving a processing element from hot to cool spots: is this en efficient method to decrease leakage power consumption in FPGAs?*, Chapter 8, Green Communications: Theoretical Fundamentals, Algorithms and Applications, CRC Press, ISBN 9781466501072, Edited by Jinsong Wu, Sundeep Rangan, and Honggang Zhang, 2012.
- [BC3] A. Nafkha, C. Moy, P. Leray, R. Seguier, J. Palicot, *Software Defined Radio Platform for Cognitive Radio: Design and Hierarchical Management*, Chapter 15, Recent Advances in Wireless Communications and Networks, Intech open, ISBN 978-953-307-274-6, Edited by Jia-Chin Lin, 2011.
- [BC2] A. Nafkha, P. Leray, C. Moy, *Implementation Platforms*, Chapter 11, Radio Engineering: from software radio to cognitive radio, WILEY, ISBN: 978-1-84821-296-1, Edited by Jacques Palicot, 2011.
- [BC1] A. Nafkha, P. Leray, C. Moy, *Plate-forme d'exécution*, Chapter 11, De la radio logicielle á la radio intelligente, Collection Télécom, Hermes, ISBN: 9782746225985, Edited by Jacques Palicot, 2010.

3.5 International Conferences

Following are listed all my publications in international conference with peer review process.

- [CI56] B. Trotobas, Y. Akourim, A. Nafkha, Y. Louet, *Adding Exploration to Tree-Based MIMO Detectors Using Insights from Bio-Inspired Firefly Algorithm*, IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), virtual conference, Apr.25 - Apr.28, 2021.
- [CI55] A. Haddad, D. Slimani, A. Nafkha and F. Bader, *Users' Power Multiplexing Limitations in NOMA System over Gaussian Channel*, IEEE International Conference on Wireless Networks and Mobile Communications (WINCOM), Reims, France, Oct. 2020.
- [CI54] A. Zeineddine, S. Paquelet, A. Nafkha, P. Jezequel and C. Moy, *Efficient Arbitrary Sample Rate Conversion for Multi-Standard Digital Front-Ends*, IEEE International New Circuits and Systems Conference (NEWCAS), Munich, Germany, Jun. 2019.
- [CI53] A. Zeineddine, S. Paquelet, M. Kanj, C. Moy, A. Nafkha, P-Y Jezequel, *Reconfigurable Newton Structure for Sample Rate Conversion*, IEEE Global Conference on Signal and Information Processing (GlobalSIP), Anaheim, California, USA, Nov. 2018.
- [CI52] B. Trotobas, A. Nafkha, *Fixed Complexity Soft-Output Detection Algorithm Through Exploration and Exploitation Processes*, The Fourteenth Advanced International Confer-

ence on Telecommunications, AICT 2018, Barcelone, Spain, Jul. 2018.

- [CI51] R. Bonnefoi, C. Moy, J. Palicot, A. Nafkha, *Low-Complexity Antenna Selection for Minimizing the Power Consumption of a MIMO Base Station*, The Fourteenth Advanced International Conference on Telecommunications, AICT 2018, Barcelone, Spain, Jul. 2018.
- [CI50] V. Gouldieff, A. Nafkha, N. Grollier, J. Palicot, S. Daumont, *Cyclic Autocorrelation based Spectrum Sensing: Theoretical Derivation Framework*, 25th International Conference on Telecommunications, ICT 2018, Saint Malo, France, Jun. 2018.
- [CI49] A. Zeineddine, A. Nafkha, C. Moy, S. Paquelet, P-Y Jezequel, *Variable Fractional Delay Filter: A Novel Architecture Based on Hermite Interpolation*, 25th International Conference on Telecommunications, ICT 2018, Saint Malo, France, Jun. 2018.
- [CI48] A. Zeineddine, S. Paquelet, A. Nafkha, C. Moy, P-Y Jezequel, *Generalization and Coefficients Optimization of the Newton Structure*, 25th International Conference on Telecommunications, ICT 2018, Saint Malo, France, Jun, 25-27,2018.
- [CI47] R. Bonnefoi, A. Nafkha, *A New Lower Bound on the Ergodic Capacity of Optical MIMO Channels*, IEEE ICC 2017 Optical Networks and Systems Symposium. ICC'17 ONS, Paris, France, May. 21-25, 2017.
- [CI46] A. Nafkha, Y. Louet, *Accurate Measurement of Power Consumption Overhead During FPGA Dynamic Partial Reconfiguration*, International Symposium on Wireless Communication Systems. ISWCS'2016, Poznan, Poland, Sept. 20-23, 2016.
- [CI45] Z. Tchobanova, A. Nafkha, G. Marinova, *Implementation of an Energy Detection Based Cooperative Spectrum Sensing on USRP Platforms for a Cognitive Radio Networks*, The Twelfth Advanced International Conference on Telecommunications, AICT'2016, Valencia, Spain, May 22-26, 2016. (Best Paper Award)
- [CI44] S. Jagdish Darak, A. Nafkha, C. Moy, J. Palicot, *Is Bayesian Multi-armed Bandit Algorithm Superior? : Proof-of-Concept for Opportunistic Spectrum Access in Decentralized Networks*, International Conference on Cognitive Radio Oriented Wireless Networks, Grenoble, France, June 1, 2016.
- [CI43] H. Kobeissi, Y. Nasser, A. Nafkha, O. Bazzi, Y. Louet, *A Simple Formulation for the Distribution of the Scaled Largest Eigenvalue and application to Spectrum Sensing*, International Conference on Cognitive Radio Oriented Wireless Networks, Grenoble, France, June 1, 2016.
- [CI42] H. Kobeissi, A. Nafkha, Y. Nasser, O. Bazzi, Y. Louet *Simple and Accurate Closed-Form Approximation of the Standard Condition Number Distribution with Application in Spectrum Sensing*, International Conference on Cognitive Radio Oriented Wireless Networks, Grenoble, France, June 1, 2016.
- [CI41] A. Nafkha, M. Naoues, K. Cichon, A. Kliks, B. Aziz, *Hybrid Scheme for Spectrum*

Sensing: Experimental Analysis, European Conference on Networks and Communications: Track 9 - Posters, EuCNC 2015, Paris, France, Jun. 29 - Jul. 2, 2015.

- [CI40] A. Nafkha, B. Aziz, M. Naoues, A. Kliks, *Experimental Study on Cyclostationary Feature and Eigenvalue based Algorithms for Spectrum Sensing*, European Conference on Networks and Communications: Track 7: Special Sessions, EuCNC 2015, Paris, France, Jun. 29 - Jul. 2, 2015.
- [CI39] C. Moy, A. Nafkha, M. Naoues, *Reinforcement Learning Demonstrator for Opportunistic Spectrum Access on Real Radio Signals*, IEEE International Symposium on Dynamic Spectrum Access Networks, DySPAN 2015, Sweden, Sept.29-Oct.2, 2015.
- [CI38] A. Nafkha, B. Aziz, M. Naoues, A. Kliks, *Cyclostationarity-Based Versus Eigenvalues-Based Algorithms for Spectrum Sensing in Cognitive Radio Systems: Experimental Evaluation Using GNU Radio and USRP*, International Workshop on Selected Topics in Mobile and Wireless Computing, IEEE WiMob'2015, Abu Dhabai, UAE, Oct. 19-21, 2015.
- [CI37] A. Nafkha, M. Naoues, K. Cichon, A. Kliks, B. Aziz, *Hybrid Spectrum Sensing Experimental Analysis Using GNU radio and USRP for Cognitive Radio*, International Symposium on Wireless Communication Systems, Track 2: Networking, protocols, cognitive radio, wireless sensor networks, services and applications, ISWCS'2015, Brussels, Belgium, Aug. 25-28, 2015.
- [CI36] B. Aziz, S. Traore, A. Nafkha, D. Le Guennec, *Spectrum Sensing for Cognitive Radio using Multi-coset Sampling*, IEEE Global Communications Conference, Globecom'2014, Austin, USA, Dec 8-12, 2014.
- [CI35] H. kobeissi, Y. Nasser, O. Bazzi, Y. Louet, A. Nafkha, *On the Performance Evaluation of Eigenvalue-Based Spectrum Sensing Detector for MIMO Systems*, URSI General Assembly and Scientific Symposium , URSI'2014, Beijing, China, Aug 16-23, 2014.
- [CI34] A. Nafkha, M. Naoues, K. Cichon , A. Kliks, *Experimental Spectrum Sensing Measurements using USRP Software Radio Platform*, International Conference on Cognitive Radio Oriented Wireless Networks, Oulu, Jun. 2-4, 2014. (Invited Paper)
- [CI33] B. Aziz, A. Nafkha, *Implementation of Blind Cyclostationary Feature Detector for Cognitive Radios Using USRP*, International Conference on Telecommunications , ICT'2014, Lisbon, Portugal, May 5-7, 2014.
- [CI32] B. Aziz, A. Nafkha, J. Palicot, H. Zhang, *Blind Wireless Standard Identification for Green Radio Communications*, International Conference on Advanced Technologies for Communications , ATC'2013, Ho Chi Minh City, Vietnam, Oct. 16-18, 2013.
- [CI31] I. Elleuch, F. Abdelkefi, A. Nafkha, M. Siala, *Efficient Limited Data Multi-Antenna Compressed Spectrum Sensing Exploiting Angular Sparsity*, International Symposium on Personal, Indoor and Mobile Radio Communications, London, UK, Sept. 8-11, 2013.
- [CI30] Z. Khalaf, J.Palicot, A. Nafkha, H. Zhang, *Blind Free Band Detector Based on the*

Sparsity of the Cyclic Auto-correlation Function, European Signal Processing Conference, EUSIPCO'2013, Marrakech, Morocco, Sep. 9-13, 2013.

- [CI29] S. Bahamou, A. Nafkha, *Noise Uncertainty Analysis of Energy Detector: Bounded and Unbounded Approximation Relationship*, European Signal Processing Conference, EUSIPCO'2013, Marrakech, Morocco, Sep. 9-13, 2013.
- [CI28] A. Nafkha, *Near Maximum Likelihood Detection Algorithm based on 1-flip Local Search over Uniformly Distributed Codes*, IEEE International Communication Conference , ICC'2013, Budapest, Hungary, Jun. 9-13, 2013.
- [CI27] A. Nafkha, J. Palicot, P. Leray, Y. Louet, *Leakage Power Consumption In FPGAs: Thermal Analysis*, International Symposium on Wireless Communication Systems, ISWCS'2012, Paris, France, Aug. 28-31, 2012. (Invited Paper)
- [CI26] Z. Khalaf, A. Nafkha, J. Palicot, *Blind Spectrum Detector for Cognitive Radio Using Compressed Sensing and Symmetry Property of the Second Order Cyclic Autocorrelation*, International Conference on Cognitive Radio Oriented Wireless Networks and Communications, Stockholm, Sweden, Jun. 18-20, 2012.
- [CI25] Z. Khalaf, A. Nafkha, J. Palicot, *Blind Spectrum Detector for Cognitive Radio using Compressed Sensing*, IEEE Global Communications Conference, Globecom'2011, Houston, Texas, USA, Dec. 5-9, 2011.
- [CI24] Z. Khalaf, A. Nafkha, J. Palicot, *Blind cyclostationary feature detector based on sparsity hypotheses for cognitive radio equipment*, IEEE International Midwest Symposium on Circuits and Systems, Seoul, South Korea, Aug. 7-10, 2011. (Invited Paper)
- [CI23] W. Jouini, R. Bollenbach, M. Guillet, C. Moy, A. Nafkha, *Reinforcement learning application scenario for Opportunistic Spectrum Access*, IEEE International Midwest Symposium on Circuits and Systems, Seoul, South Korea, Aug. 7-10, 2011.
- [CI22] Z. Khalaf, A. Nafkha, J. Palicot, *Enhanced Hybrid Spectrum Sensing Architecture for Cognitive Radio Equipment*, URSI General Assembly and Scientific Symposium, URSI GASS'2011, Istanbul, Turkey, Aug. 13-20, 2011.
- [CI21] P. Leray, A. Nafkha, C. Moy, *Implementation Scenario for Teaching Partial Reconfiguration of FPGA*, International Workshop on Reconfigurable Communication Centric Systems-on-Chip, ReCoSoC'2011, Montpellier, France, Jun. 20-22, 2011.
- [CI20] M. Hentati, A. Nafkha, X. Zhang, P. Leray, J-F. Nezan, M. Abid, *The Study of the impact of architecture design on cognitive radio*, International Multi-Conference on Systems Signals and Devices, SSD'2011, Sousse, Tunisia, Mar. 22-25, 2011.
- [CI19] Z. Khalaf, A. Nafkha, J. Palicot, M. Ghozzi, *Hybrid Spectrum Sensing Architecture for Cognitive Radio Equipment*, Advanced International Conference on Telecommunications, AICT'2010, Barcelona, May 9-15, 2010. (Best Paper Award)
- [CI18] H. Wang, W. Jouini, A. Nafkha, J. Palicot, L. Cardoso, M. Debbah, *Blind Stan-*

dard Identification with Bandwidth Shape and GI Recognition using USRP Platforms and SDR4all Tools, International Conference on Cognitive Radio Oriented Wireless Networks and Communications, Cannes, France, Jun. 9-11, 2010.

- [CI17] I. Gomez, H. Wang, A. Nafkha, V. Marojevic, C. Moy, P. Leray, A. Gelonch, *Middleware Extension for Partial Reconfiguration Management in Cognitive Radios*, Workshop on Software Radio, WSR'2010, Karlsruhe, Germany, Mar. 3-4, 2010.
- [CI16] W. Jouini, A. Nafkha, M. Lopez-Benity, J. Perez-Romero, *Joint Learning-Detection Framework: an Empirical Analysis*, Joint COST2100 and IC0902 Workshop on Cognitive Radio and Networking , IC0902'2010, Bologna, Italy, Nov. 23-25, 2010.
- [CI15] M. Mroue, A. Nafkha, J. Palicot, B. Gavalda, N. Dagonne, *An Innovative Low Complexity PAPR Reduction TR-based Technique for DVB-T2 System*, International Congress on Ultra Modern Telecommunications and Control Systems , ICUMT'2010, Moscow, Russia, Oct. 18-20, 2010.
- [CI14] X. Zhang, A. Nafkha, P. Leray, *Energy efficiency in dynamically reconfigurable SoC for data-parallel applications*, International Conference on Design and Technology of Integrated Systems in Nanoscale Era, Hammamet, Tunisia, Mar. 23-25, 2010.
- [CI13] Y. Mlayeh, F. Tlili, A. Nafkha, A. Ghazel, *2-Norm Condition Number Based Switching Algorithm in MIMO OFDM Systems*, International Conference on Communications and Networking , ComNet'2010, Tozeur, Tunisia, Nov. 4-7, 2010.
- [CI12] S. Aubert, F. Nouvel, A. Nafkha, *Complexity gain of QR Decomposition based Sphere Decoder in LTE receivers*, Vehicular Technology Conference, VTCfall'2009, Anchorage, Alaska, USA, Sep. 20-23, 2009.
- [CI11] J. Delorme, A. Nafkha, P. Leray, C. Moy, *New OPBHWICAP interface for real-time Partial reconfiguration of FPGA*, International Conference on ReConFigurable Computing and FPGAs , ReConFig'2009, Cancun, Mexico, Dec. 1-2, 2009.
- [CI10] D. Nussbaum, K. Kalfallah, R. Knopp, C. Moy, A. Nafkha, P. Leray, J. Delorme, J. Palicot, J. Martin, F. Clermidy, B. Mercier, R. Pacalet, *Open Platform for Prototyping of Advanced Software Defined Radio and Cognitive Radio Techniques*, Euro-micro Conference on Digital System Design , DSD'2009, Patras, Greece, Aug. 27-29, 2009.
- [CI9] J. Delorme, J. Martin, A. Nafkha, C. Moy, F. Clermidy, J. Palicot, *A FPGA partial reconfiguration design approach for cognitive radio based on NoC architecture*, New Circuits and Systems Conference, NEWCAS'2008, Montreal, Canada, Jun. 22-25, 2008.
- [CI8] C. Moy, A. Nafkha, P. Leray, J. Delorme, J. Palicot, D. Nussbaum, K. Kalfallah, H. Callawaert, J. Martin, F. Clermidy, B. Mercier, R. Pacalet, *IDROMel: An Open Platform Addressing Advanced SDR Challenges*, SDR Forum Technical Conference'08 , SDRForum'2008, Washington DC, USA, Nov. 27-30, 2008.

- [CI7] C. Moy, A. Nafkha, P. Leray, J. Delorme, J. Palicot, J. Martin, F. Clermidy, *FPGA Dynamic Partial Reconfiguration for very high Speed - Real-Time 4G baseband Modem Based on a Network on Chip Architecture*, Software Defined Radio Forum, SDR-Forum'2008, Washington DC, USA, Oct. 26-30, 2008.
- [CI6] A. Nafkha, J. Delorme, R. Seguier, C. MOY, J. Palicot, *A heterogeneous reconfigurable platform for cognitive radio systems*, Workshop on Software Radio, WSR'2008, Karlsruhe, Germany, Mar. 3-4, 2008.
- [CI5] A. Nafkha, E. Boutillon, C. Roland, *Near Maximum Likelihood Detection for MIMO systems using an intensification strategy over a BCH codes*, International Conference on Smart Systems and Devices, SSD'2007, Hammamet, Tunisia, Mar. 19-22, 2007.
- [CI4] A. Al Ghouwayel, Y. Louet, A. Nafkha, J. Palicot, *On the FPGA Implementation of the Fourier Transform over Finite Fields GF(2m)*, International Symposium on Communications and Information Technologies, Sydney, Australia, Oct. 16-19, 2007.
- [CI3] A. Nafkha, R. Seguier, J. Palicot, C. Moy, J-P. Delahaye, *A Reconfigurable Base-Band Transmitter for Adaptive Image Coding*, IST Mobile and Wireless Communication Summit , IST SUMMIT'2007, Budapest, Hungary, Jul. 1-5, 2007.
- [CI2] A. Nafkha, E. Boutillon, C. Roland, *A Near-Optimal Multiuser Detector for MC-CDMA systems Using Geometrical Approach*, International Conference on Acoustics, Speech and Signal Processing, ICASSP'2005, Philadelphia, PA, USA, Mar. 19-23, 2005.
- [CI1] P. Coussy, D. Gnaedig, A. Nafkha, A. Baganne, E. Boutillon, E. Martin, *A Methodology for IP integration in DSP Soc: a case study of a MAP algorithm for turbo decoder*, International Conference on Acoustics, Speech and Signal Processing, ICASSP'2004, Montreal, Canada, May 17-21, 2004.

3.6 National Conferences

Following are listed all my publications in national conference with peer review process.

- [CN3] Z. Khalaf, J. Palicot, A. Nafkha, H. Zhang, *Un détecteur aveugle de signaux de télécomunications basé sur la parcimonie de la fonction d'autocorrélation cyclique*, GRETSI Symposium on Signal and Image processing, GRETSI'2013, Brest, France, Sep. 3-6, 2013.
- [CN2] S. Aubert, F. Nouvel, A. Nafkha, *Décodeur Sphérique associé à une décomposition QR à complexité réduite dans les systèmes MIMO-OFDM*, Colloque Gretsi - Traitement du Signal et des Images , GRETSI'2009, Dijon, France, Sep. 8-11, 2009.
- [CN1] A. Nafkha, E. Boutillon, C. Roland, *Un nouveau décodeur MIMO pour des transmissions MAQ utilisant une approche géométrique: HISD*, Colloque Gretsi - Traitement du Signal et des Images, GRETSI'2005, Louvain-la-Neuve, Belgique, Sep. 8-11, 2005.

Chapter 4

Contributions on MIMO Communication Systems

In this chapter, I present the results from my various researches conducted on the field of multiple-input multiple-output (MIMO) communication systems since my PhD. In first part, I present theoretical studies of ergodic capacities of wireless MIMO Rayleigh fading and optical MIMO Jacobi fading channels. In second part, I addresses the MIMO spatial multiplexing detection problem by applying geometrical approaches or bio-inspired algorithms. For each part, I will give a brief introduction to the addressed topic, then, I will highlight my main theoretical and algorithmic contributions.

4.1 MIMO Technology Background

In the wireless communication field, multiple-input multiple-output (MIMO) technology refers to the use of multiple antennas at the transmitter and receiver side in order to improve communication performance. This technology yields a higher throughput and robustness without increasing the emitting power and allocated bandwidth. Moreover, it addressed a new data transmission concept by adding a spatial dimension to the existing time, frequency, and coding dimensions [1]. The main idea of multiple antennas systems is to exploit the spatial aspects of multipath propagation, an inherent feature of a mobile communications channel, to design high throughput transmission systems through spatial multiplexing [2] or more reliable transmission systems via spatial diversity [3], or enhanced the transmission link performance metrics using smart antennas [4, 5]. Spatial multiplexing provides high spectrum efficiency by simultaneously transmit independent information sequences over multiple transmission paths which are combined at the receiver to achieve multiplexing gain [1] [6]. Multiple antennas can also be seen as a channel coding in order to improve the error rate by transmitting and/or receiving redundant signals representing the same information sequence. Therefore, in spatial diversity, the same information-bearing signal is transmitted and/or received via different antennas where the maximum gain can be achieved when the fading occurring in the channel is independent.

1. The achieved gain in terms of bit rate with respect to a SISO system is called multiplexing gain

Typically, diversity schemes utilize space-time coding of the data stream in order to produce streams for each transmit antenna [7, 8]. In a high correlation environment (low-rank channel), multiple antennas techniques can also be used to improve the signal-to-noise ration and reduce co-channel interference in multiuser scenario by means of smart (*i.e.adaptive*) antenna arrays. Indeed, beamforming techniques are adopted to focus the energy towards certain desired directions and mitigating it towards the undesired ones [9, 10]. Recently, massive multiple-input multiple-output technology has been proposed as one of the key technologies to drive the 5G communication systems due to its promising capability of greatly improving spectral efficiency, energy efficiency, and robustness of the system. In a massive MIMO system, both the transmitter and receiver are equipped with a very large number of antenna elements. Massive MIMO offers big advantages over traditional MIMO and claims have been made about a 10-fold increase in capacity and 100-fold improvement in energy efficiency [11, 12, 13].

Over the past 10 years, various types of space division multiplexed (SDM) transmission fiber, *e.g. multi-mode fiber (MMF) or multi-core fiber(MCF)*, have been introduced and investigated as a way to provide multiple spatial channels in a single optical fiber [14, 15, 16, 17, 18]. The SDM technology achieved very high capacity through increased spatial multiplicity and offered a reduction of the cost and energy per transmitted bit. Moreover, it is one of the most promising technologies to push the capacity of optical fiber close to the 100Tb/s (the theoretical limit) of conventional single-mode fiber (SMF). Thus, SDM could be used for optical backhaul in 5G network infrastructure [19], and data center connectivity [20]. In general, multi-mode fiber or multi-core fiber technology was used to construct SDM optical communication systems. Inspired from massive MIMO in wireless communication, combining the above fiber technologies in an appropriate manner opens up the possibility of achieving a dense (massive) SDM optical communications system having a large-space multiplicity [21, 22, 23].

4.2 Considered MIMO fading channels

In this section, we present the two considered MIMO fading channel models which will be used in this chapter. In wireless MIMO communication systems, we addressed the MIMO detection problem, which has been a central research topic [24], in the context of uncorrelated MIMO Rayleigh fading channels. The main challenge resides in designing detection techniques that can recover the transmitted signals from spatial multiplexing MIMO system with low hardware computation complexity and high performance. Our contributions are as follows: **(i)** We propose two exploration-exploitation based MIMO detectors. The first algorithm was developed by adding bio-inspired exploration method to well known tree-based MIMO detectors. The second algorithm proposed a new enhanced exploration method for the geometrical-based detector. **(ii)** We present a field programmable gate array (FPGA) design and implementation of

geometrical-based algorithm for MIMO detection. In the context of MIMO optical communications [25, 26] or interference-limited multiuser MIMO [27], we investigated the channel capacity of the so-called MIMO Jacobi fading channel which is relevant channel model for the above communication systems. We derived new exact expression of the ergodic capacity for the MIMO Jacobi fading channel. Moreover, closed-form expressions of ergodic capacity upper and lower bounds are also derived.

4.2.1 MIMO Rayleigh fading channels

Consider a point-to-point MIMO system with t transmit antennas and r receive antennas in a Rayleigh fading channels as shown in Fig. 4.1. The entries of the channel matrix $\mathbf{H} \in \mathbb{C}^{r\times t}$ are assumed to be independent and identically distributed zero mean unit variance complex Gaussian random variables, *i.e.* $h_{ij} \sim \mathcal{CN}(0,1)$. The channel is assumed to be quasi-static, *i.e.* channel coefficients remain constant during one time interval, and changes independently from an interval to another. For the sake of clarity and simplification, we consider only the case where $r \geq t$. The discrete-time baseband equivalent signal model can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}, \tag{4.1}$$

where $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_t]^T \in \mathbb{C}^{t\times 1}$ represents the transmitted signal, with zero mean and covariance matrix $\mathbb{E}_{\mathbf{x}}\left[\mathbf{x}\mathbf{x}^{\dagger}\right] = \mathbf{Q}$. The received signal is denoted by $\mathbf{y} \in \mathbb{C}^{r\times 1}$, while the noise vector $\mathbf{z} \in \mathbb{C}^{r\times 1}$ is assumed to be spatially white circular Gaussian random variables with zero-mean and variance σ^2. Thus, $\mathbf{z}$ is distributed as a complex-valued multivariate Gaussian probability density function, *i.e.*, $\mathbf{z} \sim \mathcal{N}_{\mathbb{C}}\left(0, \sigma^2\mathbf{I}_r\right)$. Note that, we normalize the channel matrix $\mathbf{H}$ as $\mathbb{E}_{\mathbf{H}}\left[\mathrm{Tr}(\mathbf{H}^{\dagger}\mathbf{H})\right] = rt$, where Tr(.) denotes the trace of a matrix. The effect of the channel is reflected via the statistical properties of the instantaneous wireless MIMO correlation matrix $\mathbf{W} = \mathbf{H}^{\dagger}\mathbf{H}$. Under uncorrelated MIMO Rayleigh fading, the $t \times t$ random matrix $\mathbf{W}$ is a central uncorrelated complex Wishart matrix with r degrees of freedom and a covariance matrix $\Sigma = \mathbf{I}_t$, commonly denoted as $\mathbf{W} \sim \mathcal{CW}_t(r, \mathbf{I}_t)$ [28].

4.2.2 MIMO Jacobi fading channels

We consider an optical space division multiplexing where the multiple channels correspond to the number of excited modes and/or cores within the optical fiber. The coupling between different modes and/or cores can be described by scattering matrix formalism as reported in [29, 26]. Herein, we consider $m-$channel near lossless optical fiber with $t \leq m$ transmitting excited channels and $r \leq m$ receiving channels, as indicated in Fig. 4.2, for multi-core optical fiber scenario. The scattering matrix formalism can describe very simply the propagation through

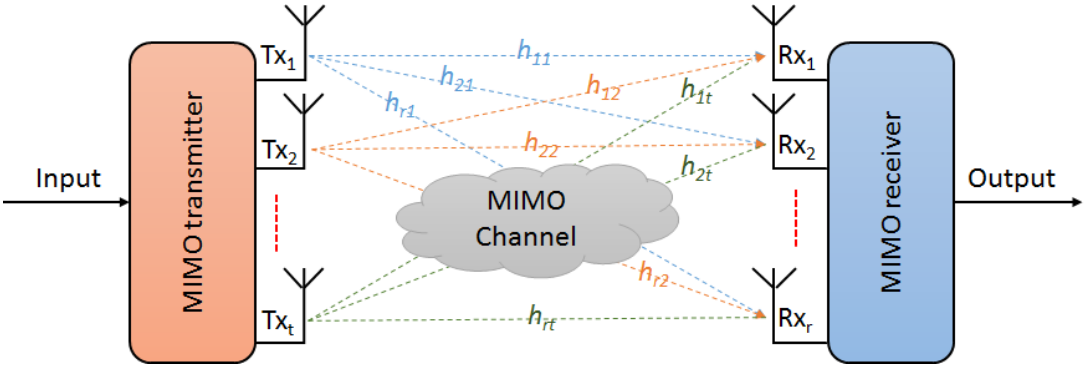


Figure 4.1 – Wireless MIMO system with t transmitter and r receiver antennas.

the fiber using $2m \times 2m$ scattering matrix $\mathbf{S}$ given as

$$\mathbf{S} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{T}_2 \\ \mathbf{T}_1 & \mathbf{R}_2 \end{bmatrix}, \tag{4.2}$$

where the $m \times m$ complex block matrices $\mathbf{R}_1$ and $\mathbf{R}_2$ describe the reflection coefficients in input and output ports of the fiber, respectively. Similarly, the $m \times m$ complex block matrices $\mathbf{T}_1$ and $\mathbf{T}_2$ stand for the transmission coefficients through the fiber from input to output sides and vice versa, respectively. We assume a strong cross-talk between cores or modes, negligible back-scattering, near-lossless propagation, and reciprocal characteristics of the fiber. Thus, we model the scattering matrix as a complex unitary symmetric matrix [26], (*i.e.* $\mathbf{S}^\dagger\mathbf{S} = \mathbf{I}_{2m}$). Therefore, the four Hermitian matrices $\mathbf{T}_1\mathbf{T}_1^\dagger$, $\mathbf{T}_2\mathbf{T}_2^\dagger$, $\mathbf{I}_m - \mathbf{R}_2\mathbf{R}_2^\dagger$, and $\mathbf{I}_m - \mathbf{R}_1\mathbf{R}_1^\dagger$ have the same set of eigenvalues $\lambda_1, \lambda_2, ..., \lambda_m$. Each of these m transmission eigenvalues is a real number belong to the interval $[0, 1]$. Assuming a unitary coupling among all transmission modes the overall transfer matrix $\mathbf{T}_1$ can be described by a $m \times m$ unitary matrix, where each matrix entry $[\mathbf{T}_1]_{ij}$ represents the complex path gain from the i^{th} transmitted mode to the j^{th} received mode. Moreover, the transmission matrix $\mathbf{T}_1$ has a Haar distribution over the group of complex unitary matrices [26, 25].

Given the fact that only $t \leq m$ and $r \leq m$ modes are addressed by the transmitter and receiver, respectively, the effective transmission channel matrix $\mathbf{H} \in \mathbb{C}^{r \times t}$ is a truncated version of $\mathbf{T}_1$ (*i.e.* without loss of generality, the effective transmission channel matrix $\mathbf{H}$ is the $r \times t$ upper-left corner of the transmission matrix $\mathbf{T}_1$ [30, 31]). As a result, the corresponding MIMO channel for this system reads

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}, \tag{4.3}$$

where $\mathbf{y} \in \mathbb{C}^{r \times 1}$ is the received signal vector, $\mathbf{x} \in \mathbb{C}^{t \times 1}$ is a $t \times 1$ transmitted signal vector with covariance matrix equal to $\frac{\mathcal{P}}{t}\mathbf{I}_t$, and $\mathbf{z} \in \mathbb{C}^{r \times 1}$ is a $r \times 1$ zero mean additive white circularly symmetric complex Gaussian noise vector with covariance matrix equal to $\sigma^2\mathbf{I}_r$ [32]. The variable

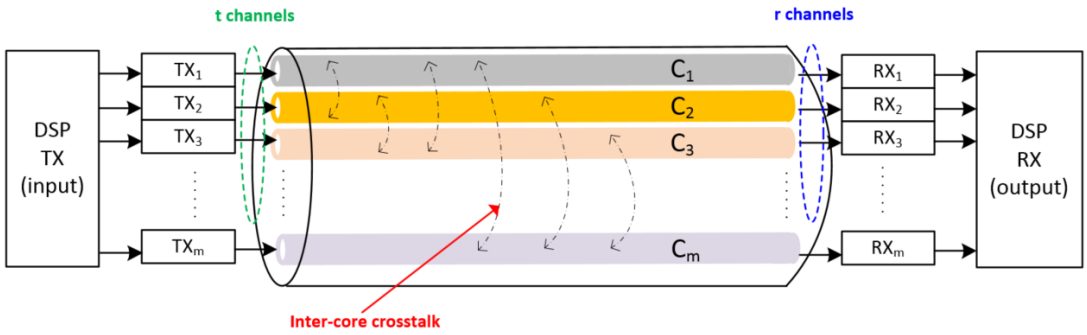


Figure 4.2 – Optical MIMO system with an MCF, C_k indicates k^{th} fiber core, with $k \in \{1, 2, \ldots, m\}$

$\mathcal{P}$ is the total transmit power across the t modes/cores, and σ^2 is the Gaussian noise variance.

4.3 Ergodic capacity of optical MIMO channel

In his seminal paper [33], C. Shannon defined the channel capacity as the maximum of the mutual information between the transmitter and the receive. In context of MIMO systems, the channel capacity depends largely on the availability of the channel state information (CSI) at the two communication ends. It is often assumed that the receiver can track the channel perfectly and thus complete CSI is possible at the receiver side (CSIR). Indeed, the instantaneous MIMO channel matrix can be estimated by using channel estimation techniques. In absence of CSI in the transmitter side the total transmitted power is divided equally between the transmit antennas (in MIMO wireless systems) or transmit modes/cores (in SDM optical fiber systems). For a MIMO system model described by (4.1) or (4.3), the channel capacity expression is given as

$$C = \max_{\mathrm{Tr}(\mathbf{Q}\leq\mathcal{P})} \ln\det\left(\mathbf{I}_r + \frac{1}{\sigma^2}\mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right) \tag{4.4}$$

where the optimization is taken on the signal covariance matrix $\mathbf{Q} = \mathbb{E}_{\mathbf{x}}\left[\mathbf{x}\mathbf{x}^\dagger\right]$ and $\mathcal{P}$ is the total transmit power. We assume that CSI is known perfectly at the receiver, and that an equal-power allocation across the transmit elements (antennas, modes or cores) is used. The ergodic capacity is defined as the statistical average of the channel capacity conditioned on channel matrix $\mathbf{H}$. Then, the Shannon ergodic capacity of MIMO channel can be expressed as

$$C_e = \begin{cases} \mathbb{E}_{\mathbf{H}}\left[\ln\det\left(\mathbf{I}_t + \frac{\rho}{t}\mathbf{H}^\dagger\mathbf{H}\right)\right] & \textbf{if} \quad t \leq r \\ \mathbb{E}_{\mathbf{H}}\left[\ln\det\left(\mathbf{I}_r + \frac{\rho}{t}\mathbf{H}\mathbf{H}^\dagger\right)\right] & \textbf{if} \quad t \geq r \end{cases} \tag{4.5}$$

where $\mathbb{E}_{\mathbf{H}}$ denotes the expectation over all channel realizations, ln is the natural logarithm

function and $\rho = \frac{\mathcal{P}}{\sigma^2}$ is the average signal-to-noise ratio (SNR). Without loss of generality, **we shall assume in the sequel that** $t \leq r$.

In the following two sections, we first develop exact ergodic capacity bounds for Jacobi MIMO channel, which is a useful model for MIMO communications over multi-mode or multi-core optical fibers, based on classical Jensen's and Minkowski's inequalities, and then we provide a new exact expression for the ergodic capacity of the MIMO Jacobi fading channel relying this time on the formula derived in [34] for the moments of the eigenvalues density of the Jacobi random matrix.

4.3.1 [C1]: Capacity Bounds of MIMO Jacobi Channels

This work was carried out in the context of the PhD of Rémi Bonnefoi. In order to obtain simplified closed-form expressions for the ergodic capacity of the Jacobi MIMO channel, we consider classical Jensen's and Minkowski's inequalities. Moreover, we used the fact that the $\ln\det(.)$ function is concave when the channel covariance matrix $\mathbf{J} = \frac{\mathbf{H}^\dagger\mathbf{H}}{t}$ is positive definite matrix.

We consider that there are $t \leq m$ excited transmitting channels and $r \leq m$ receiving channels coherently excited in the input and output side of the m-channel lossless optical fiber. We know from [25] that when the receiver has a complete knowledge of the channel matrix, the ergodic capacity is given by

$$C_{t,r}^m = C_e = \mathbb{E}_{\mathbf{H}}\left[\ln\det\left(\mathbf{I}_t + \frac{\rho}{t}\mathbf{H}^\dagger\mathbf{H}\right)\right] \tag{4.6}$$

We derive tight upper bound of the ergodic capacity for Jacobi MIMO channels, which is given in the following theorem.

Ergodic Capacity Upper Bound

Theorem 1 *: Let $t \leq r$, and $t + r \leq m$, the ergodic capacity of uncorrelated MIMO Jacobi fading channels, with only CSI at the receiver side, is upper bounded by*

$$C_{t,r}^m \leq C_{up} \triangleq t\ln\left(1 + \frac{\rho r}{m}\right) \tag{4.7}$$

In low-SNR regimes ($\rho << 1$), the derived upper bound expression (4.7) is very close to the ergodic capacity. Thus, we derive the following corollary.

Tightest capacity upper bound in low SNR regime

Corollary 1 : *For MIMO Jacobi fading channels, in the low SNR regime, the ergodic capacity upper bound C_{up} can be approximated as*

$$C_{up} \approx C_{up}^{lsnr} \triangleq \frac{rt\rho}{m} \qquad \rho \ll 1 \tag{4.8}$$

Using Kshirsagar's theorem, Minkowski's and Jensen's inequalities, the following theorem gives a tight lower bound on the ergodic capacity of MIMO Jacobi fading channels

Ergodic Capacity lower Bound

Theorem 2 : *Let $t \leq r$, and $t + r \leq m$, the ergodic capacity of uncorrelated MIMO Jacobi fading channels, with only CSI at the receiver side, is lower bounded by*

$$C_{t,r}^{m} \geq C_{lo} \triangleq t \ln \left(1 + \frac{\rho}{\sqrt[t]{\mathcal{F}_{t,r}^{m}}} \right) \tag{4.9}$$

where $\mathcal{F}_{t,r}^{m} = \prod_{p=0}^{t-1} \prod_{k=0}^{m-r-1} \exp \frac{1}{r+k-p}$.

Using the high SNR approximation of the lower bound, we derive the following corollary

Tightest capacity lower bound in high SNR regime

Corollary 2 : *For MIMO Jacobi fading channels, in the high SNR regime, the ergodic capacity lower bound C_{lo} can be approximated as*

$$C_{lo} \approx C_{lo}^{hsnr} \triangleq t \ln(\rho) - \sum_{p=0}^{t-1} \sum_{k=0}^{m-r-1} \frac{1}{r+k-p} \tag{4.10}$$

In the above theorems and corollaries, we considered the case where $t \leq r$ and $t + r \leq m$. In the case where $t + r \geq m$, it has been shown in [25, Theorem 2] that the ergodic capacity of the MIMO Jacobi fading channels can be deduced from (4.6) as follows

$$C_{t,r}^{m} = (t + r - m) \ln(1 + \rho) + C_{m-r,m-t}^{m} \tag{4.11}$$

As consequences of equation (4.11), Theorem 1 and Theorem 2 and their related corollaries can be extended to the case where $t + r \geq m$.

At a later stage, using the fact that the probability density of the eigenvalues of a random matrix from unitary ensemble can be expressed in terms of the Christoffel-Darboux kernel [35], the ergodic capacity formula given in [25] was rewritten to show that the polynomial part of the integrand consists of the Christoffel-Darboux kernel. Then, we used the new exact expression to derive a lower bound of the ergodic capacity of MIMO Jacobi fading channels. Finally, we propose an approximation to the ergodic capacity in low SNR regime. This work was published in IEEE ICC 2017 [CI47].

Ergodic Capacity lower Bound [CI47]

Theorem 3 : *Let $t \leq r$, and $t + r \leq m$, the ergodic capacity of uncorrelated MIMO Jacobi fading channels, with only CSI at the receiver side, is lower bounded by*

$$C_{t,r}^{m} \geq C_{lb} \triangleq M_t^{a,b} \int_{-1}^{1} (1-x)^a (1+x)^b \ln\left(1 + \frac{\rho(1-x)}{2}\right) \times \left[P_{t-1}^{a,b}(x) P_{t-1}^{a+1,b+1}(x) - N_t^{a,b} P_t^{a,b}(x) P_{t-2}^{a+1,b+1}(x)\right] dx \quad (4.12)$$

where $a = r - t$, $b = m - r - t$, $P_n^{\alpha,\beta}(x)$ denotes the Jacobi polynomial of degree n, order (α, β), $M_t^{a,b} = \frac{(t+a+b+1)!t!}{2^{a+b+1}(t+a-1)!(t+b-1)!(2t+a+b)}$ and $N_t^{a,b} = \frac{(t+a+b)}{(t+a+b+1)}$.

The novel expression of the ergodic capacity lower bound derived in theorem 3 can also be used to propose a low SNR approximation of the ergodic capacity. Indeed, at low SNR, the ergodic capacity can be approximated by a very simple linear expression.

Tightest capacity lower bound in low SNR regime [CI47]

Corollary 3 : *For MIMO Jacobi fading channels, in the low SNR regime, the ergodic capacity lower bound C_{lb} can be approximated as*

$$C_{lb} \approx C_{lb}^{lsnr} \triangleq \frac{rt\rho}{m} \qquad \rho \ll 1 \quad (4.13)$$

For more mathematical derivation details and simulations results about Theorem 1 and Theorem 2, our publication in OSA optics express [J13] is included hereafter.

Research Article | Vol. 25, No. 11 | 29 May 2017 | OPTICS EXPRESS 12144
Optics EXPRESS

Upper and lower bounds for the ergodic capacity of MIMO Jacobi fading channels

AMOR NAFKHA[1,2,*] AND RÉMI BONNEFOI[1]

[1] SCEE/IETR UMR CNRS 6164,CentraleSupélec campus de Rennes, 35510 Cesson-Sévigné, France
[2] B-com, 1219 Avenue des Champs Blancs, 35510 Cesson Sévigné, France
[*] amor.nafkha@centralesupelec.fr

Abstract: In multi-(core/mode) optical fiber communication, the transmission channel can be modeled as a complex sub-matrix of the Haar-distributed unitary matrix (complex Jacobi unitary ensemble). In this letter, we present new analytical expressions of the upper and lower bounds for the ergodic capacity of multiple-input multiple-output Jacobi-fading channels. Recent results on the determinant of the Jacobi unitary ensemble are employed to derive a tight lower bound on the ergodic capacity. We use Jensen's inequality to provide an analytical closed-form upper bound to the ergodic capacity at any signal-to-noise ratio (SNR). Closed-form expressions of the ergodic capacity, at low and high SNR regimes, are also derived. Simulation results are presented to validate the accuracy of the derived expressions.

OCIS codes: (110.3055) Information theoretical analysis; (000.3860) Mathematical methods in physics; (060.2330) Fiber optics communications; (060.2310) Fiber optics.

References and links

1. K. Ho, and J. Kahn, "Statistics of group delays in multimode fiber with strong mode coupling," J. Lightwave Technol. **29**(21), 3119–3128 (2011).
2. C. Lin, I. B. Djordjevic, and D. Zou, "Achievable information rates calculation for optical OFDM transmission over few-mode fiber long-haul transmission systems," Opt. Express **23**(13), 16846–16856 (2015).
3. G. M. Saridis, D. Alexandropoulos, G. Zervas and D. Simeonidou, "Survey and evaluation of space division multiplexing: from technologies to optical networks," IEEE Comm. Surveys & Tutorials **17** (4), 2136–2156 (2015).
4. V. A. J. M. Sleiffer, H. Chen, Y. Jung, P. Leoni, M. Kuschnerov, A. Simperler, H. Fabian, H. Schuh, F. Kub, D. J. Richardson, S. U. Alam, L. Grner-Nielsen, Y. Sun, A. M. J. Koonen, and H. de Waardt, "Field demonstration of mode-division multiplexing upgrade scenarios on commercial networks," Opt. Express **21**(25), 31036-31046 (2013).
5. D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres," Nat. Photonics **7**(5), 354–362 (2013).
6. R. Dar, M. Feder, M. Shtaif, "The Jacobi MIMO channel," IEEE Trans. on Information Theory **59**(4), 2426–2441 (2013).
7. P. J. Winzer, G. J. Foschini, "MIMO capacities and outage probabilities in spatially multiplexed optical transport systems," Opt. Express **19**(17), 16680-16696 (2011).
8. A. Karadimitrakis, A. L. Moustakas, P. Vivo, "Outage capacity for the optical MIMO channel," IEEE Trans. on Information Theory **60**(7), 4370–4382 (2014).
9. E. Telatar, "Capacity of multi-antenna Gaussian channels," Europ. Trans. Telecommun. **10**, 585–596 (1999).
10. J. Kaneko, "Selberg integrals and hypergeometric functions associated with Jack polynomials," SIAM J. Math. Anal. **24**, 1086–1110 (1993).
11. T. Jiang, "Approximation of Haar distributed matrices and limiting distributions of eigenvalues of Jacobi ensembles," Prob. Theory and Related Fields **144**, 221–246 (2009).
12. T. M. Cover, J. A. Thomas, *Elements of Information Theory* (John Wiley & Sons, New Jersey, 2006).
13. Z. Cvetkovski, *Inequalities: Theorems, Techniques and Selected Problems* (Springer, 2012).
14. M. E. H. Ismail, *Classical and Quantum Orthogonal Polynomials In One Variable* (Cambridge Univ. Press., 2005).
15. A. M. Kshirsagar, "The noncentral multivariate beta distribution," Ann. Math. Statist. **32**, 104–111 (1961).
16. R. J. Muirhead, *Aspects of Multivariate Statistical Theory* (Wiley, 2005).
17. A. Rouault, "Asymptotic behavior of random determinants in the Laguerre, Gram and Jacobi ensembles," ALEA Lat. Am. J. Probab. Math. Stat, https://arxiv.org/abs/math/0607767 (2007).
18. M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, 1970).

#278468 https://doi.org/10.1364/OE.25.012144
Journal © 2017 Received 11 Oct 2016; revised 16 Dec 2016; accepted 17 Dec 2016; published 15 May 2017

Optics EXPRESS

1. Introduction

To accommodate the exponential growth of data traffic over the last few years, space-division multiplexing (SDM) based on multi-core optical fiber or multi-mode optical fiber [1–4] is expected to overcome the barrier from capacity limit of single-core fiber [5]. The main challenge in SDM occurs due to in-band crosstalk between multiple parallel transmission channels (cores or modes). This non-negligible crosstalk can be dealt with by using multiple-input multiple-output (MIMO) signal processing techniques. Assuming important crosstalk between channels (cores or modes), negligible backscattering and near-lossless propagation, we can model the transmission channel as a random complex unitary matrix [6–8]. In [6], authors introduced the Jacobi unitary ensemble to model the propagation channel for fiber-optical MIMO channel and they gave analytical expression for the ergodic capacity. However, to the best of the authors' knowledge, no bounds for the ergodic capacity of the uncorrelated MIMO Jacobi-fading channels exist in the literature so far. The two main contributions of this work are: (**i**) the derivation of a lower/upper bounds on the ergodic capacity of an uncorrelated MIMO Jacobi-fading channel with identically and independently distributed input symbols, (**ii**) the derivation of simple asymptotic expressions for ergodic capacity in the low and high SNR regimes.

The rest of this paper is organized as follows: Section 2 introduces the MIMO Jacobi-fading channel model and includes the definition of ergodic capacity. We derive a lower and upper bound, at any SNR value, and an approximation, in high and low SNR regimes, to the ergodic capacity in Section 3. The theoretical and the simulation results are discussed in Section 4. Finally, Section 5 provides the conclusion.

2. Problem formulation

Consider a single segment m-channel lossless optical fiber system, the propagation through the fiber may be analyzed through its $2m \times 2m$ scattering matrix given by [8]

$$\mathbf{S} = \begin{bmatrix} \mathbf{R}_{ll} & \mathbf{T}_{rl} \\ \mathbf{T}_{lr} & \mathbf{R}_{rr} \end{bmatrix} \tag{1}$$

where $\mathbf{T}_{lr}$ and $\mathbf{T}_{rl}$ sub-matrices correspond to the transmitted from left to right and from right to left signals, respectively. The $\mathbf{R}_{ll}$ and $\mathbf{R}_{rr}$ sub-matrices present the reflected signals from left to left and from right to right. Moreover, $\mathbf{R}_{ll} = \mathbf{R}_{rr} \approx \mathbf{0}_{m\times m}$ given the fact that the backscattering in the optical fiber is negligible, and $\mathbf{T} = \mathbf{T}_{lr} = \mathbf{T}_{rl}^{\dagger}$ because the two fiber ends are not distinguishable. The notation $(.)^{\dagger}$ is used to denote the conjugate transpose matrix. Energy conservation principle implies that the scattering matrix $\mathbf{S}$ is a unitary matrix ($i.e.$ $\mathbf{S}^{-1} = \mathbf{S}^{\dagger}$ where the notation $(.)^{-1}$ is used to denote the inverse matrix.). As a consequence, the four Hermitian matrices $\mathbf{T}_{lr}\mathbf{T}_{lr}^{\dagger}$, $\mathbf{T}_{rl}\mathbf{T}_{rl}^{\dagger}$, $\mathbf{I}_m - \mathbf{R}_{ll}\mathbf{R}_{ll}^{\dagger}$, and $\mathbf{I}_m - \mathbf{R}_{rr}\mathbf{R}_{rr}^{\dagger}$ have the same set of eigenvalues $\lambda_1, \lambda_2, ..., \lambda_m$. Each of these m transmission eigenvalues is a real number between 0 and 1. Without loss of generality, the transmission matrix $\mathbf{T}$ will be modeled as a Haar-distributed unitary random matrix of dimension $m \times m$ [6].

We consider that there are $m_t \leq m$ excited transmitting channels and $m_r \leq m$ receiving channels coherently excited in the input and output side of the m-channel lossless optical fiber. Therefore, we only consider a truncated version of the transmission matrix $\mathbf{T}$, which we denote by $\mathbf{H}$, since not all transmitting or receiving channels may be available to a given link. Without loss of generality, the effective transmission channel matrix $\mathbf{H}$ is the $m_r \times m_t$ upper-left corner of the transmission matrix $\mathbf{T}$ [11]. As a result, the corresponding multiple-input multiple-output channel for this system is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} \tag{2}$$

where $\mathbf{y} \in \mathbb{C}^{m_r \times 1}$ is the received signal, $\mathbf{x} \in \mathbb{C}^{m_t \times 1}$ is the emitted signal with $\mathbb{E}[\mathbf{x}^{\dagger}\mathbf{x}] = \frac{\mathcal{P}}{m_t}\mathbf{I}_{m_t}$, and $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_{m_r})$ is circular-symmetric complex Gaussian noise. We denote $\mathbb{E}[W]$ the

Optics EXPRESS

mathematical expectation of random variable W. The variable $\mathcal{P}$ is the total transmit power across the m_t modes/cores, and σ^2 is the Gaussian noise variance. We know from [6, 9] that when the receiver has a complete knowledge of the channel matrix, the ergodic capacity is given by

$$C_{m_t,m_r}^{m,\rho} = \begin{cases} \mathbb{E}\left[\ln\det\left(\mathbf{I}_{m_t} + \frac{\rho}{m_t}\mathbf{H}^\dagger\mathbf{H}\right)\right] & \text{if } m_r \geq m_t \\ \mathbb{E}\left[\ln\det\left(\mathbf{I}_{m_r} + \frac{\rho}{m_t}\mathbf{H}\mathbf{H}^\dagger\right)\right] & \text{if } m_r < m_t \end{cases} \tag{3}$$

where ln is the natural logarithm function and $\rho = \frac{\mathcal{P}}{\sigma^2}$ is the average signal-to-noise ratio (SNR). In this paper, we consider the case where $m_r \geq m_t$ and $m_t + m_r \leq m$. The other case where $m_r < m_t$ and $m_t + m_r \leq m$ can be treated defining $m_t' = m_r$ and $m_r' = m_t$. In the case where $m_t + m_r > m$, it was shown in [6, Theorem 2,] that the ergodic capacity can be deduced from (3) as follows:

$$C_{m_t,m_r}^{m,\rho} = (m_t + m_r - m)\ln(1+\rho) + C_{m-m_r,m-m_t}^{m,\rho} \tag{4}$$

The ergodic capacity is defined as the average with respect to the joint distribution of eigenvalues of the covariance channel matrix $\mathbf{J} = \frac{1}{m_t}\mathbf{H}^\dagger\mathbf{H}$. The random matrix $\mathbf{J}$ follows the Jacobi distribution and its ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{m_t}$ have the joint density given by

$$\mathcal{F}_{a,b,m}(\lambda) = \chi^{-1} \prod_{1\leq j\leq m_t} \lambda_j^a (1-\lambda_j)^b V(\lambda)^2 \tag{5}$$

where $a = m_r - m_t$, $b = m - m_r - m_t$, $\lambda = (\lambda_1, \ldots, \lambda_{m_t})$, $V(\lambda) = \prod_{1\leq j<k\leq m_t} |\lambda_k - \lambda_j|$, χ is a normalization constant evaluated using Selberg integral formula [10], and it is given by:

$$\chi = \prod_{j=1}^{m_t} \frac{\Gamma(a+1+j)\Gamma(b+1+j)\Gamma(2+j)}{\Gamma(a+b+m_t+j+1)\Gamma(2)} \tag{6}$$

3. Tight bounds on the ergodic capacity

In order to obtain simplified closed-form expressions for the ergodic capacity of the Jacobi MIMO channel, we consider classical inequalities such as Jensen's inequality and Minkowski's inequality. Moreover, we used the concavity property of the ln det(.) function given the fact that the channel covariance matrix $\mathbf{J}$ is positive definite matrix [12, Theorem 17.9.1,].

3.1. Upper bound

The following theorem presents a new tight upper bound on the ergodic capacity of Jacobi MIMO channel.

Theorem 1 *Let $m_t \leq m_r$, and $m_t + m_r \leq m$, the ergodic capacity of uncorrelated MIMO Jacobi-fading channel, with receiver CSI and no transmitter CSI, is upper bounded by*

$$C_{m_t,m_r}^{m,\rho} \leq m_t \ln\left(1 + \frac{\rho m_r}{m}\right) \tag{7}$$

Proof of Theorem 1: We propose to use the well known Jensen's inequality [13] to obtain an upper bound for the ergodic capacity. According to this inequality and the concavity of the ln det(.) function, we can give a tight upper bound on the ergodic capacity (3) as:

$$\begin{aligned} C_{m_t,m_r}^{m,\rho} &\leq \sum_{k=1}^{m_t} \ln\left(1 + \rho\mathbb{E}\left[\lambda_k\right]\right) \\ &\leq m_t \ln\left(1 + \rho\mathbb{E}\left[\lambda_1\right]\right) \end{aligned} \tag{8}$$

Optics EXPRESS

Now, the density of λ_1 is given by [6, (67),] as

$$f_{\lambda_1}(\lambda_1) = \frac{1}{m_t} \sum_{k=0}^{m_t-1} e_{k,a,b}^{-1} \lambda_1^a (1-\lambda_1)^b \left(P_k^{(a,b)}(1-2\lambda_1)\right)^2 \tag{9}$$

where $e_{k,a,b} = \frac{\Gamma(k+a+1)\Gamma(k+b+1)}{k!(2k+a+b+1)\Gamma(k+a+b+1)}$ and $P_k^{(a,b)}(x)$ are the Jacobi polynomials [14, Theorem 4.1.1,]. They are orthogonal with respect to the Jacobi weight function $\omega^{a,b}(x) := (1-x)^a(1+x)^b$ over the interval $I = [-1, 1]$, where $a, b > -1$, and they are defined by

$$\int_{-1}^{1} (1-x)^a (1+x)^b P_n^{(a,b)}(x) P_m^{(a,b)}(x) dx = 2^{a+b+1} e_{n,a,b} \delta_{n,m} \tag{10}$$

where $\delta_{n,m}$ is the Kronecker delta function. Using (9), we can write the expectation of λ_1 as

$$\mathbb{E}[\lambda_1] = \sum_{k=0}^{m_t-1} \frac{e_{k,a,b}^{-1}}{m_t} \int_0^1 \lambda_1^{a+1} (1-\lambda_1)^b \left(P_k^{(a,b)}(1-2\lambda_1)\right)^2 d\lambda_1 \tag{11}$$

By taking $u = 1 - 2\lambda_1$, we can write

$$\mathbb{E}[\lambda_1] = \frac{1}{m_t 2^{a+b+2}} \sum_{k=0}^{m_t-1} e_{k,a,b}^{-1} \int_{-1}^{1} (1-u)^a (1+u)^b P_k^{(a,b)}(u) \left(P_k^{(a,b)}(u) - u P_k^{(a,b)}(u)\right) du \tag{12}$$

we recall from [14, (4.2.9),] the following three-term recurrence relation of Jacobi polynomials generation:

$$u P_k^{(a,b)}(u) = \frac{P_{k+1}^{(a,b)}(u)}{A_k} - \frac{C_k P_{k-1}^{(a,b)}(u)}{A_k} - \frac{B_k P_k^{(a,b)}(u)}{A_k} \quad , k > 0 \tag{13}$$

where $A_k = \frac{(2k+a+b+1)(2k+a+b+2)}{2(k+1)(k+a+b+1)}$, $B_k = \frac{(a^2-b^2)(2k+a+b+1)}{2(k+1)(k+a+b+1)(2k+a+b)}$, and $C_k = \frac{(k+a)(k+b)(2k+a+b+2)}{(k+1)(k+a+b+1)(2k+a+b)}$. Then, by employing (10), (12), and (13), the expectation of λ_1 can be expressed as

$$\mathbb{E}[\lambda_1] = \sum_{k=0}^{m_t-1} \frac{e_{k,a,b}^{-1}}{m_t 2^{a+b+2}} \int_{-1}^{1} (1-u)^a (1+u)^b P_k^{(a,b)}(u) \left(P_k^{(a,b)}(u) - u P_k^{(a,b)}(u)\right) du \tag{14}$$

thus, we can write

$$\begin{aligned} \mathbb{E}[\lambda_1] &= \frac{1}{2m_t} \sum_{k=0}^{m_t-1} \left(1 + \frac{B_k}{A_k}\right) \\ &= \frac{mr}{m} \end{aligned} \tag{15}$$

Finally, the upper bound on the ergodic capacity can be expressed as:

$$C_{m_t,m_r}^{m,\rho} \leq m_t \ln\left(1 + \frac{\rho m_r}{m}\right) \tag{16}$$

This completes the proof of Theorem 1.

In low-SNR regimes, the proposed upper bound expression is very close to the ergodic capacity. Thus, we derive the following corollary.

Optics EXPRESS

Corollary 1 *Let $m_t \leq m_r$, and $m_t + m_r \leq m$. In low-SNR regimes, the ergodic capacity for uncorrelated MIMO Jacobi-fading channel can be approximated as*

$$C_{m_t,m_r}^{m,\rho<<<1} \approx \frac{m_t m_r \rho}{m} \tag{17}$$

Proof of Corollary 1: In low-SNR regimes ($\rho <<< 1$), the function $\ln\left(1 + \frac{m_r\rho}{m}\right)$ can be approximated by $\frac{m_r\rho}{m}$.

When the sum of transmit and receive modes, $m_t + m_r$, is larger than the total available modes, m, the upper bound expression of the ergodic capacity can be deduced from (4).

3.2. Lower bound

The following theorem gives a tight lower bound on the ergodic capacity of Jacobi MIMO channels.

Theorem 2 *Let $m_t \leq m_r$, and $m_t + m_r \leq m$, the ergodic capacity of uncorrelated MIMO Jacobi-fading channel, with receiver CSI and no transmitter CSI, is lower bounded by*

$$C_{m_t,m_r}^{m,\rho} \geq m_t \ln\left(1 + \frac{\rho}{\sqrt[m_t]{F_{m_t,m_r}^{m}}}\right) \tag{18}$$

where $F_{m_t,m_r}^{m} = \prod_{j=0}^{m_t-1}\prod_{k=0}^{m-m_r-1} \exp\left(\frac{1}{m_r+k-j}\right)$

Proof of Theorem 2: We start from Minkowski's inequality [13] that we recall here for simplicity. Let $\mathbf{A}$ and $\mathbf{B}$ be two $n \times n$ positive definite matrices, then

$$[\det(\mathbf{A}+\mathbf{B})]^{\frac{1}{n}} \geq (\det(\mathbf{A}))^{\frac{1}{n}} + (\det(\mathbf{B}))^{\frac{1}{n}} \tag{19}$$

with equality iff $\mathbf{A}$ is proportional to $\mathbf{B}$. Applying this inequality to (3), a lower bound of the ergodic capacity can be obtained as

$$\begin{aligned} C_{m_t,m_r}^{m,\rho} &\geq m_t \mathbb{E}\left[\ln\left(1 + \rho\,(\det(\mathbf{J}))^{\frac{1}{m_t}}\right)\right] \\ &\geq m_t \mathbb{E}\left[\ln\left(1 + \rho \exp\left(\tfrac{1}{m_t}\ln\det(\mathbf{J})\right)\right)\right] \end{aligned} \tag{20}$$

Recalling that $\ln(1 + c\exp^x)$ is convex in x for $x > 0$, we apply Jensen's inequality [13] to further lower bound (20)

$$C_{m_t,m_r}^{m,\rho} \geq m_t \ln\left(1 + \rho \exp\left(\tfrac{1}{m_t}\mathbb{E}[\ln\det(\mathbf{J})]\right)\right) \tag{21}$$

Using the Kshirsagar's theorem [15], it has be shown in [16, Theorem 3.3.3,], and [17] that the determinant of the Jacobi ensemble can be decomposed into a product of independent beta distributed variables. We infer from [17] that

$$\ln\det(\mathbf{J}) \overset{(d)}{=} \sum_{j=1}^{m_t} \ln T_j \tag{22}$$

where $\overset{(d)}{=}$ stands for equality in distribution, T_j, $j = 1, \ldots, m_t$ are independent and

$$T_j \overset{(d)}{=} Beta(m_r - j + 1, m - m_r) \tag{23}$$

Optics EXPRESS

where $Beta(\alpha, \beta)$ is the beta distribution with shape parameters (α, β). Taking the expectation over all channel realizations of a random variable $U = \ln\det(\mathbf{J})$, we get

$$\mathbb{E}[U] = \sum_{j=0}^{m_t-1} \psi(m_r - j) - \psi(m - j) \tag{24}$$

where $\psi(n)$ is the digamma function. For positive integer n, the digamma function is also called the Psi function defined as [18]

$$\begin{cases} \psi(n) = -\gamma & n = 1 \\ \psi(n) = -\gamma + \sum_{k=1}^{n-1} \frac{1}{k} & n \geq 2 \end{cases} \tag{25}$$

where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant. Now, we can finish the proof of the Theorem 2 as follows

$$\begin{aligned} C_{m_t,m_r}^{m,\rho} &\geq m_t \ln\left(1 + \rho \exp^{\left(\frac{1}{m_t}\sum_{j=0}^{m_t-1}\psi(m_r-j)-\psi(m-j)\right)}\right) \\ &\geq m_t \ln\left(1 + \frac{\rho}{\sqrt[m_t]{\prod_{j=0}^{m_t-1}\prod_{k=0}^{m-m_r-1}\exp^{\left(\frac{1}{m_r+k-j}\right)}}}\right) \\ &\geq m_t \ln\left(1 + \frac{\rho}{\sqrt[m_t]{F_{m_t,m_r}^{m}}}\right) \end{aligned} \tag{26}$$

where $F_{m_t,m_r}^{m} = \prod_{j=0}^{m_t-1}\prod_{k=0}^{m-m_r-1}\exp^{\left(\frac{1}{m_r+k-j}\right)}$. This completes the proof of Theorem 2.

In high-SNR regimes, the proposed lower bound expression is closed to the ergodic capacity. Thus, we derive the following corollary.

Corollary 2 *Let $m_t \leq m_r$, and $m_t + m_r \leq m$. In high-SNR regimes, the ergodic capacity for uncorrelated MIMO Jacobi-fading channel can be approximated as*

$$C_{m_t,m_r}^{m,\rho>>1} \approx m_t \ln(\rho) - \sum_{j=0}^{m_t-1}\sum_{k=0}^{m-m_r-1} \frac{1}{m_r + k - j} \tag{27}$$

Proof of Corollary 2: In high-SNR regimes ($\rho >> 1$), the function $\ln\left(1 + \frac{\rho}{\sqrt[m_t]{F_{m_t,m_r}^{m}}}\right)$ can be approximated by $\ln(\rho) - \frac{1}{m_t}\ln\left(F_{m_t,m_r}^{m}\right)$.

4. Simulation results

In this section, we present numerical results to further investigate the resulting analytical equations. The tightness of the derived expressions is clearly visible in Figs. 1–3.

In Fig. 1(a), we have plotted the exact ergodic capacity obtained by computer simulation and the corresponding lower and upper bounds, for the uncorrelated MIMO Jacobi-fading channels, with $(m_t = m_r = 2, m = 6)$ and $(m_t = 4, m_r = 10, m = 16)$. At very low SNR (typically below 2 dB), the exact curves and the upper bounds are practically indistinguishable. The gaps between the exact curves of the ergodic capacity and the lower bounds considerably vanish in moderate to high SNR (typically above 20 dB). We can observe that the expression in (18) matches perfectly with the ergodic capacity expression in (3).

Figure 1(b) shows the ergodic capacities of uncorrelated MIMO Jacobi fading channels, and it proves by numerical simulations the validity of the high-SNR regimes lower-bound

Optics EXPRESS

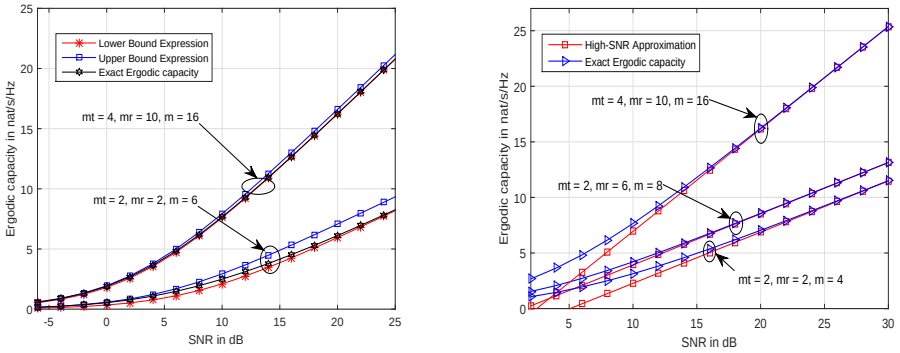


Fig. 1. (a) Comparison of the ergodic capacity and analytical lower-bound and upper-bound expressions for $(m_t = m_r = 2, m = 6)$, and $(m_t = 4, m_r = 10, m = 16)$ uncorrelated MIMO Jacobi-fading channels, (b) High-SNR lower-bound approximation of the ergodic capacity in nats per channel use versus SNR in dB.

approximation given in (27). Results are shown for different numbers of transmitted/received modes, with $m = 4$, $m = 8$, and $m = 16$. We see that the ergodic capacities approximations are accurate over a large range of high SNR values.

Figure 2(a) shows the ergodic capacity and the analytical low-SNR upper bound expression in Eq. (17) for several uncorrelated MIMO Jacobi-fading channels configurations. It is clearly seen that our expression is almost exact at very low SNR and that it gets tighter at low SNR as the number of available modes (m) increases.

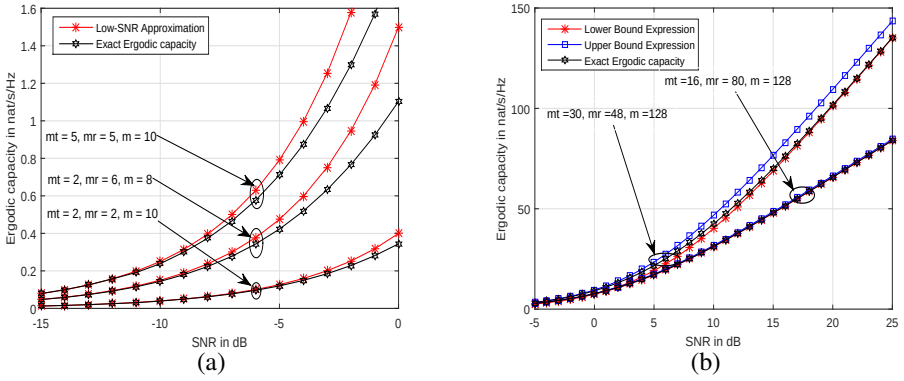


Fig. 2. (a) Low-SNR upper-bound approximation of the ergodic capacity in nats per channel use versus SNR in dB, (b) Bounds and simulation results for ergodic capacity of MIMO Jacobi channel capacity, with number available modes $m = 128$, for different numbers of transmitting and receiving channels.

Figure 2(b) shows the comparison of the ergodic capacity of the uncorrelated MIMO Jacobi-fading channels and the derived expressions of the upper and lower bounds where the number of available modes is equal to 128. As can be seen in Fig. 2(b), the derived upper and lower bounds of the ergodic capacity are close to the exact expression given in (7). We verify that our upper

Optics EXPRESS

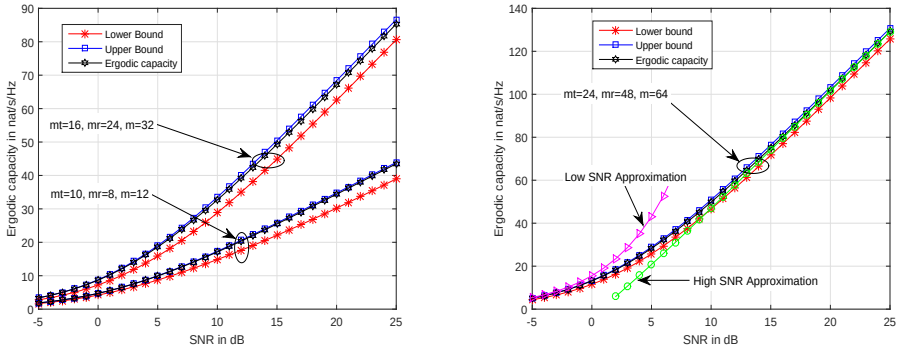


Fig. 3. (a) Comparison of upper bound, lower bound and ergodic capacity in nats per channel use versus SNR in dB when $m_t + m_r$ is larger than the available modes m (b) Bounds, upper and lower SNR approximation of the ergodic capacity of the MIMO Jacobi-fading channel where the number of available modes $m = 64$ and $m_t + m_r > m$.

and lower bounds give good approximations of the ergodic capacity even for very large number of available modes (*i.e.* $m = 128$).

In Fig. 3(a), we investigate how close the ergodic capacity is to its upper and lower bounds in cases where $m_t + m_r > m$. We address this particular case using (4). It can be observed that the proposed upper bound on the ergodic capacity is extremely tight for all SNR regimes when m_r is larger than m_t. It is important to note that there exists a constant gap between the lower bound and the exact ergodic capacity at all SNR levels. When m_t is larger than m_r, such upper and lower bounds are close to ergodic capacity at all SNR regimes. For comparison purposes, we have depicted in Fig. 3(b) the ergodic capacity of the MIMO Jacobi-fading channel obtained by computer simulation, the upper/lower bounds and the high/low SNR approximations when the sum of transmit and receive modes, $m_t + m_r$, is larger than the total available modes, m. In the high SNR regimes, the ergodic capacity and its high SNR approximation curves are almost indistinguishable. Similarly, we observe that there is almost no difference between the ergodic capacity and its low SNR approximation in the low SNR regions, while there is a significant difference in the high SNR regimes. This difference can be explained by the fact that the first order Taylor's expansion of $\ln(1 + x)$ is not valid for high values of x.

5. Conclusion

In this paper, we derive new analytical expressions of the lower-bound and upper-bound on the ergodic capacity for uncorrelated MIMO Jacobi fading channels assuming that transmitter has no knowledge of the channel state information. Moreover, we derive accurate closed-form analytical approximations of ergodic capacity in the high and low SNR regimes. The simulation results show that the lower-bound and upper-bound expressions are very close to the ergodic capacity.

Acknowledgments

The author would like to thank professor N. Demni from the institut de recherche en mathématique de Rennes (IRMAR), université de Rennes 1, for fruitful discussions.

4.3.2 [C2]: Generalized expression of the MIMO ergodic capacity

This work was carried out in collaboration with Nizar Demni from university of Rennes 1. The main idea of our research work relies on deriving a simple exact expression for the ergodic capacity of a MIMO link with uncorrelated Jacobi fading in an additive white Gaussian noise environment. Then, using a limiting transition between Jacobi and associated Laguerre polynomials, we proposed to derive a similar expression formula for the uncorrelated MIMO Rayleigh fading channels.

The exact ergodic capacity of the MIMO Jacobi fading channels has been derived in [25] by integrating the mutual information over the joint density of the eigenvalues of the hermitian channel matrix $\mathbf{HH}^{\dagger}$. Using the fact that the probability density of the eigenvalues of a random matrix from unitary ensemble can be expressed in terms of the Christoffel-Darboux kernel, we derived a new analytical expression in [CI47]. The main drawback of the new representation of the ergodic capacity is the dependence of the Christoffel-Darboux kernel on the size of the matrix. Indeed, its diagonal is written either as a sum of squares of Jacobi polynomials and the number of terms in this sum equals the size of the matrix least one, or by means of the Christoffel-Darboux formula as a difference of the product of two Jacobi polynomials whose degrees depend on the size of the matrix.

Based on the relationship between the ergodic capacity expression derived in [25] and the moments of the eigenvalues density of the Jacobi random matrix, we provide a new expression for the ergodic capacity of the MIMO Jacobi fading channel. Moreover, we show that the new formula is an average of some function over the signal to noise ratio, and it has the merit to have a simple dependence on the size of the matrix which allows for easier and more precise numerical simulations.

Ergodic Capacity of Jacobi MIMO Channels

Theorem 4 : *Assume that $t \leq r$ and $t + r < m$, then the ergodic capacity of uncorrelated MIMO Jacobi fading channels, with only CSI at the receiver side, is given by*

$$C_{t,r}^{m,\rho} = B_{t,r}^{m} \int_{0}^{1} u^{a-1}(1-u)^{b-2} P_{t-1}^{a-1,b}(1-2u) \times P_{t}^{a-1,b-2}(1-2u) Li_2(-\rho u)\, du \tag{4.14}$$

where $a = r - t + 1$, $b = m - r - t + 1$, and $B_{t,r}^{m} = \frac{t!\,(m-t)!}{\Gamma(r)\,\Gamma(m-r)}$. The function $Li_2(.)$ is the dilogarithm function.

An important asymptotic property of the Jacobi polynomial is the fact that it can be reduced

to the q^{th}-associated Laguerre polynomial of parameter $\alpha \geq 0$ through the following limit

$$L_q^{\alpha}(x) = \lim_{\beta \to \infty} P_q^{\alpha,\beta}\left(1 - \frac{2x}{\beta}\right), \; x > 0 \tag{4.15}$$

Using (4.15), we are able to give another expression for the ergodic capacity expression of the wireless MIMO Rayleigh fading channels. Indeed, the parameter b in (4) can be interpreted as the power loss through the optical fiber. Therefore, as b becomes very large, the channel matrix starts to look like a complex Gaussian matrix with independent and identically distributed entries. As a matter of fact, the MIMO Jacobi fading channel approaches the MIMO Rayleigh fading channel in the large b-limit corresponding to a huge waste of input power through the optical fiber.

Finally, the ergodic capacity expression (4) converges as $b \to \infty$ to the ergodic capacity of the uncorrelated MIMO Rayleigh fading channel, which has been already considered by Telatar [36, Theorem 2].

Ergodic Capacity of Rayleigh MIMO Channels

Theorem 5 : *Let us assume that the CSI is known perfectly at the receiver side, the ergodic capacity of the uncorrelated MIMO Rayleigh fading channel with t transmitters and r receivers, with $r \geq t$, can be expressed as*

$$C_{t,r}^{\rho} = \frac{t!}{(r-1)!} \int_0^{+\infty} u^{r-t}\, e^{-u}\, L_{t-1}^{r-t}(u)\, L_t^{r-t}(u) \times Li_2\left(-\rho u\right) du. \tag{4.16}$$

Several aspects of this work, related to the achievable sum rate for MIMO MMSE receiver in both fading channels, were further investigated. For more details, the reader is referred to our work published in IEEE Access Journal 2020 [J18] included hereafter.

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 19, 2020, accepted August 8, 2020, date of publication August 17, 2020, date of current version August 25, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3016925

Closed-Form Expressions of Ergodic Capacity and MMSE Achievable Sum Rate for MIMO Jacobi and Rayleigh Fading Channels

AMOR NAFKHA[1], (Senior Member, IEEE), AND NIZAR DEMNI[2]
[1]SCEE/IETR, CentraleSupélec, 35576 Cesson Sévigné, France
[2]IRMAR, Université de Rennes 1, 35042 Rennes, France

Corresponding author: Amor Nafkha (amor.nafkha@centralesupelec.fr)

This work was supported by the Internal Funding from CentraleSupélec.

ABSTRACT Multimode/multicore fibers are expected to provide an attractive solution to overcome the capacity limit of the current optical communication system. In the presence of strong crosstalk between modes and/or cores, the squared singular values of the input/output transfer matrix follow the law of the Jacobi ensemble of random matrices. Assuming that the channel state information is only available at the receiver, we derive a new expression for the ergodic capacity of the MIMO Jacobi fading channel. The proposed expression involves double integrals which can be easily evaluated for a high-dimensional MIMO scenario. Moreover, the method used in deriving this expression does not appeal to the classical one-point correlation function of the random matrix model. Using a limiting transition between Jacobi and associated Laguerre polynomials, we derive a similar formula for the ergodic capacity of the MIMO Rayleigh fading channel. Moreover, we derive a new exact closed-form expressions for the achievable sum rate of MIMO Jacobi and Rayleigh fading channels employing linear minimum mean squared error (MMSE) receivers. The analytical results are compared to the results obtained by Monte Carlo simulations and the related results available in the literature, which shows perfect agreement.

INDEX TERMS Additive white noise, channel capacity, detection algorithms, MIMO, optical fiber communication, optical crosstalk, probability density function, Rayleigh channels.

I. INTRODUCTION

To accommodate the exponential growth of data traffic over the last few years, the space division multiplexing (SDM) based on multicore optical fiber (MCF) or multimode optical fiber (MMF) is expected to overcome the barrier from capacity limit of single core fiber [1]–[3]. Recently, dense space division multiplexing (DSDM) with a large spatial multiplicity exceeding 30 was demonstrated with multicore technology [4], [5]. The main challenge in SDM occurs due to in-band crosstalk between multiple parallel transmission channels (cores and/or modes). This strong crosstalk can be dealt with using multiple-input multiple-output (MIMO) signal processing techniques [6]–[11]. Those techniques are widely used for wireless communication systems and they helped to drastically increase channel capacity.

The associate editor coordinating the review of this manuscript and approving it for publication was Luyu Zhao.

Assuming important crosstalk between cores and/or modes, negligible back-scattering and near lossless propagation, we can model the transmission optical channel as a random complex unitary matrix [12]–[14].

In [12], authors appealed to the Jacobi unitary ensemble (JUE) to establish the propagation channel model for MIMO communications over multimode and/or multicore optical fibers. As suggested in [17, Section I.C], the Jacobi fading channel can be used to accurately model the interference-limited multiuser MIMO system. From mathematical point of view, the JUE is a matrix-variate analogue of the beta random variable and consists of complex Hermitian random matrices which can be realized at least in two different ways [18], [22]: (**i**) We mimic the construction of a Beta distribution random variable B as a quotient of two independent Gamma random variables $B = X_1/(X1 + X_2)$ where X_1 and X_2 are replaced by two independent complex central Wishart matrices [22]. We assume that the sum $(X_1 + X_2)$ is reversible. (**ii**) We can

draw a Haar distributed unitary matrix then take the square of the radial part of an upper-left sub-matrix [18]. By a known fact for unitarily invariant-random matrices [22], the average of any symmetric function with respect to the eigenvalues density can be expressed through the one-point correlation function, also known as the single-particle density. In particular, the ergodic capacity of a matrix drawn from the JUE can be represented by an integral where the integrand involves the Christoffel-Darboux kernel associated with Jacobi polynomials ([22], p.384). The drawback of this representation is the dependence of this kernel on the size of the matrix. Indeed, its diagonal is written either as a sum of squares of Jacobi polynomials and the number of terms in this sum equals the size of the matrix least one, or by means of the Christoffel-Darboux formula as a difference of the product of two Jacobi polynomials whose degrees depend on the size of the matrix. To the best of our knowledge, this is the first study that derives exact expression of the ergodic capacity as a double integral over a suitable region. Recently in [19], [20], the authors derived expressions for the exact moments of the mutual information in the high-SNR regime for MIMO Jacobi fading channel.The obtained exact moments lead to closed-form approximations to the outage probability.

In this paper, we provide a new expression for the ergodic capacity of the MIMO Jacobi fading channel relying this time on the formula derived in [24] for the moments of the eigenvalues density of the Jacobi random matrix. The obtained expression shows that the ergodic capacity is an average of some function over the signal to noise ratio (SNR), and it has the merit to have a simple dependence on the size of the matrix which allows for easier and more precise numerical simulations. By a limiting transition between Jacobi and associated Laguerre polynomials [25], we derive a similar expression for the ergodic capacity of the MIMO Rayleigh fading channel [21]. Using the derived expressions and the work of McKay *et al.* [41], we are able to derive closed-form formulas for the achievable sum-rate of MIMO Jacobi and Rayleigh fading channels employing linear minimum mean squared error (MMSE) receivers.

The paper is organized as follows. In Section II, we recall some notations, definitions of random matrices and special functions occurring in the remainder of the paper. Section III introduces the MIMO Jacobi fading channel and the discrete-time input-output relation. In Section IV, an exact closed-form expression is derived for the ergodic capacity of MIMO Jacobi fading channel. Using the results of the previous section, we derive a new exact closed-form expression of the ergodic capacity of the MIMO Rayleigh fading channel in Section V. In both MIMO Jacobi and Rayleigh fading channels, we provide new closed-form expressions for the achievable sum rate of MIMO MMSE receivers in Section VI. In Section VII, we demonstrate the accuracy of the analytical expressions through Monte Carlo simulations. Finally, Section VIII is devoted to concluding remarks, while mathematical proofs are deferred to the appendices.

II. BASIC DEFINITIONS AND NOTATIONS

Throughout this paper, the following notations and definitions are used. We start with those concerned with special functions for which the reader is referred to the original book of Ismail [25]. The Pochhammer symbol $(x)_k$ with $x \in \mathbb{R}$ and $k \in \mathbb{N}$ is defined by

$$(x)_k = x(x+1)\dots(x+k-1); \quad (x)_0 = 1 \tag{1}$$

For $x > 0$, it is clear that

$$(x)_k = \frac{\Gamma(x+k)}{\Gamma(x)} \tag{2}$$

where $\Gamma(.)$ is the Gamma function. Note that if $x = -q$ is a non positive integer then

$$(-q)_k = \begin{cases} (-1)^k \dfrac{q!}{(q-k)!} & \text{if } k \geq q \\ 0 & \text{if } k < q \end{cases} \tag{3}$$

The Gauss hypergeometric function ${}_2F_1(.)$ is defined for complex $|z| < 1$ by the following convergent power series

$${}_2F_1(\theta, \sigma, \gamma, z) = \sum_{k=0}^{\infty} \frac{(\theta)_k (\sigma)_k}{(\gamma)_k k!} z^k \tag{4}$$

where $(.)_k$ denotes the Pochhammer symbol defined in (1) and θ, σ, γ are real parameters with $\gamma \neq \{0, -1, -2, \dots\}$. The function ${}_2F_1(.)$ has an analytic continuation to the complex plane cut along the half-line $[1, \infty[$. In particular, the Jacobi polynomials $P_q^{\alpha,\beta}(x)$ of degree q and parameters $\alpha > -1, \beta > -1$ can also be expressed in terms of the Gauss hypergeometric function (4) as follows

$$P_q^{\alpha,\beta}(x) = \frac{(\eta)_q}{q!} {}_2F_1\left(-q, q+\eta+\beta, \eta; \frac{1-x}{2}\right) \tag{5}$$

where $\eta = \alpha + 1$. An important asymptotic property of the Jacobi polynomial is the fact that it can be reduced to the q-th associated Laguerre polynomial of parameter $\alpha \geq 0$ through the following limit

$$L_q^{\alpha}(x) = \lim_{\beta \to \infty} P_q^{\alpha,\beta}\left(1 - \frac{2x}{\beta}\right), \quad x > 0 \tag{6}$$

Now, we come to the notations and the definitions related with random matrices, and refer the reader to [18], [22], [23]. Firstly, the Hermitian transpose and the determinant of a complex matrix $\mathbf{A}$ are denoted by $\mathbf{A}^\dagger$ and $\det(\mathbf{A})$ respectively. Secondly, the Laguerre unitary ensemble (LUE) is formed out of non negative definite matrices $\mathbf{A}^\dagger\mathbf{A}$ where $\mathbf{A}$ is a rectangular $m \times n$ matrix, with $m \geq n$, whose entries are complex independent Gaussian random variables. A matrix from the LUE is often referred to as a complex Wishart matrix and (m, n) are its degrees of freedom and its size respectively. Finally, let $\mathbf{X} = \mathbf{A}^\dagger\mathbf{A}$ and $\mathbf{Y} = \mathbf{B}^\dagger\mathbf{B}$ be two independent (m_1, n) and (m_2, n) complex Wishart matrices. Assume $m_1 + m_2 \geq n$, then $\mathbf{X}+\mathbf{Y}$ is positive definite and the random matrix $\mathbf{J}$, defined as $\mathbf{J} = (\mathbf{X}+\mathbf{Y})^{-1/2}\mathbf{X}(\mathbf{X}+\mathbf{Y})^{-1/2}$, belongs to the JUE. The matrix $\mathbf{J}$ is unitarily-invariant and satisfies $\mathbf{0}_n \leq \mathbf{J} \leq \mathbf{I}_n$ where $\mathbf{0}_n, \mathbf{I}_n$ stand for the null and the

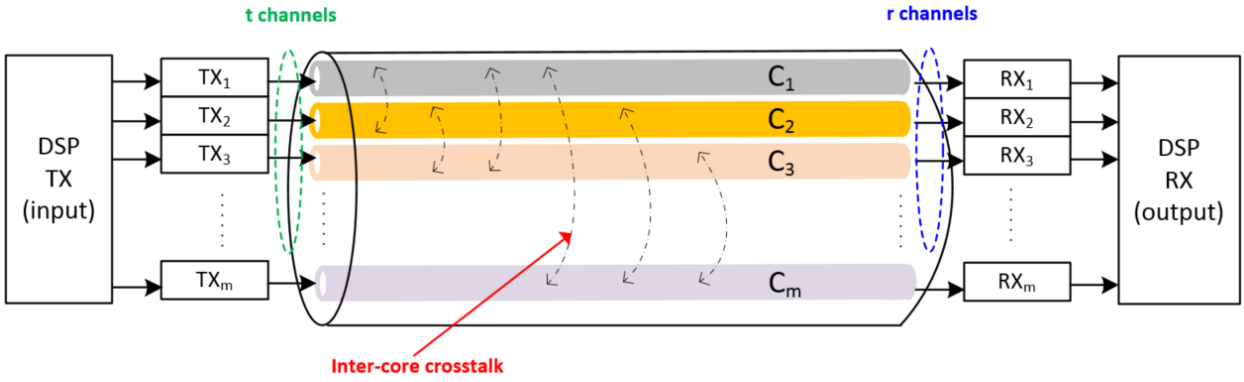


FIGURE 1. Schematic of an optical MIMO system with an MCF, C_k indicates k^{th} fiber core, with $k \in \{1, 2, \ldots, m\}$.

identity matrices respectively.[1] If $m_1, m_2 \geq n$ then the matrix $\mathbf{J}$ and the matrix $(\mathbf{I}_n - \mathbf{J})$ are positive definite and the joint distribution of the ordered eigenvalues of $\mathbf{J}$ has a probability density function given by

$$\begin{aligned}&\mathcal{F}_{a,b,n}(\lambda_1, \ldots, \lambda_n)\\ &\quad = Z_{a,b,n}^{-1} \prod_{1\leq j\leq n} \lambda_j^{a-1}(1-\lambda_j)^{b-1}\\ &\qquad \times [V(\lambda_1, \ldots, \lambda_n)]^2 1_{\{0<\lambda_1<\cdots<\lambda_n<1\}} \end{aligned} \tag{7}$$

with respect to Lebesgue measure $d\lambda = d\lambda_1 \ldots d\lambda_n$. Here, $a = m_1 - n + 1, b = m_2 - n + 1$, $Z_{a,b,n}$ is a normalization constant read off from the Selberg integral [23], [24]:

$$Z_{a,b,n} = \prod_{j=1}^{n} \frac{\Gamma(a+j-1)\Gamma(b+j-1)\Gamma(j)}{\Gamma(a+b+n+j-2)},$$

$1_{\{.\}}$ stands for the indicator function: given a set A

$$1_{\{x\in A\}} = \begin{cases} 1 & \textbf{if} \quad x \in A \\ 0 & \textbf{otherwise}, \end{cases}$$

and $V(\lambda_1, \ldots, \lambda_n) = \prod_{1\leq j<k\leq n} (\lambda_j - \lambda_k)$ is the Vandermonde polynomial. As suggested in [18], we can construct the matrix $\mathbf{J}$ from the JUE ensemble as follows: let $\mathbf{U}$ be an $m \times m$ Haar-distributed unitary matrix. Let t and r be two positive integers such that $t + r \leq m$ and $t \leq r$. Let also $\mathbf{H}$ be the $r \times t$ upper-left corner of $\mathbf{U}$, then the joint distribution of the ordered eigenvalues of matrix $\mathbf{J} = \mathbf{H}^\dagger\mathbf{H}$ is given by (7) with parameters $a = r - t + 1, b = m - r - t + 1$, and $n = t$.

In the sequel the following notation will be used. $\mathbb{E}_\nu[.]$ will denote the expectation with respect to the random variable ν. We will denote the matrix determinant by $\det(.)$, and the matrix inverse by $[.]^{-1}$. The (i, j)-th element of a matrix $\mathbf{A}$ is indicated by $[\mathbf{A}]_{i,j}$.

III. SYSTEM MODEL

We consider an optical space division multiplexing where the multiple channels correspond to the number of excited modes and/or cores within the optical fiber. The coupling between different modes and/or cores can be described by scattering matrix formalism as reported in [14], [34]–[36]. In this paper, we consider m-channel *near* lossless optical fiber with $t \leq m$ transmitting excited channels and $r \leq m$ receiving channels, as indicated in Fig. 1 for multicore optical fiber scenario. The scattering matrix formalism can describe very simply the propagation through the fiber using $2m \times 2m$ scattering matrix $\mathbf{S}$ given as

$$\mathbf{S} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{T}_2 \\ \mathbf{T}_1 & \mathbf{R}_2 \end{bmatrix}, \tag{8}$$

where the $m \times m$ complex block matrices $\mathbf{R}_1$ and $\mathbf{R}_2$ describe the reflection coefficients in input and output ports of the fiber, respectively. Similarly, the $m \times m$ complex block matrices $\mathbf{T}_1$ and $\mathbf{T}_2$ stand for the transmission coefficients through the fiber from input to output sides and vice versa, respectively. We assume a strong crosstalk between cores or modes, negligible backscattering, near-lossless propagation, and reciprocal characteristics of the fiber. Thus, we model the scattering matrix as a complex unitary symmetric matrix [16], (*i.e.* $\mathbf{S}^\dagger\mathbf{S} = \mathbf{I}_{2m}$). Therefore, the four Hermitian matrices $\mathbf{T}_1\mathbf{T}_1^\dagger, \mathbf{T}_2\mathbf{T}_2^\dagger, \mathbf{I}_m - \mathbf{R}_2\mathbf{R}_2^\dagger$, and $\mathbf{I}_m - \mathbf{R}_1\mathbf{R}_1^\dagger$ have the same set of eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$. Each of these m transmission eigenvalues is a real number belong to the interval $[0, 1]$. Assuming a unitary coupling among all transmission modes the overall transfer matrix $\mathbf{T}_1$ can be described by a $m \times m$ unitary matrix, where each matrix entry $[\mathbf{T}_1]_{ij}$ represents the complex path gain from transmitted mode i to received mode j. Moreover, the transmission matrix $\mathbf{T}_1$ has a Haar distribution over the group of complex unitary matrices [12], [14]. Given the fact that only $t \leq m$ and $r \leq m$ modes are addressed by the transmitter and receiver, respectively, the effective transmission channel matrix $\mathbf{H} \in \mathbb{C}^{r\times t}$ is a truncated[2] version of $\mathbf{T}_1$. As a result, the corresponding MIMO channel for this system reads

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} \tag{9}$$

where $\mathbf{y} \in \mathbb{C}^{r\times 1}$ is the received signal vector of dimension $r \times 1$, $\mathbf{x} \in \mathbb{C}^{t\times 1}$ is a $t \times 1$ transmitted signal vector with

[1] For two square matrices A and B, we write $A \leq B$ when $B - A$ is a non negative matrix.

[2] Without loss of generality, the effective transmission channel matrix $\mathbf{H}$ is the $r \times t$ upper-left corner of the transmission matrix $\mathbf{T}_1$ [18], [37]

TABLE 1. List of main variables.

Variables	Descriptions
m	Number of overall available modes/cores
t	Number of transmitted modes/cores
r	Number of received modes/cores
$\mathbf{S}$	$\mathbb{C}^{2m\times 2m}$ scattering matrix
$\mathbf{R}_1$	$\mathbb{C}^{m\times m}$ matrix contains the reflection coefficients of the m input modes/cores of the fiber
$\mathbf{T}_1$	$\mathbb{C}^{m\times m}$ matrix contains the transmission coefficients from the input to the output of the multi-mode/core fiber
$\mathbf{R}_2$	$\mathbb{C}^{m\times m}$ matrix contains the reflection coefficients of the m output modes/cores of the fiber
$\mathbf{T}_2$	$\mathbb{C}^{m\times m}$ matrix contains the transmission coefficients from the output to the input of the multi-mode/core fiber
$\mathbf{H}$	$\mathbb{C}^{r\times t}$ matrix contains the transmission coefficients of the effective channel when $t \leq m$ transmitting modes/cores and $r \leq m$ receiving modes/cores are used.
$C_{t,r}^{m,\rho}$	Ergodic capacity of MIMO Jacobi fading channel with t transmitting modes/cores, r receiving modes/cores, m overall available modes/cores (*i.e.* $m \geq t$, $m \geq r$), and a signal to noise ratio equal to ρ.
$C_{t,r}^{\rho}$	Ergodic capacity of MIMO Rayleigh fading channel with t transmitting antennas, r receiving antennas, and a signal to noise ratio equal to ρ.
$\mathcal{R}$	General expression of the achievable ergodic sum rate for the MMSE receiver under MIMO channel.
$\mathcal{R}_{t,r}^{m,\rho}$	Achievable ergodic sum rate of MMSE receiver under MIMO Jacobi fading channel with t transmitting modes /cores, r receiving modes/cores, m available modes/cores, and a signal to noise ratio equal to ρ.
$\mathcal{R}_{t,r}^{\rho}$	Achievable ergodic sum rate of MMSE receiver under MIMO Rayleigh fading channel with t transmitting antennas, r receiving antennas, and a signal to noise ratio equal to ρ.

covariance matrix equal to $\frac{\mathcal{P}}{t}\mathbf{I}_t$, and $\mathbf{z} \in \mathbb{C}^{r\times 1}$ is a $r \times 1$ zero mean additive white circularly symmetric complex Gaussian noise vector with covariance matrix equal to $\sigma^2\mathbf{I}_r$ [46], [47]. The variable $\mathcal{P}$ is the total transmit power across the t modes/cores, and σ^2 is the Gaussian noise variance. Table 1 provides the list of main variables used in this manuscript.

IV. ERGODIC CAPACITY OF MIMO JACOBI CHANNEL

The expression of the ergodic capacity of the MIMO Jacobi fading channel was firstly expressed in [12] as an integral over [0, 1] of the sum of squares of $min(t, r)$ Jacobi polynomials with real coefficients, is the same theoretical approach adopted by Telatar [21]. Recently, ergodic capacity bounds (upper and lower) of the MIMO Jacobi fading channel were derived in [26] and [27]. In [26], authors derived lower bound and low SNR approximation of the ergodic capacity of MIMO Jocobi fading channel by rearranging the analytical expression given in [12, Eq. (11)]. Using recent results on the determinant of the Jacobi unitary ensemble and classical Jensen's and Minkowski's inequalities, the authors, in [27], derived tight closed-form bounds for the ergodic capacity [12, Eq. (11)]. In addition, they also provided accurate closed-form analytical approximations of ergodic capacity at high and low signal to noise ratio regimes.

In this section, we provide a novel and simple closed-form expression of the ergodic capacity in the setting of MIMO Jacobi fading channel. We assume that the channel state information (CSI) is only known at the receiver, not at the transmitter. The investigation of the ergodic capacity of the MIMO Jacobi fading channel under unknown CSI at the receiver side is out of scope of the present work. Without loss of generality, in the sequel of the present paper, we shall assume that $t \leq r$ and $m \geq t + r$. The channel ergodic capacity, under a total average transmit power constraint, is then achieved by taking $\mathbf{x}$ as a vector of zero-mean circularly symmetric complex Gaussian components with covariance matrix $\mathcal{P}\mathbf{I}_t/t$, and it is given by [12, Eq. (10)]

$$C_{t,r}^{m,\rho} = \mathbb{E}_{\mathbf{H}}\left[\ln\det\left(\mathbf{I}_t + \frac{\rho\mathbf{H}^{\dagger}\mathbf{H}}{t}\right)\right], \quad t \leq r, \tag{10}$$

where $\mathbb{E}_{\mathbf{H}}[.]$ denotes the expectation over all channel realizations, ln is the natural logarithm function and $\rho = \frac{\mathcal{P}}{\sigma^2}$ is the average signal-to-noise ratio (SNR). Given the fact that both matrices $\mathbf{H}^{\dagger}\mathbf{H}$ and $\mathbf{H}\mathbf{H}^{\dagger}$ share the same non zero eigenvalues even if $m < t + r$, the authors in [12, Theorem 2] shows that the ergodic capacity is given by

$$C_{t,r}^{m,\rho} = (t + r - m)C_{1,1}^{1,\rho} + C_{m-r,m-t}^{m,\rho}, \quad t \leq r. \tag{11}$$

In the sequel of this paper, we assume further that $m > t + r \Leftrightarrow b \geq 2$ and the case $m = r + t \Leftrightarrow b = 1$ can be dealt with by a limiting procedure. Actually, our formula for the ergodic capacity derived below is valid for real $a > 0, b > 1$, and we can consider its limit as $b \to 1$. However, for ease of reading, we postpone the details of the computations relative to this limiting procedure to a future forthcoming paper.

Now, recall that the random matrix $\mathbf{H}^{\dagger}\mathbf{H}$ has the Jacobi distribution, then its ordered eigenvalues have the joint density given by (7) with parameters $a = r - t + 1$ and $b = m - t - r + 1$. Using (7), we can explicitly express the ergodic capacity (10) as

$$C_{t,r}^{m,\rho} = \int \sum_{k=1}^{t} \ln(1 + \rho\lambda_k)\, \mathcal{F}_{a,b,t}(\lambda_1, \ldots, \lambda_t)\, d\lambda_1 \ldots d\lambda_t \tag{12}$$

A major step towards our main result is the following proposition.

Proposition 1: For any $\rho \in (0, 1)$,

$$\Psi C_{t,r}^{m,\rho} = A_{t,r}\, \rho^{t-1} P_{t-1}^{r-t,m-t-r-1}\left(\frac{\rho+2}{\rho}\right) {}_2F_1(t + 1, r + 1, m + 1; -\rho) \tag{13}$$

where the operator $\Psi = \left[D_\rho(\rho D_\rho)\right]$ with D_ρ is the derivative operator with respect to ρ, and $A_{t,r} = \frac{r\,t!}{(m-t+1)_t}$.

Proof: The full proof for Proposition 1 can be found at the Appendix A. ■

Using proposition 1, we are able to derive the following new expression of the ergodic capacity of MIMO Jacobi fading channel.

Theorem 1: Assume that $r \geq t, m > t+r$, and $\rho \geq 0$, then the ergodic capacity of an uncorrelated MIMO Jacobi fading

channel is given by

$$C_{t,r}^{m,\rho} = B_{t,r}^{m} \int_0^1 u^{a-1}(1-u)^{b-2} P_{t-1}^{a-1,b}(1-2u) \times P_t^{a-1,b-2}(1-2u) Li_2(-\rho u)\, du \quad (14)$$

where $a = r - t + 1$, $b = m - r - t + 1$, and $B_{t,r}^{m} = \frac{t!\,(m-t)!}{\Gamma(r)\,\Gamma(m-r)}$. The function $Li_2(.)$ is the dilogarithm function [50] defined as

$$Li_2(z) = -\int_0^z \frac{\ln(1-v)}{v} dv,\ z \in \mathbb{C}$$

Proof: The appendix B contains proof of Theorem 1. ■

V. ERGODIC CAPACITY OF MIMO RAYLEIGH CHANNEL

The ergodic capacity of the MIMO Rayleigh fading channel was extensively examined in order to provide a compact mathematical expression in several papers [21], [28]–[32]. In [28], [29], the ergodic capacity is provided using the Christoffel-Darboux kernel, and the authors replaced the Laguerre polynomials by their expressions which is a known fact in invariant random matrix models. In [30]–[32], authors derived a closed form expression of moment generating function (MGF) so that the ergodic capacity may be derived by taking the first derivative. However, this expression of MGF relies on the Cauchy-Binet Theorem and only gives a hypergeometric function of matrix arguments [33], from which by derivatives, we can get again an alternating sum coming from the determinant. Consequently, we can not derive the proposed expression of the ergodic capacity (15) from this sum.

Using the limiting transition (6) between Jacobi and associated Laguerre polynomials, we are able to give another expression for the ergodic capacity expression of the wireless MIMO Rayleigh fading channel. Indeed, it was shown in [14], [15], that the parameter b in (14) can be interpreted as the power loss through the optical fiber. Therefore, as b becomes large, the channel matrix $\mathbf{H}$ in (9) starts to look like a complex Gaussian matrix with independent and identically distributed entries. As a matter of fact, the MIMO Jacobi fading channel approaches the MIMO Rayleigh fading channel in the large b-limit corresponding to a huge waste of input power through the optical fiber. In particular, the ergodic capacity (14) converges as $b \to \infty$ to the ergodic capacity of the uncorrelated MIMO Rayleigh fading channel already considered by Telatar in [21, Theorem 2], and we are able to derive the following new result. Note that the pioneer work of Telatar was recently revisited by Wei in [45].

Theorem 2: The ergodic capacity of the uncorrelated MIMO Rayleigh fading channel with t transmitters and r receivers, with $r \geq t$, can be expressed

$$C_{t,r}^{\rho} = \frac{t!}{(r-1)!} \int_0^{+\infty} u^{r-t} e^{-u} L_{t-1}^{r-t}(u) L_t^{r-t}(u) \times Li_2(-\rho u)\, du. \quad (15)$$

Proof: The reader can refer to Appendix C for the proof of Theorem 2. ■

VI. ACHIEVABLE SUM RATE OF MIMO MMSE RECEIVER

In this section, we are interested in the performance of linear MMSE receivers. Assuming to employ a MMSE filter, and that each filter output is independently decoded. Let ρ_k denotes the instantaneous signal to interference-plus-noise ratio (SINR) to the k^{th} MIMO subchannel.[3] Minimizing the mean squared error between the output of a linear MMSE receiver and the actually transmitted symbol $\mathbf{x}_k$ for $1 \leq k \leq t$ leads to the filter vector

$$\mathbf{g}_k = \left(\mathbf{H}\mathbf{H}^{\dagger} + \frac{t}{\rho}\mathbf{I}_r\right)^{-1} \mathbf{h}_k \quad (16)$$

where $\mathbf{h}_k$ is the k^{th} column of channel matrix $\mathbf{H}$. Applying this filter vector into (9) yields

$$\mathbf{x}_k^{mmse} = \mathbf{g}_k^{\dagger}\mathbf{y} \quad (17)$$

The achievable ergodic sum rate for the MMSE receiver can be expressed as

$$\mathcal{R} = \sum_{k=1}^{t} \mathbb{E}_{\rho_k}\left[\ln(1+\rho_k)\right] \quad (18)$$

As shown in [38], [41], [42], [44], the instantaneous received SINR for the k^{th} MMSE filter output is given by

$$\rho_k = \frac{1}{\left[\left(\mathbf{I}_t + (\rho/t)\mathbf{H}^{\dagger}\mathbf{H}\right)^{-1}\right]_{k,k}} - 1 \quad (19)$$

In general, the analytical expression of the probability density function of ρ_k is difficult to determine. This situation makes the direct evaluation of the achievable ergodic MMSE sum rate (18) very difficult.

Let $\mathbf{H}_k$ denotes the sub-matrix obtained by striking $\mathbf{h}_k$ out of $\mathbf{H}$. As shown in [49, Theorem 1.33], the k^{th} diagonal term of the matrix, $\left(\mathbf{I}_t + \frac{\rho\mathbf{H}^{\dagger}\mathbf{H}}{t}\right)^{-1}$, can be expressed as

$$\left[\left(\mathbf{I}_t + \frac{\rho\mathbf{H}^{\dagger}\mathbf{H}}{t}\right)^{-1}\right]_{k,k} = \frac{\det\left(\mathbf{I}_{t-1} + \frac{\rho\mathbf{H}_k^{\dagger}\mathbf{H}_k}{t}\right)}{\det\left(\mathbf{I}_t + \frac{\rho\mathbf{H}^{\dagger}\mathbf{H}}{t}\right)}, \quad (20)$$

where the matrix $\mathbf{H}_k^{\dagger}\mathbf{H}_k$ is the $k \times k$ principal minor of matrix $\mathbf{H}^{\dagger}\mathbf{H}$ defined by striking out the k^{th} column of $\mathbf{H}$.

Similarly to what has been developed in [41], by substituting (19) and (20) in (18), we can obtained the following expression of the achievable ergodic sum rate for the MIMO MMSE receiver.

$$\mathcal{R} = t\,\mathbb{E}_{\mathbf{H}}\left[\ln\det\left(\mathbf{I}_t + \frac{\rho\mathbf{H}^{\dagger}\mathbf{H}}{t}\right)\right] - \sum_{k=1}^{t} \mathbb{E}_{\mathbf{H}_k}\left[\ln\det\left(\mathbf{I}_{t-1} + \frac{\rho\mathbf{H}_k^{\dagger}\mathbf{H}_k}{t}\right)\right] \quad (21)$$

By employing the Haar invariant property, exchanging any two different rows or/and exchanging two different columns

[3] In our case ($t \leq r$), the MIMO channel can be decomposed into t parallel subchannels.

do not change the joint distribution of the entries, the joint probability density function of the ordered eigenvalues of $\mathbf{H}_k^{\dagger}\mathbf{H}_k$ is the same as $\mathbf{H}_j^{\dagger}\mathbf{H}_j$ for all $j \neq k$ and $j \in \{1, .., t\}$. Thus, the achievable ergodic sum rate for the MMSE receiver can be expressed as

$$\mathcal{R} = t\,\mathbb{E}_{\mathbf{H}}\left[\ln\det\left(\mathbf{I}_t + \frac{\rho\mathbf{H}^{\dagger}\mathbf{H}}{t}\right)\right] - t\,\mathbb{E}_{\mathbf{H}_1}\left[\ln\det\left(\mathbf{I}_{t-1} + \frac{\rho\mathbf{H}_t^{\dagger}\mathbf{H}_t}{t}\right)\right] \quad (22)$$

In case of MIMO Jacobi fading channel, the matrix $\mathbf{H}_t$ is the $r \times (t-1)$ left corner of the channel matrix $\mathbf{H}$. Then, the joint distribution density function of the ordered eigenvalues of $\mathbf{H}_t^{\dagger}\mathbf{H}_t$ is given by (7) with parameters $a = r-t+2$, $b = m-r-t+2$, and $n = t-1$. The following result characterizes the achievable ergodic sum rate of the MIMO Jacobi fading channel when the linear MMSE filter is used at the receiver side.

Theorem 3: For any $\rho \geq 0$, The achievable ergodic sum rate of MMSE receiver under MIMO Jacobi fading channel is given by

$$\mathcal{R}_{t,r}^{m,\rho} = t\left[C_{t,r}^{m,\rho} - C_{t-1,r}^{m,\frac{(t-1)\rho}{t}}\right] \quad (23)$$

Proof: By substituting (14) into (22). ■

Very recently, Lim and Yoon [42] proposed closed form expression of the achievable sum rate for MMSE MIMO systems in uncorrelated Rayleigh environments. However, the derived expression, [42, eq.(67)], is not closed form and does not allow a better understand of the MMSE achievable sum rate due the use of the sum of Meijer G-functions (or equivalent representation in terms of generalized hypergeometric functions). In following corollary, we presented a novel and exact closed-form formula for ergodic achievable sum rate for MMSE receiver under MIMO Rayleigh fading channels.

Corollary 1: For any $\rho \geq 0$, The achievable ergodic sum rate of MMSE receiver under MIMO Rayleigh fading channels with $t \leq r$ can be expressed as

$$\mathcal{R}_{t,r}^{\rho} = r\left[\Psi(t,r,\rho) - \Psi(t,r+1,\rho)\right] + \frac{t!}{(r-1)!} \times \int_0^{+\infty} u^{r-t+1}e^{-u}L_{t-2}^{r-t+1}(u)L_{t-1}^{r-t+1}(u) \times \left[Li_2\left(\frac{-\rho u}{t}\right) - Li_2\left(\frac{-\rho(t-1)u}{t^2}\right)\right]du. \quad (24)$$

where

$$\Psi(t,r,\rho) = \frac{t!}{(r-1)!}\int_0^{+\infty} u^{r-t}e^{-u}\left[L_{t-1}^{r-t}(u)\right]^2 \times Li_2\left(\frac{-\rho u}{t}\right)du$$

Proof: By substituting (15) into (22). ■

VII. NUMERICAL RESULTS AND DISCUSSION

In this section, we present numerical results supporting the analytical expressions derived in Section IV and Section V. All of the Monte Carlo simulation results were obtained by averaging over 10^5 independent channel realization. For MIMO Rayleigh fading channels, the entries in $\mathbf{H} \in \mathbb{C}^{r\times t}$ are independent and identically distributed complex, zero mean Gaussian random variables with normalized unit magnitude variance, and they can be obtained using a built-in MATLAB function (*i.e.* "randn"). For the MIMO Jacobi fading channels, the simulation process is initialized firstly by creating a random complex Gaussian matrix $\mathbf{G} \in \mathbb{C}^{m\times m}$ with independent and identically distributed entries that are complex circularly symmetric Gaussian with zero mean and $1/2$ variance per dimension. Then, using QR decomposition then matrix $\mathbf{G}$ can be decomposed as $\mathbf{G} = \mathbf{QR}$ where $\mathbf{Q} \in \mathbb{C}^{m\times m}$ is a unitary matrix and $\mathbf{R} \in \mathbb{C}^{m\times m}$ is upper triangular matrix. Finally, the MIMO Jacobi fading channel $\mathbf{H}$ was constructed by taking the $r \times t$ sub-matrix in the upper-left corner of matrix $\mathbf{Q}$. In both MIMO channel cases, the ergodic capacity and achievable sum rate with MMSE receivers can be obtained by averaging (10) and (22), respectively, over all realization of the channel matrix $\mathbf{H}$. Herein, we consider the case where the channel state information is available at the receiver side. Figure 2 examines the ergodic capacity of the MIMO Jacobi fading channel as a function of the SNR, when the number of parallel transmission paths is fixed to $m = 20$ and the number of transmit modes equal to the number of receive modes $r = t$. It is evident that when we increase the number of transmitted and received modes, we improve the ergodic capacity of the system. As expected, the ergodic capacity increases with SNR. Figure 2 is also shown that the two theoretical expressions curves of the ergodic capacity (14) and [12, (11)] perfectly matched the simulation results.

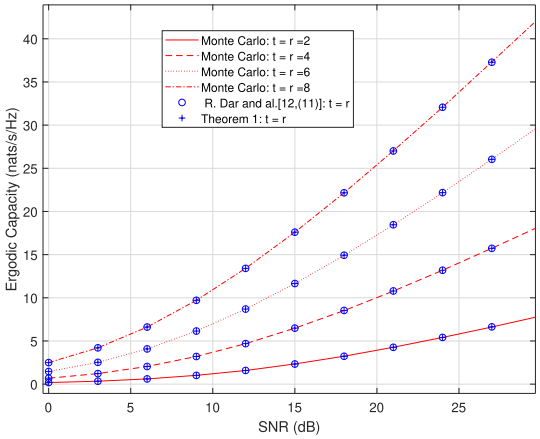


FIGURE 2. The variation of the ergodic capacity of MIMO Jacobi channel as a function of ρ for $m = 20$.

Figure 3 shows the theoretical and simulated ergodic capacity of MIMO Jacobi channel as a function of the number of received modes. Here, we fixed the number of parallel transmission paths to $m = 25$, the SNR to $\rho = 10$ dB, and the

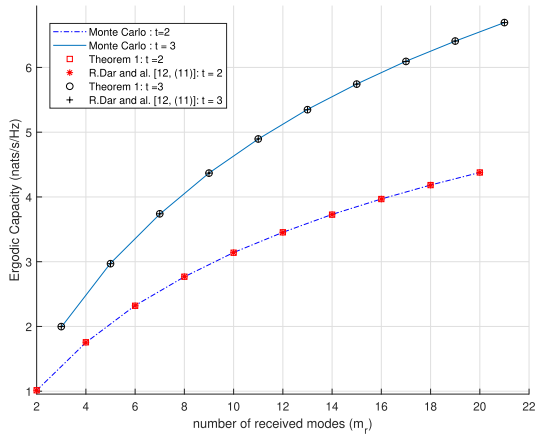


FIGURE 3. Ergodic capacity of MIMO Jacobi channel as function of receive cores and/or modes with $\rho = 10$ dB and $m = 25$.

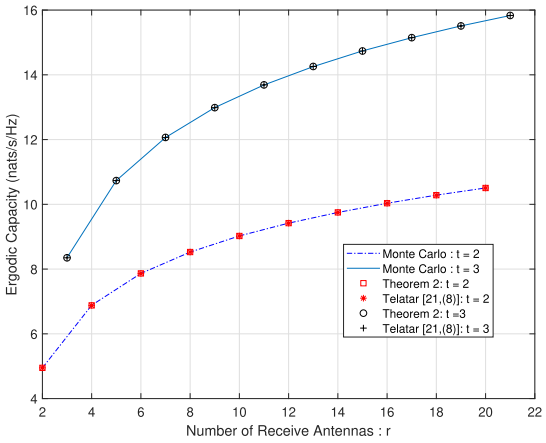


FIGURE 5. Ergodic capacity of MIMO Rayleigh channel when the number of received antennas increases and $\rho = 10$ dB.

number of transmit modes t to have following values $\{2, 3\}$. It is shown that every simulated curve is in excellent agreement with the theoretical curves calculated from (14) and [12, (11)]. The relationship between the channel capacity and the number of received modes is logarithmic. This implies that trying to improve the channel capacity by just increasing the number of received modes or cores is not efficient in the sense that the capacity increases logarithmically with r. The same relationship has been noted and discussed in the case of the uncorrelated MIMO Rayleigh fading channel (see Fig. 5, [21], and [40]).

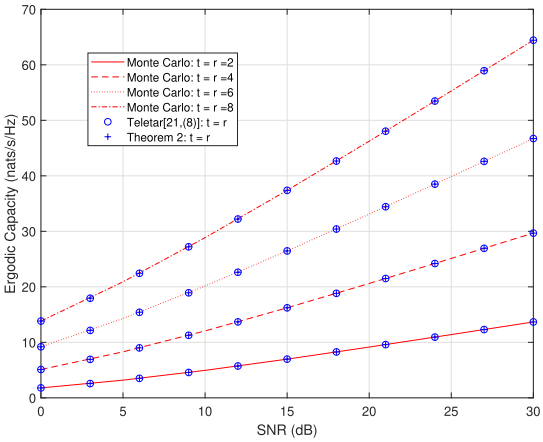


FIGURE 4. Ergodic capacity of the uncorrelated MIMO Rayleigh fading channel versus SNR for different numbers of transmit and receive antennas.

For the uncorrelated MIMO Rayleigh fading channel, the proposed expression of the ergodic capacity was verified through Monte Carlo experiments and it is shown in Fig. 4. In Fig. 4, the comparisons are shown between theoretical expressions and simulation values of the ergodic capacity as a function of the SNR. As we can observe in Fig. 4, for a given SNR, the capacity increases as the numbers of transmit and receive antennas grow. In all cases, the results demonstrate an excellent agreement between analytical expressions and Monte-Carlo simulations. Moreover, We can observe that the expression in (15) matches perfectly with the expression introduced by Telatar [21, Eq. (8)]. For the cases where $t = r = 2$ and $t = r = 4$, the obtained results are consistent with simulation results reported in [39], [40].

Figure 5 shows the ergodic capacity of uncorrelated MIMO Rayleigh fading channel of as the number of receive antennas r increases. As expected, we observe that the ergodic capacity increases in logarithmic scale with respect to r, this tallies with the result reported in [48, Eq. (6)]. As for optical MIMO channel, the three different ways to compute the uncorrelated MIMO Rayleigh fading channel capacity give the same results. These simulations were carried out to verify the mathematical derivation and no inconsistencies were noted.

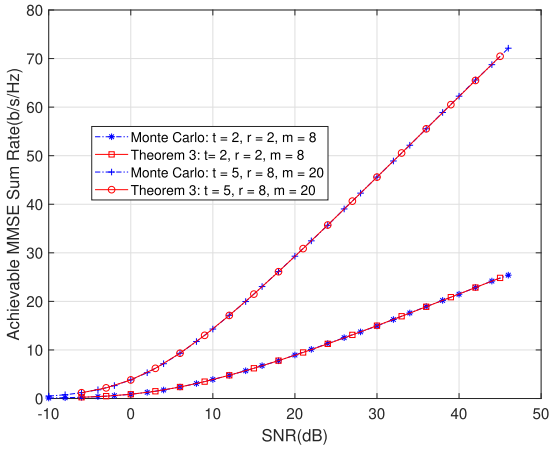


FIGURE 6. Evolution of the ergodic sum rate for the MMSE receiver over MIMO Jacobi fading channel.

We now focus on the ergodic sum rate for the MMSE receiver. We first consider the MIMO Jacobi fading channel. Fig. 6 shows the evolution of the ergodic sum rate for the MMSE receiver versus the SNR over the optical MIMO channel. For these results, we suppose that either $m = 20$ or $m = 8$. As expected, the ergodic sum rate increases with increasing SNR. Moreover, our simulation results show that

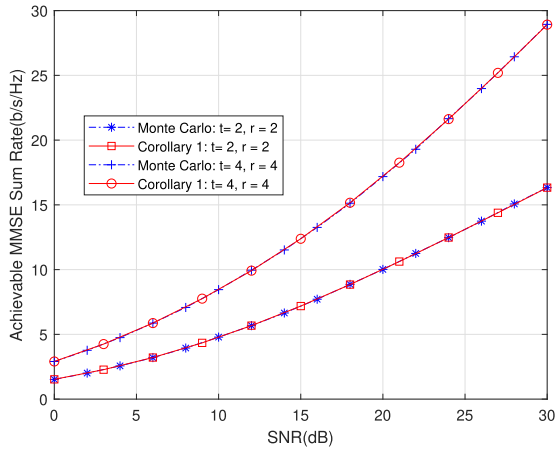


FIGURE 7. Evolution of the ergodic sum rate for the MMSE receiver over an uncorrelated MIMO Rayleigh fading channel.

the formula derived in Theorem 3 and Monte Carlo simulations provide the same results.

Finally, Fig. 7 shows the evolution of the ergodic sum rate for the MMSE receiver versus the SNR over an uncorrelated MIMO Rayleigh fading channel. We compare the sum rate obtained by means of Monte Carlo simulations and the one obtained with the formula derived in Corollary 1. Fig. 7 shows a perfect match between Monte Carlo and analytical result given in (24). It worth noting that, for $t = r = 2$ and $t = r = 4$, the obtained simulation results are the same as reported in [41], [42] and [43].

VIII. CONCLUSION

This paper has investigated the ergodic capacity of MIMO Jacobi fading channel which can be used to model accurately multimode and/or multicore optical fibers with the following characteristics: high crosstalk between modes and/or cores, negligible backscattering and near-lossless propagation. We assumed that a perfect channel state information (CSI) is only available at the receiver side, by using the joint distribution of eigenvalues of the Jacobi unitary ensemble, an exact expression of the ergodic capacity has been derived. By appealing to the limit relation between Jacobi and associated Laguerre polynomials, an exact expression of the ergodic capacity of MIMO Rayleigh fading channels has further been obtained. Furthermore, the above results led to exact expressions of the achievable sum rate for MIMO MMSE receiver in both fading channels. Monte Carlo simulations have been conducted to check the validity of the analytical results. Theoretical results show perfect matching with those obtained by simulations, and allow to derive tight bounds on the ergodic capacity for both MIMO fading channels. Considering the fact that wireless or fiber channels are subject to eavesdropping, we will address the MIMO secrecy capacity problem in future research papers.

APPENDIX A
PROOF OF PROPOSITION 1

For ease of reading, we simply denote below the ergodic capacity by $C(\rho)$. Moreover, the reader can easily check that our computations are valid for real $a > 0, b > 1$. We start by recalling from [24, Corollary 2.3] that for any $k \geq 1$,

$$\int \left(\sum_{i=1}^{t} \lambda_i^k\right) \mathcal{F}_{a,b,t}(\lambda) d\lambda = \frac{1}{k!} \sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} \prod_{j=-i}^{k-i-1} \frac{(t+j)(a+t+j-1)}{(a+b+2t+j-2)}.$$

Now, let $\rho \in [0, 1]$ and use the Taylor expansion

$$\ln(1+\rho\lambda_i) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(\rho\lambda_i)^k}{k}$$

to get

$$\sum_{i=1}^{t} \ln(1+\rho\lambda_i) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{\rho^k}{k} \left(\sum_{i=1}^{t} \lambda_i^k\right).$$

Consequently,

$$C(\rho) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \frac{\rho^k}{k!} \sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} \times \prod_{j=-i}^{k-i-1} \frac{(t+j)(a+t+j-1)}{(a+b+2t+j-2)}. \quad (25)$$

Changing the summation order and performing the index change $k \mapsto k+i+1$ in (25), we get

$$C(\rho) = \sum_{i=0}^{\infty} (-1)^i \sum_{k=0}^{\infty} \frac{(-1)^{k+i}}{(k+i+1)} \frac{\rho^{k+i+1}}{(k+i+1)!} \binom{k+i}{i} \prod_{j=-i}^{k} \frac{(t+j)(a+t+j-1)}{(a+b+2t+j-2)}.$$

Now, one can observe that the product displayed in the right hand side of the last equality vanishes whenever $i \geq t$ due to the presence of the factor $j+t, -i \leq j \leq k$. Thus, the first series terminates at $i = t-1$ and together with the index change $j \mapsto t+j$ in the product lead to

$$C(\rho) = \sum_{i=0}^{t-1} \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+i+1)} \frac{\rho^{k+i+1}}{(k+i+1)!} \binom{k+i}{i} \prod_{j=t-i}^{k+t} \frac{(j)(a+j-1)}{(a+b+t+j-2)}.$$

Next, we compute for each $t-i \leq j \leq t+k$

$$\prod_{j=t-i}^{t+k} (j) = \frac{(t+k)!}{(t-i-1)!} = \frac{(t+1)_k t!}{(t-i-1)!},$$

and similarly

$$\prod_{j=t-i}^{t+k} (a+j-1) = \frac{(a+t)_k (a)_t}{(a)_{t-i-1}}$$

$$\prod_{j=t-i}^{t+k} (a+b+t+j-2) = \frac{(a+b+t-1)_{t+k}}{(a+b+t-1)_{t-i-1}}$$

Altogether, the ergodic capacity reads

$$\frac{(a)_t}{(a+b+t-1)_t}\sum_{i=0}^{t-1}\frac{t!}{(t-1-i)!i!}\frac{(a+b+t-1)_{t-i-1}}{(a)_{t-i-1}}\sum_{k\geq 0}\frac{(-1)^k\rho^{k+i+1}}{(k+i+1)^2}\frac{(t+1)_k(a+t)_k}{(a+b+2t-1)_k k!}.$$

But the series

$$\sum_{k\geq 0}\frac{(-1)^k\rho^{k+i+1}}{(k+i+1)^2}\frac{(t+1)_k(a+t)_k}{(a+b+2t-1)_k k!}$$

as well as its derivatives with respect to ρ converge uniformly in any closed sub-interval in $]0,1[$. It follows that

$$D_\rho(\rho D_\rho)\sum_{k\geq 0}\frac{(-1)^k\rho^{k+i+1}}{(k+i+1)^2}\frac{(t+1)_k(a+t)_k}{(a+b+2t-1)_k k!} = \rho^i\,{}_2F_1(t+1,a+t,a+b+2t-1;-\rho)$$

where D_ρ is the derivative operator acting on the variable ρ. Finally, the index change $i\mapsto t-i-1$ together with

$$(1-t)_i=(-1)^i\frac{(t-1)!}{(t-1-i)!}$$

yield

$$\sum_{i=0}^{t-1}\frac{t!}{(t-1-i)!i!}\frac{(a+b+t-1)_{t-i-1}}{(a)_{t-i-1}}\rho^i = \frac{t!\rho^{t-1}}{(a)_{t-1}}P_{t-1}^{a-1,b}\left(\frac{\rho+2}{\rho}\right).$$

Since

$$\frac{t!\rho^{t-1}}{(a)_{t-1}}\frac{(a)_t}{(a+b+t-1)_t}=\frac{t!(a+t-1)\rho^{t-1}}{(a+b+t-1)_t},$$

The statement of the proposition 1 corresponds to the special parameters $a=r-t+1$ and $b=m-t-r+1$.

APPENDIX B
PROOF OF THEOREM 1

Let's $n=t$ and $\rho\in[0,1]$. From [25, Eq. (4.4.6)], we readily deduce that the hypergeometric function

$${}_2F_1(n+1,a+n,a+b+2n-1;-\rho)$$

coincides up to a multiplicative factor with the Jacobi function of the second kind $Q_n^{a-1,b-2}$ in the variable x related to ρ by

$$-\rho=\frac{2}{1-x}\quad\Leftrightarrow\quad x=\frac{\rho+2}{\rho}.$$

Consequently,

$$\left[D_\rho(\rho D_\rho)\right]C(\rho)=2\,B_{a,b,n}\frac{(1+\rho)^{b-2}}{\rho^{a+b-1}}P_{n-1}^{a-1,b}\left(\frac{\rho+2}{\rho}\right)Q_n^{a-1,b-2}\left(\frac{\rho+2}{\rho}\right)$$

where

$$B_{a,b,n}=\frac{n!\Gamma(a+b+n-1)}{\Gamma(a+n-1)\Gamma(N+n-1)}.$$

Moreover, recall from [25, Eq. (4.4.2)], that (note that $(\rho+2)/\rho>1$)

$$Q_n^{a-1,b-2}\left(\frac{\rho+2}{\rho}\right)=\frac{\rho^{a+b-3}}{2^{a+b-4}(\rho+1)^{b-2}}\int_{-1}^1(1-u)^{a-1}\times(1+u)^{b-2}\frac{P_n^{a-1,b-2}(u)}{((\rho+2)/\rho)-u}du.$$

After some mathematical manipulation and since

$$u\mapsto\frac{1}{((\rho+2)/\rho)-u}\left(P_{n-1}^{a-1,b}\left(\frac{\rho+2}{\rho}\right)-P_{n-1}^{a-1,b}(u)\right)$$

is a polynomial of degree $n-2$, then the orthogonality of the Jacobi polynomials entails

$$\left[D_\rho(\rho D_\rho)\right]C(\rho)=\frac{B_{a,b,n}}{2^{a+b-3}}\int_{-1}^1(1-u)^{a-1}(1+u)^{b-2}\times P_{n-1}^{a-1,b}(u)\frac{P_n^{a-1,b-2}(u)}{\rho(\rho+2-\rho u)}du.$$

Writing

$$\frac{1}{\rho(\rho+2-\rho u)}=\frac{1}{2}\left[\frac{1}{\rho}-\frac{(1-u)}{\rho+2-\rho u}\right],\ u\in[-1,1],$$

and using again the orthogonality of Jacobi polynomials, we get

$$\left[D_\rho(\rho D_\rho)\right]C(\rho)=-\frac{B_{a,b,n}}{2^{a+b-2}}\int_{-1}^1(1-u)^{a}(1+u)^{b-2}\times\frac{P_{n-1}^{a-1,b}(u)P_n^{a-1,b-2}(u)}{(\rho(1-u)+2)}du$$

which is still defined at $\rho=0$. A first integration with respect to ρ gives

$$[\rho D_\rho]C(\rho)=-\frac{B_{a,b,n}}{2^{a+b-2}}\int_{-1}^1(1-u)^{a-1}(1+u)^{b-2}P_{n-1}^{a-1,b}(u)\times P_n^{a-1,b-2}(u)[\ln(\rho(1-u)+2)-\ln 2]du$$

and a second integration leads to

$$C(\rho)=-\frac{B_{a,b,n}}{2^{a+b-2}}\int_{-1}^1(1-u)^{a-1}(1+u)^{b-2}P_{n-1}^{a-1,b}(u)\times P_n^{a-1,b-2}(u)\left\{\int_0^\rho\frac{\ln(v(1-u)/2+1)}{v}dv\right\}du.$$

Performing the variable changes $u\mapsto 1-2u$ in the last expression, we end up with

$$C(\rho)=-B_{a,b,n}\int_0^1u^{a-1}(1-u)^{b-2}P_{n-1}^{a-1,b}(1-2u)\times P_n^{a-1,b-2}(1-2u)\left\{\int_0^\rho\frac{\ln(vu+1)}{v}dv\right\}du$$

for any $\rho\in[0,1[$. By analytic continuation, this formula extends to the cut plane $\mathbb{C}\setminus(-\infty,0)$ and is in particular is valid for $\rho\geq 0$. Specializing it to $a=r-t+1$, and $b=m-t-r+1$ completes the proof of the Theorem 1.

APPENDIX C
PROOF OF THEOREM 2

Perform the variable change $\rho \mapsto b\rho$ in the definition of $C_{t,r}^{m,\rho}$:

$$\begin{aligned} C(b\rho) &= Z_{a,b,n}^{-1} \int \ln\left(\prod_{i=1}^{n}(1+b\rho\lambda_i)\right)\prod_{i=1}^{n}\lambda_i^{a-1}(1-\lambda_i)^{b-1} \\ &\quad \times V(\lambda)^2 1_{\{0<\lambda_1<\cdots<\lambda_n<1\}} d\lambda \\ &= \frac{Z_{a,b,n}^{-1}}{b^{(an+n(n-1))}} \int \ln\left(\prod_{i=1}^{n}(1+\rho\lambda_i)\right)\prod_{i=1}^{n}\lambda_i^{a-1} \\ &\quad \times \left(1-\frac{\lambda_i}{b}\right)^{b-1} V(\lambda)^2 1_{\{0<\lambda_1<\cdots<\lambda_n<b\}} d\lambda. \end{aligned}$$

On the other hand, our obtained expression for the ergodic capacity together with the variable change $v \mapsto bv$ entail:

$$\begin{aligned} C(b\rho) = -\frac{B_{a,b,n}}{b^a}\int_0^1 u^{a-1}\left(1-\frac{u}{b}\right)^{b-2} P_{n-1}^{a-1,b}\left(1-\frac{2u}{b}\right) \\ \times P_n^{a-1,b-2}\left(1-\frac{2u}{b}\right)\left\{\int_0^\rho \frac{\ln(vu+1)}{v}dv\right\} du \end{aligned}$$

Now

$$\lim_{b\to\infty}\frac{B_{a,b,n}}{b^a} = \frac{n!}{\Gamma(a+n-1)}$$

and similarly

$$\lim_{b\to\infty}\frac{Z_{a,b,n}^{-1}}{b^{n(a+n-1)}} = \prod_{i=1}^{n}\frac{1}{\Gamma(i)\Gamma(a+i-1)}$$

Moreover, the limiting transition (6) yields

$$\begin{aligned} \lim_{b\to\infty} P_{n-1}^{a-1,b}\left(1-\frac{2u}{b}\right) &= L_{n-1}^{a-1}(u) \\ \lim_{b\to\infty} P_n^{a-1,b-2}\left(1-\frac{2u}{b}\right) &= L_n^{a-1}(u). \end{aligned}$$

As a result,

$$\begin{aligned} \lim_{b\to\infty} C(b\rho) = -\frac{n!}{\Gamma(a+n-1)}\int_0^{+\infty} u^{a-1}e^{-u}L_{n-1}^{a-1}(u) \\ \times L_n^{a-1}(u)\left\{\int_0^\rho \frac{\ln(vu+1)}{v}dv\right\} du. \end{aligned}$$

where $\prod_{i=1}^{n}\frac{1}{\Gamma(i)\Gamma(a+i-1)}$ is the normalization constant of the density of the joint distribution of the ordered eigenvalues of a complex Wishart matrix [23]. The theorem is proved.

ACKNOWLEDGMENT

The authors gratefully acknowledge Prof. Mérouane Debbah for useful discussions and consultations, and Rémi Bonnefoi who performed part of Matlab simulations.

REFERENCES

[1] D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres," *Nature Photon.*, vol. 7, no. 5, pp. 354–362, Apr. 2013.
[2] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. C. Burrows, R. Essiambre, P. J. Winzer, D. W. Peckham, A. H. McCurdy, R. Lingle, R., "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6×6 MIMO processing," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 521–531, Feb. 15, 2012.
[3] W. Klaus, J. Sakaguchi, B. J. Puttnam, Y. Awaji, and N. Wada, "Optical technologies for space division multiplexing," in *Proc. 13th Workshop Inf. Opt. (WIO)*, Neuchatel, Switzerland, Jul. 2014, pp. 1–3.
[4] K. Shibahara, D. Lee, T. Kobayashi, T. Mizuno, H. Takara, A. Sano, H. Kawakami, Y. Miyamoto, H. Ono, M. Oguma, Y. Abe, T. Matsui, R. Fukumoto, Y. Amma, T. Hosokawa, S. Matsuo, K. Saitoh, M. Yamada, and T. Morioka, "Dense SDM (12-core × 3-mode) transmission over 527 km with 33.2-ns mode-dispersion employing low-complexity parallel MIMO frequency-domain equalization," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 196–204, Jan. 1, 2015.
[5] T. Mizuno, H. Takara, K. Shibahara, A. Sano, and Y. Miyamoto, "Dense space division multiplexed transmission over multicore and multimode fiber for long-haul transport systems," *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1484–1493, Mar. 15, 2016.
[6] H. Takahashi, D. Soma, S. Beppu, and T. Tsuritani, "Digital signal processing for space-division multiplexing (SDM) transmission," in *Proc. IEEE Photon. Conf. (IPC)*, San Antonio, TX, USA, Nov. 2019, pp. 1–2.
[7] K. Shibahara, T. Mizuno, D. Lee, and Y. Miyamoto, "Advanced MIMO signal processing techniques enabling long-haul dense SDM transmissions," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 336–348, Jan. 15, 2018.
[8] Y. Li, N. Hua, and X. Zheng, "A capacity analysis for space division multiplexing optical networks with MIMO equalization," in *Proc. Opt. Fiber Commun. Conf. Exhib. (OFC)*, Los Angeles, CA, USA, 2017, pp. 1–3.
[9] Y. Q. Hei, W. T. Li, X. C. Xu, and R. T. Chen, "Orthogonal STBC for MDL mitigation in mode division multiplexing system with MMSE channel estimation," *J. Lightw. Technol.*, vol. 35, no. 10, pp. 1858–1867, May 15, 2017.
[10] Z. G. Cao, L. Miao, S. K. Wang, W. T. Song, H. X. Gu, and Y. Q. Hei, "Soft-decision aided probabilistic data association based detection for mode division multiplexing transmission with mode-dependent loss," *IEEE Access*, vol. 7, pp. 172744–172751, Nov. 2019.
[11] S. O. Arik, J. M. Kahn, and K.-P. Ho, "MIMO signal processing for mode-division multiplexing: An overview of channel models and signal processing architectures," *IEEE Signal Process. Mag.*, vol. 31, no. 2, pp. 25–34, Mar. 2014.
[12] R. Dar, M. Feder, and M. Shtaif, "The Jacobi MIMO channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2426–2441, Apr. 2013.
[13] P. J. Winzer and G. J. Foschini, "MIMO capacities and outage probabilities in spatially multiplexed optical transport systems," *Opt. Express*, vol. 19, no. 17, pp. 16680–16696, Aug. 2011.
[14] A. Karadimitrakis, A. L. Moustakas, and P. Vivo, "Outage capacity for the optical MIMO channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4370–4382, Jul. 2014.
[15] S. H. Simon and A. L. Moustakas, "Crossover from conserving to lossy transport in circular random-matrix ensembles," *Phys. Rev. Lett.*, vol. 96, no. 13, Apr. 2006, At. no. 136805.
[16] A. Karadimitrakis, A. L. Moustakas, H. Hafermann, and A. Mueller, "Optical fiber MIMO channel model and its analysis," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Barcelona, Spain, Aug. 2016, pp. 2164–2168.
[17] Y. Chen and M. R. McKay, "Coulumb fluid, Painlevé transcendents, and the information theory of MIMO systems," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4594–4634, Jul. 2012.
[18] B. Collins, "Product of random projections, Jacobi ensembles and universality problems arising from free probability," *Probab. Theory Rel. Fields*, vol. 133, no. 3, pp. 315–344, Mar. 2005.
[19] L. Wei, Z. Zheng, and H. Gharavi, "Exact moments of mutual information of Jacobi MIMO channels in high-SNR regime," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 1–5.
[20] L. Wei, C.-H. Liu, Y.-C. Liang, and Z. Bai, "Matrix integral approach to MIMO mutual information statistics in high-SNR regime," *Entropy*, vol. 21, no. 11, p. 1071, Nov. 2019.
[21] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–596, Nov./Dec. 1999.
[22] M. L. Mehta, *Random Matrices*, 2nd ed. Boston, MA, USA: Academic, 1991.
[23] P. J. Forrester, *Log-Gases Random Matrices* (London Mathematical Society Monographs). Princeton, NJ, USA: Princeton Univ., 2007.

[24] C. Carré, M. Deneufchatel, J.-G. Luque, and P. Vivo, "Asymptotics of Selberg-like integrals: The unitary case and Newton's interpolation formula," *J. Math. Phys.*, vol. 51, no. 12, Dec. 2010, Art. no. 123516.
[25] M. E. H. Ismail, *Classical Quantum Orthogonal Polynomials One Variable*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[26] R. Bonnefoi and A. Nafkha, "A new lower bound on the ergodic capacity of optical MIMO channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, Jul. 2017, pp. 1–4.
[27] A. Nafkha and R. Bonnefoi, "Upper and lower bounds for the ergodic capacity of MIMO Jacobi fading channels," *Opt. Express*, vol. 25, no. 11, pp. 12144–12151, May 2017.
[28] H. Shin and J. Hong Lee, "Closed-form formulas for ergodic capacity of MIMO Rayleigh fading channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Anchorage, AK, USA, Jun. 2003, pp. 2996–3000.
[29] O. Oyman, R. U. Nabar, H. Bolcskei, and A. J. Paulraj, "Tight lower bounds on the ergodic capacity of Rayleigh fading MIMO channels," in *Proc. Global Telecommun. Conf. (GLOBECOM)*, Taipei, Taiwan, Nov. 2002, pp. 1172–1176.
[30] S. H. Simon, A. L. Moustakas, and L. Marinelli, "Capacity and character expansions: Moment-generating function and other exact results for MIMO correlated channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5336–5351, Dec. 2006.
[31] M. Kießling, "Unifying analysis of ergodic MIMO capacity in correlated Rayleigh fading environments," *Eur. Trans. Telecommun.*, vol. 16, no. 1, pp. 17–35, Jan./Feb. 2005.
[32] A. Maaref and S. Aissa, "Joint and marginal eigenvalue distributions of (non) central complex wishart matrices and PDF-based approach for characterizing the capacity statistics of MIMO Ricean and Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3607–3619, Oct. 2007.
[33] K. I. Gross and D. S. P. Richards, "Total positivity, spherical series, and hypergeometric functions of matrix argument," *J. Approximation Theory*, vol. 59, no. 2, pp. 224–246, Nov. 1989.
[34] C. W. J. Beenakker, "Random-matrix theory of quantum transport," *Rev. Mod. Phys.*, vol. 69, no. 3, pp. 731–808, Jul./Sep. 1997.
[35] P. J. Forrester, "Quantum conductance problems and the Jacobi ensemble," *J. Phys. A, Math. Gen.*, vol. 39, no. 22, pp. 6861–6870, May 2006.
[36] J. Carpenter, B. J. Eggleton, and J. Schröder, "Complete spatio-temporal characterization and optical transfer matrix inversion of a 420 mode fiber," *Opt. Lett.*, vol. 41, no. 23, pp. 5580–5583, Dec. 2016.
[37] T. Jiang, "Approximation of Haar distributed matrices and limiting distributions of eigenvalues of Jacobi ensembles," *Probab. Theory Rel. Fields*, vol. 144, nos. 1–2, pp. 221–246, May 2009.
[38] S. Verdu, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[39] B. Clerckx and C. Oestges, *MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems*. New York, NY, USA: Academic, 2013.
[40] X. Zhang, M. Matthaiou, E. Björnson, M. Coldrey, and M. Debbah, "On the MIMO capacity with residual transceiver hardware impairments," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 5299–5305.
[41] M. R. McKay, I. B. Collings, and A. M. Tulino, "Achievable sum rate of MIMO MMSE receivers: A general analytic framework," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 396–410, Jan. 2010.
[42] H. Lim and D. Yoon, "On the distribution of SINR for MMSE MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4035–4046, Jun. 2019.
[43] G. Alfano, C.-F. Chiasserini, A. Nordio, and S. Zhou, "Information-theoretic characterization of MIMO systems with multiple Rayleigh scattering," *IEEE Trans. Inf. Theory*, vol. 64, no. 7, pp. 5312–5325, Jul. 2018.
[44] M. Kieburg, G. Akemann, G. Alfano, and G. Caire, "Closed-form performance analysis of linear MIMO receivers in general fading scenarios," in *Proc. IEEE 23rd Int. ITG Workshop Smart Antennas (WSA)*, Vienna, Austria, Apr. 2019, pp. 1–7.
[45] L. Wei, "Ergodic MIMO mutual information: Twenty years after Emre Telatar," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 1182–1186.
[46] J. M. Kahn, K.-P. Ho, and M. B. Shemirani, "Mode coupling effects in multi-mode fibers," in *Proc. Opt. Fiber Commun. Conf. (OFC/NFOEC)*, Los Angeles, CA, USA, Mar. 2012, pp. 1–3.
[47] E. C. Song, E. Soljanin, P. Cuff, H. V. Poor, and K. Guan, "Rate-Distortion-Based physical layer secrecy with applications to multimode fiber," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 1080–1090, Mar. 2014.
[48] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.
[49] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications*. New Delhi, India: Springer-Verlag, 2014.
[50] R. Morris, "The dilogarithm function of a real argument," *Math. Comput.*, vol. 33, no. 146, pp. 778–787, 1979.

AMOR NAFKHA (Senior Member, IEEE) received the B.Sc. degree in information and communications technology engineering from the Higher School of Communications (SupCom), Tunis, Tunisia, in 2001, and the Ph.D. degree in information and communications technology from the University of South Brittany (UBS), Lorient, France, in 2006. From 2006 to 2007, he was a Postdoctoral Researcher with the Signal, Communication, and Embedded Electronics (SCEE-IETR) Research Group, CentraleSupélec, France, where he has also been an Associate Professor since January 2008. He was actively involved in the reconfigurable hardware platform implementation for software-defined radio and coauthoring several contributions on FPGA dynamic partial reconfiguration. He has published more than 70 papers in international peer-reviewed journals and conferences. His research interests include multiuser and MIMO detection, hardware implementation, information theory, sample rate conversion, and spectrum sensing techniques.

NIZAR DEMNI received the B.S. degree from Tunis El Manar II University, Tunisia, in 2003, the B.Sc. degree from the University of Aix-Marseille I, France, the Habilitation degree, in December 2013, the M.Sc. degree from the Laboratoire de Probabilites et Modèles Alèatoires, University of Paris VI, and the Ph.D. degree under the supervision of C. D. Martin. He was with Bielefeld University for a period of two years with Prof. F. Goetze. He was also with the Engineering School, Bizerte, from 2009 to 2010. He was an Associate Professor with the University of Rennes 1, where he is currently a member with the Group Probabilites et Theorie Ergodique. He has published more than 40 articles in international peer-reviewed journals. He has significative research articles on Dunkl processes and free probability theory.

4.4 Detection techniques in SM-MIMO context

This section addresses the MIMO detection problem in the context of spatial multiplexing. Indeed, one major implementation difficulty of the SM-MIMO technology is the signal separation problem at the receiving side due to the fact that the simultaneously transmitted signals interfere with each other when they arrive at the receiver. In particular, if the MIMO channel conditions are not beneficial, *i.e.*, the channel matrix $\mathbf{H}$ is bad-conditioned (a low condition number is an indicator of a good/favorable channel). The MIMO detection problem statement is simple: recover the vector of transmitted symbols $\mathbf{x}$ from the vector of received signals $\mathbf{y}$ based on the knowledge of the channel matrix $\mathbf{H}$. This problem is very challenging if the goal is to construct an MIMO detection algorithm that can achieve optimal or near-optimal detection performance with a low computation complexity cost, particularly, for large problem sizes. A review of various MIMO detection techniques for conventional and massive MIMO systems was recently provided in [24, 37, 38].

In the sequel, we consider spatial multiplexing in a MIMO system with t transmit antennas and $r \geq t$ receive antennas. The equivalent complex baseband input-output system relation is given by: $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}$, where $\mathbf{y} \in \mathbb{C}^{r\times 1}$ the received complex vector, $\mathbf{H} \in \mathbb{C}^{r\times t}$ is the channel matrix with entries that are assumed to be $\mathcal{CN}(0,1)$ independent and identically distributed random variables. The channel matrix $\mathbf{H}$ was assumed to be quasi-static and remain constant during the transmission of a block of L MIMO vectors. $\mathbf{z} \in \mathbb{C}^{r\times 1}$ is the noise vector with $\mathcal{CN}(0,\sigma^2)$ entries, *i.e.*, $\mathbb{E}_{\mathbf{z}}\left[\mathbf{z}\mathbf{z}^{\dagger}\right] = \sigma^2\mathbf{I}_r$, and $\mathbf{x} \in \mathbb{C}^{t\times 1}$ is the transmitted symbol vector. Each element $\mathbf{x}_i$ of transmitted symbol vector $\mathbf{x}$ is taken from a finite complex constellation alphabet set Ω, *i.e.*, $\mathbf{x} \in \Omega^t$. We assume that M-ary quadrature amplitude modulation (M-QAM) scheme is adopted, *i.e.*, for the 4-QAM constellation $\Omega \triangleq \{\omega_R + j\omega_I\}$ where $\omega_R, \omega_I \in \{\pm 1\}$.

The main task of a MIMO detection technique is to determine the transmitted vector $\mathbf{x}$ based on the received vector $\mathbf{y}$. When the receiver acquires perfect knowledge of $\mathbf{H}$, the maximum likelihood (ML) detection is an optimal algorithm to solve the MIMO detection problem. It calculates the a posteriori probability in terms of log-likelihood ratio for each possible transmitted vector $\mathbf{x}$ by browsing all the set Ω^t. Geometrically, the ML detector corresponds to the search for the *closest* point in a complex lattice $\mathbf{H}\mathbf{x}$ to a given r-dimensional vector $\mathbf{y}$ [39, 40, 41, 42]:

$$\hat{\mathbf{x}}_{ML} = \underset{\mathbf{x}\in\Omega^t}{\arg\min}\, ||\mathbf{y} - \mathbf{H}\mathbf{x}||_2^2 \tag{4.17}$$

Since the optimization problem (4.17) is performed over the discrete set Ω^t, the number of metric calculations required to reach the ML decision can be computed as $|\Omega|^t$, where $|\Omega|$ is the alphabet size of the used M-ary modulation scheme. Therefore, the algorithmic computation complexity of the ML decoder becomes very high for a large number of transmitted antennas and for higher-order modulation schemes. Fig. 4.3 depicted a possible classification of different

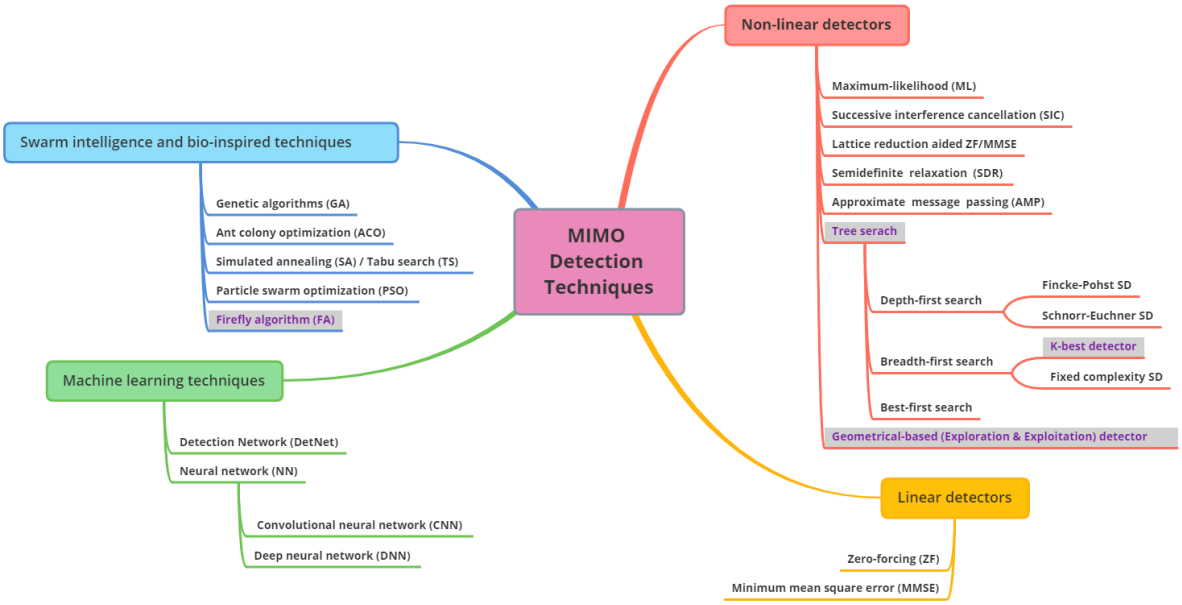


Figure 4.3 – The taxonomy of MIMO detection methods

MIMO detection schemes based on their resolution technique. All of these techniques lead to specific approaches of MIMO detection trading off between BER performance and computational complexity.

My research activities in the field of MIMO detection started during my own Phd thesis (2003-2006) with my supervisor professor Emmanuel Boutillon. In order to reduce the computational complexity, we proposed a geometrical-based algorithm to approximate ML detection. This algorithm utilizes two approaches, named exploration (*i.e diversification*) and exploitation (*i.e. intensification*) methods, to estimate a near optimal solution for the maximum likelihood detection problem with a reduced computational complexity. This work was published in [CI2] and [J2]. We kept improving computational complexity of the proposed MIMO detector during the following years. In 2007, we proposed a first improvement, which consists in using extended Bose–Chaudhuri–Hocquenghem (BCH) codes as exploration method. Comparing to other optimal or near-optimal MIMO detectors, the intensification over the BCH codes has a negligible performance degradation and a lower computational complexity for square M-QAM constellation. In 2013, we investigated a second diversification method, which consists of using uniformly distributed codes over the feasible solutions set Ω^t. The Hamming codes was studied and compared to randomized and BCH codes. For the intensification step, we performed a simple 1-flip local search algorithm where only a single feasible solution element $\mathbf{x}_i$, $i = 1, 2, \ldots, t$, was flipped in each iteration in order to reach the nearest neighbor (*i.e.* the local neighborhood of a given feasible solution $\mathbf{x}$ is the Hamming ball centred at $\mathbf{x}$ of the radius one). The above two works were respectively published at SSD 2007 [CI5] and ICC 2013 [CI28]. In the context

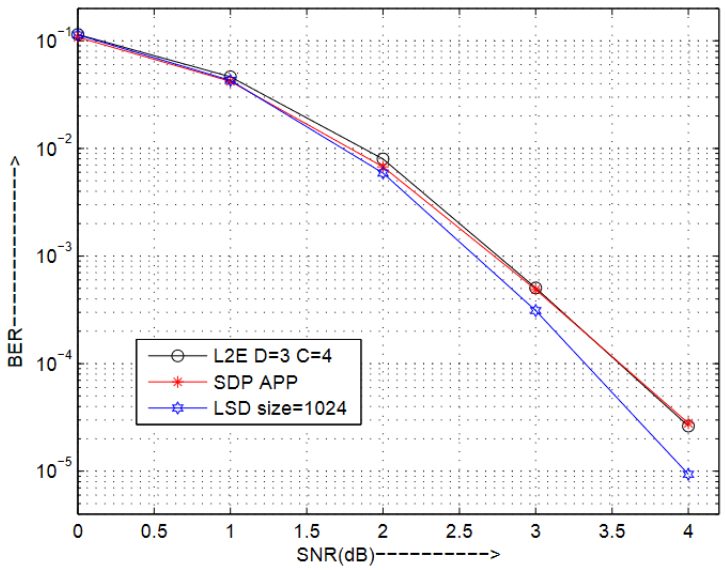


Figure 4.4 – The taxonomy of MIMO detection methods

of the PhD of Bastien Trotobas (2018-2021), we extensively investigated high performance and low complexity algorithms for the SM-MIMO detection problem. To this end, we proposed two soft-output detection algorithms for coded spatial multiplexing MIMO wireless communication systems. The first algorithm tried to explore the geometrical feature of the lattice whose basis vectors are the columns of channel matrix $\mathbf{H}$, and the second algorithm is based on a modified version of the standard firefly algorithm (FA), initially published in [43, 44]. The FA is a nature inspired swarm intelligence based optimization algorithm which is based on social behavior of lightning bugs insects. Both algorithms have been widely studied and optimal parameters (*i.e* best trade-off between performance and complexity) were extracted from Pareto front analysis. During Bastien's PhD thesis, we proposed a low-complexity pipelined hardware architecture for the geometrical-based MIMO detection algorithm published in [J16].

4.4.1 [C3]: Exploration-Exploitation trade-off based MIMO detection

The first contribution of Bastien's PhD study is a novel soft-output MIMO detector scheme, called L2E (list exploration and exploitation), to provide posteriori log-likelihood ratio (LLR) values for a soft-input error-correction-code (ECC) decoders. The proposed soft-output L2E detector is based on two stages: exploration and exploitation. The exploration stage allows the creation of a set Γ, which contains an initial LLR evaluations, and to build a subset of promising feasible solutions Γ_b that will be processed and updated by the exploitation stage.

The proposed soft-output MIMO detector has the following features: **(i)** it supports only BPSK and 4-QAM modulations, **(ii)** it is compatible with soft-input/soft-output ECC decoders

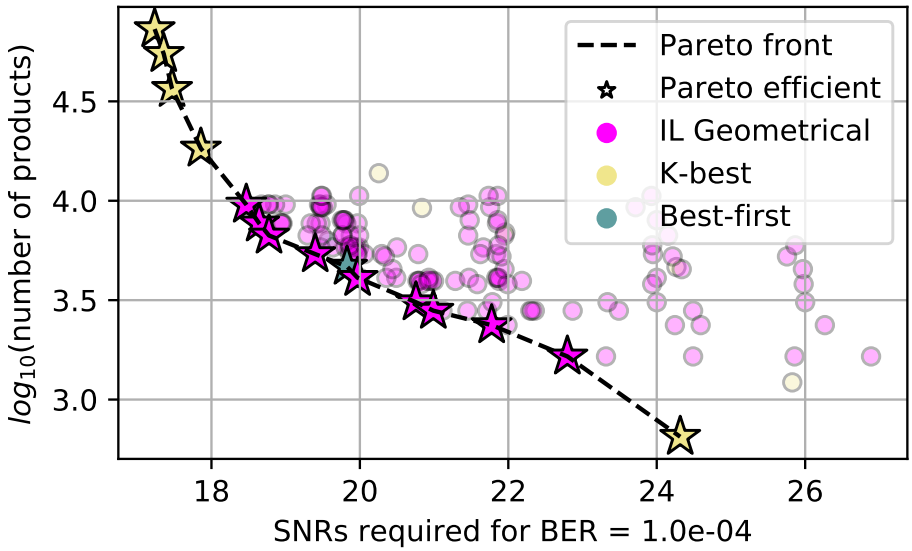


Figure 4.5 – Complexity vs. performance trade-off, Pareto plot compares the lowest required SNR (horizontal axis) versus the lowest complexity in the number of product operations (vertical axis) for 4×4 coded MIMO system with 16-QAM using IL Geometrical, soft-output K-best sphere decoding and soft-output Best-first sphere decoding algorithms.

to attain enhanced detection performance, **(iii)** it exhibits a constant computational complexity under all SNR regimes, since its complexity depends only on the number of predominant noise axis (*i.e.* D) and the number of "*good*" feasible solutions per axis (*i.e.* C). We define the predominant axis as the direction along which the values of the objective function, *i.e.* $f(\mathbf{x}) = ||\mathbf{y} - \mathbf{Hx}||_2^2$ remain constant and/or increase very slowly. In our paper [CI52], we showed that the proposed soft-output MIMO detector achieved near-optimal BER performance with low and fixed computational complexity. Fig. 4.4 shows comparative simulation results with 8×8 BICM-MIMO systems, 4-QAM, and an outer $R_c = 1/2$ convolutional code with memory 2. The parameters of L2E are $D = 3$ and $C = 4$, and hence the L2E list size has $2tCD = 192$ candidates for the soft-outputs generation. To ensure best BER performance under all SNR regimes, the size of candidate lists is set to 1024 for the two most-known soft-output MIMO detectors, *i.e.* List Sphere Decoding (LSD) [45] and soft-output Semi-Definite Programming (SDP) [46]. It is clear from Fig. 4.4 that the L2E exhibits the same error performance as the soft-output SDP detector. However, the computational complexity of the proposed soft-output detector is significantly lower compared to SDP. the proposed L2E achieves a performance with about 0.2 dB loss from optimal LSD detector at the target BER 10^{-4}.

The second contribution of Bastien's thesis is related to the enhancement of the exploration stage in order to detect high-order QAM constellations (*e.g.* 16-QAM). We proposed a novel geo-

metrical exploration technique based on the construction of variable D-dimensional box centred at the unquantized zero-forcing (ZF) solution $\mathbf{x}_{zf}^{un}$. This new algorithm variant (*i.e* improved version of L2E) will be referred as a "*IL Geometrical*" [J20]. The IL Geometrical detector provides low computational complexity and achieves near-optimal error performance for 16-QAM constellation. Two questions have been raised during the study of the soft-output L2E detector; how to choose the best couple of L2E parameters (D, C) regarding complexity-performance trade-off, and how to compare L2E algorithm with other state-of-the-art soft-output MIMO detection techniques? The answers to both questions are needed in order to build a set of acceptable and practical soft-output MIMO detectors. Thus, a multi-objective optimization technique called Pareto Optimal is used for selecting the appropriate soft-output detectors parameters offering both high BER performance and low computational complexity. Pareto front that presents the best trade-off between performance and computational complexity was depicted in Fig. 4.5 with 4×4 coded MIMO system employing a $1/2$ systematic LDPC code and 16-QAM modulation. From the simulation results, we can see that the IL Geometrical detector outperforms outperforms the K-best sphere decoder [47] and the Best-first sphere decoder [48] over a wide SNR range.

For more details on current work related to soft-output L2E detector and its improved version, the reader is referred to our recent publication in IEEE Access Journal 2020 [J20] included hereafter.

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 2, 2020, accepted October 12, 2020, date of publication October 19, 2020, date of current version October 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3032016

When Should We Use Geometrical-Based MIMO Detection Instead of Tree-Based Techniques? A Pareto Analysis

BASTIEN TROTOBAS[1,2], ADRIEN LLAVE[1,3], AMOR NAFKHA[1,2], (Senior Member, IEEE), AND YVES LOUËT[1,2], (Member, IEEE)

[1]Institut d'Électronique et des Technologies du numéRique (IETR) UMR CNRS 6164, CentraleSupélec, 35576 Cesson Sévigné, France
[2]Signal, Communication and Embedded Electronics (SCEE) Research Group, CentraleSupélec, 35576 Cesson Sévigné, France
[3]Facial Analysis Synthesis and Tracking (FAST) Research Group, CentraleSupélec, 35576 Cesson Sévigné, France

Corresponding author: Amor Nafkha (amor.nafkha@centralesupelec.fr)

ABSTRACT The soft-output multiple-input multiple-output (MIMO) detection problem has been extensively studied, and a large number of heuristics and metaheuristics have been proposed to solve it. Unlike classical tree-search based detectors, geometrical heuristic algorithms involved two consecutive steps: **(i)** an exploration step based on the geometry of the channel matrix singular vectors; **(ii)** a local exploitation step is performed in order to obtain better final solution. In this paper, new enhancements for geometrical heuristics are introduced to significantly reduce the complexity in quadrature phase-shift keying (QPSK) and allow 16 quadrature amplitude modulation (QAM) capability through new exploration techniques. The performance-complexity trade-off between the new detector and two tree-based algorithms is investigated through Pareto efficiency. The Pareto framework also allows us to select the most efficient tuning parameters based on an exhaustive search. The proposed detector can be customized on the fly using only one or two parameters to balance the trade-off between computational complexity and bit error rate performances. Moreover, the Pareto fronts demonstrate that the new geometrical heuristic is especially efficient with QPSK since it provides a significant reduction in regards to the computational complexity while preserving good bit error rate (BER) performance and ensuring high flexibility.

INDEX TERMS Geometrical detection, Pareto efficiency, soft-output MIMO detection.

I. INTRODUCTION

In the last decades, the increase in the quantity of data sent over wireless channels has led to a shortage of available frequency bands. This scarcity has driven researchers and operators to improve the spectrum efficiency of wireless communications systems and, more recently, seek new frequency bands with THz technologies. Specifically, spectrum efficiency improvement increases the data throughput and the link quality without using new frequency bands.

In this context, MIMO systems have been widely adopted for their capability to multiplex transmitted data streams over time-frequency-space dimensions. The spatial multiplexing MIMO technique is mainly used to increase the data transmission rate or spectral efficiency. The space-division multiplexing (SDM) MIMO technology can transmit several data streams in the same time-frequency slot and separate them according to spatial considerations. This multiplexing technique increases the spectrum efficiency through the addition of antennas. However, the more antennas there are, the more complex the receiver design is. Therefore, these systems require new algorithms to exploit the spatial information to separate data streams efficiently.

WiFi standards like IEEE 802.11n/ac, long-term evolution (LTE), WiMAX, and 5G, among other modern standards, rely on MIMO technologies. All of these standards use known pilot signals in order to estimate the channel state information (CSI). It is common to assume that the receiver gets a perfect CSI whereas the transmitter is CSI-agnostic. This operating regime is easier to set up as it does not require each CSI to be sent back to each emitting antenna.

The associate editor coordinating the review of this manuscript and approving it for publication was Ahmed Mohamed Ahmed Almradi.

The separation of received streams has been widely studied, and many algorithms have already been proposed in the literature. This detection problem is known to be NP-hard [1], which implies that an optimal solution cannot be computed in polynomial time (unless under the unattainable assumption of P=NP). Such optimal algorithms include the naive resolution with the maximum likelihood (ML) detector or the optimal tree-path search using sphere decoding (SD) as initiated by [2]–[4]. However, their exponential complexity does not make them suitable for hardware implementation, especially at low signal-to-noise ratio (SNR) regimes. Still, it should be noted that some SD implementations can compete with a polynomial algorithm in terms of complexity when the number of antennas and the constellation size are rather small and when the SNR is high enough [5].

Several approaches have been considered to offer heuristics and metaheuristics that provide good performance in polynomial time. Under ideal circumstances where the amount of information available is significant, linear detectors provide an acceptable result. In more difficult cases, more advanced algorithms are necessary. The earliest heuristics are based on the interference cancellation between the different received data streams. Two versions of this heuristic coexist. The succesive interference cancellation (SIC) detection scheme suppress the interference by iteratively maximizing the signal-to-interference-plus-noise ratio (SNR). This approach is well suited when the received signals present different individual quality metric corresponding to each of the received data streams [6], [7]. In the opposite scenario, when data streams have similar quality parallel interference cancellation (PIC) detectors are preferred [8], [9].

The detection problem turns out to be a complicated combinatorial optimization problem, Thus, to circumvent and simply solve this problem, tree-based heuristics algorithms can be used. These algorithms are classified according to the method of searching the tree: depth-first algorithms look for the best possible leaf through a descent, prune and backtrack process; breadth-first algorithms keep only a fixed number of paths at each step [10]–[13]; and best-first algorithms exploiting metrics to determine how to explore the tree [14], [15].

Alternative solutions have been proposed thanks to a shift in the problem perspective. For instance, detectors based on Markov chain Monte Carlo (MCMC) algorithms have been developed by addressing the problem through a probabilistic approach [16], [17]. The emergence of deep networks trained to tackle the detection problem is also studied in recent work [18]. Finally, bio-inspired metaheuristics based on ant colony optimization (ACO) [19] or on the firefly algorithm (FA) [20], [21] have also been proposed.

We previously investigated the interest of geometrical heuristics to solve the MIMO detection problem [22]. We compared the geometrical approach and the tree-based one for a bit interleaved coded modulation (BICM) scheme with QPSK modulation. The criteria of BER and complexity revealed that the geometrical method was close but not yet as good as the state-of-the-art detectors.

The geometrical approach provided in [22] was restricted to lower-order constellations (i.e. binary phase-shift keying (QPSK) and QPSK) whereas higher-order modulations were not investigated at all. Moreover, the geometrical heuristic was compared with a tree-based reference on both BER performance and complexity but these two criteria were not considered simultaneously within a trade-off perspective. This paper presents two new contributions that help to overcome the highlighted limitations present in [22]:

- Enhancements to improve geometric heuristics on QPSK and extend its use cases to high-order modulations (e.g., 16-QAM)
- Performance-complexity trade-off study using the Pareto efficiency.

The paper is organized as follows. In Section II, the MIMO system model under consideration is presented. In Section III, a detailed description of the soft-output computation is given, while in Section IV, we provide the geometrical-based MIMO detection framework. The proposed enhanced geometrical-based detection algorithm is presented in Section V. A detailed performance-complexity tradeoff analysis of the proposed algorithm through Pareto-front curves is given in Section VI for both QPSK and 16-QAM. Section VII concludes the paper.

1) NOTATIONS

In the following, bold uppercases (*resp.* lowercases) denote matrices (*resp.* vectors), whereas the other letters refer to scalars. All sets are noted with calligraphic uppercases, and the corresponding lowercase refers to their cardinality. The vector $\mathbf{h}_j$ denotes the j^{th} column of matrix $\mathbf{H}$ and $\mathbf{H}^T$, denotes the transpose $\mathbf{H}$. The natural logarithm is denoted by ln. The real and imaginary part of a complex number $a \in \mathbb{C}$ are denoted by $\Re(a)$ and $\Im(a)$, respectively. We denote the l2-norm of vector $\mathbf{y}$ as $\|\mathbf{y}\|$.

II. TRANSMISSION MODEL

Let the $N \times N$ Rayleigh-fading MIMO system transmitting N data streams to N receiving antennas. We assume a perfect CSI at the receiver (ideal channel estimator), whereas the transceiver is CSI-agnostic. From a model perspective, the receiver knows the channel matrix $\mathbf{H}_c \in \mathbb{C}^{N\times N}$ where $h_c(i,j)$ is the complex channel gain from antenna j to antenna i. The MIMO channel is modeled as quasi-static block fading where the channel path gains can be considered constant during a large block comprising hundreds of transmitted vectors [23]. The channel gains change according to a statistical model given by an independent Rayleigh-distributed envelope. Therefore, the computational complexity of any pre-processing phases, such as singular value decomposition (SVD) or QR decomposition, pseudo-inverse calculation, can be negligible. Indeed, they are performed once for several transmitted vectors.

Let $\mathcal{Q}_c$ be the set of all constellation symbols from a square QAM. We denote by $\mathbf{y}_c \in \mathbb{C}^N$ the signals received on each antenna after the propagation of the symbols $\mathbf{x}_c \in \mathcal{Q}_c^N$ through the channel and after the addition of the complex gaussian noise $\mathbf{w}_c \sim \mathcal{CN}(0, \sigma^2)$. With this notations, the system model is expressed as

$$\mathbf{y}_c = \mathbf{H}_c\mathbf{x}_c + \mathbf{w}_c. \tag{1}$$

This model can be rewritten as an equivalent real-valued expression such that

$$\mathbf{y} \triangleq \begin{bmatrix} \Re(\mathbf{y}_c) \\ \Im(\mathbf{y}_c) \end{bmatrix} \tag{2}$$

$$\mathbf{H} \triangleq \begin{bmatrix} \Re(\mathbf{H}_c) & -\Im(\mathbf{H}_c) \\ \Im(\mathbf{H}_c) & \Re(\mathbf{H}_c) \end{bmatrix} \tag{3}$$

$$\mathbf{x} \triangleq \begin{bmatrix} \Re(\mathbf{x}_c) \\ \Im(\mathbf{x}_c) \end{bmatrix} \tag{4}$$

$$\mathbf{w} \triangleq \begin{bmatrix} \Re(\mathbf{w}_c) \\ \Im(\mathbf{w}_c) \end{bmatrix} \tag{5}$$

$$\mathcal{Q} \triangleq \Re(\mathcal{Q}_c) \tag{6}$$

$$n \triangleq 2N \tag{7}$$

where the new parameter n defines the size of the real-valued matrices and vectors. It can be interpreted as the number of real-valued data streams. Ideed, switching from a complex-valued to a real-valued perspective is equivalent to process real and imaginary parts independently. In this real-valued model, (1) becomes the following equivalent system model:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}. \tag{8}$$

The BICM transceiver is composed of an encoder, an assumed-perfect interleaver, and a modulator. The receiver uses the corresponding components in a reversed order: demodulator, deinterleaver, and then decoder.

III. SOFT-OUTPUT COMPUTATIONS

Let b_{ij} be the i^{th} bit encoded in the j^{th} symbol of $\mathbf{x}$. Basic hard-output detectors provide an estimate of the vector of transmitted symbols. To improve the performance, soft-output detectors search for the log-likelihood ratio (LLR) for each bit defined as

$$L_{ij} = \ln \frac{P(b_{ij} = 1|(\mathbf{H}, \mathbf{y}))}{P(b_{ij} = 0|(\mathbf{H}, \mathbf{y}))} \tag{9}$$

with $P(b_{ij}|(\mathbf{H}, \mathbf{y}))$ the probability mass function of b_{ij} given the channel state and the received vector.

Expression (9) is not suitable for practical use as it has been shown that its computations is exponentially complex [24]. A common solution is to use the max-log approximation:

$$L_{ij} \approx \frac{1}{2\sigma^2}\left(\min_{\mathbf{x}\in\mathcal{X}_{ij}^0} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 - \min_{\mathbf{x}\in\mathcal{X}_{ij}^1} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2\right), \tag{10}$$

where $\mathcal{X}_{ij}^k = \{\mathbf{x} \in \mathcal{Q}^n : b_{ij} = k\}$ is the set of all symbols with b_{ij} equals to k [25]–[27].

This new expression (10) is still exponentially complex as the norm must be computed for each point in the constellation. Indeed, we clearly have $\mathcal{X}_{ij}^0 \cup \mathcal{X}_{ij}^1 = \mathcal{Q}^n$ which contains 2^n points. Therefore, we introduce a new subset $\mathcal{S} \subset \mathcal{Q}^n$ with a lower cardinality and approximate (10) on it. The new expression becomes

$$L_{ij} \approx \frac{1}{2\sigma^2}\left(\min_{\mathbf{x}\in\mathcal{S}\cap\mathcal{X}_{ij}^0} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 - \min_{\mathbf{x}\in\mathcal{S}\cap\mathcal{X}_{ij}^1} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2\right). \tag{11}$$

The reduced subset $\mathcal{S}$ may have no points from $\mathcal{X}_{ij}^k$ (i.e., $\mathcal{S} \cap \mathcal{X}_{ij}^k = \emptyset$). In such a configuration, we assign k to the bit $b_{i,j}$ by the fact that the probability $P(b_{ij} = k|(\mathbf{H}, \mathbf{y}))$ is considered equal to one. Therefore, L_{ij} is set to its maximum (if $b_{i,j} = 1$) or minimum (if $b_{i,j} = 0$) value to reflect this certainty.

In the remainder of this paper, the objective function denotes the squared norm involved in the LLRs, and a point is considered better than another if it has a smaller objective function.

IV. GEOMETRICAL-BASED DETECTION

Soft-output geometrical heuristics are based on three main steps: exploration, exploitation, and the LLRs computation. This section reviews the first two-step as the last phase has already been discussed in the previous section.

Exploration and exploitation steps are both designed to search for feasible solutions into $\mathcal{Q}^n$. They can be viewed as coarse and fine search methods. Coarse search step (exploration) is performed over whole solution set $\mathcal{Q}^n$ in order to find pertinent solutions, whereas exploitation step (fine search) step refines the quality of the solutions through local searches.

A. SVD-BASED EXPLORATION

The exploration step produces a set $\mathcal{P}$ of promising points to be exploited in the next step. Let $\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}$ be the SVD of $\mathbf{H}$ with $(\mathbf{U}, \mathbf{V})$ two orthogonal matrices and $\mathbf{\Lambda}$ a diagonal one containing the singular values. Without loss of generality, we assume that the singular values are sorted in ascending order:

$$\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n. \tag{12}$$

Let $\mathbf{x}^\star \in \mathbb{R}^n$ be the real vector minimizing the objective function. This point can be obtained using the Moore-Penrose inverse $\mathbf{H}^+$ through $\mathbf{x}^\star = \mathbf{H}^+\mathbf{y}$. The regular inverse can be use if the channel matrix is well conditioned, if not, the Moore-Penrose invert is required. We can rewrite the objective function as

$$\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 = \|\mathbf{U}\mathbf{\Lambda}\mathbf{V}(\mathbf{x}^\star - \mathbf{x})\|^2 \tag{13}$$

$$= \left(\mathbf{V}(\mathbf{x}^\star - \mathbf{x})\right)^T \mathbf{\Lambda}\mathbf{U}^T\mathbf{U}\mathbf{\Lambda}\left(\mathbf{V}(\mathbf{x}^\star - \mathbf{x})\right). \tag{14}$$

Given that $\mathbf{U}$ is orthogonal, the previous equation gives

$$\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 = \left(\mathbf{V}(\mathbf{x}^\star - \mathbf{x})\right)^T \mathbf{\Lambda}^2\left(\mathbf{V}(\mathbf{x}^\star - \mathbf{x})\right). \tag{15}$$

As $\mathbf{V}$ is orthogonal, its columns $\{\mathbf{v}_i\}_{i=1\ldots n}$ constitute a basis and we can introduce α_i the coordinates of $\mathbf{x}^\star - \mathbf{x}$ on this basis. The matrix $\mathbf{\Lambda}$ is diagonal, then the objective function can be expressed as

$$\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 = \sum_{i=1}^{n} \alpha_i^2 \lambda_i^2 \tag{16}$$

which provides hints on the objective function evolution. The ascending order of the singular values induces that the first α_i are less impacting than the last ones. Therefore, points with similar coordinates except for the first ones, can be considered equivalent with respect to their objective function. Exploration step aims for building the promising set $\mathcal{P}$ from points as different as possible but with equivalently low objective function. The promising set $\mathcal{P}$ is also used to initialized $\mathcal{S}$.

B. ITERATIVE LOCAL EXPLOITATION

The exploitation step gets the most of the previously selected points $\mathcal{P}$ by exploring each promising feasible solution belonging the subset $\mathcal{P}$. For each promising point in the subset, an iterative local search is perform to find better feasible solutions around the starting points. One iteration can be decomposed into three steps:

1) Generate at most $2n$ new points equal to the candidate but for one coordinate. This modified coordinate is once the previous and once the next symbol in the real-valued constellation. If the coordinate to be modified is already at an extreme value (ie. the first or last symbol in the constellation), then a single point is generated.
2) Compute the objective function for each generated feasible solution and add them to $\mathcal{S}$.
3) Select the best feasible solution between the initial candidate and the newly generated points. The best point becomes the new starting candidate for the next iteration. If there is a tie, prefer the initial point.

It is easy to prove that this algorithm reaches a stable point. First, each iteration decreases the objective function over the finite set $\mathcal{Q}^n$ of possibilities. Moreover, oscillations between two equally good points are prevented by the tie rule. Therefore, this process ends on a stable point that is not guaranteed to be the global optimum. The performance-complexity trade-off can be tuned by stopping the algorithm after a predefined number of iterations rather than waiting for a stable point.

V. ENHANCEMENTS TO GEOMETRICAL DETECTION

In this section, we present some enhancements to the geometrical-based detection framework that was presented in the previous section. Section V-A describes the previous exploitation technique and highlights its limitations, then Section V-B provides the new exploration techniques to build the set of promising feasible solutions. Section V-C introduces a lossless method to reduce the complexity and Section V-D summarizes the proposed geometrical detector and provides the algorithmic computational costs of the detection process.

A. PREVIOUS EXPLORATION TECHNIQUE

The SVD-based exploration is the core step of the geometrical heuristic since it provides a good set $\mathcal{P}$ of promising point is the key to the success of the following steps. In the previous works, $\mathcal{P}$ was build using an intersection-based process [22], [28].

Fig. 1 illustrates this exploration technique with $n = 2$ and a 16-QAM modulation scheme. The process starts by computing the line passing through $\mathbf{x}^\star$ and directed by the singular vector $\mathbf{v}_1$. The dashed blue line represents this straight line. Then, each intersection of this line with the basis axis is projected on the 16-QAM constellation to build the subset $\mathcal{P}$. This process is illustrated by the purple arrows. Fig. 1 shows the process with one direction for readability's sake, but real detectors apply this method for several directions. As an example, we used 3 directions for a 4×4 channel in [22] or 4 directions for a 30×30 channel in [28].

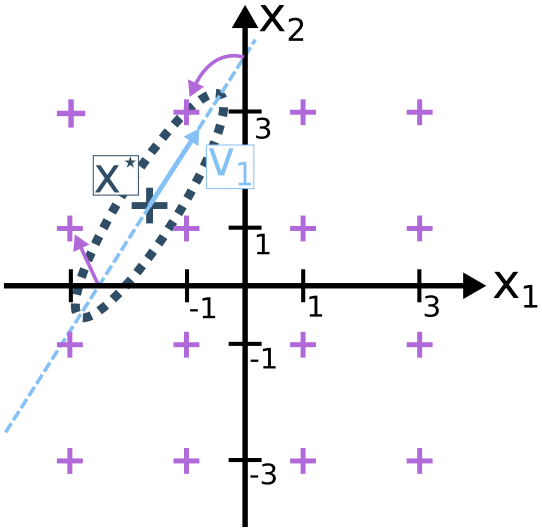


FIGURE 1. Previous exploration process with $n = 2$ and a 16-QAM.

The exploration technique is able to find promising points as the elements of $\mathcal{P}$ should have similar objective function values but it has two major drawbacks. First, the construction of $\mathcal{P}$ implies that it contains only points that are near the basis hyperplanes. For instance, the point $(-3, -3)$ cannot be reached by this exploration process, regardless of its quality, since no intersection could be projected to this point. That is why, this method performs well with QPSK, where all points are accessible but that it leads to poor results with higher-order modulation schemes. Besides, this exploration technique is very complex, with large channels. Indeed, building $\mathcal{P}$ requires to compute the intersection of several straight lines with the n basis hyperplanes. Therefore, the more n growth, the more intersections are needed. That is why the

new exploration technique should rely on steps that do not increase in complexity with additional antennas.

B. PROPOSED EXPLORATION TECHNIQUES

As exposed in the previous section, the exploration technique presented in [22] has two significant drawbacks. In this section, we propose new exploration techniques that solve these issues. Indeed, the new methods are able to select, with no preferences, each point of the constellation so that it could perform well on higher-order modulation schemes. Moreover, they are composed of steps that are independent of the number of transmitted data streams n.

1) INTERSECTION-LESS (IL) EXPLORATION

As stated in Section IV-A, the exploration builds a subset of promising points $\mathcal{P} \subset \mathcal{Q}^n$ as different as possible with equivalently low objective function. On the first hand, (16) shows that equivalent points differ only on their first coordinates when expressed on the basis $\{\mathbf{v}_i : 1 \leqslant i \leqslant n\}$. On the other hand, $\mathbf{x}^\star$ is obviously a good point since it is the global optimum on $\mathbb{R}^n$. Therefore, the sought equivalently good points can be constructed by adding to $\mathbf{x}^\star$ some linear combinations of the first $\mathbf{v}_i$. Let $n_d \leqslant n$ be the number of $\mathbf{v}_i$ used during this new exploration step.

To have equivalent points, we build them as

$$\mathbf{x}^\star + \sum_{i=1}^{n_d} \frac{\pm f}{\lambda_i} \mathbf{v}_i \tag{17}$$

with f some scalar whose value will be discussed later. Expressed in the $\mathbf{V}$ basis and after the translation to $\mathbf{x}^\star$, these points' coordinates are

$$\left(\pm\frac{f}{\lambda_1}, \ \pm\frac{f}{\lambda_2}, \ \ldots, \ \pm\frac{f}{\lambda_{n_d}}, \ 0, \ \ldots, \ 0 \right)_{(\mathbf{x}^\star, \mathbf{V})}. \tag{18}$$

Equation (16) highlights that these points have exactly the same objective function being $f^2 n_d$. However, these points do not belong to $\mathcal{Q}^n$ so that they cannot be used directly to build the subset $\mathcal{P}$. Therefore, the promising set is built by taking the nearest value in $\mathcal{Q}$ for each coordinates. This last step is the same used in the zero-forcing (ZF) detector and will be referred as the "projection on a set" in the remainder of this paper.

Fig. 2 and 3 provide examples of the intersection-less exploration process for a 16-QAM, $f = 1$ and $n = 2$. The exploration process starts at its center $\mathbf{x}^\star$. Blue arrows represent the addition of the scaled singular vectors that appear in (17) and purple ones correspond to the projection on the constellation.

Fig. 2 illustrates the case $n_d = 1$ where two different points are obtained: $(-3, -1)$ and $(-1, 3)$. Fig. 3 shows the same scenario with $n_d = 2$. In this situation, three points are generated: $(-3, -1)$, $(-3, 1)$ and $(-1, 3)$.

The points from (17) can be either within or outside the constellation regarding the magnitude of each singular value. That is why the projection can produce either good or bad points. Fig. 2 and 3 show an ideal case whereas Fig. 4 presents a poorly conditioned case with $\lambda_1 \ll \lambda_2$. In the latter, the explored points are not at all promising, which compromises the construction of $\mathcal{P}$. For this reason, we introduce the scaling coefficient f in (17) which selects the sampled iso-value. A large f induces a large iso-value and thus widely spaced points whereas a small scaling coefficient produces closer points. Thus, the proposed exploration uses a few different scales to increase the opportunity that at least one of the iso-values will generate suitable points.

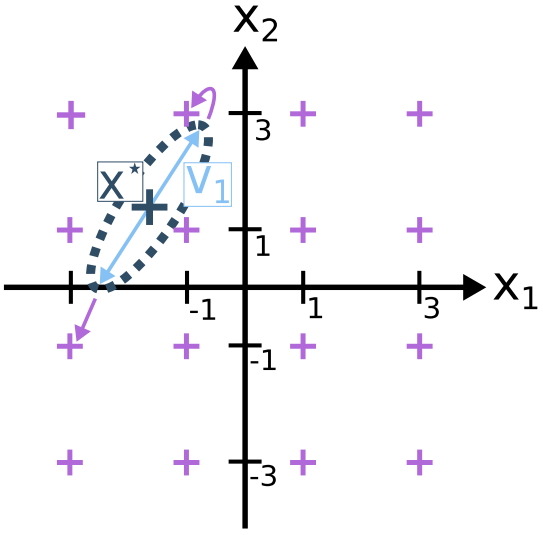


FIGURE 2. Well-conditioned intersection-less exploration with $n = 2$, $n_d = 1$ and a 16-QAM.

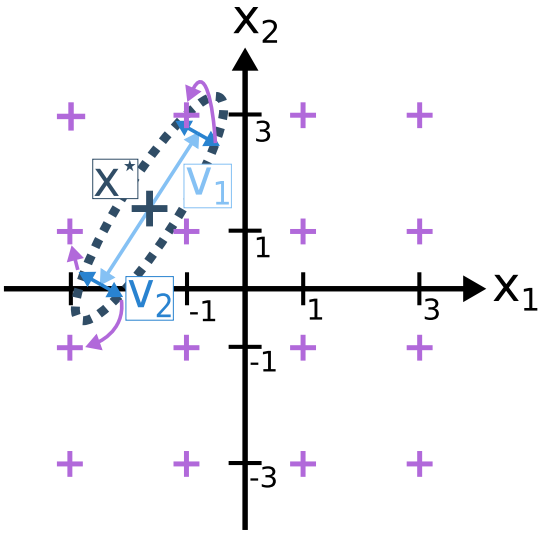


FIGURE 3. Well-conditioned intersection-less exploration with $n = 2$, $n_d = 2$ and a 16-QAM.

2) SUB-CONSTELLATION PROJECTION

It is required that $\mathcal{S}$ includes the points minimizing each of the two terms for (11) to be equivalent to reference (10). If not, the result of (11) is a close but inexact approximation of (10). The described exploration is not guaranteed to yield

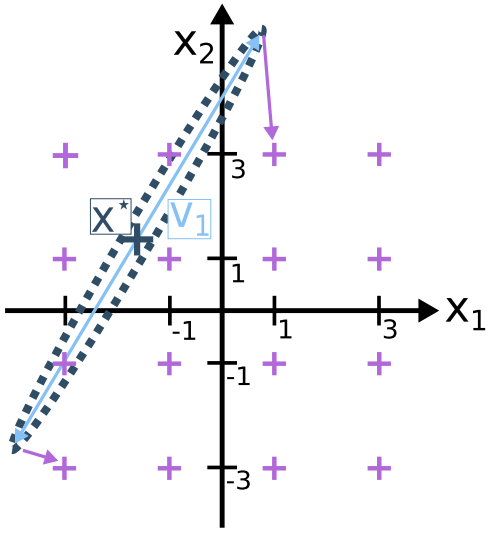


FIGURE 4. Poorly-conditioned intersection-less exploration with $n = 2$, $n_d = 1$ and a 16-QAM.

these two minima. Thus, the process can be modified to force the search for points in each sub-constellations $\mathcal{X}_{ij}^k$. This goal is easily achieved by altering the final projection. The point construction steps from $x^\star$ and singular vectors can be kept as it. Subsequently, the projection is no longer performed on $\mathcal{Q}^n$ but on every set $\mathcal{X}_{ij}^k$. The overhead of this method is to move from 1 projection to a $2nq$ projections.

C. COMPLEXITY REDUCTION

As for the majority of detectors, the overriding computational cost is the objective function evaluation. Thus, any refinement of this step greatly decreases the complexity. The efforts are focused on the number of products since this type of operation is considerably more complex than an addition.

The products involved during the objective function evaluation are due to the squares in the norm and to the computation of

$$\mathbf{Hx} = \sum_{j=1}^{n} \mathbf{h}_j x_j \tag{19}$$

with $\mathbf{h}_j$ the j^{th} column of $\mathbf{H}$. A naive computation of (19) would require n products for each $\mathbf{h}_j x_j$ and then n^2 additions to add up all the vectors. However, it is noteworthy that the values x_j are in the finite set $\mathcal{Q}$ so that there are only q possibilities for the product $\mathbf{h}_j x_j$. For instance, with $\mathcal{Q} = \{-3, -1, 1, 3\}$ (i.e., using a 16-QAM), a product $\mathbf{h}_j x_j$ can either be $-3\mathbf{h}_j$, $-\mathbf{h}_j$, $\mathbf{h}_j$ or $3\mathbf{h}_j$ but nothing else. Generally speaking, there are q different possibilities. Based on this property and remarking that $\mathbf{H}$ is known for a whole block, we can preprocess and store the q possible products $\mathbf{h}_j x_j$. This is equivalent to preprocess and store the matrices $-3\mathbf{H}$, $-\mathbf{H}$, $\mathbf{H}$ and $3\mathbf{H}$ for 16-QAM and in the general case, to store and compute the set $\{s\mathbf{H} : s \in \mathcal{Q}\}$. Then, the computation of (19) become the sum of n known vectors and does not require any product at all.

This technique dramatically reduces the amount of product required to decode the symbols at the cost of qn^2 coefficients storage. Indeed, one have to store the channel matrix $\mathbf{H}$ multiplied by each possible symbol in $\mathcal{Q}$. This storage space and the number of precalculations can be halved in ordinary situations since the usual constellations allow to obtain the negative symbols by changing the sign. Continuing the previous 16-QAM example, storing $\mathbf{H}$ and $3\mathbf{H}$ is enough to compute (19) with only additions since $-\mathbf{H}$ and $-3\mathbf{H}$ can be obtained by substracting rather than adding the preprocessed vectors.

D. SUMMARY OF THE NEW DETECTOR

Fig. 5 sums up the new geometrical algorithm presented in this paper. The upper part depicts the block-wise preprocessing. It includes the computation of the Moore-Penrose inverse $\mathbf{H}^+$ as well as the pre-processing introduced in section V-C. Since these calculations are only performed once per block of hundreds of symbol vectors, their complexity is negligible compared to the remaining part of the algorithm.

The lower part of the figure illustrates the detection process for a symbol from the received vector and the estimated noise variance. During the exploration step, the optimal vector is first derived using the pre-calculated $\mathbf{H}^+$. It is then added to the sum $\sum_{i=1}^{n_d} \frac{\pm f}{\lambda_i}\mathbf{v}_i$ that is also precomputed by the previous preprocessing. These points are then projected on the constellation and optionally on the sub-constellations, as described in Section V-B. All these points are used to initialize the subset $\mathcal{S}$. During the exploitation step, some local search iterations are performed to add new promising points to $\mathcal{S}$ (see Section IV-B). Eventually, the LLRs are estimated using the approximation of (11).

Table 1 details the complexity for each steps of the proposed geometrical detector with $p = 2^{n_d} n_f$ the number of promising points in $\mathcal{P}$ and n_f the number of scaling factor used. The preprocessed operations computed at a block level are neglected.

TABLE 1. Complexity evaluation for each step of the proposed geometrical detector.

Step	Nb prods	Nb adds
Compute $\mathbf{x}^\star = \mathbf{H}^+\mathbf{y}$	n^2	$(n-1)n$
Add prescaled $\mathbf{v}_i$ according to (17)	0	pn
Objective function evaluation	pn	$2pn$
One local search iteration	$2pn^2$	$3pn$
Max-log approximation.	$n\log_2(q)$	$n\log_2(q)$

VI. COMPARISONS WITH TREE-BASED REFERENCES

The proposed algorithm is to be compared with the reference detectors according to two criteria: BER performance and computational complexity. Indeed, highly complex detectors are required to provide the best performance in unfavorable use cases (low SNR regime). On the contrary, a simpler algorithm is preferable to increase the transmission rate under favorable conditions. Besides, the flexibility

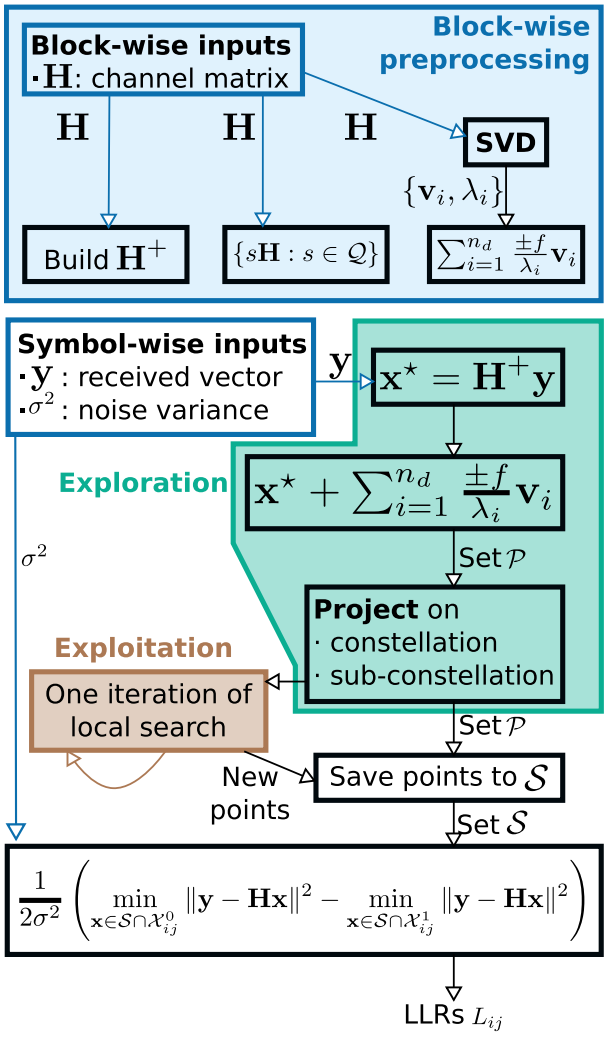


FIGURE 5. Summary of the proposed new geometrical detector.

of the algorithm is an attractive feature for matching the performance-complexity trade-off to the use cases.

Section VI-A described the tree-based reference to be simulated using the Monte-Carlo setup proposed in Section VI-B. Section VI-C introduces the Pareto efficiency and Section VI-D provides the comparisons using this framework.

A. REFERENCE TREE-BASED DETECTORS

The tree-based paradigm represents the detection problem as the search for the best path in a weighted tree. Any path starting from the root and reaching a leaf represents a decoded vector $\mathbf{x}$ whose total weight corresponds to the objective function. Each node has as many children as the number of symbols in $\mathcal{Q}$. Selecting a node at each level corresponds to detect the corresponding component of $\mathbf{x}$. In the following, we refer to "extend a node" the process of computing the children's objective function of a particular node to extend the current path.

As discussed in the introduction, tree-based detectors can be classified according to the method of searching into the tree. In this section, we describe two tree-based references: a canonical breadth-first K-best with Schnorr-Eurner enumeration [11] and a state-of-the-art best-first detector [15].

1) BREADTH-FIRST DETECTION: K-BEST ALGORITHM

The breadth-first approach builds paths from the root to the leaf with no backward step. At each level of the tree, each surviving path is extended. The K best new paths are preserved to be expended at the next level while the others are pruned. The process ends when it reaches the leaf level.

Table 2 details the computational complexity of the K-best detector as described in [11] with two metrics: the number of products and the number of additions. The complexity grows linearly with K and quadratically with n. It is assumed that the QR decomposition is performed on a block basis and therefore, can be neglected. Moreover, some preprocessing can be performed on a block basis to divide by a factor of about five the required K to reach a specific BER vs. SNR performance. As in [11], we refer by *mode 1* the algorithm without any preprocessing and by *mode 3* the one with the best preprocessing.

TABLE 2. Complexity evaluation of breadth-first K-best detector. Step numbers refer to the one described in Section III-B from [11].

Step	Nb prods	Nb adds
Compute $\mathbf{Q}^t\mathbf{y}$.	n^2	$(n-1)n$
For each level l from n to 1	repeated n times	
$\hookrightarrow$ Step 2) Compute weight for each child.	Kq	Kq
$\hookrightarrow$ Step 2) Update cumultative weight for each child.	Kq	Kq
$\hookrightarrow$ Step 4) Update path history.	$K(l-1)$.	$K(l-1)$
Sub-total for the loop	$K\left(2qn + \frac{(n-1)n}{2}\right)$	
Max-log approximation.	$K + n\log_2(q)$	$n\log_2(q)$

2) BEST-FIRST DETECTION: CROSS-LEVEL PARALLEL TREE-SEARCH

We describe in this section the algorithm from [15]. The best-first approach selects the path to extend based on the partial cumulative weight of a node rather than exploring straight to the leaf (as in depth-first paradigm) or rather to keep a specified amount of paths (as in breadth-first paradigm). The cross-level variant keeps track of the best nodes at each level through several stacks of finite length. At each iteration, the best node from each stack is popped out, extended, and the best siblings and the best child are inserted in the corresponding stack. If a stack reaches its maximal length, then the worst path is pruned. The process ends when all the stacks are empty. Moreover, the algorithm saved the cumulative weights of the best leaves so far and used them to prune paths that are already worst. This pruning criteria inspired by SD, avoids the extension of paths that are known to lead to worst solutions.

Unlike K-best detection, the best-first cross-level algorithm does not expand the same number of nodes at each run. Indeed, the number of paths pruned due to stack overflows or to poor cumulative weights depend on $\mathbf{H}$ and $\mathbf{y}$. That is

why it is not possible to find a closed-form for its complexity. Table 3 details the complexity for the three steps and neglects the QR decomposition for the same reason that for K-best. Reference [15] reports that the algorithm visits, on average, hundreds of nodes per detection with $n = 4$ and a 64-QAM. In the following, we will use the average observed complexity among all the detections as the complexity of this algorithm. This metric is computed at run time using the data from Table 3.

TABLE 3. Complexity of each step of the cross-level best-first detector. The second step is performed several times, the other are only carried out once.

Step	Nb prods	Nb adds
Compute $\mathbf{Q}^l\mathbf{y}$	n^2	$(n-1)n$
Expand a node at level l	$(n-l+1)q$	$2q$
Compute LLRs from metrics	nq	nq

B. MONTE-CARLO SIMULATION SETUP

BER performance of the two references and the geometrical detector are evaluated thanks to Monte-Carlo simulations. At each iteration, a 1440-bit random message is modulated, sent through a 4×4 MIMO channel, and received. After the reception, the number of binary errors is counted. For every SNRs, the simulation continues until it either gets 200 binary errors or reaches 5.10^5 transmitted bits.

The message is encoded in blocks of 720 bits with an irregular, systematic low-density parity-check (LDPC) code of ratio 1/2. The parity check matrix is designed according to the WiMAX standard (IEEE 802.16e) [29]. All receivers exploit the generated LLRs to decode the message with 15 iterations of a belief-passing min-sum algorithm. All the simulations are run using the CommPy framework [30].

C. PARETO FRONTS: A TOOL TO STUDY TRADE-OFFS

The detectors are compared based on the Pareto efficiency to study the performance-complexity trade-off objectively. A detector is said to be Pareto efficient if it is impossible to find an alternative detector that reduces either the complexity of the BER without losing on the other metric. The set of all Pareto-efficient detectors, called the Pareto front, represents the trade-off options. Indeed, switching from a Pareto-efficient detector to another means to favor complexity or BER in the trade-off. Conversely, a detector that is not Pareto efficient should not be selected because one can improve at least one of the metrics without losing on the other.

The computational complexicity is assessed as the number of products alone and the number of operations (products plus additions). The first situation refers to implementations on application-specific integrated circuits (ASICs) where the overriding complexity is the number of products due to its difficulty. The second comparison is better adapted to field-programmable gate arrays (FPGAs) since embedded digital signal processors (DSPs) can compute a free addition when computing a product.

Table 4 lists all the parameters tested during the described Monte-Carlo simulations for both QPSK and 16-QAM. All combinations of the listed parameters are tested for the geometrical detector. To improve the readability of the figures, only a subset of these parameters are plotted in the following sections. In any case, all the detectors claimed to be Pareto efficient are Pareto efficient for all the parameters from Table 4.

TABLE 4. List of detector parameters tested during Monte-Carlo simulations.

Parameter	Values tested
K-best K	2, 4, 8, 12, 16, 24, 32, 48, 64, 128, 192, 256
Best-first stack lengths	(1, 12, 15), (1, 7, 10), (1, 3, 5)
Sub-constellation projection	Yes, No
Number of iterations for exploration step	0, 1, 2, 3, 4, 5, 6, 7, 8
Number of dimensions n_d	1, 2, 3, 4, 5, 6, 7, 8
Scaling factors f	From 1 to 5 factors among 0.125, 0.25, 0.5, 1, 2, 4

D. SIMULATION RESULTS

The preprocessing described in Section V-C could be adapted to simplify the tree-based references. The results will be provided in two scenarios to permit a meaningful comparison. Firstly with the strict implementation of references and using preprocessing only for geometric detectors in Section VI-D1 and VI-D2. Then by extending the preprocessing to all the compared algorithms in Section VI-D3.

1) COMPARISONS FOR QPSK

Fig. 6 and 7 present the Pareto efficient algorithms for a QPSK. The first one approximates the complexity by the number of products, whereas the second one estimates it by the number of operations (products plus additions). Table 5 details the parameters corresponding to the Pareto efficient detectors for the number of product/BER trade-off.

For readability reasons, only a subset of the geometric detectors tested is shown. The plots are restricted to detectors using the sub-constellation projection presented in Section V-B2 with two scale factors ($f = 0.25, f = 4.0$) and no iterations. Indeed, these parameters include the geometric Pareto efficient detectors.

These parameters highlight that for a very small constellation, the geometrical detector should handle the performance-complexity trade-off using the number of dimensions n_d. Indeed, the receivers can select the number of dimensions based on the measured SNR, build the promising set $\mathcal{P}$ using the sub-constellation variant, and then apply the max-log approximation without further exploitation.

Fig. 6 shows that the proposed algorithm requires ten times fewer products that the canonical K-best *mode 1* at the cost of 0.7 dB. Moreover, the geometrical allow for an

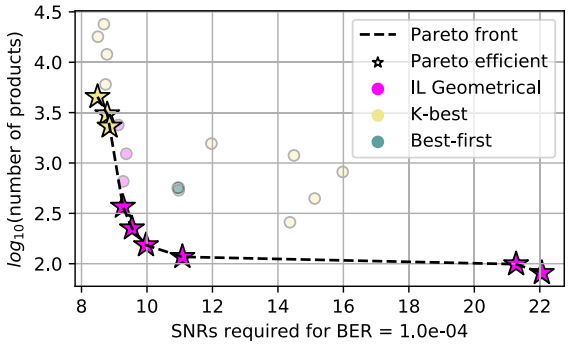


FIGURE 6. Pareto front for a QPSK: number of products/BER trade-off.

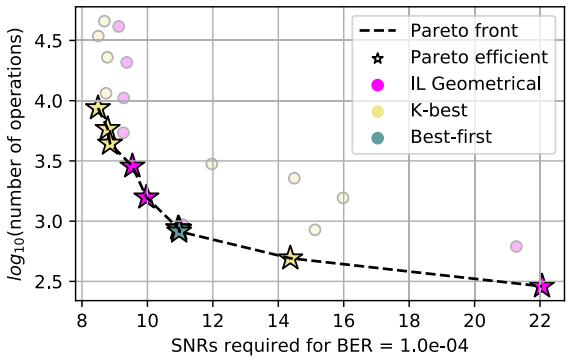


FIGURE 7. Pareto front for a QPSK: number of operations/BER trade-off.

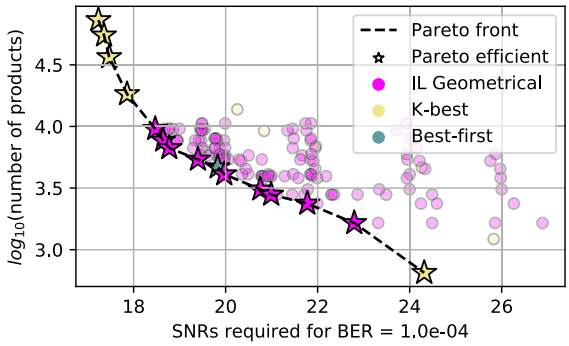


FIGURE 8. Pareto front for a 16QAM: number of products/BER trade-off.

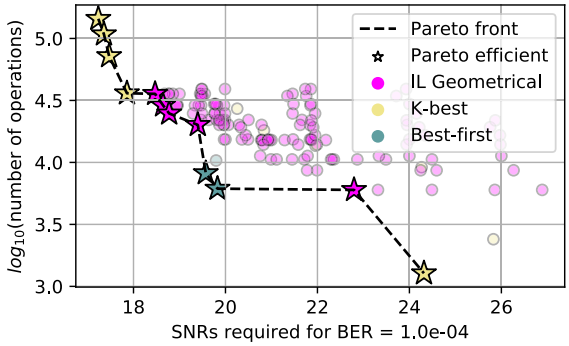


FIGURE 9. Pareto front for a 16QAM: number of operations/BER trade-off.

TABLE 5. Pareto efficient detectors for number of products/BER trade-off with QPSK. Geometrical detectors reported in this table are all tuned with $f \in \{0.25, 4\}$ and no iterations at all. The sub-constellation projection is enabled.

Detector	$\log_{10}$(prods)	$\log_{10}$(ops)	SNRs for 1.10^{-4}
K-Best $K = 48$	3.66	3.94	8.5 dB
K-Best $K = 32$	3.48	3.77	8.8 dB
K-Best $K = 24$	3.36	3.64	8.9 dB
Geom $n_d = 5$	2.57	3.73	9.3 dB
Geom $n_d = 4$	2.35	3.46	9.6 dB
Geom $n_d = 3$	2.18	3.20	10.1 dB
Geom $n_d = 2$	2.07	2.98	11.1 dB
Geom $n_d = 1$	2.00	2.79	21 dB
Geom $n_d = 0$	1.90	2.46	22 dB

on-the-fly configuration since changing n_d is enough to select the working point. Fig. 7 shows tighter spreads with the three detectors being efficient at some SNR range.

2) COMPARISONS FOR 16-QAM

Fig. 8 and 9 provides the same trade-off analysis for a 16-QAM and Table 6 details the parameters of the Pareto efficient detectors. In this section, we only plot the geometrical detectors with $n_d = 2$ and no sub-constellation projection. It is no more efficient to project on each sub-constellation as the number of projection required grows linearly with the constellation size. Therefore, the overhead is no more neglected when switching from QPSK to 16-QAM.

Table 6 highlights that the on-the-fly configuration relies on the number of iterations and the number of scaling factors when using a 16-QAM. Both Pareto fronts show that the geometrical detector is efficient at some SNR ranges. However, the complexity gap between tree-based and geometrical detectors is not as important as for QPSK.

The hierarchy observed in QPSK remains the same for 16-QAM scheme:

- canonical K-best remains effective for the worst SNRs with very high K,
- best-first is effective but not flexible,
- the geometrical algorithm provides good results and adaptability for moderate SNRs.

3) COMPARISONS WITH EQUIVALENT PREPROCESSING

Fig. 10 and 11 represent the same comparison when all detectors are using the same preprocessing as described in section V-C. For tree-based algorithms, this results in a halving of the number of products. Indeed, the norm computation now requires only products for the squaring. In this scenario, the Pareto efficient detectors were all Pareto efficient in the previous one, and we refer to the previous table for the parameters.

We do not know any tree-based implementation featuring this kind of preprocessing. However, [15] reports in section IV-C1 a computation method leading to an equivalent decrease in complexity. Indeed, the products **Hx**, are computed only by shifts and additions, which restricts similarly the number of products required.

Fig. 11 illustrates that although being Pareto efficient, the geometrical detector is not a wise choice in 16-QAM

TABLE 6. Pareto efficient detectors for number of products/BER trade-off with 16-QAM. Geometrical detectors reported in this table are all tuned with $n_d = 2$. The sub-constellation projection is disabled.

Detector	$\log_{10}$(prods)	$\log_{10}$(ops)	SNRs for 1.10^{-4}
K-Best $K = 256$	4.86	5.16	17.2 dB
K-Best $K = 192$	4.74	5.03	17.4 dB
K-Best $K = 128$	4.56	4.86	17.5 dB
K-Best $K = 64$	4.26	4.56	17.9 dB
Geom 5 its, $f \in \{0.25, 1, 2\}$	3.98	4.55	18.5 dB
Geom 4 its, $f \in \{0.5, 1, 2\}$	3.86	4.46	18.7 dB
Geom 5 its, $f \in \{0.5, 2\}$	3.82	4.40	18.8 dB
Geom 4 its, $f \in \{0.5, 2\}$	3.73	4.30	19.4 dB
Best-first (1, 3, 5)	3.67	3.79	19.8 dB
Geom 3 its, $f \in \{0.5, 2\}$	3.61	4.19	20 dB
Geom 4 its, $f = 1$	3.49	4.05	20.7 dB
Geom 2 its, $f \in \{0.25, 1\}$	3.45	4.02	21 dB
Geom 3 its, $f = 1$	3.74	3.94	21.9 dB
Geom 2 its, $f = 0.5$	3.21	3.78	22.8 dB
K-Best $K = 2$	2.81	3.11	24.3 dB

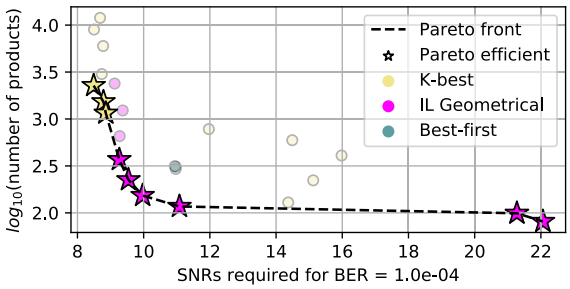


FIGURE 10. Pareto front for a QPSK: number of products/BER trade-off.

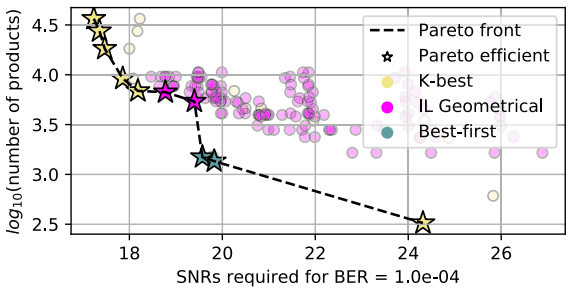


FIGURE 11. Pareto front for a 16QAM: number of operations/BER trade-off.

when preprocessing is possible for the tree-based detectors. K-best is preferable to save in SNRs, and best-first provides a significantly simpler option.

Fig. 10 shows that the proposed geometrical detector is particularly relevant in QPSK. The complexity gain is considerable, and the trade-off between computational complexicity and BER peformance can be easily achieved with few parameters.

VII. CONCLUSION

In this paper, we proposed an extensive comparison of a new geometrical detector with tree-based references. The performance-complexity trade-off is studied using the Pareto efficiency framework. We presented new exploration techniques to improve the performance-complexity trade-off and extend the geometrical detectors to higher-order modulation schemes. Moreover, a preprocessing method is introduced to reduce further the number of products required.

The Pareto fronts show that K-best is suitable in the worst SNR regimes, whereas the geometrical detector and best-first are efficient when SNRs are moderate. Besides, the proposed detector outperforms the two references in the QPSK scenario by providing a significant gap in complexity and by allowing a simple on-the-fly configuration. Moreover, the gap is large enough so that switching from *mode 1* to *mode 3* is not enough for K-best to reach the same performance as the proposed algorithm.

REFERENCES

[1] D. Micciancio, "The hardness of the closest vector problem with preprocessing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1212–1215, Mar. 2001.
[2] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. Comput.*, vol. 44, no. 170, pp. 463–471, Apr. 1985.
[3] C. P. Schnorr and M. Euchner, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," *Math. Program.*, vol. 66, nos. 1–3, pp. 181–199, Aug. 1994.
[4] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.
[5] J. Jalden and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.
[6] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Int. Symp. Signals, Syst., Electron. Conf.*, Oct. 1998, pp. 295–300.
[7] A. Riadi, M. Boulouird, and M. M. Hassani, "ZF/MMSE and OSIC detectors for UpLink OFDM massive MIMO systems," in *Proc. IEEE Jordan Int. Joint Conf. Electr. Eng. Inf. Technol. (JEEIT)*, Apr. 2019, pp. 767–772.
[8] W. Chin, A. Constantinides, and D. Ward, "Parallel multistage detection for multiple antenna wireless systems," *Electron. Lett.*, vol. 38, no. 12, p. 597, 2002.
[9] C. Studer, S. Fateh, and D. Seethaler, "ASIC implementation of soft-input soft-output MIMO detection using MMSE parallel interference cancellation," *IEEE J. Solid-State Circuits*, vol. 46, no. 7, pp. 1754–1765, Jul. 2011.
[10] K. K. Y. Wong and P. J. McLane, "A low-complexity iterative MIMO detection scheme using the soft-output M-algorithm," *IST Mobile Summit*, vol. 76, p. 5, Jun. 2005.
[11] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-best sphere decoding for MIMO detection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 491–503, Mar. 2006.
[12] Z. Yan, G. He, Y. Ren, W. He, J. Jiang, and Z. Mao, "Design and implementation of flexible dual-mode soft-output MIMO detector with channel preprocessing," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 11, pp. 2706–2717, Nov. 2015.
[13] S.-J. Choi, S.-J. Shim, Y.-H. You, J. Cha, and H.-K. Song, "Novel MIMO detection with improved complexity for near-ML detection in MIMO-OFDM systems," *IEEE Access*, vol. 7, pp. 60389–60398, 2019.
[14] C.-H. Liao, T.-P. Wang, and T.-D. Chiueh, "A 74.8 mW soft-output detector IC for 8×8 spatial-multiplexing MIMO communications," *IEEE J. Solid-State Circuits*, vol. 45, no. 2, pp. 411–421, Feb. 2010.
[15] G. He, X. Zhang, and Z. Liang, "Algorithm and architecture of an efficient MIMO detector with cross-level parallel tree-search," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 28, no. 2, pp. 467–479, Feb. 2020.

[16] J. C. Hedstrom, C. H. Yuen, R.-R. Chen, and B. Farhang-Boroujeny, "Achieving near MAP performance with an excited Markov chain Monte Carlo MIMO detector," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 7718–7732, Dec. 2017.
[17] G. C. G. Fernandes and M. G. S. Bruno, "Complexity-reduced suboptimal equalization with Monte Carlo based MIMO detectors," in *Proc. 27th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2019, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/8902499/
[18] N. Samuel, T. Diskin, and A. Wiesel, "Deep MIMO detection," in *Proc. IEEE 18th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2017, pp. 1–5.
[19] J. C. Marinello and T. Abrão, "Lattice reduction aided detector for MIMO communication via ant colony optimisation," *Wireless Pers. Commun.*, vol. 77, no. 1, pp. 63–85, Jul. 2014, doi: 10.1007/s11277-013-1495-z.
[20] B. Nasiri and M. R. Meybodi, "History-driven firefly algorithm for optimisation in dynamic and uncertain environments," *Int. J. Bio-Inspired Comput.*, vol. 8, no. 5, p. 326, 2016. [Online]. Available: http://www.inderscience.com/link.php?id=10000417
[21] A. Datta and V. Bhatia, "A near maximum likelihood performance modified firefly algorithm for large MIMO detection," *Swarm Evol. Comput.*, vol. 44, pp. 828–839, Feb. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2210650217304571
[22] B. Trotobas, Y. Akourim, A. Nafkha, Y. Louet, and J. Weiss, "Evaluation of the complexity, performance and implementability of geometrical MIMO detectors: The example of the exploration and exploitation list detector," *Int. J. Adv. Telecommun.*, vol. 13, nos. 1–2, pp. 1–9, 2019.
[23] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016. [Online]. Available: http://arxiv.org/abs/1503.06854
[24] C. Xu, D. Liang, S. Sugiura, S. Xin Ng, and L. Hanzo, "Reduced-complexity approx-log-MAP and max-log-MAP soft PSK/QAM detection algorithms," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1415–1425, Apr. 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6461035/
[25] W. Koch and A. Baier, "Optimum and sub-optimum detection of coded data disturbed by time-varying intersymbol interference (applicable to digital mobile radio receivers)," in *Proc. GLOBECOM IEEE Global Telecommun. Conf. Exhib.*, Dec. 1990, pp. 1679–1684.
[26] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proc. IEEE Int. Conf. Commun. ICC*, Jun. 1995, pp. 1009–1013.
[27] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.
[28] A. Nafkha, E. Boutillon, and C. Roland, "Quasi-maximum-likelihood detector based on geometrical diversification greedy intensification," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 926–929, Apr. 2009.
[29] *IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems—Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, Standard 802.16e-2005, 2005.
[30] V. Taranalli and B. Trotobas. (Mar. 2020). *CommPy: Digital Communication With Python, Revision f4c68b*. [Online]. Available: github.com/veeresht/CommPy

BASTIEN TROTOBAS received the B.S. degree from the University of Rennes, in 2015, and the M.S. degree from the École Normale Supérieur de Rennes, France, in 2017. He is currently pursuing the Ph.D. degree with the Signal, Communication, and Embedded Electronics (SCEE-IETR) Research Group, CentraleSupélec, Rennes, France. His Ph.D. thesis deals with the development of MIMO detectors from both an algorithmic and a VLSI implementation perspective.

ADRIEN LLAVE received the M.Sc. degree from the École Nationale Supérieure Louis-Lumière, Saint-Denis, France, in 2015, and the M.Sc. degree from the Grenoble-INP Phelma Engineering School, Grenoble, France, in 2016, all in audio engineering and signal processing. He is currently pursuing the Ph.D. degree with the CentraleSupélec/IETR, Rennes. He worked in 2016 for 3DSoundLabs, Rennes, France, on the binaural synthesis individualization based on acoustical simulation. His research interests include statistical signal processing, optimization, and psychoacoustics applied to hearing aids.

AMOR NAFKHA (Senior Member, IEEE) received the B.Sc. (Eng.) degree from the Higher School of Communications (SupCom), Tunis, Tunisia, in 2001, and the Ph.D. degree from the University of South Brittany (UBS), Lorient, France, in 2006, all in information and communications technology. From 2006 to 2007, he was a Postdoctoral Researcher with the Signal, Communication, and Embedded Electronics (SCEE-IETR) Research Group, CentraleSupélec, France. During this time at SCEE, he was actively involved in the reconfigurable hardware platform implementation for software-defined radio, coauthoring several contributions on FPGA dynamic partial reconfiguration. Since January 2008, he has been an Associate Professor with CentraleSupélec. He has published more than 70 papers in international peer-reviewed journals and conferences. His research interests include multiuser and MIMO detection, hardware implementation, information theory, sample rate conversion, and spectrum sensing techniques.

YVES LOUËT (Member, IEEE) received the Ph.D. degree in digital communications and the Habilitation (HDR) degree from Rennes University, France, in 2000 and 2010, respectively. The topic of his Ph.D. thesis regarded peak to average power reduction in OFDM modulation with channel coding. He was a Research Engineer in 2000 with SIRADEL Company, Rennes, and involved in channel propagation modeling for cell planning. He was involved in French collaborative research projects, including COMMINDOR, ERASME, and ERMITAGES about channel modeling in many frequency bands, especially, 5 and 60 GHz for further telecommunication systems. In 2002, he became an Associate Professor with Supélec, Rennes. He also became a Full Professor with Supélec. He is currently with CentraleSupélec. He is also the Head of the Signal Communication Embedded Electronics Research Group, Institute of Electronics and Telecommunications, Rennes Lab (CNRS), and the Vice Chair of URSI Commission C. His research interests include signal processing and digital communications applied to software and cognitive radio systems. He was involved in many collaborative European projects, including FP7E2R, CELTIC B21C, CELTIC SHARING, NoE Newcom, and COST and French projects, including ANR PROFIL, ANR INFOP, FUI AMBRUN, TEPN, and WINOCOD. His research interests include new waveforms design for green cognitive radio and energy efficiency enhancement.

4.4.2 [C4]: Well-distributed Regions based MIMO detection

Maximum likelihood (ML) based detection provides minimum bit error performance but it is infeasible for larger MIMO systems because its computational complexity grows exponentially with the number of transmit antennas t. This is also true for efficient implementations of ML detection through the sphere decoding algorithm. In this study, we adopt the same approach as the L2E detector which consist of two stages: exploration stage followed by an exploitation stage. The first stage provides a subset which contains a good initial feasible solutions that are well distributed in the search space Ω^t. In the second stage, a local search algorithm starts from each initial candidate point, iteratively moves to a neighbor feasible solutions and then provides the best optimal solution visited. In contrast to the geometrical approach, to design the initial feasible solutions, performed by the L2E detector during the exploration stage, in the present work, we adopted promising techniques borrowed from coding theory with the aid of maximum distance separable codes.

The first attempt to find well distributed feasible solutions, which can act as starting points for exploitation stage (*i.e.* greedy search), dates back to my own PhD and post-doctoral fellowship. In IEEE SSD 2007 [CI5], a MIMO detector based on BCH codes was proposed and its performance was investigated through simulations. Fig. 4.6 shows that comparing to random-based strategy, the BCH-based MIMO detector achieves an improved bit error performance. Indeed, at BER of about 10^{-4}, the BCH-based algorithm is only 0.4 dB worse than the sphere decoder, while the random-based MIMO detector exhibits a loss of 1.8 dB compared to the optimal detector. At the IEEE ICC 2013 conference [CI28], I proposed a near-maximum likelihood detection algorithm based on 1-flip local search (*i.e* also called intensification or exploitation strategy) over initial feasible solutions subset ξ_{start} drawn from Hamming codes. This work extended the concept of constructing well-distributed initial feasible points through the use of coding theory. In the following, we denote the BCH-based and the Hamming-based MIMO detectors by intensification over well-distributed regions (IWDR) detector.

The overall complexity of the considered MIMO detectors during this work on the well-distributed regions approach is summarized in Fig. 4.7. Compared to the sphere decoding (SD) and the optimum maximum likelihood (ML) detectors, it can be observed that IWDR based MIMO detectors are able to take advantage from their simple and parallel structure and offer low computational complexity. Currently, we are working on the development of improved variants of IWDR-based detectors through the use of other mathematical approaches such as locality-sensitive hashing or oriented nonuniform distributed codes.

4.4.3 [C5]: Improved tree-search detection via bio-inspired firefly algorithm

Recently, bio-inspired algorithms have shown promising performance and gained tremendous interest as an attractive solution to ML detection problem (4.17). Some of these algorithms be-

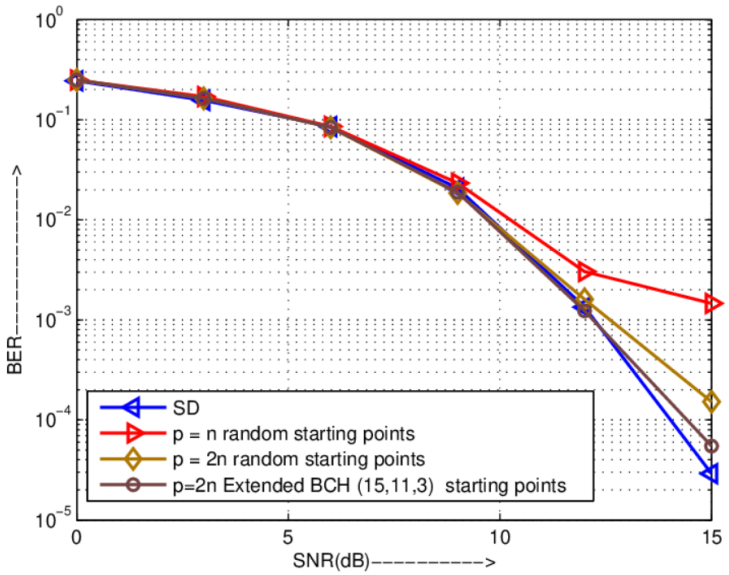


Figure 4.6 – BER Performance of hard-output sphere decoding (SD), random-based (RRS), and BCH-based MIMO detectors for a 8×8 MIMO systems with 4-QAM modulation and $n = 16$.

MIMO detector	Number of Flops
ZF	$4t^3 + 8rt^2 - t^2 + 3rt + t - r$
MMSE	$12t^3 + 8rt^2 + 2t^2 + 3rt + t - r$
ML	$(4rt + 2r)M^t$
SD [49]	$4t^3 + 7t^2 - \frac{t}{2} + (2t + 2)\frac{M^{\eta t}-1}{M-1}$, where $\eta = \frac{3\sigma^2}{c^2(M^2-1)+6\sigma^2}$, and $c^2 = \mathbb{E}\left[\lvert\mathbf{h}_i\rvert^2\right]$
IWDR	$8t^3 + [3\theta(L+1)+2]t^2 - 2r$, where $\theta \leq 3$ and L the cardinality of ξ_{start}

Figure 4.7 – MIMO Detectors Complexity for $t \times r$ MIMO system with M-QAM modulation.

came paradigms in conventional and massive MIMO detection problem, such as ant colony optimization (ACO), genetic algorithms (GA), particle swarm optimization (PSO) and firefly algorithm (FA) [24]. Since the maximum likelihood detection problem is NP-hard, due to the discrete nature of the transmitted signal constellation, a large number of papers have been published in order to provide near optimal bio-inspired algorithm with a reasonable computational complexity (see papers [43, 44, 49, 50, 51, 52] and references therein).

In order to find near-optimal solutions for ML problem, bio-inspired algorithms have the ability to efficiently explore (the worst-case scenario is *random walk*) the search space Ω^t without getting trapped in a limited area region. On the opposite, the tree-based detectors are capable to exploit (the worst-case scenario is *local search*) efficiently the search space Ω^t with a reasonable complexity considering early stopping of the search process through pruning paths. Therefore, the balance of exploration and exploitation stages is one of the major factors that contribute to the success of an algorithm.

In the context of the PhD of Bastien Trotobas, we investigated the possibility of incorporating a probabilistic behavior in classical tree-search MIMO detectors in order to strike a balance between intensification and diversification of the search space. In this regard, we have chosen to use the firefly algorithm version published in [52]. The major steps of FA can be summarized by the following three main rules (see Fig. 4.8): **(i)** Fireflies are attracted to one another in despite of their gender. **(ii)** Attractiveness is proportional to their brightness, which is declined with increase in the distance between them. When there is no brighter firefly, then they move randomly. **(iii)** Brightness is proportional to the value of the considered objective function.

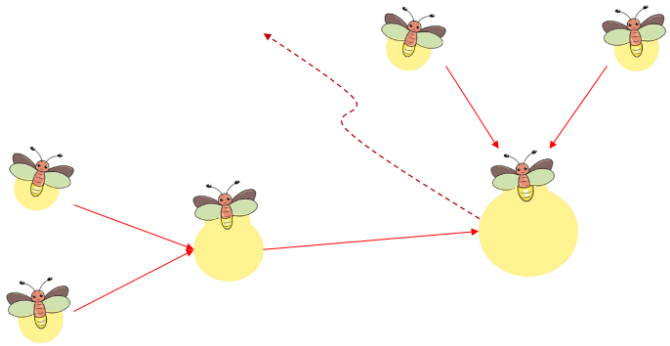


Figure 4.8 – Conceptual view of the natural behaviors of fireflies

We proposed an improvement of tree-based MIMO detectors by the introduction of a stochastic exploration mechanism operating at each tree-level (*i.e* layer). The firefly algorithm is employed to obtain the above stochastic behavior. Therefore, each stationary firefly can be interpreted as a tree node, and the run of a given firefly represents a path from the tree root to a leaf, as depicted in Fig. 4.9.

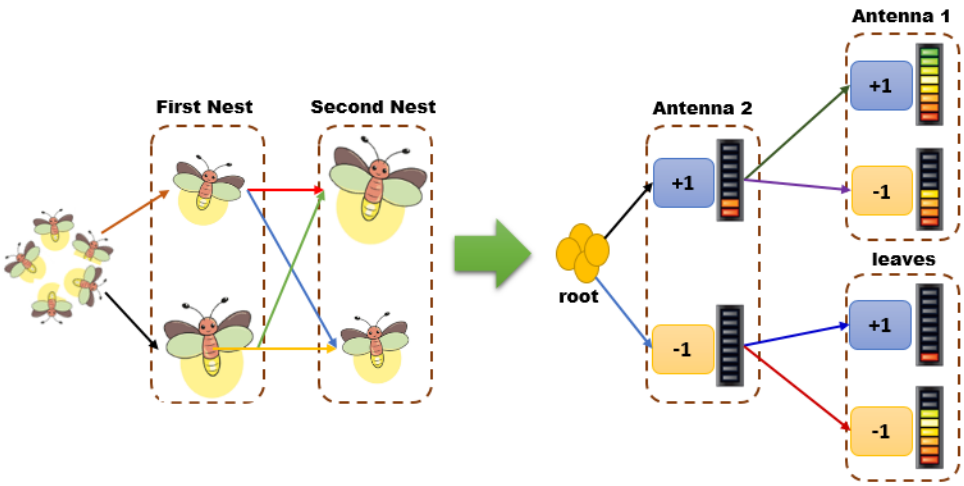


Figure 4.9 – Firefly algorithm interpreted as a tree search for 2×2 MIMO system with BPSK

The resulting new tree-search detector is demonstrated to be highly efficient in terms of computational complexity and bit error performance compared to classical tree-based, IL Geometrical and firefly-based MIMO detectors. This work was recently published in IEEE VTC2020-Spring [CI56], included hereafter.

Adding Exploration to Tree-Based MIMO Detectors Using Insights from Bio-Inspired Firefly Algorithm

Bastien Trotobas*, Youness Akourim†, Amor Nafkha* and Yves Louët*
*SCEE/IETR UMR CNRS 6164, CentraleSupélec Avenue de la boulaie, 35576 Cesson Sévigné, France
Email: {bastien.trotobas,amor.nafkha,yves.louet}@centralesupelec.fr
†École normale supérieure de Rennes, Campus de Ker Lann, Avenue Robert Schuman, 35170 Bruz, France
Email: youness.akourim@ens-rennes.fr

***Abstract*—The standard multiple-input multiple-output (MIMO) detectors exploit the available information to resolve the detection problem. Alternative algorithms, such as bio-inspired or geometrical detectors, mix exploitation with exploration to bypass local minima and enhance the results. This paper examines the benefits of adding exploration to the traditional tree-based detectors. For this purpose, a new interpretation of the bio-inspired detector based on the firefly algorithm (FA) is proposed. It is studied in a tree search paradigm and extended to soft-outputs. The findings suggest that the addition of a stochastic exploration to tree-based detectors significantly improves performance with a small computational overhead.**

***Index Terms*—MIMO detection, firefly algorithm, soft-output, exploration vs. exploitation, tree-search, Pareto analysis**

I. Introduction

Multiple-input multiple-output systems have made their way into most of the standards such as WiFi (IEEE 802.11n/ac), WiMAX, long-term evolution (LTE), and 5G. This technology exploits the spatial components of the wireless channel to increase link robustness and provide significant capacity gain via transmit diversity and space-division multiplexing (SDM). The latter technology improves throughput and link quality without requiring new frequency bands. However, the substantial increase in data rates comes at the expense of a more complex receiver design. Indeed, all data streams add up in the same time-frequency slot. New detection algorithms are therefore required to separate the data streams and retrieve signals transmitted from different antennas. The MIMO receivers typically rely on pilot signals to get information about the channel state information (CSI).

In spatial multiplexing MIMO systems, the optimum maximum likelihood decoding (MLD) problem leads to an integer least-squares (ILS) problem, which is equivalent to finding the closest lattice point to a given point. Due to the discrete search space of the ILS problem, it is NP-hard [1]. Therefore, it is impossible to find an optimal detection algorithm in reasonable polynomial calculation time on a Turing machine. Such an algorithm could exist if P=NP, but this possibility is considered unlikely, and no algorithm has been found so far despite intensive researches. Thus, several optimal algorithms, as well as polynomial heuristics, have been developed to tackle the MLD detection problem. Both types of algorithms provide hard and soft outputs versions. Hence, they can both be coupled with channel coding techniques to improve the MIMO link quality.

The MLD detection problem is an instance of combinatorial optimization problem, which seeks to find the best solution to a problem out of a large discrete set (the feasible set). In general, combinatorial optimization can be solved using classical algorithms that rely on a tree representation. Branch and bound approaches, known as sphere decoding (SD) or depth-first tree-search, provides an optimal solution in a non-polynomial time [2]–[6]. These algorithms can detect in an acceptable time when the number of antennas is limited [7]. The application of Dijkstra's algorithm, called best-first (BF) tree-search, is another method providing an optimal result in a reasonably short time [8]–[10]. However, this type of detector requires a lot of storage space and is harmed by input dependent run-time. Breadth-first detectors have been introduced as approximations of the previous algorithms to produce results in a polynomial time. Thus, the M-algorithm [11] and its successor, the K-best detector [12]–[14], benefit from predictable the time and storage utilization, still providing near-optimal performance.

The above detectors are based just on tree exploitation step given the available information. Other techniques are available in literature based on mixing exploitation with exploration in order to escape from local minima and produce more reliable results. In particular, bio-inspired metaheuristics such as ant colony optimization (ACO) [15] or FA [16] have been proposed. These algorithms are much more complex than tree-based techniques. However, they are also more flexible. Indeed, they add an exploration component instead of the pure exploitation used in tree-based detectors. Another method used a geometrical approach to perform exploration and exploitation steps has been recently proposed in [17].

This paper investigates the use of an exploratory element in common tree-based algorithms. To achieve this goal, we reformulate the FA detector described in [16] as a tree-path algorithm associated with an exploratory factor. We propose a new tree-based detector with parameters that can be adjusted to allow an exploitation-exploration trade-off. The proposed detector can provide soft outputs for higher-order modulation schemes. Moreover, this detector is compared with well-known tree-based detectors without exploration step and the geometrical exploration-exploitation detector.

The paper is organized as follows. The MIMO system model under consideration, the MLD detection problems, and the QR decomposition are presented in Section II. In Section III, we introduce the hard-output FA detector, and we extend it to produce the soft-output in Section IV. In Section VI, we present the simulation results of the proposed algorithm. Finally, Section VII concludes the paper.

II. System Model and Mathematical Definitions

The MIMO model is introduced in Section II-A. Section II-B defines the hard-output, and soft-output detection problems and Section II-C rewrites these problems introducing the QR decomposition.

A. MIMO Model

We model the $n \times n$ MIMO system for a single time/frequency slot. This model is suitable for any link such that the interferences between frequency slots and between time slots are negligible. Let $\mathbf{H} \in \mathbb{C}^{n\times n}$ be the channel matrix such that H_{ij} is the complex gain corresponding to the path from antenna j to antenna i. Let $\mathcal{Q}$ be the set of all the constellation symbols available. Let $\mathbf{y} \in \mathbb{C}^n$ be the signal received on each antenna that corresponds to the symbols transmitted $\mathbf{x} \in \mathcal{Q}^n$ after propagation through the channel added to the circularly-symmetric Gaussian noise $\mathbf{w} \sim \mathcal{CN}(0, \sigma^2)$. The link model is then expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}. \tag{1}$$

B. Definition of the Detection Problems

On the one hand, the hard-output detection problem refers searching for the most reliable transmitted symbols given the channel state and the received vector. This process is equivalent to solve the combinatorial optimization problem

$$\underset{\hat{\mathbf{x}}\in\mathcal{Q}^n}{\arg\min} \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2. \tag{2}$$

We further denote by the objective function the expression

$$E(\hat{\mathbf{x}}) = \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2. \tag{3}$$

On the other hand, the soft-output detection problem denotes the computation of the log-likelihood ratios (LLRs) of each transmitted bit b_{ij} with b_{ij} the i^{th} bit encoded in the j^{th} symbol emitted. The LLRs are commonly approximated using the max-log approximation

$$L_{ij} \approx \frac{1}{2\sigma^2}\left(\min_{\hat{\mathbf{x}}\in\mathcal{X}_{ij}^0} \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2 - \min_{\hat{\mathbf{x}}\in\mathcal{X}_{ij}^1} \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2\right), \tag{4}$$

where $\mathcal{X}_{ij}^k = \{\hat{\mathbf{x}} \in \mathcal{Q}^n : b_{ij} = k\}$ is the set of all symbols with b_{ij} equals to k [5], [18], [19]. Most detectors approximate (4) using a list $\mathcal{L} \subset \mathcal{Q}^n$ rather than computing the 2^n objective functions. Therefore, the LLRs expression becomes

$$L_{ij} \approx \frac{1}{2\sigma^2}\left(\min_{\hat{\mathbf{x}}\in\mathcal{L}\cap\mathcal{X}_{ij}^0} \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2 - \min_{\hat{\mathbf{x}}\in\mathcal{L}\cap\mathcal{X}_{ij}^1} \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2\right). \tag{5}$$

If $\mathcal{L}\cap\mathcal{X}_{ij}^k = \emptyset$, it is assumed that $b_{ij} = k$ and L_{ij} is set to its maximum or minimum value to express the reliability on this bit.

C. Rewriting using QR Decomposition

Let $\mathbf{H} = \mathbf{QR}$ be the QR decomposition of the channel matrix with $\mathbf{Q}$ a unitary matrix and $\mathbf{R}$ an upper triangular one. Since $\mathbf{Q}$ is an isometry, we can rewrite the objective function as

$$E(\hat{\mathbf{x}}) = \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\|^2 = \|\mathbf{Q}^*\mathbf{y} - \mathbf{Q}^*\mathbf{Q}\mathbf{R}\hat{\mathbf{x}}\|^2 = \|\tilde{\mathbf{y}} - \mathbf{R}\hat{\mathbf{x}}\|^2 \tag{6}$$

with $\tilde{\mathbf{y}} \triangleq \mathbf{Q}^*\mathbf{y}$ the rotated and reflected received vector. We can then introduce the distance d_i such that

$$E(\hat{\mathbf{x}}) = \|\tilde{\mathbf{y}} - \mathbf{R}\hat{\mathbf{x}}\|^2 = \sum_{i=1}^{n} d_i \tag{7}$$

where

$$d_i \triangleq \left|\tilde{y}_i - \sum_{j=i}^{n} R_{i,j}\hat{x}_j\right|^2 \tag{8}$$

This new expression allows for the computation of d_i using only the i last components of $\mathbf{x}$. Therefore, the detector can decide on the last component (i.e., the symbol of the last antenna) without any assumption on the other symbols. Then, each symbol can be detected from the last to the first without the need for cross-assumptions.

III. FA-Based As a Randomized Tree-Based Detector

The firefly algorithm is a bio-inspired metaheuristic that can handle the majority of optimization problems. A particular variant of it has been proposed in [16] to tackle the combinatorial optimization in the hard-output detection problem (2). Section III-A reminds the main steps of this detector in the common swarm-based framework. Section III-B revisits the FA-based detector as a randomized tree-search to highlight that the swarm-based framework and the tree-based one overlap.

A. FA as a Swarm-Based Detector

The FA detector, as described in [16], is a swarm-based bio-inspired meta-heuristic detector. As an illustration, Fig. 1 shows the firefly algorithm implemented for a MIMO system with 2×2 antennas using a BPSK modulation scheme. From a bio-inspired perspective, a transmit antenna can be referred to as a nest, and each constellation symbol is depicted by a stationary firefly ($i.e$ the stars). The FA detector simulates moving bugs ($i.e$ the disks) that have to choose a mate in each nest, based on their attractiveness. Once a moving firefly reaches the last nest, its chosen mates correspond to a decoded symbol vector. For instance, a firefly selecting each time the biggest star on the figure will detect $\hat{\mathbf{x}} = (+1, -1)$. The swarm is composed of F bugs that travel at once, and the final result is the best path selected among all the achieved ones ($i.e$. the one with the lower objective function value).

The attractiveness of the stationary firefly i to the moving firefly m is computed as

$$\beta_{i,m} = e^{-\gamma d_{i,m}^k}, \tag{9}$$

where $\gamma \geqslant 0$, $k \geqslant 1$ are two parameters, and $d_{i,m}$ the distance computed using (8) assuming the previous choices of the firefly m. As discussed in Section II-C, the QR decomposition allows to compute the attractiveness of a firefly based on the mates chosen on the previously visited nests but no assumptions on the next choices are required. The parameters γ and k control the exploration-exploitation trade-off. The bigger these parameters are, the more the difference in the objective function are amplified, and therefore, the more the detector focus on good children. Conversely, when these parameters are small, the algorithm is more likely to explore paths rather than focusing on a few promising ones.

A moving firefly selects a mate in each nest based on its attractiveness using a non-uniform random choice. For each nest, the moving firefly m will choose the mate i with probability

$$P_{i,m} = \frac{\beta_{i,m}}{\sum_j \beta_{j,m}} \tag{10}$$

with $\sum_j \beta_{j,m}$ the sum of the attractiveness of all the fireflies in this nest. Therefore, the FA swarm explores several possibilities thanks to the stochastic term rather than using only an exploitation process as in most MIMO detectors.

B. FA as a Randomized Tree-Based Detector

The FA detector may be interpreted as a stochastic tree-path search. In this case, each stationary firefly in a nest represents a node in a tree. Besides, a moving firefly run corresponds to a path from the root to a tree leaf. Fig. 2 shows the same example as Fig. 1, highlighting the tree structure. Consequently, decode a message corresponds to build as many tree paths as the number of moving fireflies. The partial path is extended following a stochastic process defined by the probability of (10). The random process can be considered an exploration since it allows extending the path to nodes that are not the best, hoping to be compensated for the over-cost and obtained a better final result.

This paradigm shift provides a more straightforward comparison with many other typical detectors such as K-best or BF. Indeed, the complexity of these algorithms is often measured through the number of visited nodes. This metric is now easily accessible for FA. To illustrate, K-Best visits roughly Kn nodes with K the algorithm parameter whereas FA accesses Fn nodes. Therefore, this paradigm shift highlights that K and F have exactly the same role in evaluating the complexity of these two detectors. Section VI-A will introduce a re-evaluation of the FA complexity by exploiting the tree approach.

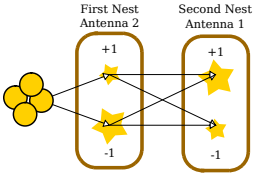


Fig. 1. Representation of FA as a firefly swarm with $n = 2$ and a BPSK.

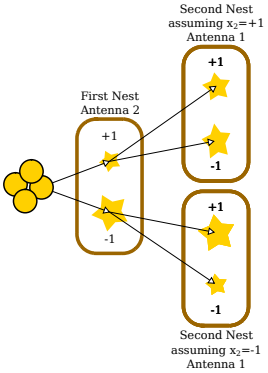


Fig. 2. Representation of FA as tree-search algorithm ($n = 2$ and BPSK).

IV. Proposed Soft-Output FA-Based Detector

In this section, we propose a new detector based on the FA. Section IV-A introduces the new algorithm that tackles the soft-output detection problem defines in II-B. Then, Section IV-B provides a simplification for the attractiveness computation.

A. Description of the Soft-Output FA-Based Detector

Most soft-output MIMO detectors rely on the max-log approximation computed on a restricted list of candidates as described in Section II-B. The FA-based detector can easily be adapted to generate a list of rather than a symbol vector $\mathbf{x}$. Indeed, its detection process is based on the parallel construction of F paths. Therefore, it is possible to build the list $\mathcal{L}$ from the aggregation of the F paths and derive the LLRs from equation (5).

Fig. 3 sums up all the steps of the proposed soft-output detector. This pseudo-code is similar to the hard-output detector except for the aggregation of all the built paths in the list $\mathcal{L}$. This algorithm can be used as such for any modulation scheme, including higher-order modulations.

B. Padé Approximant for the Attractiveness Computation

The attractiveness computation requires the evaluation of an exponential function, which may represent a high cost. Nevertheless, any approximation of (9) has to keep its major characteristics, i.e., $\beta_{i,m}$ is strictly positive and strictly decreasing with $d_{i,m}$. Hence, we propose to estimate (9) by the simplest Padé approximant satisfying these two properties. The Padé approximant of order (0,1) of the exponential function is given by

$$e^x \approx \frac{1}{1-x}, \tag{11}$$

the attractiveness expression becomes

$$\beta_{i,m} \approx \frac{1}{1+\gamma d_{i,m}^k} \tag{12}$$

This new formulation replaced the exponential with addition and an inversion, which is easier to compute. In the following, the proposed detector is tested using this attractiveness expression.

Inputs:

- Received vector: $\mathbf{y}$
- QR decomposition of the channel matrix: $\mathbf{Q}, \mathbf{R}$
- Noise variance: σ^2

Compute $\tilde{\mathbf{y}} = \mathbf{Q}^*\mathbf{y}$.
Initialize an empty candidate list $\mathcal{L} = \emptyset$.
for all firefly/path $m \in \{0 \dots F\}$ **do**
  Start at the root node with objective function $E_m = 0$.
  for coordinate $c = n$ **to** 1 **do**
    Evaluate the attractiveness $\beta_{i,m}$ of each child i.
    Compute the probability to select each child using (10).
    Select a child randomly according to the probabilities.
    Increment the objective function E_m by d_c from (8).
  end for
  Add the completed path to the list $\mathcal{L}$.
end for
Approximate LLRs based on σ^2, $\mathcal{L}$ and (5).
return the approximated LLRs.

Fig. 3. Pseudo-code of the proposed soft-output FA-based detector. The number of fireflies/paths $F \in \mathbb{N}_*$ is set for all the detections.

V. Evaluation of the Proposed Detector

This section introduces the evaluation criteria and the simulation setup in Section V-A and the reference detectors in Section V-B.

A. Simulation Settings and Evaluation Criteria

The addition of exploration in the tree search and the proposed algorithm are evaluated according to two opposing criteria: computational complexity and bit error rate (BER) performance. The complexity is assessed through the number of product operations required per detected vector. The Padé approximant presented in Section IV-B simplifies the computation of the exponential into a product and division. The complexity is investigated as a statistic since the detector behavior is stochastic. The statistics of number of the product operations and the BER performance are based on Monte-Carlo simulations. For each SNR, the simulation runs until 200 binary errors occur or 5.10^5 bits are transmitted. The data message is encoded by the irregular, systematic LDPC code of ratio 1/2 from the WiMAX standard IEEE 802.16e. Each block contains 720 bits of information, and decoding is performed by 15 iterations of the belief-passing min-sum algorithm. For simulation codes, see [20]. The balance between these two competing criteria is explored through a Pareto analysis. A detector is considered Pareto efficient if no algorithm can enhance one criterion without degrading the others. The set of Pareto efficient detectors constitutes the Pareto front, and selecting one or the other depends on the complexity-performance trade-off.

B. Reference Detection Algorithms

The new detector will be compared to several well-known detectors to evaluate its suitability and determine the relevance of the exploration in the tree-based algorithm. Three detectors are considered for comparison: two tree-based algorithms without exploration and a geometrical detector mixing exploration and exploitation. These detectors are selected for their differences regarding the exploration step and their well-suited structure for hardware implementation. Those algorithms create a list and use the max-log approximation described in (5) to produce the soft-output values. The remainder of this section briefly discusses the characteristics of each algorithm. The number of product operations of the tree-based detector and the geometrical algorithm was already given in [17].

1) Breadth-first K-best [12]: K-best is a breadth-first tree-based detector with sub-optimal performance. This algorithm searches the tree from root to leaf, retaining only a given number of paths at each tree level. This approximation allows it to achieve near-optimal performance with a polynomial-time and storage space. Furthermore, a single parameter adjusts the complexity-performance trade-off being the number of paths retained.

2) Best-first tree search [10]: The approximation-free best-first detectors achieve optimal performance but require a significant memory space. Reference [10] introduces an approximated best-first tree search reducing memory usage and maintaining near-optimal performance. This is achieved by constraining the size of the node heaps.

3) Geometrical-based detector [17]: The geometrical detector explores based on the channel matrix singular value decomposition (SVD) and then exploits the intermediate result through local search. This combination of exploration and exploitation makes it more efficient than tree-based detectors in some scenarios. Nevertheless, the lack of hindsight on this technique, especially on the mathematical aspect, complicates the refinement of this class of algorithms.

VI. Results on Complexity and Performance

Section VI-A evaluates the impact of the exploration on the complexity compared to the usual tree-search pattern. Then, Section VI-B explores the performance-complexity trade-off to assess the proposed detector's relevance in comparison with state-of-the-art algorithms.

A. Complexity Cost of Exploration vs. Usual Tree-Search

The conventional tree-based detectors build paths starting from the root and extending some partial paths. By contrast, the proposed FA-based detector independently builds F paths from the root to a leaf. Thus, the same node can be visited several times when constructing the F paths. It is then interesting to avoid recomputing the metric and to reuse the values already calculated. Fig. 4 shows the ratio of the number of unique nodes to the total number of visited nodes for a 4x4 MIMO system and a 16-QAM. The three boxplots show the minimum, maximum and quartiles for parameters settings. For the three tests, $k = 3$ and only F and γ change. The first two lines only differ in the number of fireflies/paths. They show that, as F decreases, the ratio of the number of unique nodes increases since the total number of nodes decreases. The last two lines display the evolution of the ratio as a function of

γ. As reported in Section III-A, the more γ dicreases, the more the detector explores more different nodes. The three boxplots in Fig. 4 are typical cases of parameter settings. The boxplots point out that there are always less than 50% unique nodes. Furthermore, the threshold of 20% unique nodes is rarely crossed with two third quartiles below this level and one with a ratio of 24%. This descriptive statistic confirms that it is interesting to share the results rather than to recompute the metric each time the node is encountered. We further assume that each node is computed once for all the paths that visited it. This assumption change neither the pseudo-code from Fig. 3 nor the results of the detector.

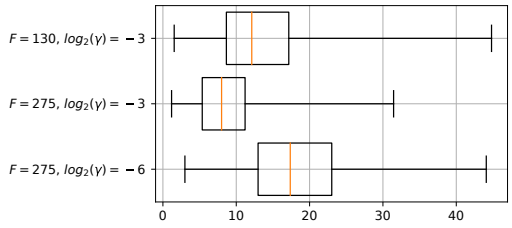


Fig. 4. Boxplot of the ratio of the number of unique to total amount of visited nodes. The ratio is displayed as percents, whiskers represent the minimum and maximum, $k = 3$ and SNR is set to 17dB.

B. Performance-Complexity Trade-off Analysis

Fig. 5 displays the Pareto analysis of the performance-complexity trade-off. An exhaustive search sweeps all the tuning parameters of each algorithm to explore the potential of each detector. The algorithms are compared based on the number of products per detected vector and according to the SNR required for a BER of 1.10^{-4}. The Pareto analysis demonstrates that the proposed detector is efficient across a wide operating range from the worst SNRs to the average regimes where it outperforms K-Best and the geometrical detector. Moreover, the proposed detector shows better flexibility than Best-First detection. The previous references are only efficient when complexity is the critical aspect. Indeed, the proposed detector needs a fair number of fireflies/paths to achieve reasonable performance and is therefore not adapted when the complexity must be extremely low.

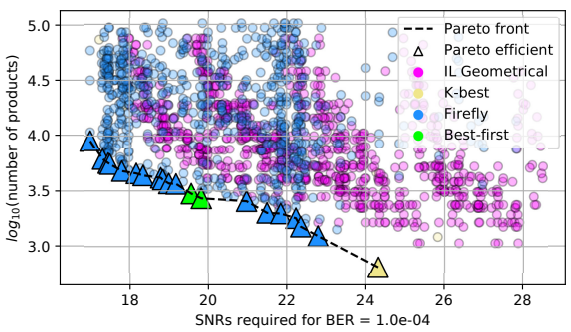


Fig. 5. Pareto analysis of Performance versus Complexity for a 16-QAM.

VII. Conclusion

In this paper, the bio-inspired FA-based detector has been redesigned in a tree path paradigm and extended to soft-output. A Pareto analysis demonstrated its interest compared to several references. Besides, this result indicates that exploration an advantage for tree-based detectors. Furthermore, the exploration-exploitation approach presented is more efficient than the one used in geometrical detectors.

References

[1] D. Micciancio, "The hardness of the closest vector problem with preprocessing," *IEEE T. Inform. Theory*, vol. 47, no. 3, pp. 1212–1215, Mar. 2001.

[2] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. of Computation*, vol. 44, no. 170, pp. 463–471, 1985.

[3] C. P. Schnorr and M. Euchner, "Lattice Basis Reduction: Improved Practical Algorithms and Solving Subset Sum Problems." *Math. prog.*, vol. 66, no. 1-3, pp. 181–199, 1994.

[4] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE T. Inform. Theory*, no. 8, pp. 2201–2214, Aug. 2002.

[5] B. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE T. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[6] C. Studer and H. Bolcskei, "Soft–input soft–output single tree-search sphere decoding," *IEEE T. Inform. Theory*, vol. 56, no. 10, pp. 4827–4842, Oct. 2010.

[7] J. Jalden and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE T. Signal Proces.*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[8] C.-H. Liao, T.-P. Wang, and T.-D. Chiueh, "A 74.8 mw soft-output detector ic for 8 x 8 spatial-multiplexing mimo communications," *IEEE J. Solid-ST. Circ.*, vol. 45, no. 2, pp. 411–421, Feb. 2010.

[9] R. Y. Chang and W.-H. Chung, "Best-first tree search with probabilistic node ordering for mimo detection," *IEEE Trans. Wirel. Commun.*, vol. 11, no. 2, pp. 780–789, Feb. 2012.

[10] G. He, X. Zhang, and Z. Liang, "Algorithm and architecture of an efficient mimo detector with cross-level parallel tree-search," *IEEE T. VLSI Syst.*, vol. 28, no. 2, pp. 467–479, Feb. 2020.

[11] K. K. Y. Wong and P. J. McLane, "A low-complexity iterative mimo detection scheme using the soft-output m-algorithm," *IST Mob. Sum.*, p. 5, Jun. 2005.

[12] Z. Guo and P. Nilsson, "Algorithm and implementation of the k-best sphere decoding for mimo detection," *IEEE J. Sel. Area Comm.*, vol. 24, no. 3, pp. 491–503, Mar. 2006.

[13] Z. Yan, G. He, Y. Ren, W. He, J. Jiang, and Z. Mao, "Design and implementation of flexible dual-mode soft-output mimo detector with channel preprocessing," *IEEE T. Circuits-I*, vol. 62, no. 11, pp. 2706–2717, Nov. 2015.

[14] S.-J. Choi, S.-J. Shim, Y.-H. You, J. Cha, and H.-K. Song, "Novel mimo detection with improved complexity for near-ml detection in mimo-ofdm systems," *IEEE Access*, vol. 7, pp. 60 389–60 398, 2019.

[15] J. C. Marinello and T. Abrão, "Lattice reduction aided detector for mimo communication via ant colony optimisation," *Wireless Pers. Commun.*, vol. 77, no. 1, pp. 63–85, Jul. 2014.

[16] A. Datta and V. Bhatia, "A near maximum likelihood performance modified firefly algorithm for large mimo detection," *Swa. and Evol. Computation*, vol. 44, pp. 828–839, Feb. 2019.

[17] B. Trotobas, A. Llave, A. Nafkha, and Y. Louët, "When Should We Use Geometrical-Based MIMO Detection Instead of Tree-Based Techniques? A Pareto Analysis," *IEEE Access*, vol. 8, pp. 191 163–191 173, 2020, conference Name: IEEE Access.

[18] W. Koch and A. Baier, "Optimum and sub-optimum detection of coded data disturbed by time-varying intersymbol interference," in *IEEE Global Telecommu. Conf. and Exhib.*, Dec. 1990, pp. 1679–1684.

[19] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal map decoding algorithms operating in the log domain," in *IEEE Int. Conf. on Commu.*, vol. 2, Jun. 1995, pp. 1009–1013.

[20] B. Trotobas and Y. Akourim, "Simulation codes," 2020. [Online]. Available: bitbucket.org/scee_ietr/exploration_tree_search

4.5 Summary

The work on the MIMO system capacity and MIMO detection techniques was performed mainly in the context of two PhDs theses, two Master theses, and a collaboration with Nizar DEMNI, associate professor at the University of Rennes 1 (UR1). A remarkable result was the derivation of new exact and compact expressions of the ergodic capacity of both Rayleigh and Jacobi fading MIMO channels with the assumption that CSI is available only at the receiver side. Moreover, simplified closed-form expressions for high/low SNR regime and tight closed-form upper and lower bounds are derived. In the context of MIMO detection topic, our main contributions can be summarized as depicted in Fig. 4.10, which shows the trade-off relation between exploration and exploitation of proposed near-optimal MIMO detection algorithms. The stochastic tree-search based detector consists of adding an exploration stage to the well-known tree-search based sphere detection techniques, however, the geometrical based approach allows an efficient exploration of the search space followed by a simple exploitation stage (*i.e.* local search). We perform a series of Matlab simulations in order to test the effectiveness of both proposed detectors in terms of computational complexity and performance. We believe that combine exploration and exploitation strategies would be the best way to design high-performance and low-complexity MIMO detection algorithms.

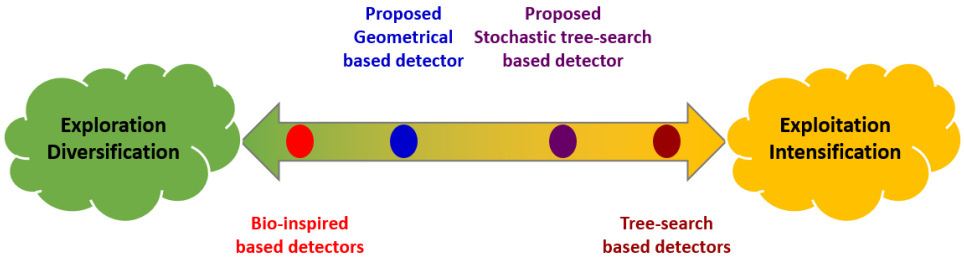


Figure 4.10 – MIMO detection techniques: Exploration vs Exploitation

We resume below the combined outputs of the work on MIMO communication systems described in the present chapter.

Research Outputs of MIMO Communication Systems

- **Publications**: 1 Patent + 5 Journals + 10 Conferences + 1 Chapter.
- **Collaborators**: Demni N., Boutillon E., Louet Y., Roland C.
- **Postdoctoral researcher**:
- **PhD Students**: Trotobas B., Bonnefoi R., Aubert S.
- **Master Students**: Akourim Y.
- **Supported by**: ENS grant, IETR grant

BIBLIOGRAPHY

[1] J. Mietzner, R. Schober, L. Lampe, W. H. Gerstacker, and P. A. Hoeher, "Multiple-antenna techniques for wireless communications - a comprehensive literature survey," *IEEE Communications Surveys Tutorials*, vol. 11, no. 2, pp. 87–105, 2009.

[2] G. J. Foshini and M. J. Gans, "On the limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[3] J. Winters, "Smart antennas for wireless systems," *IEEE Personal Communications*, vol. 5, no. 1, pp. 23–27, 1998.

[4] L. Godara, "Applications of antenna arrays to mobile communications, part i: Performance improvement, feasibility, and system considerations," *Proceedings of the IEEE*, vol. 85, no. 7, pp. 1029–1030, 1997.

[5] ——, "Application of antenna arrays to mobile communications. ii. beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.

[6] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-blast: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *1998 URSI International Symposium on Signals, Systems, and Electronics. Conference Proceedings (Cat. No.98EX167)*, 1998, pp. 295–300.

[7] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[8] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.

[9] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277–1294, 2002.

[10] D. Love, R. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2735–2747, 2003.

[11] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[12] E. Bjornson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive mimo systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7112–7139, 2014.

[13] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[14] P. M. Krummrich, "Spatial multiplexing for high capacity transport," *Optical Fiber Technology*, vol. 17, no. 5, pp. 480–489, 2011.

[15] P. J. Winzer, "Energy-efficient optical transport capacity scaling through spatial multiplexing," *IEEE Photonics Technology Letters*, vol. 23, no. 13, pp. 851–853, 2011.

[16] P. J. Winzer and G. J. Foschini, "Mimo capacities and outage probabilities in spatially multiplexed optical transport systems," *Opt. Express*, vol. 19, no. 17, pp. 16 680–16 696, 2011.

[17] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. C. Burrows, R.-J. Essiambre, P. J. Winzer, D. W. Peckham, A. H. McCurdy, and R. Lingle, "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6×6 mimo processing," *Journal of Lightwave Technology*, vol. 30, no. 4, pp. 521–531, 2012.

[18] D. J. Richardson, J. M. Fini, and N. L. E., "Space-division multiplexing in optical fibres," *Nature Photon*, vol. 7, no. 5, pp. 345–362, 2013.

[19] S. Rommel, D. Perez-Galacho, J. M. Fabrega, R. Munoz, S. Sales, and I. Tafur Monroy, "High-capacity 5g fronthaul networks based on optical space division multiplexing," *IEEE Transactions on Broadcasting*, vol. 65, no. 2, pp. 434–443, 2019.

[20] L. Zhang, J. Chen, E. Agrell, R. Lin, and L. Wosinska, "Enabling technologies for optical data center networks: Spatial division multiplexing," *Journal of Lightwave Technology*, vol. 38, no. 1, pp. 18–30, 2020.

[21] T. Mizuno, T. Kobayashi, H. Takara, A. Sano, H. Kawakami, T. Nakagawa, Y. Miyamoto, Y. Abe, T. Goh, M. Oguma, T. Sakamoto, Y. Sasaki, I. Ishida, K. Takenaga, S. Matsuo, K. Saitoh, and T. Morioka, "12-core $\times$ 3-mode dense space division multiplexed transmission over 40 km employing multi-carrier signals with parallel mimo equalization," in *OFC 2014*, 2014, pp. 1–3.

[22] K. Shibahara, D. Lee, T. Kobayashi, T. Mizuno, H. Takara, A. Sano, H. Kawakami, Y. Miyamoto, H. Ono, M. Oguma, Y. Abe, T. Matsui, R. Fukumoto, Y. Amma, T. Hosokawa, S. Matsuo, K. Saitoh, M. Yamada, and T. Morioka, "Dense sdm (12-core × 3-mode) transmission over 527 km with 33.2-ns mode-dispersion employing low-complexity parallel mimo frequency-domain equalization," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 196–204, 2016.

[23] K. Shibahara, T. Mizuno, D. Lee, Y. Miyamoto, H. Ono, K. Nakajima, Y. Amma, K. Takenaga, and K. Saitoh, "Iterative unreplicated parallel interference canceler for mdl-tolerant dense sdm (12-core × 3-mode) transmission over 3000 km," *Journal of Lightwave Technology*, vol. 37, no. 6, pp. 1560–1569, 2019.

[24] S. Yang and L. Hanzo, "Fifty years of mimo detection: The road to large-scale mimos," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 1941–1988, 2015.

[25] R. Dar, M. Feder, and M. Shtaif, "The jacobi mimo channel," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2426–2441, 2013.

[26] A. Karadimitrakis, A. L. Moustakas, and P. Vivo, "Outage capacity for the optical mimo channel," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 4370–4382, 2014.

[27] Y. Chen and M. R. McKay, "Coulumb fluid, painlevé transcendents, and the information theory of mimo systems," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4594–4634, 2012.

[28] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, pp. 475–501, Jun 1964.

[29] C. W. J. Beenakker, "Random-matrix theory of quantum transport," *Rev. Mod. Phys.*, vol. 69, pp. 731–808, Jul 1997.

[30] B. Collins, "Product of random projections, jacobi ensembles and universality problems arising from free probability," *Probability Theory and Related Fields*, vol. 133, p. 315–344, 2005.

[31] T. Jiang, "Approximation of haar distributed matrices and limiting distributions of eigenvalues of jacobi ensembles," *Probability Theory and Related Fields*, vol. 144, p. 221–246, 2009.

[32] E. C. Song, E. Soljanin, P. Cuff, H. V. Poor, and K. Guan, "Rate-distortion-based physical layer secrecy with applications to multimode fiber," *IEEE Transactions on Communications*, vol. 62, no. 3, pp. 1080–1090, 2014.

[33] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[34] C. Carré, M. Deneufchatel, J. G. Luque, and P.Vivo, "Asymptotics of selberg-like integrals : The unitary case and newton's interpolation formula," *Journal of Mathematical Physics*, vol. 51, no. 22, p. 123516, 2010.

[35] G. E. Andrews, R. Askey, and R. Roy, *Special functions (Encyclopedia of mathematics and its applications)*. Cambridge University Press, 1999.

[36] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, 1999.

[37] L. Bai and J. Choi, *Low Complexity MIMO Detection*. Springer Publishing, 2014.

[38] M. A. Albreem, M. Juntti, and S. Shahabuddin, "Massive mimo detection techniques: A survey," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3109–3132, 2019.

[39] W. van Etten, "Maximum likelihood receiver for multiple channel transmission systems," *IEEE Transactions on Communications*, vol. 24, no. 2, pp. 276–283, 1976.

[40] S. Verdu, "Computational complexity of optimum multiuser detection," *Algorithmica*, no. 4, p. 303–312, 1989.

[41] M. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2389–2402, 2003.

[42] K. V. Vardhan, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "A low-complexity detector for large mimo systems and multicarrier cdma systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 3, pp. 473–485, 2008.

[43] X.-S. Yang, *Nature-Inspired Metaheuristic Algorithms: First Edition*. Luniver Press, 2008.

[44] ——, "Firefly algorithms for multimodal optimization," in *Stochastic Algorithms: Foundations and Applications*, O. Watanabe and T. Zeugmann, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 169–178.

[45] B. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 389–399, 2003.

[46] A. Mobasher, M. Taherzadeh, R. Sotirov, and A. K. Khandani, "A near-maximum-likelihood decoding algorithm for mimo systems based on semi-definite programming," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 3869–3886, 2007.

[47] G. Zhan and P. Nilsson, "Algorithm and implementation of the k-best sphere decoding for mimo detection," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 491–503, 2006.

[48] G. He, X. Zhang, and Z. Liang, "Algorithm and architecture of an efficient mimo detector with cross-level parallel tree-search," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 28, no. 2, pp. 467–479, 2020.

[49] M. Jiang and L. Hanzo, "Multiuser mimo-ofdm for next-generation wireless systems," *Proceedings of the IEEE*, vol. 95, no. 7, pp. 1430–1469, 2007.

[50] K. K. Soo, Y. M. Siu, W. S. Chan, L. Yang, and R. S. Chen, "Particle-swarm-optimization-based multiuser detector for cdma communications," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 3006–3013, 2007.

[51] J.-K. Lain and J.-Y. Chen, "Near-mld mimo detection based on a modified ant colony optimization," *IEEE Communications Letters*, vol. 14, no. 8, pp. 722–724, 2010.

[52] D. Arijit and B. Vimal, "A near maximum likelihood performance modified firefly algorithm for large mimo detection," *Swarm and Evolutionary Computation*, vol. 44, pp. 828–839, 2019.

Chapter 5

Contributions on Spectrum Sensing for Cognitive Radio

Cognitive radios have been proposed as the means to promote the efficient use of the limited available spectrum by exploiting the unused frequency channels in licensed band. Cognitive radio networks involve *primary users* of the spectrum, who are incumbent licensees, and *secondary users* who seek to opportunistically use the spectrum when the primary users are absent. Therefore, secondary users must be able to detect very weak signals in very noisy environments in order to avoid interfering with the primary users. In the literature, several spectrum sensing techniques that differ in their contexts and employed strategies have been developed and extensively investigated during the last two decades. Thus, sensing techniques can be classified in many different ways: narrowband/wideband with respect to the bandwidth size, or cooperative/non-cooperative according to the number of secondary users involved in the sensing process, or blind/non-blind in regard to prior knowledge of signals. The interested reader can refer to survey papers [1, 2, 3, 4, 5, 6, 7], and references therein.

In this chapter, only non-cooperative spectrum sensing techniques are considered. We investigate the spectrum sensing problem in finite-sample size and asymptotic single-antennas or multiple antenna cognitive radio equipment under synchronous and asynchronous scenarios. In synchronous sensing, both primary and secondary users are synchronized in time. However, in asynchronous sensing, the secondary user is no longer synchronized in time with the primary user. Finite-size random matrix theory, compressed sensing framework, and moment matching approximation are extensively used in order to analytically derive the false-alarm and detection probabilities.

5.1 Noncooperative narrowband spectrum sensing

The problem of signal detection is to decide whether a particular spectrum band is vacant (*i.e.* available) or not. That is, in its simplest form we want to distinguish between the two hypotheses: null hypothesis $\mathcal{H}_0$ where the interested channel bandwidth contains only noise component, and alternative hypothesis $\mathcal{H}_1$ where the primary user's signal is present. In the sequel, we address the spectrum sensing problem through a single cognitive radio that is equipped

with M receive antennas. Let N denotes the number of samples at each antenna of the secondary user receiver (*i.e.* N is assumed to be greater than M), and $\mathbf{y}_m(k)$ is the k^{th} received sample at the m^{th} receiving antenna. Therefore, we could write the received sample at the m^{th} antenna as:

$$\mathbf{y}_m(k) = \mathbf{h}_m(k)\mathbf{x}(k) + \mathbf{w}_m(k) \tag{5.1}$$

where $\mathbf{x}(k)$, $k \in [1, \ldots, N]$, represents the primary user signal to be detected, $\mathbf{h}_m(k)$ denotes the channel fading coefficient between the primary user and the m^{th} receive antenna, and $\mathbf{w}_m(k)$ is the background noise and various interference sources, including signals which are not of interest to detect. We assume that the additive noise $\mathbf{w} = vect\,[\mathbf{y}_1, \ldots, \mathbf{y}_M]$ is white, zero-mean, and circularly symmetric complex Gaussian (*i.e.* $\mathbf{w} \sim \mathcal{N}(0, \sigma_w^2 \mathbf{I}_{MN})$, where σ_w^2 is the noise variance). When the primary user is absent, *i.e.* $\mathbf{x}(k) = 0$, the equation (5.1) can be reduced to $\mathbf{y}_m(k) = \mathbf{w}_m(k)$, $m \in [1, \ldots, M]$. Thus, analytically, the signal detection can be formalized as a binary hypothesis test problem:

$$\begin{cases} \mathcal{H}_0 : & \mathbf{y}_m(k) = \mathbf{w}_m(k) \\ \mathcal{H}_1 : & \mathbf{y}_m(k) = \mathbf{h}_m(k)\mathbf{x}(k) + \mathbf{w}_m(k) \end{cases} \tag{5.2}$$

Let the vector $\mathbf{y} = vect\,[\mathbf{y}_1, \ldots, \mathbf{y}_M]$ of length MN contains all received samples stacked in one vector. Without loss of generality, we assume that the status of the primary user remains unchanged during the sensing time. Moreover, we assume that the sensing time is smaller than the coherence time of the channel. Then, the channel gain $\mathbf{h}_m(k)$ can be viewed as a constant during the spectrum sensing process. The sensing task must choose one of the two hypotheses (*i.e.* null or alternative) based on the test statistic $\mathcal{T}(\mathbf{y})$, resulting from the processing of all entries of $\mathbf{y}$, which is compared to a predetermined decision threshold λ as follows:

$$\mathcal{T}(\mathbf{y}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda \tag{5.3}$$

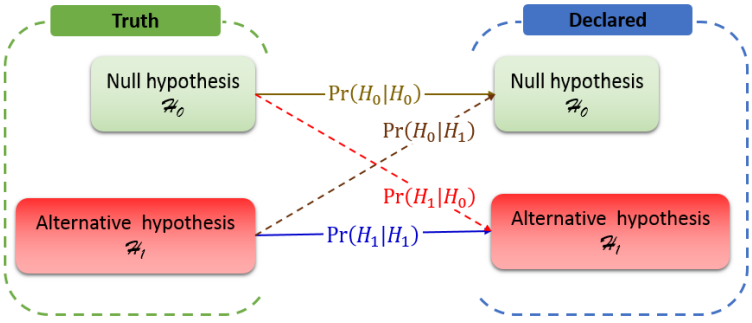


Figure 5.1 – Hypothesis test and possible outcomes with their corresponding probabilities.

Clearly, the problem facing a researcher is to choose the best test statistic $\mathcal{T}(\mathbf{y})$ and the

corresponding decision threshold λ that provide good detection performance. In accord with (5.3), one can define four probabilities related to the hypothesis test as depicted in Fig. 5.1 . Under alternative hypothesis, we define the probability of detection as $P_d = Pr\left(\mathcal{T}(\mathbf{y}) > \lambda|\mathcal{H}_1\right)$ and the probability of missed detection as $P_{mi} = Pr\left(\mathcal{T}(\mathbf{y}) < \lambda|\mathcal{H}_1\right)$. In similar way, under null hypothesis, we define the probability of false alarm as $P_{fa} = Pr\left(\mathcal{T}(\mathbf{y}) > \lambda|\mathcal{H}_0\right)$ and the probability of correct detection of idle channel as $P_i = Pr\left(\mathcal{T}(\mathbf{y}) < \lambda|\mathcal{H}_0\right)$. Note that it is not possible to reduce both error probabilities (*i.e.* P_{fa} and P_{mi}) simultaneously. In general, the performance of a given spectrum sensing technique is quantified in terms of its receiver operating characteristics (ROC), which gives the probability of detection P_d as a function of the probability of false alarm P_{fa} for a fixed SNR value.

Under the common detection criteria [4, 8], the likelihood ratio test (LRT) is the optimal decision solution for binary hypothesis testing problem (5.2). Based on the Neyman-Pearson lemma, when P_{fa} and the noise variance σ_w^2 is given, the LRT will maximize the detection probability. The decision statistic of LRT is determined using the following form:

$$\mathcal{T}_{LRT}(\mathbf{y}) = \frac{Pr(\mathbf{y}|\mathcal{H}_1)}{Pr(\mathbf{y}|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda \tag{5.4}$$

Recall that the threshold λ is determined by the nominal value of the false alarm probability P_{fa}. Because the noise is a white Gaussian random process, the two distributions $Pr(\mathbf{y}|\mathcal{H}_0)$ and $Pr(\mathbf{y}|\mathcal{H}_1)$ are given by:

$$\begin{cases} Pr(\mathbf{y}|\mathcal{H}_0) = \prod_{k=1}^{M} \frac{1}{(2\pi\sigma_w^2)^{(N/2)}} \exp\left(-\frac{\|\mathbf{y}_k\|^2}{2\sigma_w^2}\right) \\ Pr(\mathbf{y}|\mathcal{H}_1) = \prod_{k=1}^{M} \frac{1}{(2\pi)^{(N/2)}\sqrt{\det(\mathbf{R}_\mathbf{y})}} \exp\left(-\mathbf{y}_k\mathbf{R}_\mathbf{y}^{-1}\mathbf{y}_k^\dagger\right) \end{cases} \tag{5.5}$$

where $\mathbf{R}_\mathbf{y} = \mathbb{E}\left[\mathbf{y}^\dagger\mathbf{y}\right]$. When the primary signal entries are complex Gaussian with zero mean and covariance $\mathbf{R}_\mathbf{x} = \sigma_x^2\mathbf{I}_N$, the covariance matrix of $\mathbf{y}$, *i.e.* $\mathbf{R}_\mathbf{y}$, is a scaled identity matrix. Hence, the LRT can be reduces to classical energy detector technique for signal detection.

5.2 Energy detection

Without loss of generality and unless otherwise stated, we consider that $M = 1$. The energy detector measures the energy of the received signal $\mathbf{y}$ during an observation interval of N time samples and declares the presence of the primary signal if the measured energy is greater than the predefined threshold λ. It is the most commonly used technique for the purpose of spectrum sensing due to its low computational complexity, and the fact that it does not need a priori knowledge about the primary user signal [9, 10, 11]. The test statistic for the energy detector is

defined as [9]:

$$\mathcal{T}_{ED}(\mathbf{y}) = \sum_{k=1}^{N} |\mathbf{y}_1(k)|^2 \tag{5.6}$$

The exact performance of the energy detector can be derived by statistical analysis of the test statistic (5.6). Under the null hypothesis $\mathcal{H}_0$, $2\mathcal{T}_{ED}(\mathbf{y})/\sigma_w^2$ is a chi-squared random variable with $2N$ degrees of freedom, this leads to a probability of false alarm as

$$P_{fa} = Pr\left(\mathcal{T}(\mathbf{y}) > \lambda | \mathcal{H}_0\right) = 1 - \mathcal{G}_{2N}\left(2\lambda/\sigma_w^2\right) \tag{5.7}$$

where $\mathcal{G}_n(t) = [\Gamma(n)]^{-1} \int_0^{t/2} u^{n-1} \exp(-u) du$. Similarly, under the alternative hypothesis $\mathcal{H}_1$, $2\mathcal{T}_{ED}(\mathbf{y})/(\sigma_x^2 + \sigma_w^2)$ is distributed as a chi-squared random variable with $2N$ degrees of freedom. Therefore, the probability of detection is given by

$$P_d = Pr\left(\mathcal{T}(\mathbf{y}) > \lambda | \mathcal{H}_1\right) = 1 - \mathcal{G}_{2N}\left(2\lambda/\sigma_w^2(1+\rho)\right) \tag{5.8}$$

where $\rho = \|\mathbf{h}_1\|^2 \sigma_x^2/\sigma_w^2$ denotes the signal to noise ratio. According to the central limit theorem (CLT), when N is sufficiently large, the probability distribution of the test statistic $\mathcal{T}(\mathbf{y})$ can be approximated by Gaussian distributions under both hypothesis. Thus, the expressions of false alarm probability (5.7) and detection probability (5.8) can be simplified as follows

$$\begin{cases} P_{fa} = \mathcal{Q}\left(\frac{\frac{2\lambda}{\sigma_w^2} - 2N}{s\sqrt{N}}\right) \\ P_d = \mathcal{Q}\left(\frac{\frac{2\lambda}{\sigma_w^2(1+\rho)} - 2N}{s\sqrt{N}}\right) \end{cases} \tag{5.9}$$

where $\mathcal{Q}(u)$ denotes the Q-function associated with the standard Gaussian distribution. As indicated in equation(5.7), the energy detector threshold λ is a function of the noise variance σ_w^2 which is assumed to be estimated without error. The lack of knowledge of the exact noise power is called noise uncertainty [12]. In practical systems, the exact value of σ_w^2 is generally unavailable due to the fact that the total noise consists of noise from RF environment, non-linearity present in the RF front end, quantization noise, inherent time-varying thermal noise, *etc.* Thus, the energy detector performance degrades drastically in very low SNR regime with the emergence of the SNR wall problem [12, 13, 14, 15]. In following sections 5.2.1, 5.2.2, and 5.2.3, we give some of our contributions related to energy detection based spectrum sensing.

5.2.1 [C1]: Noise uncertainty analysis

In the work performed during the Master of Sara Bahamou, we investigated the relationship between the bounded noise uncertainty model proposed by Tandra et *al.* [12] and the unbounded model proposed by Jouini [13]. In the bounded uncertainty model, the exact noise power σ_w^2

lies in a finite range around the nominal noise power σ_r^2, and we assume that σ_w^2 is uniformly distributed in the dB domain as follows

$$\sigma_{w,dB}^2 \sim \mathcal{U}\left(\sigma_{r,dB}^2 - \alpha_{dB}, \sigma_{r,dB}^2 + \alpha_{dB}\right) \tag{5.10}$$

where $\sigma_{w,dB}^2 = 10\log_{10}(\sigma_w^2)$, $\sigma_{r,dB}^2 = 10\log_{10}(\sigma_r^2)$ and $\alpha_{dB} = 10\log_{10}(1+\alpha)$. In practical systems, the noise uncertainty factor of 1 dB to 2 dB is normally observed [16]. In the unbounded uncertainty model, the exact noise variance σ_w^2 does no longer belong to a finite range. In [13], author proposed a normal distribution to describe, in dB domain, the difference between the exact noise variance σ_w^2 and the nominal noise variance σ_r^2.

$$\delta_{dB} = \left(\sigma_{w,dB}^2 - \sigma_{r,dB}^2\right) \sim \mathcal{N}\left(0, \sigma_{\delta_{dB}}^2\right) \tag{5.11}$$

In our paper published in IEEE EUSIPCO conference 2013 [CI29], we have investigated the energy detector performance under bounded and unbounded noise uncertainty models. Moreover, we have obtained the relationship between both models using confidence interval. Finally, We proved that the Tandra's bounded model was a particular case of the unbounded noise uncertainty log-normal distribution approximation.

5.2.2 [C2]: Finite-sample size correlated multiple antennas energy detector

Still targeting energy detector in practical systems, in this section, we investigated the detection performances under spatial correlation between receive antennas and finite number of received samples conditions. Thus, we considered that the number of received antennas M is strictly greater than 1, and a small number of received samples N. We assumed that the received signal energy at different antennas are combined with equal gain.

In this context, by exploiting a result on the approximation of the distribution of linear combination of correlated chi-square random variables with an uncorrelated chi-square distribution [17], we derived closed-form expressions for the detection and the false alarm probabilities. We evaluated the accuracy of the proposed expressions through Matlab and correspondingly compared them with their closest related work, [18], while varying the number of samples N and the number of receive antennas M. This work was done in collaboration with my former postdoc Babar Aziz, and it was published in IEEE Wireless Communications Letters 2014 which is included hereafter [J8].

Closed-Form Approximation for the Performance of Finite Sample-Based Energy Detection Using Correlated Receiving Antennas

Amor Nafkha, *Member, IEEE*, and Babar Aziz, *Member, IEEE*

***Abstract*—In this letter, we present a new accurate closed-form approximation of the performance of a finite number of samples based on energy detector (ED) using multiple correlated receiving antennas. Herein, recent results on sum of correlated chi-squared random variables are employed to derive the probability density function of the ED statistic under both the noise only and the signal plus noise hypotheses. This allows us to derive accurate closed-form expressions of the detection and false alarm probabilities for finite sample-based ED using correlated receiving antennas. Simulation results are presented to validate the accuracy of the derived expressions.**

***Index Terms*—Cognitive radio, spectrum sensing, energy detection, correlated multiple antennas, performance analysis.**

I. Introduction

In cognitive radio networks, spectrum sensing is an important task to detect licensed users (primary users) and to prevent any harmful interference with them [1]. Various spectrum sensing techniques have been proposed including matched filter, energy detection and cyclostationary feature detection, as well as some emerging methods such as eigenvalue-based sensing, wavelet-based sensing, covariance-based sensing etc [2], [3]. In general, the performance analysis of a spectrum sensing technique is based on the detection and false alarm probabilities at different signal-to-noise ratio levels. The false alarm probability, P_{fa}, is defined as the probability that the detector declares the presence of primary user, when it is actually absent. The probability of detection, P_d, is defined as the probability that the cognitive radio correctly detects the presence of the primary user.

The matched filter is the optimal way to detect any signal. However, it requires perfect knowledge of primary user's signal. The simplest spectrum sensing technique is the energy detection which does not need any prior information about the primary user signal. The main drawback of the energy detector is its sensitivity to the noise power uncertainty [4]. As an alternative, cyclostationary feature detection shows good performance in low SNR scenarios, but it has a large amount of computational complexity which makes it unsuitable for real-time spectrum sensing. The main advantage of eigenvalue-based spectrum sensing is that it does not require any prior information about the primary user's signal and outperforms the energy detector technique especially in the presence of noise uncertainty. However, the main drawback of eigenvalue-based spectrum sensing techniques is the requirement of large numbers of samples and sensors, which are not practical in a real cognitive network. Recently, Zhang *et al.* [9] presented an interesting spectrum sensing technique by using the finite random matrix theory and by invoking novel results on exact distribution of standard condition number of dual Wishart random matrices. This sensing technique requires only a few samples and keeps the blindness feature intact. However, for a smaller number of receiving antennas $(M = 3, 4, 5 \ldots)$, to the best knowledge of the authors, no explicit form to derive the performance of finite random matrix based spectrum sensing exisits, neither for the correlated nor for the uncorrelated cases. In [5], the performance of the energy detector with multiple correlated antennas was analyzed for Rayleigh fading channel. Kim *et al.* [5] approximate the energy detector statistics by a Gaussian by invoking the central limit theorem (CLT). However, to reduce the detection time and increase the agility of the cognitive radios, only a finite number of signal samples may be available and the conditions for validity of the CLT may not hold. Here, we probe the question of what will be the expressions for P_d and P_{fa} when the number of received samples size is finite. Exploiting a recent result on the approximation of the distribution of linear combination of correlated chi-square random variables with an uncorrelated chi-square distribution, we are able to derive approximate expressions for the P_d and P_{fa} probabilities.

The main contribution of this letter is the derivation of new formulas for the detection and the false alarm probabilities when the cognitive radio employs finite-samples based energy detector under correlated receiving antennas. The rest of this letter is organized as follows. In Section II, the system model and some statistical results are introduced. Performance metrics and the theoretical analysis of the energy detector under spatial correlation constitute Section III. Section IV details the numerical results and Section V provides the conclusion.

Notation: Bold-face letters are used to denote matrices and vectors. Let us define the Hermitian of vector $\mathbf{x}$ as $\mathbf{x}^H$. An identity matrix of size L is denoted as $\mathbf{I}_L$. A superscript $[.]^T$ denotes the vector or matrix transpose operation. We use $\mathcal{CN}(\mu, \sigma^2)$ to represent the complex Gaussian distribution with mean μ and variance σ^2. $E[\mathbf{z}]$ and $V[\mathbf{z}]$ are the expectation and the variance of random variable $\mathbf{z}$. The $\sim$ symbol means "*distributed according to*." The chi-squared distribution with q degrees of freedom is denoted as χ^2_q.

II. System Model and Problem Formulation

In this letter, we suppose that unlicensed (secondary) user is equipped with M receiving antennas and each antenna receives N samples. We consider a linear array at the SU node with equally spaced antenna elements. We denote the hypotheses of the presence and absence of the primary user's signal by $\mathbb{H}_1$ and

Manuscript received May 10, 2014; revised August 3, 2014; accepted August 5, 2014. Date of publication August 18, 2014; date of current version December 17, 2014. The associate editor coordinating the review of this paper and approving it for publication was D. Hong.

A. Nafkha is with Supélec/IETR, Cesson-Sévigné Cedex 35576, France (e-mail: amor.nafkha@supelec.fr).

B. Aziz is with IFSTTAR, LEOST, Villeneuve d'Ascq 59666, France (e-mail: babar.aziz@ifsttar.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LWC.2014.2348998

$\mathbb{H}_0$, respectively. At the mth antenna, the received signals can be represented as

$$\begin{cases} \mathbb{H}_0 : \mathbf{y}_m(k) = \mathbf{w}_m(k) \\ \mathbb{H}_1 : \mathbf{y}_m(k) = \mathbf{h}_m(k)\mathbf{x}(k) + \mathbf{w}_m(k) \end{cases} \quad (1)$$

where $k = 1, \ldots, N,\ m = 1, \ldots, M,\ \mathbf{y}_m(k)$ is the received signal at the mth antenna of the secondary user, $\mathbf{h}_m(k) \sim \mathcal{CN}(0, \sigma_h^2)$ is the channel coefficient from the PU to the mth received antenna, $\mathbf{x}(k)$ denotes the complex phase-shift-keying (PSK) modulated signal transmitted by the primary user with the received signal power $\boldsymbol{P}$ at the mth antenna, and $\mathbf{w}_m(k) \sim \mathcal{CN}(0, \sigma_w^2)$ is the additive noise sample at mth antenna which is assumed to be zero-mean circular complex Gaussian random variable. $\mathbf{x}(k)$, $\mathbf{h}_m(k)$ and $\mathbf{w}_m(k)$ are assumed to be independent of each other. As proposed in paper [5], we assumed that the channel coefficients, $\mathbf{h}_m(k)$, are correlated between antennas but independent over time. The primary user signal, $\mathbf{x}(k)$, has an equal probability of transmission for each modulated symbol.

In this letter, the correlation in the wireless channel is modelled by an exponentially correlated model [7], and it is given as,

$$[\mathbf{h}_1(k), \ldots, \mathbf{h}_M(k)]^T = \mathbf{R}^{1/2} [\mathbf{g}_1(k), \ldots, \mathbf{g}_M(k)]^T \quad (2)$$

where $\{\mathbf{g}_1(k)\}_{m=1,\ldots,M}^{k=1,\ldots,N}$ has identically independent complex Gaussian distribution with zero mean and unit variance and the elements of $\mathbf{R}$ are represented through a single real correlation parameter ρ, which is the coefficient of correlation between two adjacent antenna array elements

$$\boldsymbol{R}_{ij} = \begin{cases} \rho^{i-j}, & i \le j \\ \boldsymbol{R}_{ji}^*, & i > j \end{cases} ; i, j = 1, \ldots, M \quad (3)$$

where * denotes the complex conjugate. This model is a physically-reasonable model in the sense that the correlation decreases as distance between antennas increases. The exponential model can approximate the correlation in a uniform linear array under rich scattering conditions [7].

III. Sensing Performance of Energy Detector

The secondary user senses the same primary user's signal using M receiving antennas and over N samples at each antennas. Then, the received signal matrix $\mathbf{Y}$ of size $M \times N$ can be written in terms of the $M \times N$ complex channel matrix $\mathbf{H}$ and the noise matrix $\mathbf{W}$ of the same size as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W} \quad (4)$$

where $\mathbf{H}$ is a zero mean Gaussian matrix with covariance matrix $\mathbf{R}$ and $\mathbf{W}$ is a zero mean Gaussian matrix of covariance $\sigma_w^2 \mathbf{I}_M$, respectively. Using vectorization operation,[1] one can rewrite

$$\mathbf{y} = \mathbf{h} \odot \mathbf{x} + \mathbf{w} \quad (5)$$

where $\odot$ denotes the Hadamard (element-by-element) product, $\mathbf{y} = [\mathbf{y}_1(1), \ldots, \mathbf{y}_M(1), \ldots, \mathbf{y}_1(N), \ldots, \mathbf{y}_M(N)]^T \in \mathbb{C}^{MN \times 1}$ is the observation vector, $\mathbf{h} = \mathrm{vec}(\mathbf{H})$, $\mathbf{x} = \mathrm{vec}(\mathbf{X})$, and $\mathbf{w} = \mathrm{vec}(\mathbf{W})$. $\boldsymbol{\Sigma_0}$ and $\boldsymbol{\Sigma_1}$ denote the covariance matrices of $\mathbf{y}$ under hypothesis $\mathbb{H}_0$ and $\mathbb{H}_1$, respectively. Therefore, we can write:

$$\begin{cases} \mathbb{H}_0 : \boldsymbol{\Sigma_0} = \mathbf{E}[\mathbf{y}\mathbf{y}^{\mathbf{H}} | \mathbb{H}_0] = \sigma_{\mathbf{w}}^{\mathbf{2}} \mathbf{I}_{\mathbf{MN}} \\ \mathbb{H}_1 : \boldsymbol{\Sigma_1} = \mathbf{E}[\mathbf{y}\mathbf{y}^{\mathbf{H}} | \mathbb{H}_1] = \boldsymbol{P}\sigma_h^2 \mathbf{I}_N \otimes \mathbf{R} + \sigma_w^2 \mathbf{I}_{MN} \end{cases} \quad (6)$$

where $\otimes$ denotes the Kronecker product. In this letter, the received signal energy at different antennas are combined with equal gain. Thus, when the N received signal samples over each of the M received antennas are available, the observed energy value at the secondary user is given by

$$\mathcal{T} = \sum_{i=1}^{N} \sum_{m=1}^{M} |\mathbf{y}_m(i)|^2 = \sum_{i=1}^{N} \mathcal{T}_i \quad (7)$$

where $\mathcal{T}_i = \sum_{m=1}^{M} |\mathbf{y}_m(i)|^2$ denotes the partial test statistics. Then, under the spatial correlated antennas, we have the following theorem for the energy statistic.

Theorem 1: Consider a secondary user equipped with M receive antennas where each of them collects N samples. Let $\mathbf{R} \in \mathbb{C}^{M \times M}$ denotes the receive antenna correlation matrix. Then, the hypothesis testing for energy statistic (7) is approximated as

$$\begin{cases} \mathbb{H}_0 : \mathcal{T} & \sim \sigma_w^2 \chi_{MN}^2 \\ \mathbb{H}_1 : \mathcal{T} & \sim \frac{\sum_{k=1}^{M} \left(P\sigma_h^2 \lambda_k + \sigma_w^2\right)^2}{2M\left(P\sigma_h^2 + \sigma_w^2\right)} \chi_{Nf}^2 \end{cases} \quad (8)$$

where $\{\lambda_k\}_{k=1}^{M}$ are the eigenvalues of the matrix $\mathbf{R}$, and the parameter $f = (2M^2(P\sigma_h^2 + \sigma_w^2)^2)/(\sum_{k=1}^{M}(P\sigma_h^2\lambda_k + \sigma_w^2)^2)$.

Proof: Please see the Appendix ■

The energy detection outcome can be seen as the output of decision policy π that maps the received energy $\mathcal{T}$ into a binary value $d = \pi(\mathcal{T}),\ d \in \{0, 1\}$. To evaluate the performance of the decision policy π under the binary hypothesis test, we can define the following probabilities:

$$\begin{cases} P_{fa} = \dfrac{\Gamma\left(\frac{MN}{2}, \frac{\tau}{2\sigma_w^2}\right)}{\Gamma\left(\frac{MN}{2}\right)} \\ P_d = \dfrac{\Gamma\left(\frac{Nf}{2}, \frac{\tau M\left(P\sigma_h^2 + \sigma_w^2\right)}{\sum_{k=1}^{M}\left(P\sigma_h^2\lambda_k + \sigma_w^2\right)^2}\right)}{\Gamma\left(\frac{Nf}{2}\right)} \end{cases} \quad (9)$$

where $\Gamma(k)$ is the gamma function evaluated at k and $\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t}\, dt$ is the upper incomplete gamma function. For Neyman-Pearson detection with a fixed false alarm probability $P_{fa} = \nu$, we can choose τ as follows:

$$\tau = \mathcal{F}^{-1}\left(1 - \nu, \frac{MN}{2}, 2\sigma_w^2\right) \quad (10)$$

where $\mathcal{F}^{-1}(1 - \nu, ((MN)/2), 2\sigma_w^2)$ is the gamma inverse cumulative distribution function with shape parameter $(MN)/2$ and scale parameter $2\sigma_w^2$ for the probability $1 - \nu$.

IV. Numerical Results

In this section, we analyze the proposed approximation through Matlab simulations. We compare our results with the approximation method proposed in [5]. We study these approximations for correlated and uncorrelated antennas cases and for various number of received samples. Simulation parameters are

[1] If $\mathbf{B}$ is an $M \times N$ matrix, $\mathrm{vec}(\mathbf{B})$ is the vector $[\mathbf{b}_1; \mathbf{b}_2; \ldots; \mathbf{b}_N]$, where $\mathbf{b}_i$ is the ith column of $\mathbf{B}$ and ; denotes change of row

TABLE I
SIMULATION PARAMETERS

Parameters	Value
The angular spread (Λ), the correlated antenna array elements	$0.5°$
The angular spread (Λ), the uncorrelated antenna array elements	$180°$
Distance between antenna elements (d), the correlated antenna array elements	$d = \lambda/8$
Distance between antenna elements (d), the uncorrelated antenna array elements	$d = \lambda/2$
Number of antenna array elements (M)	2 to 10
SNR [dB]	-12 to 0
Coefficient of correlation (ρ)	0 to 1
Channel parameter (σ_h^2)	1
Number of samples per antenna (N)	10, 20, 50, 100, 200

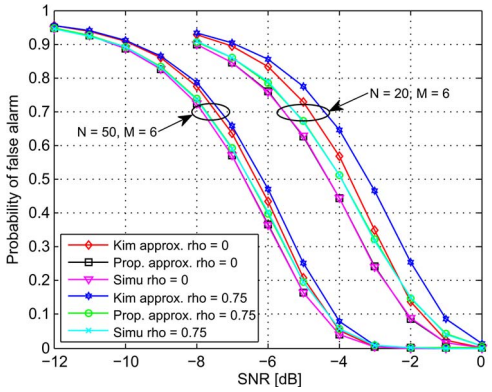


Fig. 1. Probability of false alarm plotted against SNR values.

established according to IEEE 802.22 [10] (see Table I). In our simulations, we consider that at least 100 samples of received signal are available for the decision process [11], [12]. In Fig. 1, the probability of false alarm P_{fa} is plotted as a function of SNR when the antenna array elements are correlated with the correlation coefficient $\rho = 0$ and $\rho = 0.75$. The P_{fa} results are plotted for $P_d = 0.99$. The number of samples per antenna element are $N = 20$ and $N = 50$ and the number of antenna array elements is $M = 6$. It is observed that our proposed approximation is closer to the Matlab simulations compared to the Kim approximation method proposed in [5]. This is true for both uncorrelated and correlated antenna elements case. As N is increased from 20 samples to 50 samples, Kim's approximation approaches the simulation results. This is because it is based on the Central Limit Theorem (CLT) and, therefore, the accuracy of Kim's approximation is improved when using a larger N. However, even with $N = 50$ samples available per antenna, a significant difference between [5] and Matlab simulations can be observed. On the other hand, our proposed approximation matches the Matlab simulation P_{fa} curves for $N = 20$ and 50.

Fig. 2 presents the comparison of the two approximations in terms of probability of correct detection (P_d) is plotted against varying SNR values for $P_{fa} = 0.001$. Results are plotted for $\rho = 0$ and $\rho = 0.75$. Note that we fixed the number of receiving antennas $M = 6$ and we increase the number of received samples per antenna from $N = 10$ to $N = 200$. We observe that the P_d curves obtained using proposed analytical match the Matlab simulation results for all values $N = 20$. Also, we can note that the Kim approximation error decays with the increase of N, and it will entirely vanish to zero which can be explained by the CLT.

In Fig. 3 we examine the approximations in terms of the receiver operating characteristics (ROC) curves. The ROC

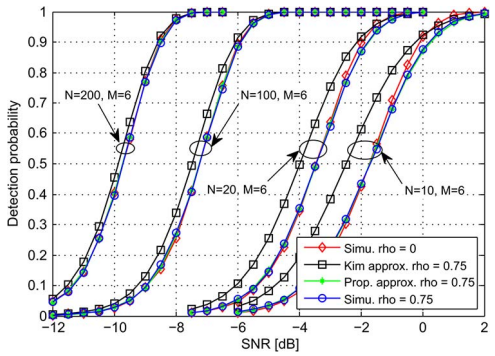


Fig. 2. Probability of correct detection plotted against SNR values.

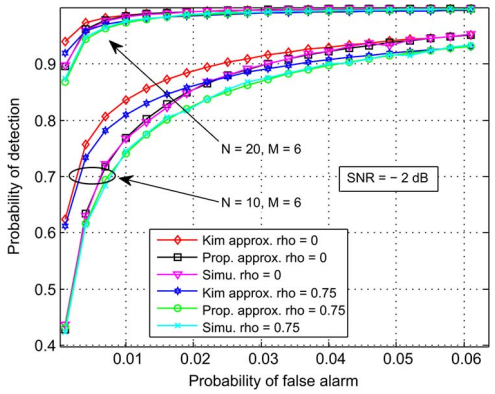


Fig. 3. Receiver operating characteristics (ROC) curves for the two approximations at SNR $= -2$ dB.

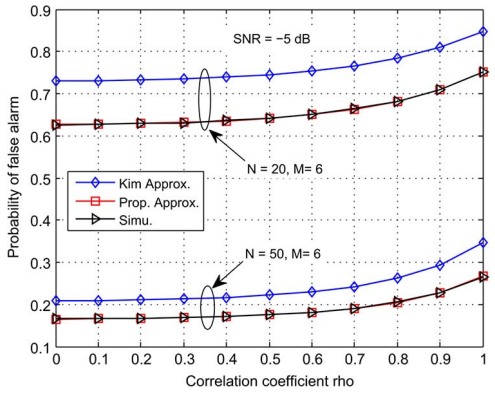


Fig. 4. Probability of false alarm P_{fa} plotted against correlation coefficient ρ values for SNR $= -5$ dB and $P_d = 0.99$.

curves are plotted for SNR $= -2$ dB, $\rho = 0$, and $\rho = 0.75$. It is observed that our approximation is very close to the Matlab simulation results. However, the Kim approximation deviates from the simulated results. Furthermore, it is observed that the difference between the simulation results and approximation in [5] increases as N is decreased from 20 to 10 samples per antenna. This confirms that the approximation in [5] is no longer valid especially when the number of samples is very small, the CLT does not hold. Finally, in Fig. 4, we plot the probability of false alarm against varying correlation coefficient ρ for SNR $= -5$ dB and $P_d = 0.99$. For a given SNR and P_d, P_{fa} increases with increase in correlation coefficient ρ. It is observed that our approximation matches the simulation results for all values of ρ for both $N = 20$ and 50. However, a significant difference exists between the simulated results and Kim approximation [5].

V. Conclusion

In this letter, we have presented a new accurate closed-form approximation of the performances of the energy detector under multiple correlated antennas and finite number of received samples conditions. We derived accurate closed-form expressions of the probability of detection (P_d) and false alarm (P_{fa}) for finite sample-based energy detector using correlated two receiving antennas. We have shown that the simulation results validate the accuracy of the derived expressions.

Appendix

First, we introduce the following two lemmas from Hou [6] and Alouini [8], which will be used in the proof of Theorem 1

Lemma 1 (Alouini et al. [8]): Let $\{\mathbf{z}_k\}_{k=1}^{M}$ be a set of M correlated not necessarily identically distributed gamma variates with shape and scale parameters α and β_k, respectively, i.e., $\mathbf{z}_k \sim \gamma(\alpha, \beta_k)$, and let r_{ij} denote the correlation coefficient between $\mathbf{z}_i$ and $\mathbf{z}_j$, then the moment generating function (MGF) of $\mathbf{u} = \sum_{k=1}^{M} \mathbf{z}_k$, $\mathcal{M}_{\mathbf{u}}(s) = E[e^{s\mathbf{u}}]$, can be expressed as

$$\mathcal{M}_{\mathbf{u}}(s) = \prod_{k=1}^{M} (1 - s\lambda_k)^{-\alpha} \tag{11}$$

where $\{\lambda_k\}_{k=1}^{M}$ are the eigenvalues of the matrix $\mathbf{A} = \mathbf{DC}$. $\mathbf{D}$ is the $M \times M$ diagonal matrix with the entries $\{\beta_k\}_{k=1}^{M}$ and $\mathbf{C}$ is the $M \times M$ positive definite matrix defined by

$$\mathbf{C} = \begin{bmatrix} 1 & \sqrt{r_{12}} & \cdots & \sqrt{r_{1M}} \\ \sqrt{r_{21}} & 1 & \cdots & \sqrt{r_{2M}} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \sqrt{r_{M1}} & \cdot & \cdots & 1 \end{bmatrix}$$

Lemma 2 (Hou [6]): Let $\{\mathbf{v}_k\}_{k=1}^{M}$ be a set of M correlated chi-square random variables with 2 degrees of freedom and $\{\omega_k\}_{k=1}^{M}$ are positive constants. By matching the first two moments, The distribution of the random variable $\mathcal{R} = \sum_{k=1}^{M} \omega_k \mathbf{v}_k$ can be accurately approximated by that of $c.\chi_f^2$, where χ_f^2 is a chi-square variate with f degrees of freedom. It follows that:

$$c = \frac{V[\mathcal{R}]}{2E[\mathcal{R}]}, \text{ and } f = \frac{2E[\mathcal{R}]^2}{V[\mathcal{R}]} \tag{12}$$

The value of f is often not an integer, so the χ_f^2 distribution is actually a gamma distribution $\gamma(\alpha, \beta)$ with shape and scale parameters $\alpha = f/2$ and $\beta = 2$, respectively.

We can now prove Theorem 1

Proof: Under the hypothesis $\mathbb{H}_0$, for a given $i \in \{1, 2, \ldots, N\}$, the partial test statistics $\mathcal{T}_i$ is the sum of the squares of M Gaussian random variables with zero mean and variance σ_w^2. Then, under the assumption of independent and identically distributed (i.i.d.) noise components, the distribution of $\mathcal{T}_i$ is the central chi-square distribution $\sigma_w^2 . \chi_M^2$. For the energy detector, the test statistics $\mathcal{T}$ is the sum of N independent partial test statistics. Hence, the distribution of $\mathcal{T}$ follows a central chi-square distribution $\sigma_w^2 . \chi_{MN}^2$ with MN degrees of freedom. Under the hypothesis $\mathbb{H}_1$, the partial test statistic $\mathcal{T}_i$ is the sum of M correlated chi-square random variables each have 2 degrees of freedom. Using the fact that the chi-square distribution is a special case of the gamma distribution, $\mathcal{T}_i$ follows the distribution of the sum of M correlated gamma random variates. From Lemma 1, for a given $i \in \{1, 2, \ldots, N\}$, the mean and the variance of the partial test statistic $\mathcal{T}_i$ can be expressed as

$$E[\mathcal{T}_i|\mathbb{H}_1] = M\left(P\sigma_h^2 + \sigma_w^2\right) \tag{13}$$

$$V[\mathcal{T}_i|\mathbb{H}_1] = \sum_{k=1}^{M} \left(P\sigma_h^2\lambda_k + \sigma_w^2\right)^2 \tag{14}$$

Then, by employing Lemma 2, (13) and (14), we can approximate each of partial test statistic $\mathcal{T}_i,\ i \in \{1, 2, \ldots, N\}$, under the hypothesis $\mathbb{H}_1$ by $c\chi_f^2$, where c and f are obtained by

$$c = \frac{\sum_{k=1}^{M} \left(P\sigma_h^2\lambda_k + \sigma_w^2\right)^2}{2M\left(P\sigma_h^2 + \sigma_w^2\right)} \tag{15}$$

$$f = \frac{2M^2\left(P\sigma_h^2 + \sigma_w^2\right)^2}{\sum_{k=1}^{M} \left(P\sigma_h^2\lambda_k + \sigma_w^2\right)^2} \tag{16}$$

$\{\mathcal{T}_i\}_{i=1}^{N}$ are independent and identically distributed random variables that follow $c\chi_f^2$ distribution. Accordingly, under $\mathbb{H}_1$, the energy detector's test statistic, $\mathcal{T} = \sum_{i=1}^{N} \mathcal{T}_i$, is distributed as chi-square distribution, $c\chi_{Nf}^2$ with Nf degrees of freedom. The variables c and f are given in (15) and (16). This completes the proof. ■

References

[1] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proc. IEEE*, vol. 97, no. 5, pp. 849–877, May 2009.

[2] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101–116, May 2012.

[3] Y. Zeng, Y. C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing techniques for cognitive radio: Challenges and solutions," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, pp. 381 465-1–381 465-15, Jan. 2010.

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

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

[6] C. D. Hou, "A simple approximation for the distribution of the weighted combination of nonindependent or independent probabilities," *Stat. Probab. Lett.*, vol. 73, no. 2, pp. 179–187, Jun. 2005.

[7] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sep. 2001.

[8] M.-S. Alouini, A. Abdi, and M. Kaveh, "Sum of gamma variates and performance of wireless communication systems over Nakagami-fading channels," *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, pp. 1471–1480, Nov. 2001.

[9] W. Zhang, G. Abreu, M. Inamori, and Y. Sanada, "Spectrum sensing algorithms via finite random matrices," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 164–175, Jan. 2012.

[10] IEEE 802.22 Working Group on Wireless Regional Area Networks (WRAN). [Online]. Available: http://grouper.ieee.org/groups/802/22

[11] G. Noh *et al.*, "Throughput analysis and optimization of sensing-based cognitive radio systems with Markovian traffic," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 4163, 4169, Oct. 2010.

[12] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326, 1337, Apr. 2008.

5.2.3 [C3]: Hybrid spectrum sensing architecture

To overcome the noise uncertainty problem of energy detector and to increase the detection performance, we proposed a hybrid spectrum sensing architecture which consists of double threshold energy detector (DTED) and cyclostationary feature detector (CFD) as depicted in Fig 5.2. This new architecture has significantly improved the performance of classical energy detector by introducing two thresholds, λ_0 and λ_1, that are used to define a "*no decision*" region. The use of lower and higher thresholds allows the probabilities of false alarm and missed detection to be set arbitrarily low at the cost of an increased uncertainty region. If the value of the test statistic falls within the no-decision region, then no decision is made by the energy detector, and the adaptation block orients the received input samples to a secondary detector, called cyclostationary feature detector, in order to make a decision on the presence or absence of primary user's signal. The process of the proposed hybrid detector involves a successive adjustment of thresholds λ_0 and λ_1, through which an improved energy detection scheme could be achieved. Thus, if the final decision is that the primary channel is idle, then the lower threshold λ_0 will be updated through the adaptation block, otherwise the hybrid spectrum sensing detector updates the value of the higher threshold λ_1.

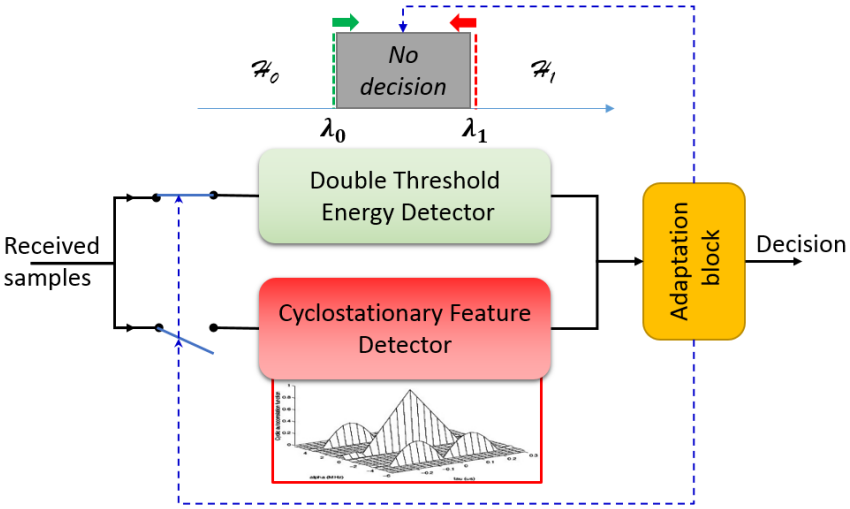


Figure 5.2 – Block diagram of hybrid spectrum sensing detector.

Given the decision of the CFD detector, two methods have been proposed to update the lower and higher threshold values. In the first method, the lower and higher thresholds are taken as the averages of buffer $\mathcal{B}_0$ and $\mathcal{B}_1$ values, respectively. In the second method, the last order statistic of buffer $\mathcal{B}_0$ and the first order statistic of buffer $\mathcal{B}_1$ are used to update the lower

and the higher threshold respectively.

Algorithm 1: Average-based Hybrid spectrum sensing detector algorithm

Result: Final decision $\mathcal{H}_0$ or $\mathcal{H}_1$
Initialize: $\lambda_0 \leftarrow 0$, $\lambda_1 \leftarrow +\infty$, and the desired false alarm rate P_{fa};
Create two circular buffers: $\mathcal{B}_0$ and $\mathcal{B}_1$;
while *True* **do**
  Calculate $\mathcal{T}(\mathbf{y})$;
  if $\mathcal{T}(\boldsymbol{y}) \in [\lambda_0, \lambda_1]$ **then**
    Get the **Final decision** using the CFD detector;
    if ***Final decision*** $= \mathcal{H}_0$ **then**
      Save $\mathcal{T}(\mathbf{y}) \in \mathcal{B}_0$;
      Update: λ_0 as the mean of buffered values in $\mathcal{B}_0$;
    else
      Save $\mathcal{T}(\mathbf{y}) \in \mathcal{B}_1$;
      Update: λ_1 as the mean of buffered values in $\mathcal{B}_1$;
    end
  else
    IF $\mathcal{T}(\mathbf{y}) \geq \lambda_1$, then **Final decision** $= \mathcal{H}_1$;
    IF $\mathcal{T}(\mathbf{y}) \leq \lambda_0$, then **Final decision** $= \mathcal{H}_0$;
  end
end

Algorithm 2: Order-Statistic-based Hybrid spectrum sensing detector algorithm

Result: Final decision $\mathcal{H}_0$ or $\mathcal{H}_1$
Initialize: $\lambda_0 \leftarrow 0$, $\lambda_1 \leftarrow +\infty$, and the desired false alarm rate P_{fa};
Create two circular buffers: $\mathcal{B}_0$ and $\mathcal{B}_1$;
while *True* **do**
  Calculate $\mathcal{T}(\mathbf{y})$;
  if $\mathcal{T}(\boldsymbol{y}) \in [\lambda_0, \lambda_1]$ **then**
    Get the **Final decision** using the CFD detector;
    if ***Final decision*** $= \mathcal{H}_0$ **then**
      Save $\mathcal{T}(\mathbf{y}) \in \mathcal{B}_0$;
      Update: λ_0 as the last order statistic of the buffer $\mathcal{B}_0$;
    else
      Save $\mathcal{T}(\mathbf{y}) \in \mathcal{B}_1$;
      Update: λ_1 as the first order statistic of the buffer $\mathcal{B}_1$;
    end
  else
    IF $\mathcal{T}(\mathbf{y}) \geq \lambda_1$, then **Final decision** $= \mathcal{H}_1$;
    IF $\mathcal{T}(\mathbf{y}) \leq \lambda_0$, then **Final decision** $= \mathcal{H}_0$;
  end
end

The first method is used by the average-based hybrid spectrum sensing detector (AHSD) as shown in Algorithm 1, while the second one is used by the order-statistic-based hybrid spectrum sensing detector (OSHSD) as depicted in Algorithm 2. The AHSD technique has been enhanced through the use of buffer $\mathcal{B}_0$ to estimate the noise variance σ_w^2, this variant will be referred to as EAHSD. The performances of the proposed hybrid spectrum sensing detectors (*i.e* AHSD and EAHSD), ideal energy detector (IED), and the cyclostationary feature detector (CFD) are extensively evaluated using Matlab. We considered that the probabilities that the primary transmission is active and inactive are equal, *i.e.*, $P(\mathcal{H}_0) = P(\mathcal{H}_1)$. A band pass quadrature phase-shift keying (QPSK) with cyclic frequency equal to $\zeta = 1/T_0$, where T_0 is the symbol period, is adopted as the primary user signal. If the value of the cyclic autocorrelation function of the signals received by the secondary user at the cyclic frequency ζ is nonzero, then the CFD detector decide that the channel is busy (*i.e.* $\mathcal{H}_1$); otherwise, the channel is idle (*i.e.* $\mathcal{H}_0$). We set the size of buffer $\mathcal{B}_0$ (resp. $\mathcal{B}_1$) to a size of 100 (resp. 30). Fig. 5.3 depicts the probability of detection versus signal to noise ratio of proposed schemes, CFD and IED with $M = 1$ and $N = 4500$ received samples. Regarding performance, it is clear from the plot that EAHSD-based sensing outperforms CFD sensing and initial AHSD sensing, however, it shows a loss of 3 dB when compared to the ideal energy detector scheme. Typically, a loss of 3 dB in the detection performance of ideal energy detector corresponds to a noise uncertainty of 0.5 dB, which can be considered as a very small value in practical systems.

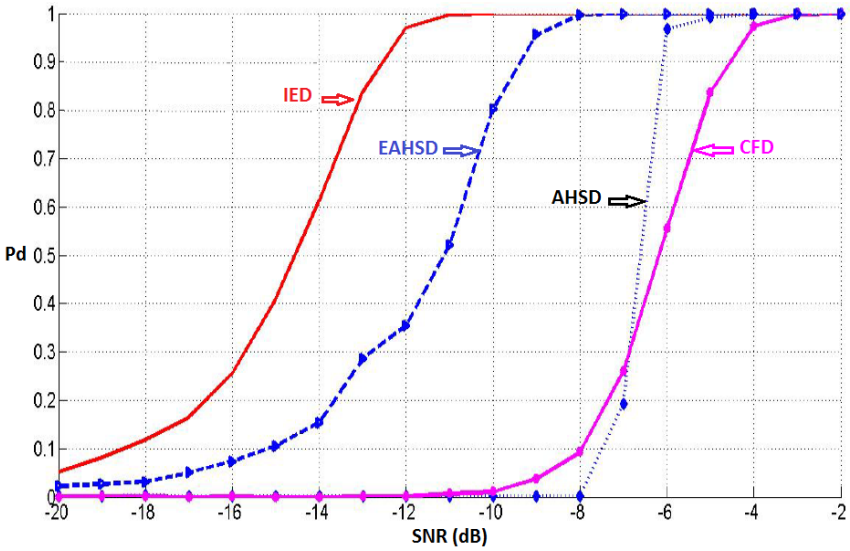


Figure 5.3 – Detection performance for IED, CFD, AHSD and EAHSD detectors with $P_{fa} = 0.01$.

This work was carried out in the context of Ziad Khalaf's thesis. Moreover, it received the best paper award at the sixth Advanced International Conference on Telecommunications in 2010 [CI19]. Interested readers can be referred to our journal paper [J8] and references therein.

5.3 Cyclostationary feature detection

Given that the energy detector method can be highly sensitive to unknown and changing noise level, cyclostationary feature detection was first introduced in [19]. The cyclostationarity feature of primary signals may be produced by modulation format or coding scheme or simply introduced in order to facilitate channel estimation or synchronization [20, 21, 22]. As in most communication systems, the additive noise is considered as wide sense stationary random process with no correlation, while modulated signals are cyclostationary with spectral correlation due to embedded spectral redundancy of signal periodicities, cyclostationary feature detector can be utilized to differentiate primary user's signal from noise [23, 24, 25]. Let us recall the definition of the cyclic autocorrelation function of a given received signal $\mathbf{y}$.

Cyclic Autocorrelation Function

Definition 1 *A discrete time zero-mean cyclostationary process, $\boldsymbol{y}$, is characterized by the property that its time-varying autocorrelation $r_{\boldsymbol{yy}}(t,\tau) = \mathbb{E}_{\boldsymbol{y}}\left[\boldsymbol{y}(t)\boldsymbol{y}^{\dagger}(t+\tau)\right]$ exhibits periodicity, T_0 called the cyclic period, in time domain. Thus, the Fourier series expansion with respect to time t is given as [23]:*

$$r_{\boldsymbol{yy}}(t,\tau) = \sum_{\alpha\in\mathcal{A}_\alpha} R_{\boldsymbol{yy}}(\alpha,\tau)e^{j2\pi\alpha t} \tag{5.12}$$

where $\mathcal{A}_\alpha = \{\alpha = k/T_0, k\in\mathbb{Z}\}$, and the Fourier coefficient $R_{\boldsymbol{yy}}(\alpha,\tau)$, also called the cyclic autocorrelation function, is expressed as:

$$R_{\boldsymbol{yy}}(\alpha,\tau) = \lim_{T\to+\infty}\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} r_{\boldsymbol{yy}}(t,\tau)e^{-j2\pi\alpha t}dt \tag{5.13}$$

A discrete cyclic autocorrelation function of discrete time signal $\mathbf{y}(n)$ with a fixed lag $\tau\in\mathbb{Z}/\{0\}$ is defined in the similar manner as (5.13)

$$R_{\mathbf{yy}}(\alpha,\tau) = \lim_{N\to+\infty}\frac{1}{N}\sum_{k=0}^{N-1}\mathbf{y}(k)\mathbf{y}^{*}(k+\tau)e^{-j2\pi\alpha k} \tag{5.14}$$

where N is the number of observation. In practical, given finite length of received data, we used the unbaised consistent estimator of the cyclic autocorrelation function, which is defined as:

$$\hat{R}_{\mathbf{yy}}(\alpha,\tau) \triangleq \frac{1}{N}\sum_{k=0}^{N-1}\mathbf{y}(k)\mathbf{y}^{*}(k+\tau)e^{-j2\pi\alpha k} \tag{5.15}$$

In the sequel of this section, we assumed that $\mathbf{y}$ has zero mean and it was sufficiently over-sampled (*i.e.* $f_s \geq 2\theta B$, where B is the the mono-lateral bandwidth, θ is the order of cyclo-stationarity of the received signal and f_s is frequency sampling rate). Now, for cyclostationary feature based spectrum sensing, the binary hypothesis test problem can be expressed as:

$$\begin{cases} \mathcal{H}_0: & \hat{R}_{\mathbf{yy}}(\alpha,\tau) \approx 0; \quad \forall \alpha \notin \mathcal{A} \text{ and } \tau \in \{1,\ldots,N\} \\ \mathcal{H}_1: & \hat{R}_{\mathbf{yy}}(\alpha,\tau) \neq 0; \quad \forall \alpha \in \mathcal{A} \text{ and for some } \tau \in \{1,\ldots,N\} \end{cases} \tag{5.16}$$

Therefore, the binary hypothesis test, given in (5.16), turns to a problem of examining whether $\hat{R}_{\mathbf{yy}}(\alpha,\tau)$ is small value or not for appropriately selected α and τ. Let us consider a N-dimensional vector $\mathbf{u}_\tau = [\mathbf{u}_\tau(0),\ldots,\mathbf{u}_\tau(N-1)]^T$ with entries $\mathbf{u}_\tau(k) = \mathbf{y}(k)\mathbf{y}^*(k+\tau)$, then the cyclic auto-correlation function (5.15) can be reformulated as follows

$$\hat{R}_{\mathbf{yy}}(\alpha,\tau) = \frac{1}{N}\sum_{k=0}^{N-1} \mathbf{u}_\tau(k) e^{-j2\pi\alpha k} \tag{5.17}$$

Now, let us fix the value of $\tau = \tau_0$ and define a N-dimensional vector $\hat{R}^{(\tau_0)}_{\mathbf{yy}}(\alpha)$ as follows:

$$\hat{R}^{(\tau_0)}_{\mathbf{yy}}(\alpha) = [\hat{R}_{\mathbf{yy}}(-\alpha_m,\tau_0), \hat{R}_{\mathbf{yy}}(-\alpha_m+\delta,\tau_0),\ldots,\hat{R}_{\mathbf{yy}}(\alpha_m-\delta,\tau_0), \hat{R}_{\mathbf{yy}}(\alpha_m,\tau_0)]^T, \tag{5.18}$$

where $\delta = 2\alpha_m/N$, α_m is the maximum observed cyclic frequency, then the equation (5.18) is nothing but a scaled version of the discrete Fourier transform of the vector $\mathbf{u}_{\tau_0}$ (*i.e.* $\hat{R}^{(\tau_0)}_{\mathbf{yy}}(\alpha) = (1/N)\mathcal{D}\mathbf{u}_{\tau_0}$, where $\mathcal{D}$ is the discrete Fourier transform matrix [26]).

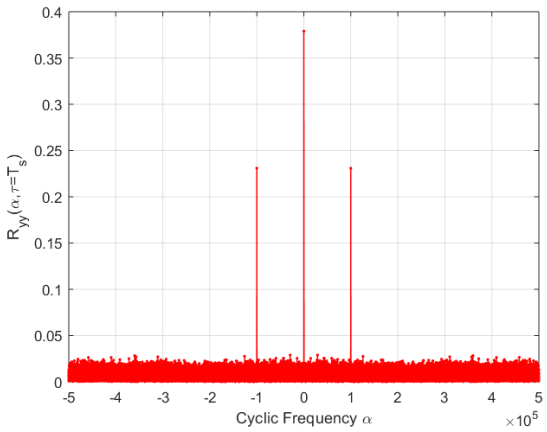


Figure 5.4 – Magnitude of cyclic auto-correlation function $\hat{R}^{(\tau)}_{\mathbf{yy}}(\alpha)$ at $\tau_0 = T_s$ for a BPSK modulated signal and $SNR = -3$ dB where T_s is the sampling period.

We denote that $\hat{R}^{(\tau_0)}_{\mathbf{yy}}(\alpha)$ is an N-dimensional vector with a specific structure depending on the null or alternative hypothesis. In Figure 5.4, we depict the magnitude of the cyclic auto-

correlation of a BPSK modulated signal with the fundamental cyclic frequency $\alpha = 1/T_s$ where T_s refers to the symbol period of the BPSK. Under alternative hypothesis $\mathcal{H}_1$, the cyclic auto-correlation is obtained using $N = 4096$ received samples. We denote that for $\alpha = 0$, we have the classical classical auto-correlation of the received samples. Under null hypothesis $\mathcal{H}_0$, the magnitude of the cyclic auto-correlation function is shown in Figure 5.5.

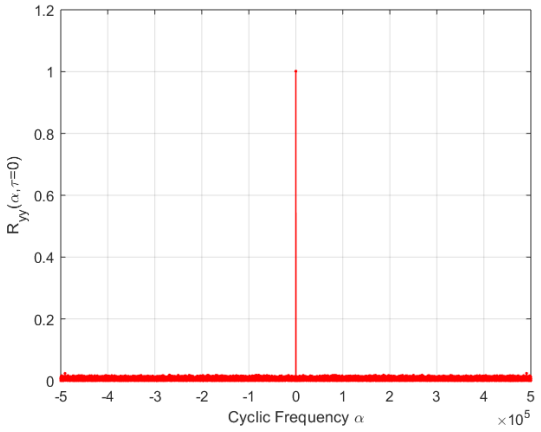


Figure 5.5 – Magnitude of cyclic auto-correlation function $\hat{R}^{(\tau)}_{\mathbf{yy}}(\alpha)$ at $\tau_0 = 0$ for noise only.

We can remark that, under both hypothesis, almost all entries of the cyclic auto-correlation vector $\hat{R}^{(\tau)}_{\mathbf{yy}}(\alpha)$ are very close to zeros. Thus, $\hat{R}^{(\tau)}_{\mathbf{yy}}(\alpha)$ can be estimated through a sparse reconstruction techniques in the cyclic frequency domain using compressive sensing theory.

Let us consider a n-dimensional vector $\mathbf{u}_{\tau_0,n} = [\mathbf{u}_{\tau_0}(0), \ldots, \mathbf{u}_{\tau 0}(n-1)]^T$ and an $n \times N$ matrix $\mathcal{A}$ which is the sub-matrix of the discrete Fourier transform matrix made of the n upper rows of $\mathcal{D}$ with $n << N$. In order to reconstruct $\hat{R}^{(\tau_0)}_{\mathbf{yy}}(\alpha)$, the main idea is to solve the following problem knowing that the N-dimensional cyclic auto-correlation vector is a sparse vector in the cyclic frequency domain.

$$\textbf{P1:} \qquad \mathbf{u}_{\tau_0,n} = \mathcal{A}^* \hat{R}^{(\tau_0)}_{\mathbf{yy}}(\alpha) \tag{5.19}$$

The above problem **P1** is a linear inverse problem with sparseness constraint, the equivalent sparse problem is given by:

$$\textbf{P2:} \qquad \min \|\hat{R}^{(\tau_0)}_{\mathbf{yy}}(\alpha)\|_0, \quad \text{s.t.} \quad \mathbf{u}_{\tau_0,n} = \mathcal{A}^* \hat{R}^{(\tau_0)}_{\mathbf{yy}}(\alpha) \tag{5.20}$$

In the following, we will use orthogonal matching pursuit (OMP) algorithm to find a sparse solution of problem **P2**. The OMP algorithm is a greedy compressed sensing recovery algorithm which selects the best fitting column of the sub-matrix $\mathcal{A}^*$ in each iteration. In the next two sections, we will present two blind cyclostationary feature detection methods, which have been proposed in the context of the PhD thesis of Ziad Khalaf.

5.3.1 [C4]: Successive-difference based cyclostationnary feature defector

Using the fact that the cyclic autocorrelation function of a linearly modulated signal is a sparse function in the cyclic frequency domain, we proposed a new blind spectrum sensing detector, which is based on the fact that two estimated cyclic auto-correlation vectors of two successive packets of samples have almost the same cyclic frequencies. The main advantages of the new proposed spectrum sensing detector compared to the classical cyclo-stationary feature detector developed by Dandawate and Giannakis in [23] are: **(i)** the limited number of received samples used in the CAF estimation process (*i.e.* $n << N$), **(ii)** there is no prior knowledge needed (*e.g.* cyclic frequencies), **(iii)** higher performance and lower complexity. This work was firstly published as an invited paper in IEEE MWSCAS 2011 [CI24]. An enhanced version of the proposed spectrum sensing algorithm has been published in GLOBECOM 2011 [CI25].

5.3.2 [C5]: Symmetry property based cyclostationary feature detection

Let us recall from [27] that the cyclic auto-correlation function $R_{\mathbf{yy}}(\alpha, \tau)$ presents the following symmetry properties:

$$R_{\mathbf{yy}}(\alpha, -\tau) = R_{\mathbf{yy}}(\alpha, \tau) \quad \text{and} \quad R_{\mathbf{yy}}(-\alpha, \tau) = R^*_{\mathbf{yy}}(\alpha, \tau) \tag{5.21}$$

By taking the complex magnitude of (5.21), we get the equality $\|R_{\mathbf{yy}}(-\alpha, \tau)\| = \|R_{\mathbf{yy}}(\alpha, \tau)\|$. Motivated by our previous contribution [C4], we proposed a novel spectrum sensing technique based on the symmetric property (5.21) of the complex magnitude of cyclic auto-correlation function. The main idea is to develop a simple scheme to detect the presence or not of the symmetry property, with respect to the axis $\alpha = 0$ for a given lag $\tau = \tau_0$, of the the estimated sparse CAF vector. Thus, if the estimated CAF, reconstruction done through limited iteration numbers of OMP algorithm, verifies approximately (5.21) then the alternative hypothesis $\mathcal{H}_1$ is declared, otherwise the null hypothesis $\mathcal{H}_0$ is true. It is important to note that the probability of obtaining a sparse CAF vector with a symmetry property under $\mathcal{H}_0$ is very low ($< 1/N$). It is shown by simulations that the proposed detection algorithm outperforms other existing approaches with a limited computation complexity. Further details on the theoretical framework and simulation performance are published in CROWNCOM 2012 [CI26] and ICT 2018 [CI50].

5.4 Eigenvalue based spectrum sensing techniques

The eigenvalue-based spectrum sensing techniques were mainly investigated in the context of the PhD thesis of Hussein Kobeissi (2013-2016) in collaboration with Lebanese University and American University of Beirut.

Technique based	Symmetry	Successive-difference	Classical[23]
Complexity	$\mathcal{O}(nMN)$	$\mathcal{O}(npMN)$	$\mathcal{O}(MNL+4ML^2)$
Samples used	100	1255	3170
$\log(Operations)$	13.55	16.08	13.45
Prior info.	none	none	Cyclic frequency α

Table 5.1 – Computation complexity comparison among different cyclostationary feature detectors for a desired (P_{fa},P_d) point of (0.1,0.9) at $SNR = 0$ dB

5.4.1 [C6]: Eigenvalue based detection in finite and asymptotic dimensions

In a first effort, we consider the performance probabilities and the decision threshold of the standard condition number (SCN) based detector in finite and asymptotic cases and we discussed the statistics of the SCN decision metric. Main contributions can be summarized as:

- We generalized the joint distribution of the central semi-correlated Wishart matrices to cover equal eigenvalues of the correlation matrix. Moreover, we studied the impact of system parameters on non-central/central approximation.
- For finite dimension, we derived an exact distribution of the SCN distribution and we also provided a close simple approximation of the SCN distribution through the use of the generalized extreme value (GEV) distribution and moments matching techniques.
- For asymptotic dimension, we proposed the GEV distribution as the SCN distribution and approximated the SCN moments.
- We derived new closed-form expressions for the detection and false-alarm probabilities of SCN-based spectrum sensing detector. We verified the analytical derivation results through Monte-Carlo simulations (please refer to Figure 5.6 for more details).

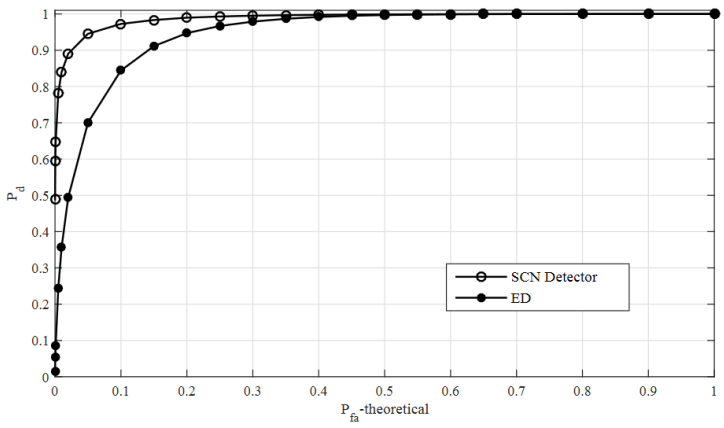


Figure 5.6 – ROC of the SCN-based detector vs. ROC of the ED where $K = 3$, $N = 500$, $SNR = -10$ dB and 0.1 dB of noise uncertainty.

A detailed paper entitled "*On the detection probability of the standard condition number*

detector in finite-dimensional cognitive radio context" dedicated to the theoretical and simulation work was published in EURASIP Journal on Wireless Communications and Networking [J9].

5.4.2 [C7]: Simple Formulation for Scaled Largest Eigenvalue Based Detector

In this contribution, we considered the scaled largest eigenvalue (SLE) based detector which is an optimal for detecting a single primary user in an uncertain noise environment. We proved that the SLE could be modeled using standard Gaussian distribution under some constraints. Moreover, we showed that the trace of central semi-correlated complex Wishart matrix follows a Gaussian distribution through the use of the central limit theorem (CLT). We derived the cumulative distribution function and the probability density function of the SLE based on the distributions of the largest eigenvalue and the trace. The false alarm probability, the detection probability and the decision threshold were also considered as we derived new simple and accurate forms. These forms are simple functions of the means and variances of the largest eigenvalue and the trace as well as the correlation coefficient between them. The correlation between the largest eigenvalue and the trace is studied and simple expressions are provided. In the following, we summarize the contributions of this work:

- Derivation of the distribution of the trace of the complex received covariance matrix for both $\mathcal{H}_0$ and $\mathcal{H}_1$ hypothesis.Derivation of the distribution of the scaled largest eigenvalue based detector for both hypotheses.
- Derivation of a simple form for the correlation coefficient between the largest eigenvalue and the trace under both hypotheses.
- Derivation of a simple form for the probability of false-alarm, P_{fa}, the detection probability, P_d, and the detection threshold. Theoretical findings are validated through Monte-Carlo simulations (please refer to Figure 5.7 for more details).

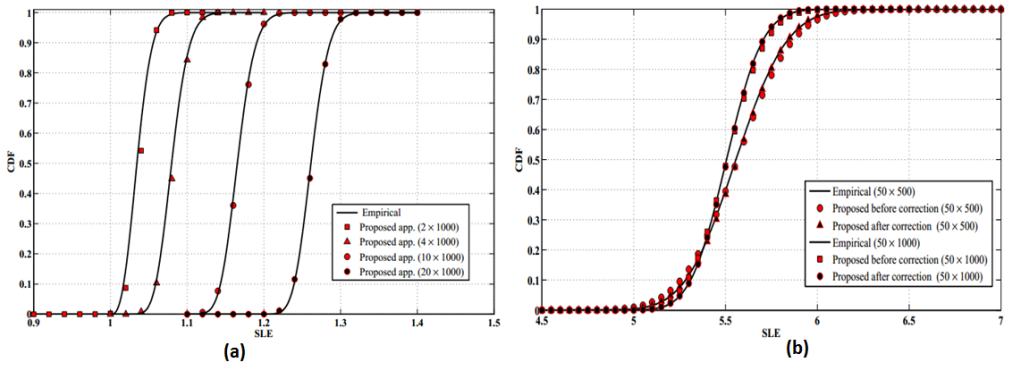


Figure 5.7 – Empirical CDF of the SLE under: **(a)** $\mathcal{H}_0$ hypothesis and its corresponding Gaussian approximation for different values of K, **(b)** $\mathcal{H}_0$ hypothesis and its corresponding proposed approximation for different values of N with $K = 50$ and $SNR = -10$ dB.

The major part of this work was published as a journal paper entitled "*On the Performance Analysis and Evaluation of Scaled Largest Eigenvalue in Spectrum Sensing: A Simple Form Approach*" in EAI Endorsed Transactions on Cognitive Communications in 2017 [J11].

5.4.3 [C8]: Full vs partial multi-antenna exploitation for spectrum sensing

Cognitive radio system with massive MIMO technology could use all antennas for the spectrum sensing process and achieve an enormous performance enhancement. However, it might be enough to use a fewer number of antennas for the sensing scheme and thus define a more efficient way for antenna exploitation. In this regard,two scenarios could be considered in this section:

- **Full antenna exploitation scenario**: The cognitive radio system may use all of its antennas in the spectrum sensing process and hence may reduce the number of samples required and decrease the sensing time.
- **Partial antenna exploitation scenario**: The cognitive radio system may fix certain number of antennas to a given spectrum sensing process and use the others for other purposes (such as transmission) or it may dynamically change the number of antennas K and/or the number of samples N according to predefined performance and also use the rest for other purposes.

In the first scenario, we considered that the cognitive radio will use all received antennas for spectrum sensing process and thus we approximate the largest eigenvalue (LE) based detector [28, 29] using the generalized extreme value distribution approximation under null hypothesis $\mathcal{H}_0$ while keeping the Gaussian approximation under alternative hypothesis $\mathcal{H}_1$. Accordingly, the threshold could be set. In the second scenario, we considered two approaches: **(i)** Fixed resources and **(ii)** dynamic resources. For this scenario, an optimal threshold for the largest eigenvalue based detector was derived and the minimum required resources are discussed if noise power is perfectly known and for standard condition number based and scaled large eigenvalue based detectors when noise uncertainty is considered. We showed that the dynamic approach is the best solution for an efficient antenna exploitation while maintaining on a target performance. Simulation tests have been done to validate the derived expressions and to study the different approaches. In addition, comparison between largest eigenvalue based, standard condition number based and scaled largest eigenvalue based detectors is illustrated in terms of detector performance and minimum requirements. The work is detailed in the journal publication entitled "*ELASTIC-Enabling Massive-Antenna for Joint Spectrum Sensing and Sharing: How Many Antennas Do We Need ?*" in IEEE Transactions on Cognitive Communications and Networking included hereafter.

ELASTIC- Enabling Massive-Antenna for Joint Spectrum Sensing and Sharing: How Many Antennas Do We Need?

Hussein Kobeissi, Youssef Nasser, *Senior Member, IEEE*, Oussama Bazzi, Amor Nafkha, *Senior Member, IEEE*, and Yves Louët, *Member, IEEE*

***Abstract*—Massive antenna and cognitive radio (CR) technologies have attracted many research interests due to the additional resources offered in striving against the spectrum crisis. In this paper, we propose a general framework to enlable massive antenna exploitation for spectrum sensing and sharing in CR. Using random matrix theory and moment matching method, we derived a simple approximation of the distributions of three eigen-value-based detectors namely the largest eigen-value (LE), the scaled LE, and standard condition number. This has led to a simple analytical formulation and optimization of the number of antennas and the number of samples to reach a target performance. To exploit the large number of antennas, we proposed two sensing scenarios. The first is based on full antenna exploitation and guarantees an optimal performance despite the transmission conditions. The second is based on partial antenna exploitation, which determines the exact number of antennas and samples required to reach the target performance. We have shown that the framework offers an additional degree of freedom in the selection of the optimal system parameters, namely the number of antennas, while the remaining antennas are exploited for sharing in other dimensions of the spectrum hypercube.**

***Index Terms*—Cognitive radio, MIMO systems, eigenvalue based detector, spectrum sensing, Wishart matrix.**

I. Introduction

5G, THE fifth generation of mobile networks, is expected to accommodate new demands and high data-rates that are growing at an unprecedented pace requiring a large amount of radio frequency resources. To fulfill these requirements, the Cognitive Radio (CR) technology has been introduced to offer a more efficient use of spectrum and thus reduce its scarcity [1]. A CR device can sense an unused channel and adjust its transceiver parameters accordingly.

In CR networks, Spectrum Sensing (SS) is the task of obtaining awareness about the spectrum usage. It concerns two scenarios of detection: (i) detecting the absence of the Primary User (PU) in a licensed spectrum and (ii) detecting the presence of the PU to avoid interference. Hence, SS plays a major role in the performance of the CR technology for both the PU and Secondary User (SU). In this regard, several SS algorithms have been proposed [2]. These techniques include energy detector, matched filter detector, cyclostationary feature detector, etc [3]. SS techniques with superior performance and robustness were also designed using the eigenvalues of the received signal's covariance matrix. These detectors, classified under the name *eigenvalue based detector* (EBD), rely on the use of random matrix theory (RMT) and different eigenvalue properties of the sample covariance matrix in decision making. The key advantage of the EBD lies in the fact that it can reach a high sensing performance without necessarily requiring knowledge about the primary signal and the noise power. EBD techniques include, but are not limited to, the largest eigenvalue (LE) detector [4]–[7], the scaled largest eigenvalue (SLE) detector [7]–[15], and the standard condition number detector (SCN) [4], [6], [16]–[20].

In the context of massive antenna deployment (up to few hundred in the literature), it is very likely for CRs to exploit the advantages offered by the multiple-input multiple-output (MIMO) technologies to improve secondary communications [21]–[26]. For instance, massive antenna technology allows the secondary users exploiting the angle-of-arrival dimension in the transmission hyperspace using beamforming [27]–[30]. As such, CR could be combined with massive MIMO through the additional degree of freedom offered by the large number of antennas to identify the unused channels while achieving a significant increase in the performance of the SS detector. It might be also enough to use a fewer number of antennas for the sensing process while the rest could be used for other purposes. Hence, two scenarios could be considered:

1. *Full Antenna Exploitation (FAE) Scenario:* Therein, the CR module may use all of its antennas to detect primary users and hence it reduces the number of samples

Manuscript received July 19, 2018; revised December 30, 2018 and February 15, 2019; accepted February 15, 2019. Date of publication February 26, 2019; date of current version June 7, 2019. The associate editor coordinating the review of this paper and approving it for publication was D. Niyato. *(Corresponding authors: Hussein Kobeissi; Youssef Nasser.)*

H. Kobeissi is with the Signal, Communication, and Embedded Electronics Research Group, CentraleSupélec, 35510 Rennes, France, also with the Doctoral School of Science and Technology, Lebanese University, Beirut 00961, Lebanon, and also with the Computer and Communication Engineering Department, International University of Beirut, Beirut 00961, Lebanon (e-mail: hussein.kobeissi@liu.edu.lb).

Y. Nasser is with the ECE Department, American University of Beirut, Beirut 1107 2020, Lebanon (e-mail: youssef.nasser@ieee.org).

O. Bazzi is with the Department of Physics and Electronics, Faculty of Science 1, Lebanese University, Beirut 961, Lebanon (email: obazzi@ul.edu.lb).

A. Nafkha and Y. Louët are with the Signal, Communication, and Embedded Electronics Research Group, CentraleSupélec of Rennes, 35510 Rennes, France (e-mail: amor.nafkha@centralesupelec.fr; yves.louet@centralesupelec.fr).

Digital Object Identifier 10.1109/TCCN.2019.2901847

required to perform spectrum sensing. Thus, the sensing time overhead is reduced.

2. *Partial Antenna Exploitation (PAE) Scenario:* the CR module may set a certain number of antennas to perform primary users detection and use the rest of the antennas for other purposes (mainly transmission). It may also dynamically change the number of antennas *K* and/or the number of samples *N* according to predefined performance, and use the remaining antennas for other purposes.

Accordingly, the calculation of the minimum number of antennas and/or the minimum number of samples required to reach a given target performance is very important. However, these values are directly related to the statistics of the spectrum sensing approach and its corresponding performance metrics and mechanisms.

In this paper, we mainly consider the framework to EnlabLe mASsive anTenna exploItation for spectrum sensing and sharing CR (ELASTIC). We firstly introduce the LE, SLE and SCN detectors, and we propose new approximations for their distribution based on the generalized extreme value (GEV) probability density function (PDF). These detectors are then considered and studied in the two antenna exploitation scenarios, i.e., FAE and PAE. The main contributions of this paper can be summarized as follows:

- Proposition of a new approximation for the cumulative distribution function (CDF) and the PDF of the LE detector.The proposed approximation provides a simple formulation of the proposed analytical framework and allows an optimization of the number of antennas as stated by ELASTIC. To the best of the authors knowledge, this approximation has not been proposed in literature including our previous works.
- Extension of our previous work in [14], [19], and [20] to provide key design metrics for the operators to optimize the CR network. Indeed, in our previous works, we have derived new approximations of the PDF EBD detectors. These approximations led to simple formulations that could be exploited in the design of a system with multiple antennas. In ELASTIC, we exploit these formulations to set and derive key design metrics depending on the required system performance given in terms of probability of false alarm and detection. In other words, this work proposes a simple analytical tool to tune these metrics (such as detection threshold) in terms of the network transmission parameters (as the number of antennas).[1]
- Derivation of the optimal decision thresholds for the three detectors in both FAE and PAE scenarios. The latter is given in terms of probability of false alarm P_{fa} and probability of detection P_d.
- Derivations of the minimum requirements in number of antennas *K* and number of samples *N* that achieve a target detection performance.

[1]The paper does not reproduce our previous work but provides a reminder on the main derivations that have been stated to assure a smooth readability of the current work. These derivations are then extended to optimize the Massive-MIMO CR parameters.

Note 1: This work shows that a reduced number of antennas is needed to sense a channel. Accordingly, in a massive antenna environment, a large number of channels can be sensed simultaneously.

Note 2: It is worthy that the exact distributions of the three detectors have been derived in literature. However, these distributions are too complicated to be used in the analytical optimization of the system metrics. For instance, a Marchenko-Pastur (MP) law is used for the SCN distribution [4] but its expression is hard to be analytically exploited.

To the best of the authors knowledge, this work is a first of its kind in literature and will definitely open new research pathways to exploit the analytical derivations provided in this paper to other applications such as spectrum sharing, cognitive user capacity analysis, mmWave base station massive sensing capabilities in vehicular networks [31], etc.

The rest of this paper is organized as follows. In Section II, we present the main rationale behind ELASTIC framework and give the benefits of these approaches. Section III introduces the system framework, system model and defines the main parameters for EBD. In Section IV, the largest eigenvalue detector is firstly introduced and then followed by the proposed approximation of its PDF in different scenarios. In Sections V and VI, the SLE and and SCN detectors are introduced and then followed by the proposed PDF distributions respectively. Full and partial antenna exploitation scenarios are analytically considered in Section VII. Threshold derivation and optimization are considered along with optimal system requirements in Section VIII. Results are validated and discussed in Section IX while the conclusion is drawn in Section X.

Notations: Vectors and Matrices are represented, respectively, by lower and upper case boldface. The symbols $|.|$ and $tr(.)$ indicate, respectively, the determinant and trace of a matrix while $(.)^{1/2}$, $(.)^T$, and $(.)^\dagger$ are the square root, transpose, and Hermitian symbols respectively. $\boldsymbol{I}_n$ is the $n \times n$ identity matrix and $\mathbf{1}_{KN}$ is a $K \times N$ ones matrix. Symbol $\sim$ stands for "distributed as" and $E[.]$ stands for the expected value.

II. The Rationale Behind ELASTIC

As stated earlier, two main scenarios could be considered: the FAE and the PAE. While the full antenna exploitation scenario is using all of the antennas at the same time for spectrum sensing, the partial antenna exploitation scenario could be decomposed into two different approaches: (i) exploit a fixed number of antennas for SS and thus use the rest for other purposes, it will be labeled fixed PAE (FPAE) and (ii) a dynamic approach in which the CR does not set a predefined number of antennas but it dynamically allocates a certain number of antennas for SS according to certain constraints, labeled dynamic PAE (DPAE).

The division into approaches could be justified by the simple fact that a trade-off between spectrum sensing performance and exploitation of available antennas for other targets could exist. This depends on the different metrics and performance

objectives. In this work, the following approaches could be considered:

- Sensing multi-channels simultaneously [32], [33]: Within the framework of ELASTIC, a new approach is considered: the antennas are clustered so that each set is used to sense certain channel and/or certain direction. Hence, several spectrum holes could be found at the same time.
- Sense and transmit simultaneously: Contrarily to the conventional sensing techniques [34], simultaneous sensing and transmission could be considered as a full duplex approach in which two RF chains are required at the CR receiver [35]. Likewise, the CR could maintain a set of antennas for SS and the others for the transmission in a full duplex scenario.
- Green radio: ELASTIC framework offers an energy efficiency perspective in which a limited number of antennas are switched-on while the others are in sleep mode [36].
- Reduce sensing time: The FAE scenario may not be the optimal scheme to reduce the sensing time since PAE could offer the same performance with less antennas. Hence, an optimal trade-off between the number of antennas and the required number of samples to make a decision may decrease the total sensing time.
- Increase system throughput: In general, using some or all of these methods would increase the total system throughput as the system will be able to transmit data over the remaining time/antennas.

Consequently, the use of a fixed number or dynamic number of antennas for spectrum sensing is extremely important for CR with massive antenna technology. In both cases, the CR technology has to optimize the number of antenna K and samples N to perform primary users detection so that to maintain a target performance for any transmission conditions (given for, e.g., in terms of Signal-to-Noise Ratio SNR). In both cases, the remaining number of antennas will be used for other transmission purposes.

III. ELASTIC FRAMEWORK

A. System Model

Consider a multiple-antenna CR system, say K antennas, aiming to detect the presence/absence of a single PU during a sensing period equivalent to NT_s where T_s is the sampling period. For this detection problem, there are two hypotheses: the null hypothesis $\mathcal{H}_0$ corresponds to the absence of PU transmission (i.e., spectrum hole); and the alternative hypothesis $\mathcal{H}_1$ corresponds to the presence of the PU transmission (i.e., spectrum being used). The received vector, at instant nT_s, under both hypotheses is given by:

$$\mathcal{H}_0 : \boldsymbol{y}(n) = \boldsymbol{b}(n), \tag{1}$$

$$\mathcal{H}_1 : \boldsymbol{y}(n) = \boldsymbol{h}(n)s(n) + \boldsymbol{b}(n), \tag{2}$$

where $\boldsymbol{y}(n) = [y_1(n), \ldots, y_K(n)]^T$ is the observed $K \times 1$ complex samples received from different antennas at instant n. $\boldsymbol{b}(n)$ is a $K \times 1$ complex circular white Gaussian noise with zero mean and variance σ_b^2. $\boldsymbol{h}(n)$ is a $K \times 1$ vector that represents the channels' coefficients between the PU and each antenna at the CR receiver and $s(n)$ stands for the primary signal sample at instant nT_s, having a Gaussian distribution with zero mean and variance σ_s^2. In this paper, under $\mathcal{H}_1$, the model represents a single PU transmitted over an independent and identically distributed (i.i.d.) channel coefficients with zero mean and variance σ_h^2 or equivalently the PU signal is sampled subject to uncorrelated (fast) Rayleigh fading. This assumption is in line with the literature works such [18], [24]. Moreover, the channel coefficients $h(n)$ are assumed to be i.i.d. at each time instant n. After collecting N samples from each antenna, the received signal matrix $\boldsymbol{Y}$ is written as:

$$\boldsymbol{Y} = \begin{pmatrix} y_1(1) & y_1(2) & \cdots & y_1(N) \\ y_2(1) & y_2(2) & \cdots & y_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ y_K(1) & y_K(2) & \cdots & y_K(N) \end{pmatrix} \tag{3}$$

Without loss of generality, we suppose that $K \leq N$ and we define the received sample covariance matrix as $\boldsymbol{W} = \boldsymbol{Y}\boldsymbol{Y}^{\dagger}$.

1) $\mathcal{H}_0$ Hypothesis: Under $\mathcal{H}_0$ hypothesis, the entries of the matrix $\boldsymbol{Y}$ are complex Gaussian with zero mean and variance σ_b^2. Therefore, the sample covariance matrix, $\boldsymbol{W}$, is a $K \times K$ central uncorrelated complex Wishart matrix with N degrees of freedom (DoF) and with statistical covariance matrix:

$$\boldsymbol{\Sigma} = \sigma_b^2 \mathbf{I}_K \tag{4}$$

and is denoted by $\mathbf{W} \sim \mathcal{CW}_K(N, \sigma_b^2 \mathbf{I}_K)$.

2) $\mathcal{H}_1$ Hypothesis: Under $\mathcal{H}_1$, we assume the presence of a single PU. Consequently, $\mathbf{W}$ follows a central semi-correlated complex Wishart distribution, denoted by $\mathbf{W} \sim \mathcal{CW}_K(N, \boldsymbol{\Sigma})$, with N DoF and statistical covariance matrix $\boldsymbol{\Sigma}$ given by [37]:

$$\boldsymbol{\Sigma} = \sigma_s^2 \boldsymbol{h}\boldsymbol{h}^{\dagger} + \sigma_b^2 \mathbf{I}_K \tag{5}$$

The statistical covariance matrix $\boldsymbol{\Sigma}$ is a rank 1 perturbation of the identity matrix, then it belongs to the class of spiked population model first introduced by [38]. Since the rank of $\sigma_s^2 \boldsymbol{h}\boldsymbol{h}^{\dagger}$ is 1, then all but one of the eigenvalues of $\boldsymbol{\Sigma}$ are still σ_b^2. Denoting the average SNR by:

$$\rho = \frac{\sigma_s^2 \sigma_h^2}{\sigma_b^2}, \tag{6}$$

and by using the property that the trace of a matrix equals the sum of its eigenvalues, the eigenvalues of $\boldsymbol{\Sigma}$ are given by:

$$\boldsymbol{\sigma} = \sigma_b^2 \big[K\rho + 1, \mathbf{1}_{1,K-1}\big]. \tag{7}$$

B. Eigenvalue Based Detector

EBD could be divided into two classes: detectors that require the knowledge of noise variance and detectors that do not require this knowledge. LE detector is the optimal detector when noise variance is perfectly known while SLE detector is the optimal under the generalized likelihood ratio (GLR) criterion when noise variance is unknown [7], [10]. Other EBD techniques such as the SCN detector provide good performance while they don't require channel knowledge or primary user information. For a given decision threshold $\hat{\lambda}_{EBD}$, the general EBD algorithm is given by: where $[\lambda_1, \ldots, \lambda_K]$ is the vector of eigenvalues of the covariance matrix $\boldsymbol{W}$, X_{EBD} is the EBD metric used for detection and D_{EBD} is its decision.

Algorithm 1: Eigenvalue Based Detector

Input: $\boldsymbol{Y}$, $\hat{\lambda}_{EBD}$
Output: D_{EBD}
compute $\boldsymbol{W} = \boldsymbol{Y}\,\boldsymbol{Y}^{\dagger}$;
get $[\lambda_1, \cdots, \lambda_K]$ of $\boldsymbol{W}$ using EVD;
evaluate X_{EBD};
decide $D_{EBD} = X_{EBD} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \hat{\lambda}_{EBD}$;

C. Performance Probabilities

The performance of any spectrum sensing technique is usually expressed in terms of its false-alarm probability and its missed-detection probability, both known as the error probabilities. The probability of false-alarm probability (P_{fa}) and the probability of missed-detection (P_{md}) are defined as follows:

$$P_{fa} = Pr(X \geq \alpha/\mathcal{H}_0) \tag{8}$$

$$P_{md} = Pr(X < \alpha/\mathcal{H}_1) \tag{9}$$

where *X* denotes the statistical EBD metric.

D. Operation Conditions of the EBD Derivations

In this paper, two different conditions will be distinguished.

1) Asymptotic Condition (AC): The AC is given by:

$$(K, N) \to \infty \text{ with } K/N \to r \in (0, 1), \tag{10}$$

2) Critical Condition (CC): The critical condition is related with the SNR requirement by:

$$\rho > \rho_c = \frac{1}{\sqrt{KN}}. \tag{11}$$

IV. Largest Eigenvalue Detector

The LE statistical metric is defined as the ratio between the largest eigenvalue of the received covariance matrix and the noise power. It can be expressed as follows:

$$X_{LE} = \frac{\lambda_1}{\sigma_b^2}. \tag{12}$$

Its exact distribution is derived in literature in the form of matrix determinant (see [39], [40]).

A. LE Detector in Asymptotic Regime

The asymptotic regime is defined in this paper when (10) is verified. Asymptotically, X_{LE} metric follows a Tracy-Widom (TW) distribution for central uncorrelated Wishart matrices (i.e., $\mathcal{H}_0$) and a Gaussian distribution for sample covariance matrices of spiked population model (i.e., $\mathcal{H}_1$) [41]. However, for a fixed *K* and as $N \to +\infty$ then the metric X_{LE} converges to a Gaussian distribution [42].

1) Null Hypothesis Case: By considering the AC in (10), X_{LE}, properly centered and scaled, follows asymptotically a TW distribution of order 2 (TW2) as follows:

$$X'_{LE} = \frac{X_{LE} - a_1(K, N)}{b_1(K, N)} \sim TW2, \tag{13}$$

where $a_1(K, N)$ and $b_1(K, N)$ are respectively the centering and scaling coefficients defined by:

$$a_1(K, N) = \left(\sqrt{K} + \sqrt{N}\right)^2, \tag{14}$$

$$b_1(K, N) = \left(\sqrt{K} + \sqrt{N}\right)\left(K^{-1/2} + N^{-1/2}\right)^{\frac{1}{3}}. \tag{15}$$

2) $\mathcal{H}_1$ Hypothesis Case: Under the asymptotic condition in (10) and the critical condition in (11), it has been shown that X_{LE} follows a normal distribution such that [41]:

$$P\left(\frac{X_{LE} - a_2(K, N, \sigma)}{\sqrt{b_2(K, N, \sigma)}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du, \tag{16}$$

with

$$a_2(K, N, \sigma) = \sigma_1\left(N + \frac{K}{\sigma_1 - 1}\right) \tag{17}$$

$$b_2(K, N, \sigma) = \sigma_1^2\left(N - \frac{K}{(\sigma_1 - 1)^2}\right) \tag{18}$$

where σ_1 is the largest eigenvalue of (7). Definitely, if $\rho < \rho_c$, the null hypothesis case stands in which X_{LE}, under (10), follows the TW2 distribution [41]. Accordingly, the PU signal has no effect on the eigenvalues and could not be detected.

B. Approximating X_{LE} Distribution Under $\mathcal{H}_0$

It has been shown that, for a fixed *K* and as $N \to \infty$, X_{LE} follows a normal distribution [42]. However, from a FAE perspective, i.e., both *K* and *N* are large and AC is satisfied, X_{LE} follows the TW2 distribution rather than the normal distribution. TW2 distribution requires, from one side, the calculation of complex derivations that include special functions and, from the other side, these derivations depend on many transmission parameters. Hence, these derivations are not analytically tractable for the CR system metrics optimization (for instance optimal *K* and *N*). Accordingly, two alternatives may be used to avoid this expensive calculation: (i) use a lookup table (LUT) approach or (ii) approximate its PDF by simpler expression. The LUT is interesting as it provides an off-line calculation of the system metrics however it might be too expensive in case of dynamic antenna exploitation. In this paper, we propose a new approximation for X_{LE} distribution by using the GEV distribution and the moment matching method. Our objective is to find a simpler form that could be analytically used in the optimization problem of both FAE and PAE. Now, considering the asymptotic condition in (10), the mean, variance and skewness of the X_{LE} are given by the following Corollary.

Corollary 1: Let $\mathbf{W} \sim \mathcal{CW}_K(N, \sigma_b^2\mathbf{I}_K)$ be a central uncorrelated complex Wishart matrix with *N* DoF and correlation matrix $\sigma_b^2\mathbf{I}_K$. If *K* and *N* obey (10), then the mean, variance and skewness of X_{LE} associated to $\mathbf{W}$ are given, respectively, by:

$$\mu_{X_{LE}} = b_1(K, N)\mu_{TW2} + a_1(K, N), \tag{19}$$

$$\sigma^2{}_{X_{LE}} = b_1^2(K, N)\sigma_{TW2}^2, \tag{20}$$

$$\mathcal{S}_{X_{LE}} = \mathcal{S}_{TW2}, \tag{21}$$

with $\mu_{TW2} = -1.7710868074$, $\sigma^2_{TW2} = 0.8131947928$ and $\mathcal{S}_{TW2} = 0.2240842036$ are, respectively, the mean, variance and skewness of the TW2 distribution [43].

Proof: This result comes directly from (13). ■

Fortunately, the mean, variance and skewness expressions of X_{LE} have simple forms when asymptotic condition in (10) is achieved. In this regard, the following proposition is useful.

Proposition 1: Let *N* and *K* obey (10), i.e., under AC, then the CDF and PDF of the LE under null hypothesis can be accurately approximated respectively by:

$$F(x;\theta,\beta,\xi) = e^{-\left(1+\left(\frac{x-\theta}{\beta}\right)\xi\right)^{-1/\xi}} \tag{22}$$

$$f(x;\theta,\beta,\xi) = \frac{1}{\beta}\left(1+\left(\frac{x-\theta}{\beta}\right)\xi\right)^{\frac{-1}{\xi}-1} e^{-\left(1+\left(\frac{x-\theta}{\beta}\right)\xi\right)^{-1/\xi}} \tag{23}$$

where ξ, β, and θ are respectively the shape, scale and location parameters of the GEV distribution and are given by:

$$\xi = -0.06393\mathcal{S}^2_{X_{LE}} + 0.3173\mathcal{S}_{X_{LE}} - 0.2771 \tag{24}$$

$$\beta = \sqrt{\frac{\sigma^2_{X_{LE}}\xi^2}{g_2 - g_1^2}} \tag{25}$$

$$\theta = \mu_{X_{LE}} - \frac{(g_1-1)\beta}{\xi} \tag{26}$$

and the mean, the variance and the skewness of X_{LE} are given by Corollary 1.

Proof: The result can be found using [44, Lemma 5]. ■

It follows from Proposition 1 that TW2 distribution, itself, could be approximated using GEV distribution by considering the mean, the variance and the skewness of TW2 mentioned in Corollary 1. Moreover, it will be shown through simulation results that this approximation is valid even if *K* and *N* are not asymptotically large.

C. Approximating X_{LE} Distribution Under $\mathcal{H}_1$

For $\mathcal{H}_1$ hypothesis, X_{LE} follows the Gaussian distribution [42]; it is indeed simple to exploited and any approximation in this case is useless.

V. SLE Detector

The SLE detector is a blind detector that does not require information about the noise power. The statistical metric is defined as the ratio of the largest eigenvalue to the normalized trace of the received covariance matrix, and it is given by:

$$X_{SLE} = \frac{\lambda_1}{T_n} = \frac{\lambda_1}{\frac{1}{K}\sum_{i=1}^{K}\lambda_i} \tag{27}$$

In the literature, the results on the statistics of the SLE are relatively limited. They are based on tools from RMT [8], [10], [45] and Mellin transform [11], [13], [45]. X_{SLE} was considered, asymptotically, to follow the same distribution as the X_{LE} (i.e., TW distribution) [10]. However, as a non-negligible error still exists in this approximation, a new form was provided based on the TW distribution and its second derivative in [8].

A. Null Hypothesis Case

Under $\mathcal{H}_0$, both the X_{LE} and the normalized trace of the matrix $\mathbf{W}$ follow the Gaussian distribution as $N \to +\infty$ which is realistic in practical spectrum sensing scenarios. Accordingly, we have shown that the SLE detector can be formulated using standard Gaussian function as follows [14], [15].

Theorem 1: Let X_{SLE} be the statistical metric associated to $\mathbf{W} \sim \mathcal{CW}_K(N, \sigma_b^2\mathbf{I}_K)$. Then, for a fixed *K* and as $N \to +\infty$, the CDF and the PDF of X_{SLE} are, respectively, given by:

$$F(x) = \Phi\left(\frac{x\mu_{T_n} - \mu_{\lambda_1}}{\sqrt{\sigma^2{}_{\lambda_1} - 2xc + x^2\sigma^2{}_{T_n}}}\right) \tag{28}$$

$$f(x) = \frac{\mu_{T_n}\sigma^2{}_{\lambda_1} - c\mu_{\lambda_1} + \left(\mu_{\lambda_1}\sigma^2{}_{T_n} - c\mu_{T_n}\right)x}{\left(\sigma^2{}_{\lambda_1} - 2xc + x^2\sigma^2{}_{T_n}\right)^{\frac{3}{2}}} \times \phi\left(\frac{x\mu_{T_n} - \mu_{\lambda_1}}{\sqrt{\sigma^2{}_{\lambda_1} - 2xc + x^2\sigma^2{}_{T_n}}}\right) \tag{29}$$

with

$$\Phi(v) = \int_{-\infty}^{v}\phi(u)du \text{ and } \phi(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}} \tag{30}$$

where μ_{λ_1}, μ_{T_n} and $\sigma^2{}_{\lambda_1}$, $\sigma^2{}_{T_n}$ are, respectively, the mean and the variance of λ_1 given by (19), (31) and (20), and T_n given by (32). The parameter c is given by $c = \sigma_{\lambda_1}\sigma_{T_n}\gamma$ where γ is the correlation coefficient between λ_1 and T_n provided by (33).

$$\mu_{T_n} = N, \tag{31}$$

$$\sigma^2{}_{T_n} = N/K, \tag{32}$$

$$\sigma = \frac{\sigma_{T_n}}{\sigma_{\lambda_1}} \cdot \frac{\theta\ \mu_{X_{SLE}} - \mu_{\lambda_1}}{\theta + \mu_{T_n}}. \tag{33}$$

and $\mu_{X_{SLE}}$ is the mean of X_{SLE} provided by (34) and θ is given by (35).

$$\mu_{X_{SLE}} = \frac{\mu_{\lambda_1}}{\mu_{T_n}} \tag{34}$$

$$\theta = 1.01\mu_{T_n} - 0.2713\sigma_{T_n}. \tag{35}$$

Proof: The CDF is given by (28) and the PDF is its derivative in (29) [46]. The reader can refer to [14] and [15] for more details. ■

B. $\mathcal{H}_1$ Hypothesis Case

Under $\mathcal{H}_1$, the normalized trace follows the Gaussian distribution as $N \to +\infty$ whereas the LE follows the Gaussian distribution as $(K, N) \to +\infty$ with $K/N \to r \in (0,1)$ and $\rho > \rho_c = 1/\sqrt{KN}$. Accordingly, the distribution of the SLE is given by Theorem 2 [15].

Theorem 2: Let X_{SLE} be the statistical metric associated to $\mathbf{W} \sim \mathcal{CW}_K(N, \mathbf{\Sigma})$. Then, as $(K, N) \to +\infty$ with $K/N \to r \in (0,1)$ and $\rho > \rho_c = 1/\sqrt{KN}$, the CDF and PDF of X_{SLE} are, respectively, given by (28) and (29). Moreover, μ_{λ_1}, μ_{T_n} and $\sigma^2{}_{\lambda_1}$, $\sigma^2{}_{T_n}$ are, respectively, the mean and the variance of λ_1 and T_n given by (36), (38) and (37), (39) respectively where σ_1 is the largest eigenvalue of (7). The parameter c is

defined by $c = \sigma_{\lambda_1}\sigma_{T_n}\gamma$ where γ is the correlation coefficient between λ_1 and T_n given by (40).

$$\mu_{\lambda_1} = \sigma_1\left(N + \frac{K}{\sigma_1 - 1}\right), \tag{36}$$

$$\sigma^2{}_{\lambda_1} = \sigma_1^2\left(N - \frac{K}{(\sigma_1 - 1)^2}\right), \tag{37}$$

$$\mu_{T_n} = \frac{N}{K}(\sigma_1 + K - 1), \tag{38}$$

$$\sigma^2{}_{T_n} = \frac{N}{K^2}\left(\sigma_1^2 + K - 1\right), \tag{39}$$

$$\gamma = \frac{\sigma_{T_n}}{\sigma_{\lambda_1}} \cdot \frac{\theta\ (\mu_X + \epsilon) - \mu_{\lambda_1}}{\theta + \mu_{T_n}} \tag{40}$$

and $\mu_{X_{SLE}}$ is the mean of X_{SLE} provided by (34), θ is given by (35) and ϵ is a variable used to model the mean error.

Proof: Same as the proof of Theorem 1 ■

VI. SCN Detector

The SCN detector is another blind detector that uses the eigenvalues of the sample covariance matrix in decision making [6], [18], [47]–[49]. It is defined as the ratio of the largest to the smallest eigenvalue of the covariance matrix as follows:

$$X_{SCN} = \frac{\lambda_1}{\lambda_K} \tag{41}$$

In literature, the SCN metric was studied asymptotically in [4] and the threshold was presented according to MP law. Zeng and Liang [6] improved the accuracy of the asymptotic statistical distribution of the SCN by using the TW distribution. This work was further extended in [16] by using Curtiss formula where both the largest and the smallest eigenvalues converge to TW distributions when $(K, N) \to +\infty$ as shown in [50] and [51]. However, all these distributions are not analytically tractable. Hence, a simplification of these distributions is required.

A. Null Hypothesis Case

The SCN metric distribution can be approximated using a simple and accurate GEV PDF given by Theorem 3 [19]:

Theorem 3: Let X_{SCN} be the statistical metric associated to $\boldsymbol{W} \sim \mathcal{CW}_K(N, \sigma_b^2 \boldsymbol{I}_K)$. If AC condition is satisfied, then the CDF and PDF of X_{SCN} can be asymptotically and tightly approximated respectively by:

$$F(x;\theta_0,\beta_0,\xi_0) = e^{-\left(1+\left(\frac{x-\theta_0}{\beta_0}\right)\xi_0\right)^{-1/\xi_0}} \tag{42}$$

$$f(x;\theta_0,\beta_0,\xi_0) = \frac{1}{\beta_0}\left(1 + \left(\frac{x-\theta_0}{\beta_0}\right)\xi_0\right)^{\frac{-1}{\xi_0}-1} \times\ e^{-\left(1+\left(\frac{x-\theta_0}{\beta_0}\right)\xi_0\right)^{-1/\xi_0}} \tag{43}$$

where ξ_0, β_0 and θ_0 are defined respectively by:

$$\xi_0 = -0.06393\mathcal{S}_{X_{SCN}}^2 + 0.3173\mathcal{S}_{X_{SCN}} - 0.2771 \tag{44}$$

$$\beta_0 = \sqrt{\frac{\sigma_{X_{SCN}}^2 \xi^2}{g_2 - g_1^2}} \tag{45}$$

$$\theta_0 = \mu_{X_{SCN}} - \frac{(g_1 - 1)\beta}{\xi} \tag{46}$$

where $\mu_{X_{SCN}}$, $\sigma^2_{X_{SCN}}$ and $\mathcal{S}_{X_{SCN}}$ are defined in [19, Th. 1] and $g_i = \Gamma(1 - i\xi)$.

Proof: Refer to [19] for the detailed proof. ■

B. $\mathcal{H}_1$ Hypothesis Case

Likewise, we introduce the following theorem which approximates the distribution of X_{SCN} using the simple GEV distribution under $\mathcal{H}_1$ hypothesis.

Theorem 4: Let X_{SCN} be the statistical metric associated to $\boldsymbol{W} \sim \mathcal{CW}_K(N, \boldsymbol{\Sigma})$. If AC and CC conditions are satisfied, then the CDF and PDF of X_{SCN} can be asymptotically and tightly approximated by:

$$F(x;\theta_1,\beta_1,\xi_1) = e^{-\left(1+\left(\frac{x-\theta_1}{\beta_1}\right)\xi_1\right)^{-1/\xi_1}} \tag{47}$$

$$f(x;\theta_1,\beta_1,\xi_1) = \frac{1}{\beta_1}\left(1 + \left(\frac{x-\theta_1}{\beta_1}\right)\xi_1\right)^{\frac{-1}{\xi_1}-1} \times\ e^{-\left(1+\left(\frac{x-\theta_1}{\beta_1}\right)\xi_1\right)^{-1/\xi_1}} \tag{48}$$

where ξ_1, β_1 and θ_1 are defined respectively by:

$$\xi_1 = -0.06393\mathcal{S}_{X_{SCN}}^2 + 0.3173\mathcal{S}_{X_{SCN}} - 0.2771 \tag{49}$$

$$\beta_1 = \sqrt{\frac{\sigma_{X_{SCN}}^2 \xi^2}{g_2 - g_1^2}} \tag{50}$$

$$\theta_1 = \mu_{X_{SCN}} - \frac{(g_1 - 1)\beta}{\xi} \tag{51}$$

where $\mu_{X_{SCN}}$, $\sigma^2_{X_{SCN}}$ and $\mathcal{S}_{X_{SCN}}$ are defined in [19, Th. 2] and $g_i = \Gamma(1 - i\xi)$.

Proof: Refer to [19] for the detailed proof. ■

VII. FAE and PAE: Performance Probabilities

The target of this section is to summarize the different cases analyzed in this paper and provide the corresponding performance metrics given here in terms of P_{fa} and P_d. The latter will be used in the threshold optimization and the system specifications within the ELASTIC framework.

A. Summary on the Proposed Analytical Derivations

It is very clear from the previous sections that the proposed yet simple approximations could be directly applied within the framework of ELASTIC. For the sake of simplicity, we give in Table I, a summary of the proposed approximations. They will be used in the optimization of the CR system metrics in both the PAE and FAE scenarios with any transmission conditions.

B. Performance Probabilities of the LE Detector

In this section, we present the two mentioned scenarios of antenna exploitation within the ELASTIC framework. In the FAE scenario, the CR will use all of its antennas for the SS process. Accordingly, the LE detector will be working in the asymptotic regime of both *K* and *N*. On the other hand, in the PAE scenario the CR will use a fewer number of antennas

TABLE I
SUMMARY OF THE PROPOSED PDFS

Metric	Existing Dist. (under $\mathcal{H}_0$)	Existing Dist. (under $\mathcal{H}_1$)	Proposed Dist. (under $\mathcal{H}_0$)	Proposed Dist. (under $\mathcal{H}_1$)
LE	TW2 or Normal	Normal	**Proposition 1** - GEV	Normal
SLE	using TW2 or Mellin Trans.	using Mellin Trans.	**Theorem 1** - Gaussian function	**Theorem 2** - Gaussian function
SCN	TW2 with Curtiss Formula	TW2 with Curtiss Formula	**Theorem 3** - GEV	**Theorem 4** - GEV

for the SS process and thus the LE detector will be working in the asymptotic regime of *N* only (i.e., *N* is relatively large w.r.t *K*).

1) FAE Scenario: According to Proposition 1 and using (8), the expression of P_{fa} is given by:

$$P_{fa} = 1 - e^{-\left(1+\left(\frac{\hat{\lambda}_{X_{LE}}-\theta}{\beta}\right)\xi\right)^{-1/\xi}}. \tag{52}$$

Similarly, by using (16) and (9) the expression of P_{md} is given by:

$$P_{md} = \Phi\left(\frac{\hat{\lambda}_{X_{LE}} - \mu_1}{\sigma_1}\right) \tag{53}$$

where Φ is the CDF of the standard normal distribution, $\hat{\lambda}_{X_{LE}}$ is the threshold of the LE detector, μ_1 and σ_1 are respectively the mean and the standard deviation of the LE metric under $\mathcal{H}_1$ given in (16).

2) PAE Scenario: Here, we consider the following:

- FPAE approach: in this case, the CR will use certain fixed number *K* and $N >> K$.
- DPAE approach: here, the CR will use dynamic number of antenna *K* and $N >> K$.

In both cases, *K* is finite and *N* is relatively large. Under $\mathcal{H}_0$ hypothesis, X_{LE} is approximated by Normal distribution where the mean and the variance are approximated using (19) and (20) respectively. Under $\mathcal{H}_1$ hypothesis, results in [42] show that X_{LE} could be approximated by Gaussian distribution in (16) even for small values of *K*. Then, P_{fa} and P_{md} are expressed as follows:

$$P_{fa} = 1 - \Phi\left(\frac{\hat{\lambda}_{X_{LE}} - \mu_0}{\sigma_0}\right) \tag{54}$$

$$P_{md} = \Phi\left(\frac{\hat{\lambda}_{X_{LE}} - \mu_1}{\sigma_1}\right) \tag{55}$$

where $\Phi(.)$ is the CDF of the standard normal distribution; μ_0 and σ_0 are the mean and the standard deviation of X_{LE} under $\mathcal{H}_0$ hypothesis and are given by (19) and (20) respectively; μ_1 and σ_1 are the mean and the standard deviation of the X_{LE} under $\mathcal{H}_1$ hypothesis and are given by (17) and (18).

C. Performance Probabilities of the SLE and SCN Detectors

Using (8) and (28), the P_{fa} of the SLE detector is given by:

$$P_{fa}(\alpha) = Q\left(\frac{\hat{\lambda}_{X_{SLE}}\mu_{T_n} - \mu_{\lambda_1}}{\sqrt{\sigma^2{}_{\lambda_1} - 2\alpha c + \alpha^2\sigma^2{}_{T_n}}}\right) \tag{56}$$

where $Q(.)$ is the Q-function. μ_{λ_1}, $\sigma^2{}_{\lambda_1}$, μ_{T_n} and $\sigma^2{}_{T_n}$ are given respectively (19), (20), (31) and (32). P_{md} is derived the same way using $\mathcal{H}_1$ hypothesis. Using Theorems 3 and 4, and using (8) and (9), the P_{fa} and P_{md} of the SCN detector are respectively expressed as:

$$P_{fa} = 1 - e^{-\left(1+\left(\frac{\hat{\lambda}_{X_{SCN}}-\theta_0}{\beta_0}\right)\xi_0\right)^{-1/\xi_0}}, \tag{57}$$

$$P_{md} = 1 - P_d = e^{-\left(1+\left(\frac{\hat{\lambda}_{X_{SCN}}-\theta_1}{\beta_1}\right)\xi_1\right)^{-1/\xi_1}}. \tag{58}$$

VIII. THRESHOLD OPTIMIZATION

The target of this section is to derive the optimal decision threshold for a desired P_{fa} and/or P_{md}. However, as these metrics depend on *K*, *N*, and maybe the SNR (in case of LE), the optimization of this threshold will depend on these main parameters. Hence, finding the optimal values of *K*, *N* for a target P_{fa} and/or P_{md} is analytically not straightforward. To do so, the designer has to select a key performance or run an algorithmic search as shown in Algorithm 2. In both cases, the optimization is still easier than using of the exact distribution (such TW2) of the sensing metric or using the LUTs. Definitely, if the system parameters are to be changed dynamically (i.e., DPAE), then the optimization problem turns out to be finding the optimal value of *K* and/or *N* that fit the performance requirements. It is worth mentioning that the execution time of Algorithm 2 is too small in our simulations. Despite this fact, the search of the optimal solution could be made faster by specifying a range of value for *K* and *N*. In this section, the optimal threshold is given for a given *K* and *N*.

A. Threshold Optimization for the LE Detector

The performance probabilities depend on the decision threshold ($\hat{\lambda}_{X_{LE}}$), hence it is necessary to choose an appropriate value based on system requirements. The typical approach for setting the threshold is given by the constant false-alarm rate (CFAR) strategy in which the threshold is chosen in order to guarantee a target false-alarm rate ($\hat{p}_{fa}$). Based on the CFAR scenario, then the decision threshold is expressed using the inverse of P_{fa} as follows.

1) FAE Scenario:

$$\hat{\lambda}_{X_{LE}} = \mu + \frac{\sigma}{\xi}\left(-1 + \left[-\ln\left(1 - \hat{P}_{fa}\right)\right]^{-\xi}\right). \tag{59}$$

2) PAE Scenario: For finite *K* and relatively large *N* and based on the CFAR scenario, the decision threshold is expressed using the inverse of P_{fa} as follows:

$$\hat{\lambda}_{X_{LE}} = \mu_0 + \sigma_0\Phi^{-1}\left(1 - \hat{P}_{fa}\right) \tag{60}$$

The CFAR threshold selection strategy is not optimal since it ensures a fixed false-alarm probability rather than a required SS performance given in terms of both P_{fa} and P_{md}. An optimal threshold could be selected as to minimize the total error probability of the system such that:

$$\hat{\lambda}_{X_{LE}} = \underset{\alpha}{\operatorname{argmin}}\big(w_0 P_{fa} + w_1 P_{md}\big) \tag{61}$$

where w_0 and w_1 are weighting coefficients that are chosen according to system priority. In order to solve this minimization problem, one can simply take the derivative equals to zero and the second derivative positive (i.e., concave). One can choose $w_0 = 0$ or $w_1 = 0$ to minimize one of the error probabilities. However, it is typical to choose $w_0 = w_1 = 0.5$ to minimize the sum of the error probabilities. Then $\lambda_{X_{LE}}$ should be selected such that it minimizes $P_{fa} + P_{md}$, i.e., its derivative is equal to zero:

$$\left(\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}\right)\lambda_{X_{LE}}^2 + \left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_0}{\sigma_0^2}\right)\lambda_{X_{LE}} + \left(\frac{\mu_0^2}{2\sigma_0^2} - \frac{\mu_1^2}{2\sigma_1^2} - ln\left(\frac{w_0\sigma_1}{w_1\sigma_0}\right)\right) = 0 \tag{62}$$

The optimal threshold is then given by finding the roots of (62):

$$\hat{\lambda}_{X_{LE}} = \frac{\mu_1\sigma_0^2 - \mu_0\sigma_1^2 + \sqrt{\sigma_0^2\sigma_1^2\Big((\mu_0-\mu_1)^2 - 2(\sigma_0^2-\sigma_1^2)ln\Big(\frac{w_0\sigma_1}{w_1\sigma_0}\Big)\Big)}}{\sigma_0^2-\sigma_1^2} \tag{63}$$

Observations on (63)*:*

- The optimal threshold in (63) requires the knowledge of the SNR value.
- The weights w_0 and w_1 could be tuned according to the required P_{fa} and P_{md}. A reduced w_0 means an improved P_{fa} which requires a larger $\hat{\lambda}_{X_{LE}}$. This is in line with (63). The same conclusion could be derived on w_1.
- It could be used in any of the scenarios to minimize the error probabilities. However, in the dynamic case, i.e., when *K* and/or *N* are to be selected dynamically, the designer should set up the target P_{fa} and P_{md} to be obtained while the optimization will consist in finding the minimal *K* and *N*. Accordingly, a fixed $(\hat{P}_{fa}, \hat{P}_d)$ is selected to evaluate the required *K* and *N* to achieve this performance. This will be discussed next.

B. Minimum Requirements for the LE Detector With Dynamic Parameters

For a target $(\hat{P}_{fa}, \hat{P}_d)$ and at a given SNR ρ, the CR system should optimize certain number of antennas for a certain number of samples. By eliminating $\lambda_{X_{LE}}$ from P_{fa} and P_d in (54) and (55) respectively, one can solve for *K* (or *N*) the following equation:

$$\sigma_0\Phi^{-1}\Big(1-\hat{P}_{fa}\Big) - \sigma_1\Phi^{-1}\Big(1-\hat{P}_d\Big) + \mu_0 - \mu_1 = 0 \tag{64}$$

TABLE II
REQUIRED *K* FOR A GIVEN *N* IN EXAMPLE 1

N	200	300	350	400	450	500	600	1000
K	14	10	8	8	7	6	5	4

TABLE III
REQUIRED *N* FOR A GIVEN *K* IN EXAMPLE 1

K	20	18	15	10	8
N	132	147	177	273	349

Hence, at a certain SNR value and for a target detection performance $(\hat{P}_{fa}, \hat{P}_d)$ the system can dynamically choose the couple (*K*, *N*) that most enhances its global performance (i.e., throughput, power saving etc.). Note that finding a general solution for (64) is not straightforward and thus we solve for numerical values.

Example 1: Consider the following example, $\hat{P}_{fa} = 0.1$, $\hat{P}_d = 0.9$ and $\rho = -15dB$, then we get Tables II and III.

Table II provides, for each value of *N*, the corresponding number *K* of antennas required in the sensing process to achieve the target performance. On the other hand, Table III provides for each value of *K* the corresponding number of samples *N* that should be acquired by each antenna to achieve the considered performance. It is worth mentioning that the values of *K* and *N* evaluated using (64) are real valued numbers and thus are rounded to $+\infty$. The designer can choose for instance $K = 8$ and $N = 350$ configuration or $K = 6$ and $N = 500$. The system may also be expected to have dynamic behavior in case of a change in the SNR values. This could be also achieved using (64).

C. Extension to SLE and SCN Detectors

This section extends the dynamic antenna exploitation on the LE detector and finds the minimum requirements for both SLE and SCN detectors within ELASTIC framework.

1) SCN Detector: The threshold could be computed using (57) and (58) according to a required error constraint. For example, for a target P_{fa}, the threshold is given by:

$$\hat{\lambda}_{X_{SCN}} = \theta_0 + \frac{\beta_0}{\xi_0}\Big(-1 + \big[-\ln\big(1-P_{fa}\big)\big]^{-\xi_0}\Big). \tag{65}$$

For a target $(\hat{P}_{fa}, \hat{P}_d)$, by considering the SCN detector and eliminating $\lambda_{X_{SCN}}$ from both (57) and (58) we get:

$$\theta_0 - \frac{\beta_0}{\xi_0} - \theta_1 + \frac{\beta_1}{\xi_1} + \frac{\beta_0}{\xi_0}\Big[-\ln\Big(1-\hat{P}_{fa}\Big)\Big]^{-\xi_0} - \frac{\beta_1}{\xi_1}\Big[-\ln\Big(1-\hat{P}_d\Big)\Big]^{-\xi_1} = 0 \tag{66}$$

where θ_i, β_i and ξ_i are the location, scale and shape parameters of the GEV distribution where $i = 0$ refers to $\mathcal{H}_0$ hypothesis and $i = 1$ refers to $\mathcal{H}_1$ hypothesis. Their expressions are provided by Theorems 3 and 4.

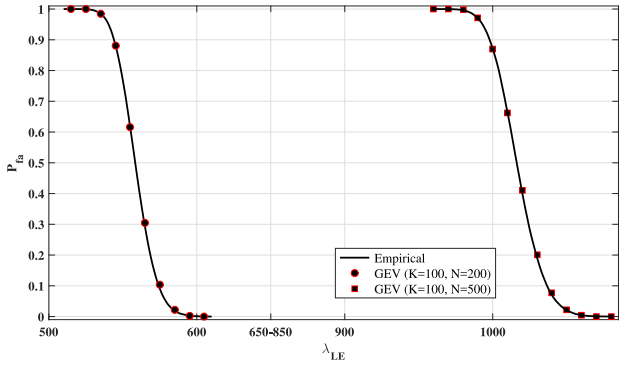


Fig. 1. Empirical P_{fa} of the LE detector in the asymptotic case and its corresponding GEV approximation for $K = 100$ and different values of N.

2) SLE Detector: Here, same derivations are provided. Hence, we can write:

$$\frac{\mu_{12}^{\mathcal{H}_0} - \Delta^2 r_0 \sigma_{12}^{\mathcal{H}_0}}{\mu^2_{T_n^{\mathcal{H}_0}} - \Delta^2 \sigma^2_{T_n^{\mathcal{H}_0}}} - \frac{\mu_{12}^{\mathcal{H}_1} - \Lambda^2 r_1 \sigma_{12}^{\mathcal{H}_1}}{\mu^2_{T_n^{\mathcal{H}_1}} - \Lambda^2 \sigma^2_{T_n^{\mathcal{H}_1}}} + \frac{\Delta\sqrt{m_v^{\mathcal{H}_0} - 2r_0\mu_{12}^{\mathcal{H}_0}\sigma_{12}^{\mathcal{H}_0} + \Delta^2\left[\sigma_{12}^{\mathcal{H}_0}\right]^2\left(r_0^2-1\right)}}{\mu^2_{T_n^{\mathcal{H}_0}} - \Delta^2\sigma^2_{T_n^{\mathcal{H}_0}}} - \frac{\Lambda\sqrt{m_v^{\mathcal{H}_1} - 2r_1\mu_{12}^{\mathcal{H}_1}\sigma_{12}^{\mathcal{H}_1} + \Lambda^2\left[\sigma_{12}^{\mathcal{H}_1}\right]^2\left(r_1^2-1\right)}}{\mu^2_{T_n^{\mathcal{H}_1}} - \Lambda^2\sigma^2_{T_n^{\mathcal{H}_1}}} = 0 \quad (67)$$

where $\Delta = Q^{-1}(\hat{P}_{fa})$ and $\Lambda = Q^{-1}(\hat{P}_d)$ with $Q^{-1}(.)$ is the inverse Q-function; the expressions $\mu_{12}^{\mathcal{H}_i} = \mu_{\lambda_1^{\mathcal{H}_i}}\mu_{T_n^{\mathcal{H}_i}}$, $\sigma_{12}^{\mathcal{H}_i} = \sigma_{\lambda_1^{\mathcal{H}_i}}\sigma_{T_n^{\mathcal{H}_i}}$, $m_v^{\mathcal{H}_i} = \mu^2_{T_n^{\mathcal{H}_i}}\sigma^2_{\lambda_1^{\mathcal{H}_i}} + \mu^2_{\lambda_1^{\mathcal{H}_i}}\sigma^2_{T_n^{\mathcal{H}_i}}$ and $\mathcal{H}_i$ refers to one of the two hypothesis $\mathcal{H}_0$ and $\mathcal{H}_1$ under which the expressions are calculated. Similar to the LE detector case, using (66) and (67) we can determine the minimum requirements of the system that could be used to achieve a target performance.

IX. Simulation and Discussion

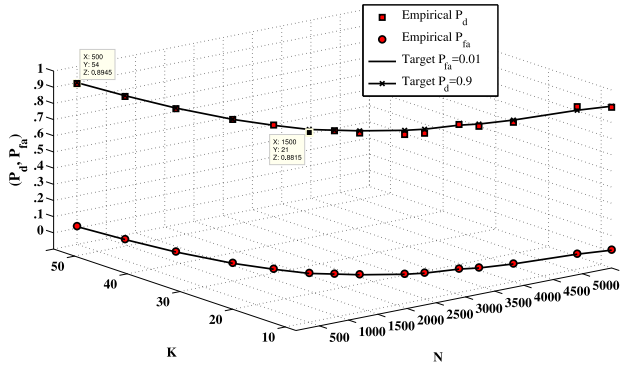
In this section, we verify the analytical derivation results through Monte-Carlo simulations. We validate the theoretical analysis presented in Sections IV-B–VIII-C. The simulation results are obtained by generating 10^5 random realizations of $\boldsymbol{Y}$.

A. Validation of Analytical Results

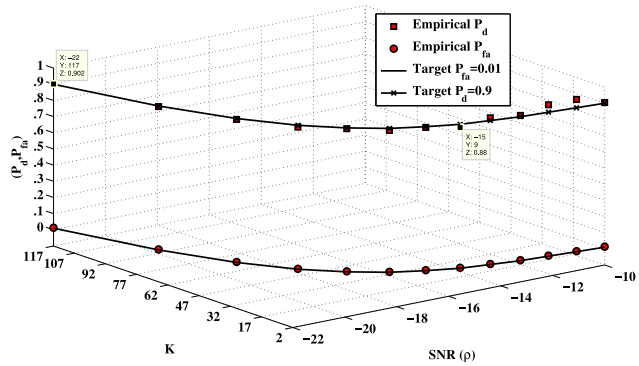
The asymptotic approximation provided by Proposition 1 is validated in Fig. 1. The results are taken for two different values of N while fixing $K = 100$ and $\rho = -10dB$. The results show a perfect match between empirical results and the proposed approximation.

B. Full and Partial Antenna Exploitation Scenarios

Figure 2 shows the empirical P_{fa} and P_d of the LE detector and its corresponding target values $(\hat{P}_{fa} = 0.01, \hat{P}_d = 0.9)$ while changing the other parameters. In Fig. 2(a) we set



(a) Fixed $\rho = -20dB$ and variable N



(b) Fixed $N = 500$ and variable ρ

Fig. 2. Empirical P_{fa} and P_d of the LE detector and its corresponding target values by using dynamic method.

$\rho = -20dB$ and we consider a variable N while in Fig. 2(b) we set $N = 500$ and we consider a variable ρ, as summarized by Algorithm 2. Simulation results show high accuracy of the analytical results evaluated using (64). The empirical P_{fa} is indeed 0.01 while the accuracy of the P_d increases as K increases which reflects the effect of AC under $\mathcal{H}_1$ hypothesis. The very small difference between the empirical and the target performance is due to the rounding of K to $+\infty$. The effect of rounding of K could be clearly noticed in Fig. 2(b) for $\rho = -12dB$ where the exact value is $K = 3.3$ and the rounded value is $K = 4$. Moreover, the results show that as N or ρ increases the number of antennas required to achieve the target performance decreases and hence these antennas could be exploited for other use.

Figure 3 shows a comparison between different scenarios, full, dynamic and fixed antenna exploitation. Therein, we consider a CR with $K = 200$ antennas. In the case of FPAE, only 5 antennas are used while K is dynamic in DPAE. In the latter, the value of K is calculated using (64) and depends on the SNR. It is obvious in this figure that the full exploitation of the antenna scenario achieves the best performance, however it is exploiting all the antennas all the time even if it not necessary. In this case, the threshold is calculated using the GEV approximation in Section IV-A. The use of a fixed number of antennas leads to a worst performance as the transmission conditions get worse. However, if the target performance is

Algorithm 2: Dynamic K Simulation Algorithm for LE Detector

Input: $\boldsymbol{Y}$, σ_η^2, $(\hat{P}_{fa}0.01, \hat{P}_d = 0.9)$, ρ, N
Output: (P_{fa}, P_d)
evaluate K w.r.t. ρ or N;
compute $\hat{\lambda}_{LE}$ w.r.t. K and N ;
generate $(K \times N)$ matrix $\boldsymbol{Y}$ for $\mathcal{H}_0$ and $\mathcal{H}_1$;
get λ_1 of $\boldsymbol{W} = \boldsymbol{Y}\boldsymbol{Y}^\dagger$ for $\mathcal{H}_0$ and $\mathcal{H}_1$;
evaluate $X_{LE_i} = \frac{\lambda_1}{\sigma_\eta^2}$ for $\mathcal{H}_0$ and $\mathcal{H}_1$;
if $X_{LE_0} > \hat{\lambda}_{LE} \rightarrow P_{fa}$;
if $X_{LE_1} > \hat{\lambda}_{LE} \rightarrow P_d$;
repeat;

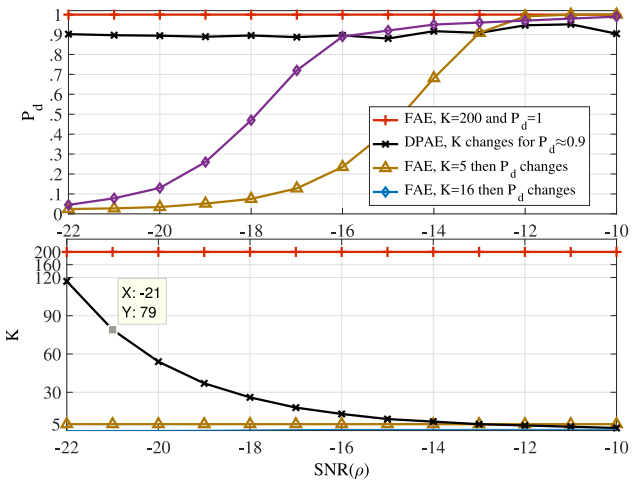
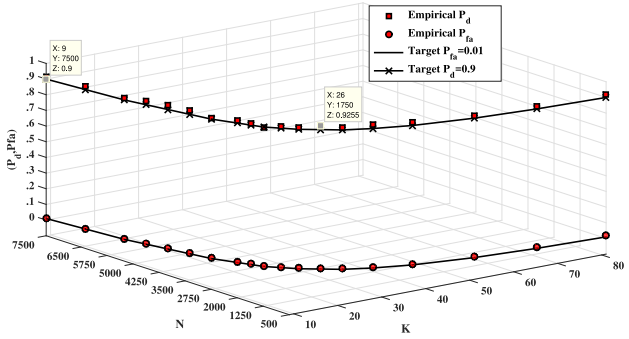


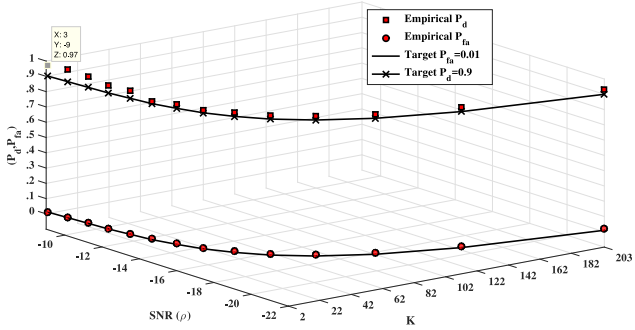
Fig. 3. Empirical P_d of the LE detector and the corresponding number of antennas, K, used for sensing in Full, Dynamic and Fixed methods w.r.t ρ and fixing $N = 500$.

well defined then it can be achieved all the time through an efficient use of the antennas. From the figure, the results show a tremendous decrease in the value of K as the SNR increases with approximately stable $P_d = 0.9$. Using dynamic exploitation scenario, the CR system will gain around 115 antennas that could be used for other purposes. In addition, the computational complexity of the sample covariance matrix and the eigenvalues in the detection algorithm will be decreased since the received matrix size, $(K \times N)$, is decreased. Moreover, results show that it is enough to use $K = 2$ for $N = 500$ starting from $\rho = -10dB$. Indeed, since $K = 2$ is the smallest value for the EBD then, for $\rho \leq -10dB$, the designer can fix $K = 2$ and starts to minimize N accordingly. In this case, as N decreases the threshold could be computed using the GEV approximation in the finite case (see Table I).

The SCN and SLE detectors are considered next using similar approach as in Algorithm 2. The results are shown in Figures 4 through 7. Figures 4 and 6 show P_d and P_{fa} of the SCN and SLE detectors respectively for different values of K. The latter is changed according to the target $(\hat{P}_{fa} = 0.01, \hat{P}_d = 0.9)$ and, the variation of SNR or N. Figures 5 and 7 show the variation of P_d with respect to SNR



(a) Fixed $\rho = -20dB$ and variable N



(b) Fixed $N = 500$ and variable ρ

Fig. 4. Empirical P_{fa} and P_d of the SCN detector and its corresponding target values by using dynamic method.

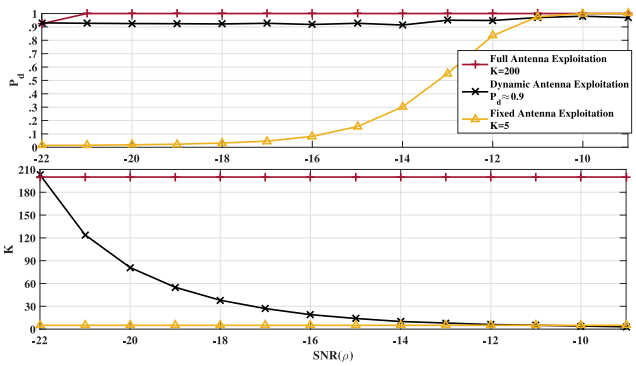
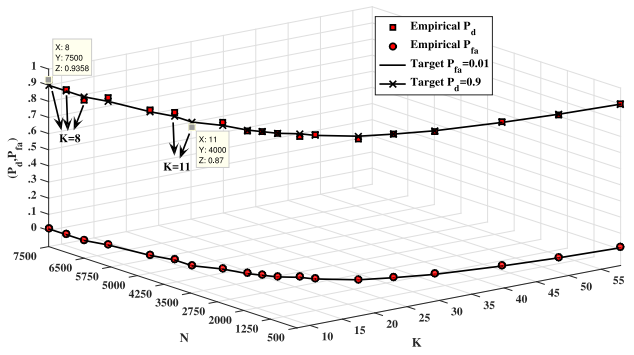
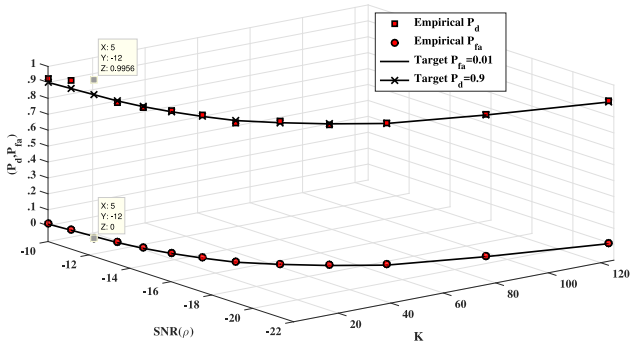


Fig. 5. Empirical P_d of the SCN detector and the corresponding number of antennas, K, used for sensing in Full, Dynamic and Fixed methods w.r.t ρ and fixing $N = 500$.

and K in the aforementioned scenarios. In the SLE detector case, we suppose that the mean error $\epsilon = 0$. Results show high accuracy of the analytical results evaluated using (66) and (67). Like the LE case, when K takes small values, the negligible difference between the empirical and target P_d is mainly due the rounding of K. Moreover, it could be noticed from Fig. 5 that at $\rho = -22dB$ and using $N = 500$ then the required value of K is 203. In this case and since the CR is equipped with 200 antennas, then the designer must fix $K = 200$ and starts to increase N accordingly to achieve the target performance. For larger ρ, these expressions are very useful and accurate to



(a) Fixed $\rho = -20dB$ and variable N



(b) Fixed $N = 500$ and variable ρ

Fig. 6. Empirical P_{fa} and P_d of the SLE detector and its corresponding target values by using dynamic method.

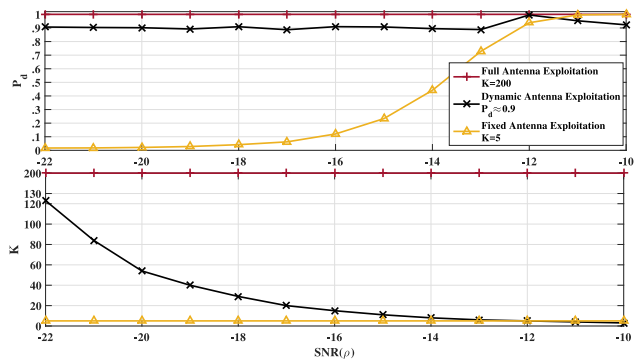


Fig. 7. Empirical P_d of the SLE detector and the corresponding number of antennas, K, used for sensing in Full, Dynamic and Fixed methods w.r.t ρ and fixing $N = 500$.

make a dynamic system in which the antennas are efficiently utilized.

C. LE, SCN and SLE Comparison

In this section, we provide a comparison between the considered detectors. For the LE detector, the noise power is supposed to be perfectly known while SCN and SLE detectors are totally-blind and do not require this knowledge. Figure 8 plots the Region of Convergence (ROC) of these detectors for $N = 500$, $K = 5$ and $\rho = -15dB$. Simulation results show that the LE detector outperforms the SLE detector and the

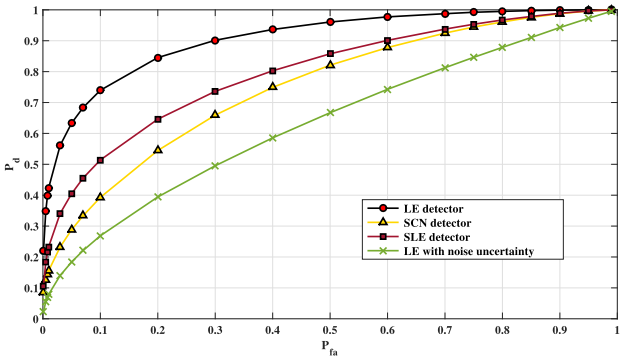


Fig. 8. ROC of the LE, SCN and SLE detectors when $k = 5$, $N = 500$ and $\rho = -15dB$.

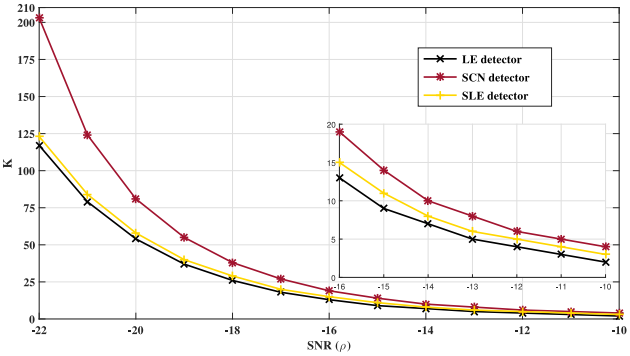


Fig. 9. Required K of the LE, SCN and SLE detectors for dynamic antenna exploitation w.r.t ρ and fixing $N = 500$, $P_d = 0.9$ and $P_{fa} = 0.01$.

SLE detector in turn outperforms the SCN detector. Indeed, LE is the optimal detector when noise variance is perfectly known while SLE detector is the optimal under the generalized likelihood ratio (GLR) criterion when noise variance is unknown [7], [10]. However, if noise power is not perfectly known, the performance of the LE detector will degrade and it might be even worse than the performance of SLE and SCN detectors as shown in Fig. 8 where $0.2dB$ noise variance uncertainty is considered.

Figures 9 and 10 show the minimum required number of antennas K for the LE, SCN and SLE detectors to achieve a target $(\hat{P}_{fa} = 0.01, \hat{P}_d = 0.9)$ when changing ρ and N respectively. Thee results are aligned with those of 9. In addition, it is also noticeable that the difference in K for LE and SLE are close whereas K required of the SCN detector is larger. Indeed, LE and SLE detectors are optimal when the noise power is perfectly known and noise power uncertain environments respectively. In this regard, it is worth mentioning that by considering LE detector with noise uncertain environment its performance will degrade and thus the number K of required antennas will increase.

D. Experimental Validation

The target of this section is to provide some experimental results done in a laboratory environment. In our experiment, we decided to adopt an off-the-shelf prototyping hardware,

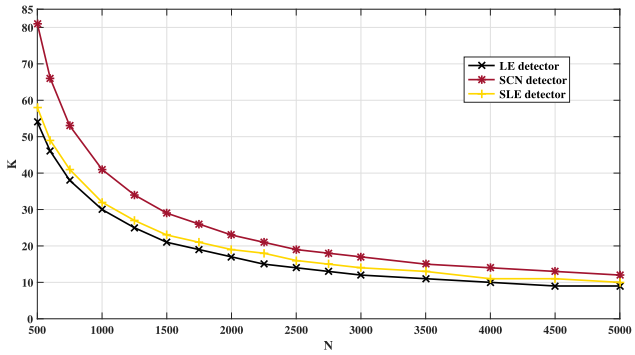
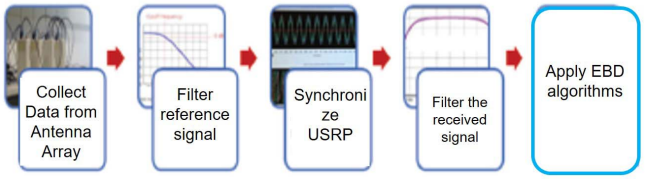
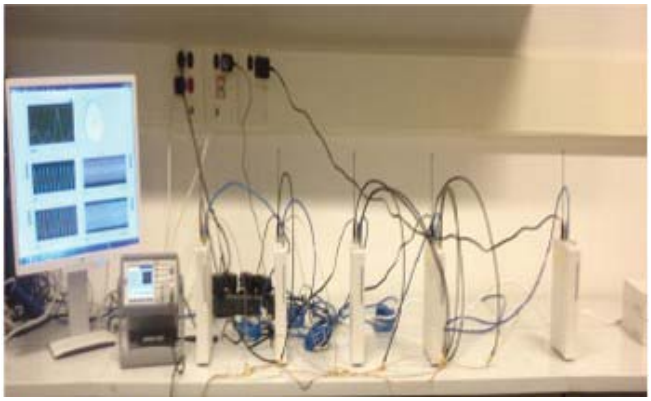


Fig. 10. Required K of the LE, SCN and SLE detectors for dynamic antenna exploitation w.r.t N and fixing $\rho = -20dB$, $P_d = 0.9$ and $P_{fa} = 0.01$.

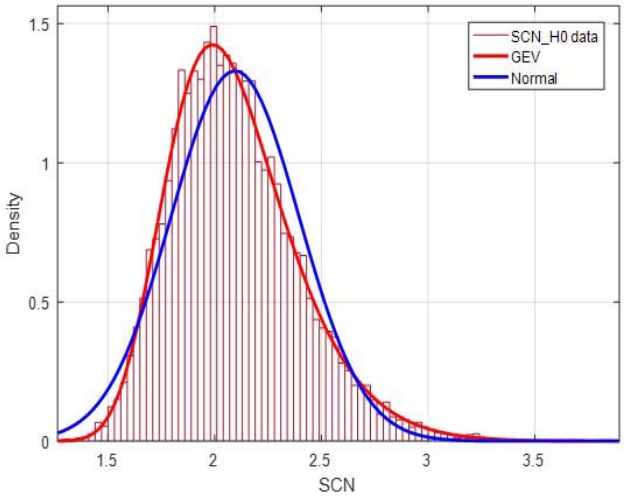


(a) Test Bench Process

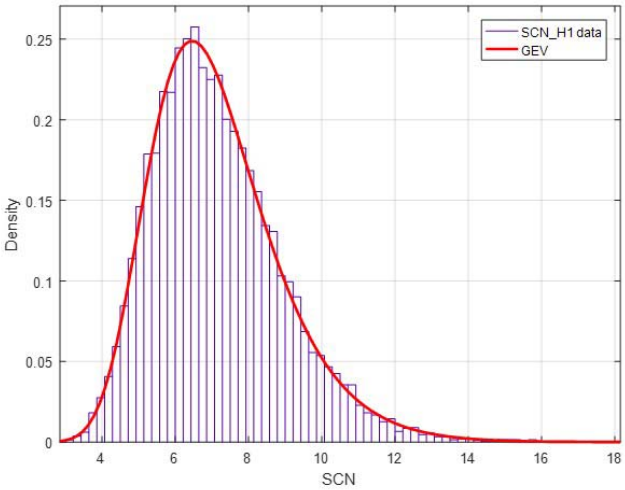


(b) Test Bench Experiment

Fig. 11. Lab experiment setup.



(a) SCN under H0



(b) SCN under H1

Fig. 12. SCN measurements and PDF fitting curves, $\rho = -10dB$, $K = 4$, $N = 500$.

called Universal Software Radio Peripheral (USRP), provided by National Instruments to design a multi-antenna system. However, as the number of available devices is limited to 6, the proof-of-concept has been done only in a selected showcase. Moreover, as these devices require synchronization, one USRP should be completely devoted for this purpose. Other synchronization problems (time stamp misalignment, common clock reference) have been also fixed. Finally, one USRP has been selected to emulate the transmitter.

To realize the test bench shown in Fig. 11(a), the following items have been used:

- 4 NI-USRP 2920 used as array elements
- 1 NI-USRP 2920 for synchronization
- 1 NI-USRP 2920 serving as transmitter
- A waveform generator

In this test bench, a 10 MHz sinusoidal signal modulated a signal with carrier frequency of 915 MHz. The signal was generated in intermittent times, each of 0.1 sec duration.

Fig. 12 provides the histogram of the SCN measurements under H0 and H1. It is very clear that the proposed GEV approximation fits very well the SCN measurements. Moreover, GEV approximation outperforms the Gaussian distribution proposed in literature. Finally, Fig. 13 shows that the experimental ROC curve is aligned with the theoretical derivations provided in terms of GEV.

X. Conclusion

CR equipped with massive antenna technology will achieve a significant increase in the performance of the multi-antenna SS detector. However, it might be enough to use fewer number of antennas for the sensing process and achieve a desired performance. In this paper, we have considered the LE, SCN and SLE detectors with two exploitation scenarios: FAE and PAE. The latter is further decomposed into two options: (i) fixed use and (ii) dynamic use. We extended our previous

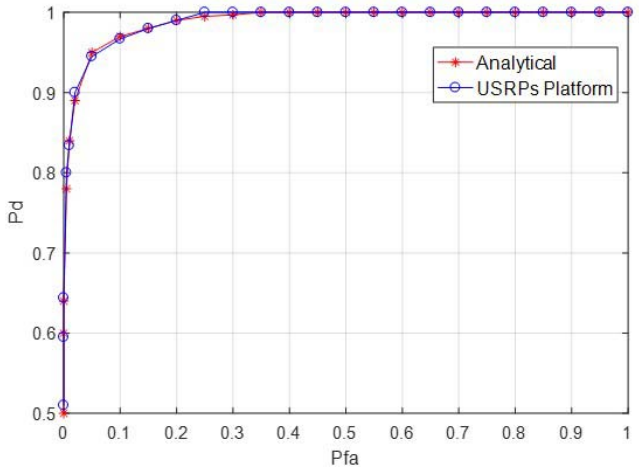


Fig. 13. Experimental SCN ROC curve.

work on the approximation of the distribution of the LE detector as a GEV distribution for both finite and asymptotic cases. Then, an optimized decision threshold that minimizes the error probabilities of the LE detector has been derived in both fixed and dynamic cases. Likewise, we used our recent work on SLE and SCN to derive the necessary sensing metrics and design parameters. This paper provided a mathematical framework to compute the minimum requirements of the CR system to achieve the desired performance of all exploitation scenarios in a massive antennas environment. Finally, it has been shown that the dynamic approach is the best solution for an efficient antenna exploitation. In this case, the CR design parameters should be dynamically tuned to meet the predefined performance and specifications.

REFERENCES

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

[2] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, 1st Quart., 2009.

[3] A. Ali and W. Hamouda, "Advances on spectrum sensing for cognitive radio networks: Theory and applications," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1277–1304, 2nd Quart., 2017.

[4] L. S. Cardoso, M. Debbah, P. Bianchi, and J. Najim, "Cooperative spectrum sensing using random matrix theory," in *Proc. IEEE Int. Symp. Wireless Pervasive Comput. (ISWPC)*, May 2008, pp. 334–338.

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

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

[7] B. Nadler, F. Penna, and R. Garello, "Performance of eigenvalue-based signal detectors with known and unknown noise level," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2011, pp. 1–5.

[8] B. Nadler, "On the distribution of the ratio of the largest eigenvalue to the trace of a Wishart matrix," *J. Multivariate Anal.*, vol. 102, no. 2, pp. 363–371, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0047259X10002113

[9] P. Bianchi, J. Najim, G. Alfano, and M. Debbah, "Asymptotics of eigen-based collaborative sensing," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Taormina, Italy, Oct. 2009, pp. 515–519.

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

[11] L. Wei, "Non-asymptotic analysis of scaled largest eigenvalue based spectrum sensing," in *Proc. 4th Int. Congr. Ultra Mod. Telecommun. Control Syst. Workshops (ICUMT)*, Oct. 2012, pp. 955–958.

[12] L. Wei, M. R. McKay, and O. Tirkkonen, "Exact Demmel condition number distribution of complex Wishart matrices via the Mellin transform," *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 175–177, Feb. 2011.

[13] L. Wei, O. Tirkkonen, P. Dharmawansa, and M. McKay, "On the exact distribution of the scaled largest eigenvalue," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Ottawa, ON, Canada, 2012, pp. 2422–2426. doi: 10.1109/ICC.2012.6364410.

[14] H. Kobeissi, Y. Nasser, A. Nafkha, O. Bazzi, and Y. Louët, "A simple formulation for the distribution of the scaled largest eigenvalue and application to spectrum sensing," in *Proc. CROWNCOM*, May 2016, pp. 284–293.

[15] H. Kobeissi, A. Nafkha, Y. Nasser, O. Bazzi, and Y. Louët, "Scaled largest eigenvalue in spectrum sensing: A simple form approach," *EAI Endorsed Trans. Cogn. Commun.*, vol. 3, no. 10, pp. 1–8, 2017.

[16] F. Penna, R. Garello, D. Figlioli, and M. A. Spirito, "Exact non-asymptotic threshold for eigenvalue-based spectrum sensing," in *Proc. IEEE 4th Int. Conf. CROWNCOM*, Hannover, Germany, Jun. 2009, pp. 1–5.

[17] F. Penna, R. Garello, and M. A. Spirito, "Probability of missed detection in eigenvalue ratio spectrum sensing," in *Proc. IEEE Int. Conf. Wireless Mobile Comput. Netw. Commun. (WIMOB)*, Oct. 2009, pp. 117–122.

[18] W. Zhang, G. Abreu, M. Inamori, and Y. Sanada, "Spectrum sensing algorithms via finite random matrices," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 164–175, Jan. 2012.

[19] H. Kobeissi, A. Nafkha, Y. Nasser, O. Bazzi, and Y. Louët, "Simple and accurate closed-form approximation of the standard condition number distribution with application in spectrum sensing," in *Proc. CROWNCOM*, 2016, pp. 351–362.

[20] H. Kobeissi, Y. Nasser, A. Nafkha, O. Bazzi, and Y. Louët, "On the detection probability of the standard condition number detector in finite-dimensional cognitive radio context," *EURASIP J. Wireless Commun. Netw.*, vol. 2016, no. 1, pp. 1–11, 2016. doi: 10.1186/s13638-016-0634-0.

[21] L. Bixio, G. Oliveri, M. Ottonello, M. Raffetto, and C. S. Regazzoni, "Cognitive radios with multiple antennas exploiting spatial opportunities," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4453–4459, Aug. 2010.

[22] H. Sarvanko, M. Höyhtyä, M. Matinmikko, and A. Mämmelä, "Exploiting spatial dimension in cognitive radios and networks," in *Proc. 6th Int. ICST Conf. Cogn. Radio Orient. Wireless Netw. Commun. (CROWNCOM)*, Osaka, Japan, Jun. 2011, pp. 360–364.

[23] H. Islam, Y.-C. Liang, and A. T. Hoang, "Joint power control and beamforming for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2415–2419, Jul. 2008.

[24] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 88–102, Feb. 2008.

[25] R. Zhang, F. Gao, and Y.-C. Liang, "Cognitive beamforming made practical: Effective interference channel and learning-throughput tradeoff," in *Proc. IEEE 10th Workshop Signal Process. Adv. Wireless Commun.*, Jun. 2009, pp. 588–592.

[26] T. Chen, H. Yuan, T. Zhao, Z. Zhang, and X. Ao, "Joint beamforming and power allocation for secure communication in cognitive radio networks," *IET Commun.*, vol. 10, no. 10, pp. 1156–1162, Jul. 2016.

[27] N. Seifi, R. W. Heath, M. Coldrey, and T. Svensson, "Adaptive multicell 3-D beamforming in multiantenna cellular networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6217–6231, Aug. 2016.

[28] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[29] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[30] W. H. Chin, Z. Fan, and R. Haines, "Emerging technologies and research challenges for 5G wireless networks," *IEEE Wireless Commun.*, vol. 21, no. 2, pp. 106–112, Apr. 2014.

[31] J. Choi *et al.*, "Millimeter-wave vehicular communication to support massive automotive sensing," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 160–167, Dec. 2016.

[32] H. Sun, A. Nallanathan, C.-X. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: A survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 74–81, Apr. 2013.

[33] K. Wanuga, N. Gulati, H. Saarnisaari, and K. R. Dandekar, "Online learning for spectrum sensing and reconfigurable antenna control," in *Proc. 9th Int. Conf. Cogn. Radio Orient. Wireless Netw. Commun. (CROWNCOM)*, Jun. 2014, pp. 508–513.

[34] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[35] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Rate gain region and design tradeoffs for full-duplex wireless communications," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3556–3565, Jul. 2013.
[36] E. Björnson, M. Kountouris, and M. Debbah, "Massive MIMO and small cells: Improving energy efficiency by optimal soft-cell coordination," in *Proc. ICT*, May 2013, pp. 1–5.
[37] L. Wei, P. Dharmawansa, and O. Tirkkonen, "Locally best invariant test for multiple primary user spectrum sensing," in *Proc. 7th Int. Cogn. ICST Radio Orient. Wireless Netw. Commun. (CROWNCOM)*, Jun. 2012, pp. 367–372.
[38] I. M. Johnstone, "On the distribution of the largest eigenvalue in principal components analysis," *Ann. Stat.*, vol. 29, no. 2, pp. 295–327, 2001.
[39] C. G. Khatri, "Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations," *Ann. Math. Stat.*, vol. 35, no. 4, pp. 1807–1810, Dec. 1964. doi: 10.1214/aoms/1177700403.
[40] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1050–1060, Apr. 2009.
[41] J. Baik and J. W. Silverstein, "Eigenvalues of large sample covariance matrices of spiked population models," *J. Multivariate Anal.*, vol. 97, no. 6, pp. 1382–1408, 2006.
[42] O. Tirkkonen and L. Wei, "Exact and asymptotic analysis of largest eigenvalue based spectrum sensing," in *Foundation of Cognitive Radio Systems*. Rijeka, Croatia: InTech, 2012.
[43] F. Bornemann, "On the numerical evaluation of distributions in random matrix theory: A review with an invitation to experimental mathematics," *Markov Processes Related Fields*, vol. 16, pp. 803–866, Apr. 2009.
[44] H. Kobeissi, A. Nafkha, Y. Nasser, Y. Louët, and O. Bazzi, "Approximating the standard condition number for cognitive radio spectrum sensing with finite number of sensors," *IET Signal Process.*, vol. 11, no. 2, pp. 145–154, 2017.
[45] L. Wei and O. Tirkkonen, "Analysis of scaled largest eigenvalue based detection for spectrum sensing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2011, pp. 1–5.
[46] G. Marsaglia, "Ratios of normal variables," *J. Stat. Softw.*, vol. 16, no. 4, pp. 1–10, May 2006.
[47] H. Artes, D. Seethaler, and F. Hlawatsch, "Efficient detection algorithms for MIMO channels: A geometrical approach to approximate ML detection," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2808–2820, Nov. 2003.
[48] J. Maurer, G. Matz, and D. Seethaler, "Low-complexity and full-diversity MIMO detection based on condition number thresholding," in *Proc. IEEE Int. Conf. (ICASSP)*, vol. 3, Apr. 2007, pp. 61–64.
[49] V. Erceg, P. Soma, D. S. Baum, and A. J. Paulraj, "Capacity obtained from multiple-input multiple-output channel measurements in fixed wireless environments at 2.5 GHz," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 1, May 2002, pp. 396–400.
[50] C. A. Tracy and H. Widom, "On orthogonal and symplectic matrix ensembles," *Commun. Math. Phys.*, vol. 177, no. 3, pp. 727–754, 1996. [Online]. Available: http://projecteuclid.org/euclid.cmp/1104286442
[51] O. N. Feldheim and S. Sodin, "A universality result for the smallest eigenvalues of certain sample covariance matrices," *Geom. Functional Anal.*, vol. 20, no. 1, pp. 88–123, 2010. doi: 10.1007/s00039-010-0055-x.

Hussein Kobeissi received the B.Eng. degree in electrical and electronic engineering (specialization telecommunications) and the Master of Research degree in signal, telecom, image and speech processing from Lebanese University (LU) in 2011 and 2012, respectively, and the Ph.D. degree in telecommunications and signal processing from LU and CentraleSupélec in 2016. He is an Assistant Professor with the CCE Department, International University of Beirut. His research interests include cognitive radios, spectrum sensing, massive MIMOS, and SDR.

Youssef Nasser (S'06–M'10–SM'11) received the master's and Ph.D. degrees in signal processing and communications from the National Polytechnic Institute of Grenoble, France, in 2003 and 2006, respectively, and the Habilitation á Diriger la Recherche degree, the highest education degree a scholar can achieve, in 2015. From 2003 to 2006, he was with the Laboratory of Electronics and Information Technologies (Laboratoire d'Electronique et de Technologies de l'Information -LETI), Grenoble, as a Research and Development Engineer. From 2006 to 2010, he was a Senior Research Engineer with the Institute of Electronics and Telecommunications of Rennes, France. He is currently with the Department of Electrical and Computer Engineering, American University of Beirut. He has published over 130 papers in international peer-reviewed journals and conferences.

Oussama Bazzi received the Bachelor of Engineering degree in electrical engineering from the American University of Beirut in 1987, the master's degree in electronics from the University of Valenciennes, France, in 1988, and the Ph.D. degree in electronics from Hainaut Cambresis, France, in 1992. He joined Lebanese University in 1996, where he is currently a Full Professor and the Head of the Research Group TSPI (Telecoms, Signal Processing and Images), Faculty of Sciences. His research interests are in the areas of signal processing, wireless and mobile radio communications, radio resource management, hybrid broadcast techniques, cooperative communication techniques, and cognitive radio networks.

Amor Nafkha (S'03–M'08–SM'16) received the B.Sc. (Eng.) degree in information and communications technology from the Higher School of Communications, Tunis, Tunisia, in 2001, and the Ph.D. degree in information and communications technology from the University of South Brittany, Lorient, France, in 2006. From 2006 to 2007, he was a Post-Doctoral Researcher with the Signal, Communication, and Embedded Electronics Research Group, CentraleSupélec, France, where he was actively involved in the reconfigurable hardware platform implementation for software-defined radio, co-authoring several contributions on FPGA dynamic partial reconfiguration. Since 2008, he has been an Associate Professor with CentraleSupélec. His research interests include multiuser and MIMO detection, hardware implementation, optical MIMO capacity, and spectrum sensing techniques.

Yves Louët (M'06) received the Ph.D. degree in digital communications and the HDR (Research Habilitation) degree from Rennes University, France, in 2000 and 2010, respectively. His Ph.D. thesis was focused on peak-to-average power reduction in OFDM with channel coding. He was a Research Engineer with SIRADEL Company, Rennes, France, from 2000 to 2002, where he was involved in channel propagation modeling for cell planning. In 2002, he became an Associate Professor with Supélec, Rennes, where he became a Full Professor in 2010. He is currently with CentraleSupélec. He is the Head of the Signal Communication Embedded Electronics Research Group with the Institute of Electronics and Telecommunications, Rennes, and the Vice-Chair of URSI Commission C. His research activities regard signal processing and digital communications applied to software and cognitive radio systems.

5.5 Asynchronous spectrum sensing in cognitive radio networks

Most of the studies on spectrum sensing techniques were conducted assuming a perfect synchronization between the secondary and the primary users. However, in practice, the synchronous assumption is very hard to satisfy in the context of real cognitive radio networks environment. Hereafter, we present a theoretical formulation of the standard condition number (SCN) based spectrum sensing technique under an asynchronous cognitive radio networks.

5.5.1 [C9]: Asynchronous standard condition number based detector

We consider a scenario where secondary user is equipped with two antennas and performed a spectrum sensing process under Rayleigh flat fading channels. The main contributions of this work are as follows:

- Under asynchronous scenario and unknown primary user activities, we derived new analytical expressions for the cumulative distribution function and probability density function of the SCN random variable. Moreover, we give the closed form expressions of detection and false alarm probabilities.
- The probability density function of the SCN significantly changes with increasing the mean number of samples received before transition from null to alternative hypothesis and vice versa during the sensing interval.
- As noise uncertainty problem in synchronous energy based spectrum sensing detection, under asynchronous PU traffic, we established the existence of the false-alarm probability wall, $P_{fa,wall}^{Asy}$ and we derived its analytic expression. Specifically, we showed that for SCN based detector there exists an P_{fa} below which detection becomes impossible in the absence of synchronization between primary and secondary users.
- Computer simulations are carried out to verify the analytical results, such as SCN-based detection performances and the false-alarm probability wall.

This work was published in IEEE Access in 2020 and included hereafter [J19].

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 2, 2020, accepted August 27, 2020, date of publication August 31, 2020, date of current version September 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3020500

Standard Condition Number Based Spectrum Sensing Under Asynchronous Primary User Activity

AMOR NAFKHA, (Senior Member, IEEE)
SCEE/IETR, CentraleSupélec, 35576 Cesson Sévigné, France
e-mail: amor.nafkha@centralesupelec.fr

This work was supported by internal funding from CentraleSupélec.

ABSTRACT Cognitive radio (CR) is a promising technology which enables the secondary user (SU) to sense and detect the presence or absence of the primary user (PU) in the frequency band of interest. Therefore, high detection probability is needed to ensure that primary user is adequately protected, while low false-alarm probability provides the opportunity of using free channel and a better throughput performance of secondary user. Most of the studies on spectrum sensing techniques were conducted assuming a perfect synchronization between the secondary and the primary users. However, in practice, the synchronous assumption is very hard to satisfy in the context of real cognitive radio networks environment. In this paper, we present a theoretical formulation of the standard condition number (SCN) based spectrum sensing technique under an asynchronous cognitive radio networks. We assume the case where the primary users generate asynchronous slotted traffic. The standard condition number distributions under null and alternative hypotheses are derived where the secondary user is equipped with two receive antennas. Under asynchronous primary user traffic, we establish the existence of a lower limit of the false-alarm probability below which we are unable to derive a decision threshold for the SCN-based detector.

INDEX TERMS Spectrum sensing, Markov process, standard condition number, detection probability, false-alarm probability, asynchronous/synchronous traffic.

I. INTRODUCTION

Due to the under-utilization problem of frequency spectrum, the concept of cognitive radio system has been developed as a reliable and effective solution [1]. The cognitive radio exploited unoccupied frequency bands owned by primary (*i.e* licensed) users. In order to benefit from the unused parts of the spectrum, a secondary (*i.e* unlicensed) user needs to monitor it to determine which portions are being unused by the primary user. Sensing an unused frequency band is a difficult task due to noise uncertainty, noise correlation, and the unknown primary user traffic.

Several spectrum sensing techniques like matched filter detector, energy detector, cyclostationary feature detector, high-order statistics based detector, covariance-based detector, and eigenvalue-based detector have been proposed and discussed in terms of their complexity-performance trade-offs in literature [2]–[4]. The energy detector (ED) is a non-coherent, blind and optimal detector for independent and identically distributed signal samples. It gives good detection performance with low computational complexity. However, ED is limited by its dependency on noise uncertainty [5]. Multiple antenna techniques currently are used in communications and their effectiveness has been proven in several aspects. In the context of cognitive radio, multiple antenna at the secondary user can effectively enhance signal transmission and spectrum sensing. Multiple antenna spectrum sensing can improve the primary user detection performance by exploring the spatial diversity [6], [7].

In most studies in literature, spectrum sensing is carried out when the behaviour of primary user remains constant either absent or present during the entire sensing time. This constraint implies that the secondary user and the primary user must be perfectly synchronized which may be difficult to achieve particularly in very low signal to noise ratio regimes. Recently, there have been several works which address the

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Feng.

effect of unknown primary user traffic patterns on the spectrum sensing [8]–[11]. Due to timing misalignment of the primary signal, the primary user may arrive or leave the licensed band at any time during the secondary user sensing time. In fact, this scenario appears when the primary user network has a high traffic rate or when a short secondary frame interval is used. In [10], authors address the sensing performance of energy detector against the random arrival and departure of primary user, while [11] study the effect of a single primary user traffic on secondary user's sensing performance as well as on the throughput. In [8] authors study the effect of multiple primary users traffic on the joint sensing-throughput performance of secondary user. In [9], the authors investigate the performance of the largest eigenvalue based detector under an unknown primary user traffic scenario, and where the secondary user is equipped with multiple receive antennas. They show that the performance gain due the spatial diversity is significantly reduced by the effect of a high transition probability of the primary user traffic.

To the best of authors' knowledge, the performance of standard condition number based spectrum sensing with unknown primary user traffic has not been investigated in the literature, partially due to the difficulty of deriving a simple closed-form expression of the probability density function of standard condition number in finite dimension. Our study is of real interest for two reasons. First, the used standard condition number detector is an efficient spectrum sensing technique in multi-dimensional cognitive radio systems since no a priori knowledge is needed. In fact, in contrast to energy detector and largest eigenvalue based detector, the standard condition number based detector is robust against noise uncertainty. Secondly, the SCN based detector is particularly cost-effective, which makes it a great tool for embedded cognitive receivers. In this paper, we consider a scenario where secondary user is equipped with two antennas and performs spectrum sensing under Rayleigh flat fading channels. The main contributions of this paper are as follows: **(i)** Under asynchronous scenario and unknown primary user activities, we derive new analytical expressions for the cumulative distribution function and probability density function of the SCN random variable. Moreover, we give the closed form expressions of detection and false alarm probabilities. **(ii)** The probability density function of the SCN significantly changes with increasing the mean number of samples received before transition from null to alternative hypothesis and vice versa during the sensing interval. **(iii)** Under asynchronous PU traffic, we establish the existence of the false-alarm probability wall, $P_{fa}^{Asy,wall}$, and we derive its analytic expression.

The remainder of this paper is organized as follows: In Section II, we describe the system model, and we derive mathematical expressions of SCN probability density functions for null and alternative hypotheses under synchronous and asynchronous PU traffic. Matlab simulations and performances analysis under synchronous and asynchronous primary user are provided to validate the accuracy of the derived expressions in Section III. Moreover, we establish the existence of the false-alarm probability wall under asynchronous PU traffic scenario. Finally, Section IV concludes the paper.

Notations: $\mathcal{CN}(0, \sigma^2)$ represents the complex Gaussian distribution with zero mean and σ^2 variance. $\mathbf{I}_x$ means the identity matrix of order x. $\mathbb{E}_x[.]$ denotes the expectation with respect to the random variable x. $\Gamma(.)$ defines the Gamma function as referred in [17, eq. 8.310.1], ${}_1F_1(.)$ is the confluent hypergeometric function as defined in [17, eq. 9.210.1], and ${}_2F_1(.)$ denotes the Gauss hypergeometric function as given in [17, eq. 9.14.2]. χ^2_{2N} represents a central chi-square distribution with $2N$ degrees of freedom and $\chi^2_{2N}(c)$ represents a non-central chi-square distribution with $2N$ degrees of freedom and a non-centrality parameter c.

II. SYSTEM MODEL AND THEORETICAL DERIVATIONS

To theoretically derive the impact of the traffic behavior of primary user on the SCN-based spectrum sensing detector. We start by analytically deriving the probability density functions of the SCN of the received covariance matrix when the time is slotted in both primary and secondary networks and slots of secondary network are synchronous to the primary network. Then, we derive the distributions of SCN of the received covariance matrix when the primary network and the secondary network are not aligned in their timing (*i.e.* asynchronous behavior).

A. SYNCHRONOUS SCN-BASED SPECTRUM SENSING

Let us consider a single secondary user equipped with two uncorrelated receive antennas which is aiming to detect the presence/absence of a single primary user during a sensing time τ_s. During this sensing interval, each of two antennas in secondary user receives N samples (*i.e.* $\tau_s = N/F_s$, where F_s is the sampling frequency). We model the spectrum sensing problem under a synchronous PU traffic, where the PU did not randomly appear or disappear during the sensing time, as a binary hypothesis testing problem: Given the received signal during τ_s, a decision is to be made between the null hypothesis $\mathcal{H}_0$ (*i.e. absence of a PU*) and the alternative hypothesis $\mathcal{H}_1$ (*i.e. presence of a PU*). Then, the received vector $\boldsymbol{y}$ at given instant $n \in [1, 2, \ldots, N]$ can be written, under both hypothesis, as:

$$\begin{cases} \mathcal{H}_0: & \boldsymbol{y}(n) = \boldsymbol{\eta}(n) \\ \mathcal{H}_1: & \boldsymbol{y}(n) = \boldsymbol{h}(n)\boldsymbol{s}(n) + \boldsymbol{\eta}(n). \end{cases} \tag{1}$$

where $\boldsymbol{\eta}(n)$ is a 2×1 complex circular white Gaussian noise, $\boldsymbol{h}(n)$ is a 2×1 complex vector that represents the channel's coefficient between the PU and each antenna at the cognitive radio receiver, modelled like a flat channel as a complex additive white Gaussian noise (AWGN), and $\boldsymbol{s}(n)$ stands for the primary user signal to be detected. After collection N sample at each antenna, the received signal matrix can be

written as:

$$\mathbf{Y} = \begin{pmatrix} y_1(1) & y_1(2) & \dots & y_1(n) \\ y_2(1) & y_2(2) & \dots & y_2(n) \end{pmatrix}, \quad (2)$$

and we define the received samples covariance matrix as: $\mathbf{W} = \frac{1}{2}\mathbf{Y}\mathbf{Y}^H$, where $(.)^H$ denotes the Hermitian complex conjugate. According to binary hypothesis testing, the covariance matrix $\mathbf{W}$ follows the Wishart distribution with two dimension and N degrees of freedom ($i.e.\mathbf{W} \sim \mathcal{CW}_2(N, \Sigma)$). Under the $\mathcal{H}_0$ null hypothesis, the covariance matrix $\mathbf{W}$ follows an uncorrelated central Wishart distribution. However, under the $\mathcal{H}_1$ hypothesis, the covariance matrix $\mathbf{W}$ follows an uncorrelated non-central Wishart distribution.

$$\begin{cases} \mathcal{H}_0: & \mathbf{W} \sim \mathcal{CW}_2(N, \mathbf{I}_2) \\ \mathcal{H}_1: & \mathbf{W} \sim \mathcal{CW}_2(N, \mathbf{I}_2, \Omega_2). \end{cases} \quad (3)$$

where $\Omega_2 = \mathbf{Z}\mathbf{Z}^H$ is called non-centrality matrix, with $\mathbf{Z} = [\boldsymbol{h}(1)\boldsymbol{s}(1), \dots, \boldsymbol{h}(N)\boldsymbol{s}(N)]$.

Let's consider the SCN-based statistic for decision process. The SCN-based spectrum sensing technique is a blind detector that uses the eigenvalues of the covariance matrix $\mathbf{W}$. Denoting by $\lambda_1 \geq \lambda_2 > 0$, the ordered eigenvalues of $\mathbf{W}$, the SCN is defined as the ratio of the largest to the smallest eigenvalue of the covariance matrix as follows: $\kappa = \frac{\lambda_1}{\lambda_2}$. In order to make decisions about the hypothesis ($\mathcal{H}_0$ or $\mathcal{H}_1$), we need to derive the distribution of the statistical metric, κ, under null and alternative hypotheses.

Theorem 1: Let $\kappa_N^{\mathcal{H}_0}$ be the SCN associated to the dual complex uncorrelated central Wishart matrix $\mathbf{W} \sim \mathcal{CW}_2(N, \mathbf{I}_2)$ with an arbitrary N, then the PDF and the CDF of $\kappa_N^{\mathcal{H}_0}$ can be expressed, respectively, as

$$f_{\kappa_N^{\mathcal{H}_0}}(x) = \frac{8\Gamma(N+\frac{1}{2})(x-1)^2(4x)^{N-2}}{\sqrt{\pi}\Gamma(N-1)(x+1)^{2N}}; \quad x \in [1, +\infty[\quad (4)$$

and

$$F_{\kappa_N^{\mathcal{H}_0}}(x) = \frac{N(1-x) + 2\,[{}_2F_1(1, -N, N, -x) - 1]}{(4x)^{1-N}(1+x)^{2N-1}} \times \frac{2\Gamma(N+\frac{1}{2})}{\sqrt{\pi}\Gamma(N+1)} - 1; \quad x \in [1, +\infty[\quad (5)$$

Proof: Under null hypothesis $\mathcal{H}_0$, the sample covariance matrix $\boldsymbol{W}$ follows a dual uncorrelated complex central Wishart distribution $\mathcal{CW}_2(N, \boldsymbol{I}_2)$. Let's $X = \kappa_N^{\mathcal{H}_0} = \lambda_1/\lambda_2$ be the standard condition number of $\boldsymbol{W}$. Then, using the characteristic polynomial of $\boldsymbol{W}$, we can express the standard condition number x as following

$$x = \frac{1 + \sqrt{1 - 4D/T^2}}{1 - \sqrt{1 - 4D/T^2}} \quad (6)$$

where D and T denote the determinant and the trace of the received covariance matrix $\mathbf{W}$, respectively. The exact probability density function of the random variable $A = 4D/T^2$ can be found in [13]. Making a change of variable A to $U = \sqrt{1-A}$ with Jacobian $J = 2u$, the probability density of U reads

$$f_U(u) = \frac{4\Gamma(N+\frac{1}{2})u^2(1-u^2)^{N-2}}{\sqrt{\pi}\Gamma(N-1)}; \quad u \in [0, 1] \quad (7)$$

Let X be the random variable defined as $X = \Phi(U) = \frac{1+U}{1-U}$. The function Φ is increasing for all U. Then, we can find the inverse function Φ^{-1} as follows:

$$u = \frac{x-1}{x+1} = \Phi^{-1}(x), \quad x \in [1, +\infty[\quad (8)$$

We can then find the derivative of Φ^{-1} with respect to x as

$$\frac{d\Phi^{-1}(x)}{dx} = \frac{2}{(x+1)^2} \quad (9)$$

Finally, the probability density of the standard condition number X can be expressed as:

$$\begin{aligned} f_X(x) &= f_U[\Phi^{-1}(x)] \quad . \quad |\frac{d\Phi^{-1}(x)}{dx}| \\ &= \frac{8\Gamma(N+\frac{1}{2})(x-1)^2(4x)^{N-2}}{\sqrt{\pi}\Gamma(N-1)(x+1)^{2N}}, \end{aligned} \quad (10)$$

where $x \in [1, +\infty[$. The cumulative distribution function of the standard condition number under null hypothesis $\mathcal{H}_0$ can be obtained by integrating the PDF of X. Thus, the CDF of X can be given as

$$F_X(x) = \frac{N(1-x) + 2\,[{}_2F_1(1, -N, N, -x) - 1]}{(4x)^{1-N}(1+x)^{2N-1}} \times \frac{2\Gamma(N+\frac{1}{2})}{\sqrt{\pi}\Gamma(N+1)} - 1 \quad (11)$$

We notice that the distribution density of the SCN under null hypothesis, $\kappa_N^{\mathcal{H}_0}$, only depends on the number of received samples N. Finally, the theorem 1 is proved. ■

Likewise, we introduce the following theorem which gives the exact probability density function of the standard condition number under the alternative hypothesis $\mathcal{H}_1$.

Theorem 2: Let $\kappa_{N,\omega_1}^{\mathcal{H}_1}$ be the SCN associated to the dual complex uncorrelated non-central Wishart matrix $\mathbf{W} \sim \mathcal{CW}_2(N, \mathbf{I}_2, \Omega_2)$ with an arbitrary N and rank one matrix Ω_2, then the PDF of $\kappa_{N,\omega_1}^{\mathcal{H}_1}$ is expressed as

$$\begin{aligned} f_{\kappa_{N,\omega_1}^{\mathcal{H}_1}}(x) = &\frac{2\Gamma(2N-1)(x-1)x^{N-2}e^{\frac{-\omega_1}{2}}}{\omega_1\Gamma(N-1)^2(x+1)^{2N-1}} \\ &\times [{}_1F_1(2N-1, N-1, \frac{\omega_1 x}{2(x+1)}) \\ &- {}_1F_1(2N-1, N-1, \frac{\omega_1}{2(x+1)})] \end{aligned} \quad (12)$$

where $\omega_1 = 2\rho N$ is non-zero eigenvalue of the non-centrality matrix Ω_2. The signal to noise ratio is defined by $\rho = \mathbb{E}_h\left[||\boldsymbol{hs}||^2\right]/\mathbb{E}_\eta\left[||\boldsymbol{\eta}||^2\right]$.

Proof: Let's denote by $x = \kappa_{\mathbf{W}}^{\mathcal{H}_1}$ the standard condition number of the dual complex uncorrelated central Wishart matrix $\boldsymbol{W} \sim \mathcal{CW}_2(N, \boldsymbol{I}_2, \boldsymbol{\Omega}_2)$, with ω_1 and ω_2 the eigenvalues of the non-centrality matrix $\boldsymbol{\Omega}_2$. For the rank one

matrix $\boldsymbol{\Omega}_2$, they are readily given as $\omega_1 = 2\rho N \neq 0$, ρ is the signal-to-noise ratio, while $\omega_2 = 0$. The rank one matrix assumption is justified by the fact that there is at most one primary user. Using the generalized Bartlett's decomposition when the non-centrality matrix $\boldsymbol{\Omega}_2$ has rank one [18, Th. 10.3.8], Then, let the Cholesky decomposition of $\boldsymbol{W}$ be $\boldsymbol{W} = \boldsymbol{Q}\boldsymbol{Q}^H$,where the lower triangular matrix $\boldsymbol{Q}$ has real positive entries $\boldsymbol{q}_{ii}$ on the diagonal. The entries $\boldsymbol{q}_{ij}$ of $\boldsymbol{Q}$ are all statistically independent and their distributions as follows:

$$\begin{cases} \boldsymbol{q}_{11}^2 \sim \chi_{2N}^2(\omega_1) \\ \boldsymbol{q}_{22}^2 \sim \chi_{2N-2}^2 \\ \boldsymbol{q}_{21}^2 \sim \chi_2^2. \end{cases} \tag{13}$$

Since $\boldsymbol{W} = \boldsymbol{Q}\boldsymbol{Q}^H$, we have

$$\begin{cases} T = \mathrm{tr}(\boldsymbol{W}) = \mathrm{tr}(\boldsymbol{Q}\boldsymbol{Q}^H) = \sum_{i\leq j}^{2} \boldsymbol{q}_{ij}^2 \\ D = \det(\boldsymbol{W}) = \prod_{i=1}^{2} \boldsymbol{q}_{ii}^2. \end{cases} \tag{14}$$

The first step of the proof is to derive the probability density function of the random variable $A = \frac{4D}{T^2}$. By applying a change of variables as follows

$$A = \frac{4\boldsymbol{q}_{11}^2\boldsymbol{q}_{22}^2}{(\boldsymbol{q}_{11}+\boldsymbol{q}_{22}+\boldsymbol{q}_{12})^2} \tag{15}$$

$$X_1 = \boldsymbol{q}_{11} \tag{16}$$

$$X_2 = \boldsymbol{q}_{22} \tag{17}$$

we can derive the joint distribution function of random variables A, X_1, and X_2 as

$$f_{A,X_1,X_2}(a,x_1,x_2) = \frac{x_1^{N-1}x_2^{N-2}\sqrt{\frac{x_1x_2}{a}}e^{-\sqrt{\frac{x_1x_2}{a}}}}{2^{2N}\Gamma(N)\Gamma(N-1)a} \times e^{-\frac{\omega_1}{2}}{}_0F_1(N,\frac{\omega_1 x_1}{4}) \tag{18}$$

where ${}_0F_1(\mu,x) = \sum_{j=0}^{+\infty}\frac{\Gamma(\mu)}{\Gamma(\mu+j)}\frac{x^j}{j!}$. Since we are interested in a only, $x_1 \in [0,+\infty[$ and $x_2 \in [\frac{x_1}{a}(1-\sqrt{1-a})^2, \frac{x_1}{a}(1+\sqrt{1-a})^2]$ are integrated out of the above expression; the following result is then the marginal probability density function corresponding to a:

$$f_A(a) = \frac{2\Gamma(2N-1)a^{N-2}e^{\frac{-\omega_1}{2}}}{\omega_1\Gamma(N-1)^2 4^{N-1}} \times[{}_1F_1(2N-1,N-1,\omega_1\frac{1+\sqrt{1-a)}}{4}) - {}_1F_1(2N-1,N-1,\omega_1\frac{1-\sqrt{1-a)}}{4})]; \tag{19}$$

where $a \in [0,1]$. The above equation completes the first step of the proof. Then, similar to proof of Theorem 1, we now proceed with the second step of the proof. Making a change of variable A to $X = \frac{1+\sqrt{1-A}}{1-\sqrt{1-A}}$ with Jacobian $J = \frac{4(x-1)}{(x+1)^3}$, the probability density function of the condition number under alternative hypothesis $\mathcal{H}_1$ reads

$$f_{\kappa_{N,\omega_1}^{\mathcal{H}_1}}(x) = \frac{2\Gamma(2N-1)(x-1)x^{N-2}e^{\frac{-\omega_1}{2}}}{\omega_1\Gamma(N-1)^2(x+1)^{2N-1}} \times[{}_1F_1(2N-1,N-1,\frac{\omega_1 x}{2(x+1)}) - {}_1F_1(2N-1,N-1,\frac{\omega_1}{2(x+1)})]; \tag{20}$$

where $x \in [1,+\infty[$. Finally, we obtain the desired result. ■

B. ASYNCHRONOUS SCN-BASED SPECTRUM SENSING

We consider here that the gathered samples can follow a mixture of the null and alternative hypotheses. We will rely on a model of the primary user traffic proposed in [9]. In the sequel, the traffic behavior of primary user is modeled by a two-state Markov process (or binary Markov process) as shown in Fig. 1. The first state is the busy state when the primary user is emitting a signal whereas the second state is the idle state when the primary user is absent. We denote by α the probability of going from idle to busy state, and β the probability of going from busy to idle state.

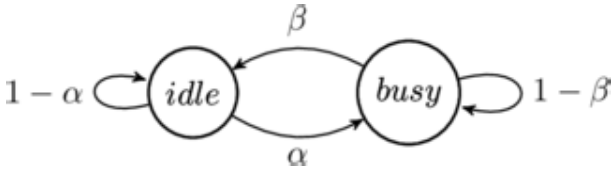


FIGURE 1. Markov chain representing the transitions between idle and busy states.

Then, we can express the mean length of idle and busy periods as, respectively, $M_i = 1/\alpha$ and $M_b = 1/\beta$. For simplicity, we assume that there is at most one state transition during the sensing time τ_s. This is backed up by the fact that, for usual mean length of idle and busy state and length of sensing interval, the probability of observing more than one transition is very small. Fig. 2 shows that as soon as the number of received samples N becomes greater than 20, the probability of having at most one transition becomes close to one and stable when $M = M_i = M_b$ is greater than N. Interested readers can refer to [14]–[16] for more details.

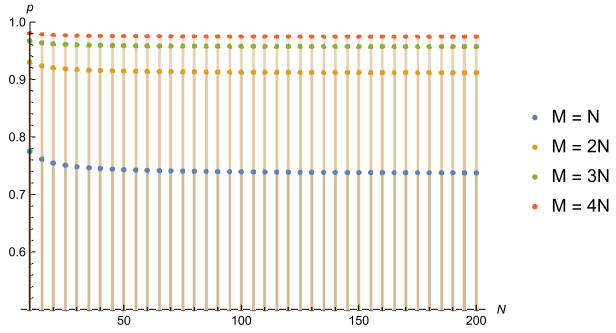


FIGURE 2. Probability of having at most one transition during the sensing time as function of the number of received samples N.

For the asynchronous primary user traffic model, we change the definitions of $\mathcal{H}_0$ and $\mathcal{H}_1$ slightly as presented in table 1.

Denoting by steady state (SS) the situation when the received matrix entries come from the same primary user's

TABLE 1. Hypothesis definitions under asynchronous PU traffic.

$SS \mid \mathcal{H}_0$	All received samples are collected during the $idle$ state.
$TS \mid \mathcal{H}_0$	The first received samples are collected during the $busy$ state and the last received sample belongs the $idle$ state.
$SS \mid \mathcal{H}_1$	All received samples are collected during the $busy$ state.
$TS \mid \mathcal{H}_1$	The first received samples are collected during the $idle$ state and the last received sample belongs the $busy$ state.

state (*i.e* idle or busy), and transient state (TS) the situation where the entries of received matrix are a mix between primary user's states.

Theorem 3: The probabilities that all received samples belong to the same steady state ($\mathcal{H}_0$ resp. $\mathcal{H}_1$) are given by

$$P_{SS|\mathcal{H}_0} = \frac{1}{1+\alpha\sum_{d=1}^{N}\frac{(1-\beta)^{N-d-1}}{(1-\alpha)^{N-d}}} \tag{21}$$

$$P_{SS|\mathcal{H}_1} = \frac{1}{1+\beta\sum_{d=1}^{N}\frac{(1-\alpha)^{N-d-1}}{(1-\beta)^{N-d}}} \tag{22}$$

Proof: Let $P_i = \frac{M_i}{M_i+M_b}$ and $P_b = \frac{M_b}{M_i+M_b}$ denote the probability of a sample being from a idle and busy states, respectively. Then, the joint probability mass of having no transition event and a idle state is written as:

$$P(SS, \mathcal{H}_0) = P_i(1-\alpha)^N \tag{23}$$

Similarly, the joint probability mass of having a transition from $\mathcal{H}_1$ to $\mathcal{H}_0$ and the $\mathcal{H}_0$ state is given by

$$P(TS, \mathcal{H}_0) = P_b\sum_{d=1}^{N-1}(1-\beta)^{N-d-1}\beta(1-\alpha)^d \tag{24}$$

Consequently, we obtain:

$$\begin{aligned} P_{SS|\mathcal{H}_0} &= \frac{P(SS,\mathcal{H}_0)}{P(SS,\mathcal{H}_0)+P(TS,\mathcal{H}_0)} \\ &= \frac{1}{1+\alpha\sum_{d=1}^{N}(1-\beta)^{N-d-1}(1-\alpha)^{d-N}} \end{aligned} \tag{25}$$

For $P_{SS|\mathcal{H}_1}$, a similar reasoning yields:

$$\begin{aligned} P_{SS|\mathcal{H}_1} &= \frac{P(SS,\mathcal{H}_1)}{P(SS,\mathcal{H}_1)+P(TS,\mathcal{H}_1)} \\ &= \frac{1}{1+\beta\sum_{d=1}^{N}(1-\alpha)^{N-d-1}(1-\beta)^{d-N}} \end{aligned} \tag{26}$$

■

From the equations (21) and (22), we can note that if M_i and M_b are equal to a given value M, (*i.e* $\alpha = \beta$), then $P_{SS|\mathcal{H}_0} = P_{SS|\mathcal{H}_1} = \frac{M}{M+N}$.

In order to study the transient states, we depict the timing misalignment among PU and SU in Fig. 3. Let us define two random variables: D_0, the number of received samples belonging the null hypothesis $\mathcal{H}_0$ during a transition from busy state to idle state, and D_1, the number of received samples belonging the alternative hypothesis $\mathcal{H}_1$ during a transition from idle state to busy state.

Theorem 4: The probabilities of receiving d_0 (*resp.* d_1) samples belonging the null hypothesis $\mathcal{H}_0$ (*resp.* the alternative hypothesis $\mathcal{H}_1$) during a transition from busy to idle state (*resp.* from idle to busy state) are given respectively by:

$$P_{D_0}(d_0) = \frac{(1-\beta)^{N-d_0-1}(1-\alpha)^{d_0-N}}{\sum_{d=1}^{N-1}(1-\beta)^{N-d-1}(1-\alpha)^{d-N}} \tag{27}$$

$$P_{D_1}(d_1) = \frac{(1-\alpha)^{N-d_1-1}(1-\beta)^{d_1-N}}{\sum_{d=1}^{N-1}(1-\alpha)^{N-d-1}(1-\beta)^{d-N}} \tag{28}$$

Proof: During a transient state from a busy to a idle state, the probability to have d_0 samples belonging to a null hypothesis $\mathcal{H}_0$ is given by:

$$P(TS, d_0, \mathcal{H}_0) = P_b(1-\beta)^{N-d_0-1}\beta(1-\alpha)^{d_0} \tag{29}$$

where $P_b = \frac{M_b}{M_i+M_b}$. As a consequence, the probability of receiving d_0 samples belonging the null hypothesis $\mathcal{H}_0$ during a transition from busy to idle state can be expressed as:

$$\begin{aligned} P_{D_0}(d_0) &= \frac{P(TS,d_0,\mathcal{H}_0)}{\sum_{d=1}^{N-1}P(TS,d,\mathcal{H}_0)} \\ &= \frac{(1-\beta)^{N-d_0-1}(1-\alpha)^{d_0-N}}{\sum_{d=1}^{N-1}(1-\beta)^{N-d-1}(1-\alpha)^{d-N}} \end{aligned} \tag{30}$$

Similarly, we derive $P_{D_1}(d_1)$ as

$$\begin{aligned} P_{D_1}(d_1) &= \frac{P(TS,d_1,\mathcal{H}_1)}{\sum_{d=1}^{N-1}P(TS,d,\mathcal{H}_1)} \\ &= \frac{(1-\alpha)^{N-d_1-1}(1-\beta)^{d_1-N}}{\sum_{d=1}^{N-1}(1-\alpha)^{N-d-1}(1-\beta)^{d-N}} \end{aligned} \tag{31}$$

■

From equations (27) and (28), we can denote that if M_i is equal to M_b, then the analytical expression of $P_{D_0}(d_0)$ and $P_{D_1}(d_1)$ can be reduced to: $P_{D_0}(d_0) = P_{D_1}(d_1) = \frac{1}{N-1}$, with $1 \leq d_0, d_1 < N$.

C. SCN DISTRIBUTIONS UNDER AN ASYNCHRONOUS PU TRAFFIC

Before deriving the distributions of the SCN under an asynchronous primary user traffic, the notations of several probability density functions are defined for clarity and understanding in table 2.

Using the fact that $f^{TS}_{\mathcal{H}_0,d} = f^{TS}_{\mathcal{H}_1,N-d}$, we can write the probability density functions, for a transient state (from busy to idle state, or vice versa) with unknown number of received samples coming from a busy state, as:

$$f^{TS}_{\mathcal{H}_0} = \sum_{d=1}^{N-1}P_{D_0}(d)f^{TS}_{\mathcal{H}_1,N-d} \tag{32}$$

$$f^{TS}_{\mathcal{H}_1} = \sum_{d=1}^{N-1}P_{D_1}(d)f^{TS}_{\mathcal{H}_1,d} \tag{33}$$

Consequently, we are able to derive the probability density function of the SCN under the null hypothesis $\mathcal{H}_0$ (resp. alternative hypothesis $\mathcal{H}_1$), given the considered asynchronous

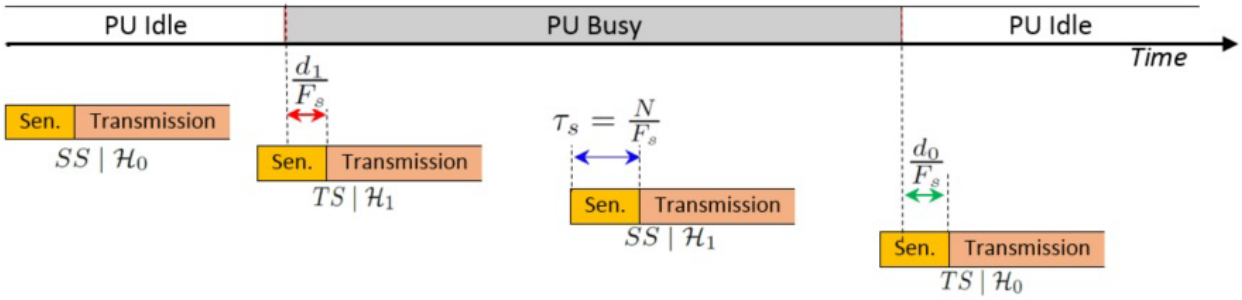


FIGURE 3. Timing misalignment model for four traffic patterns. The PU is slotted, whereas the SU is asynchronous to the PU. τ_S is the sensing time and F_S is the sampling frequency at the secondary user receiver.

TABLE 2. Probability density functions under transient and steady states.

$f^{SS}_{\mathcal{H}_0} = f_{\kappa_N^{\mathcal{H}_0}}$	Probability density function of the SCN when the N received samples belong idle state.
$f^{SS}_{\mathcal{H}_1} = f_{\kappa_{N,2\rho N}^{\mathcal{H}_1}}$	Probability density function of the SCN when the N received samples belong busy state.
$f^{TS}_{\mathcal{H}_0,d} = f_{\kappa_{N,2\rho(N-d)}^{\mathcal{H}_1}}$	Probability density function of the SCN for a transient from busy to idle state with d samples belonging the idle state.
$f^{TS}_{\mathcal{H}_1,d} = f_{\kappa_{N,2\rho d}^{\mathcal{H}_1}}$	Probability density function of the SCN for a transient from idle to busy state with d samples belonging the busy state.

PU traffic, as following:

$$f^{Asy}_{\kappa_N^{\mathcal{H}_0}} = P_{SS|\mathcal{H}_0} f_{\kappa_N^{\mathcal{H}_0}} + P_{TS|\mathcal{H}_0} f^{TS}_{\mathcal{H}_0} \tag{34}$$

$$f^{Asy}_{\kappa_N^{\mathcal{H}_1}} = P_{SS|\mathcal{H}_1} f_{\kappa_{N,\omega_1}^{\mathcal{H}_1}} + P_{TS|\mathcal{H}_1} f^{TS}_{\mathcal{H}_1} \tag{35}$$

If we assume that $M_i = M_b = M$ ($i.e.\ \alpha = \beta$), then Eq.(34) and Eq.(35) can be written as follows:

$$f^{Asy}_{\kappa_N^{\mathcal{H}_0}} = \frac{M f_{\kappa_N^{\mathcal{H}_0}} + \frac{N}{N-1}\sum_{d=1}^{N-1} f_{\kappa_{N,2\rho(N-d)}^{\mathcal{H}_1}}}{N+M} \tag{36}$$

$$f^{Asy}_{\kappa_N^{\mathcal{H}_1}} = \frac{M f_{\kappa_{N,2\rho N}^{\mathcal{H}_1}} + \frac{N}{N-1}\sum_{d=1}^{N-1} f_{\kappa_{N,2\rho(N-d)}^{\mathcal{H}_1}}}{N+M} \tag{37}$$

III. PERFORMANCES ANALYSIS UNDER SYNCHRONOUS/ASYNCHRONOUS TRAFFIC

In this section, we analyzed the gap in term of false alarm probability, detection probability, and receiver operating characteristics between the classical synchronous spectrum sensing schema and the asynchronous spectrum sensing schema as described in Section II-B. In the sequel of this paper, for simplicity, we assume that the mean length of idle PU state M_i and busy PU state M_b are equal to a given parameter M.

A. FALSE-ALARM PROBABILITY

Let us denote by λ^{syn} the decision threshold under a synchronous PU traffic, then the false alarm probability P^{Syn}_{fa}, defined as the probability of detecting the presence of primary user while it does not exist, is given by:

$$\begin{aligned} P^{Syn}_{fa} &= Prob(\kappa_N^{\mathcal{H}_0} \geq \lambda^{syn} \mid \mathcal{H}_0) \\ &= 1 - F_{\kappa_N^{\mathcal{H}_0}}(\lambda^{syn}) \end{aligned} \tag{38}$$

where $F_{\kappa_N^{\mathcal{H}_0}}(.)$ is the cumulative distribution function of the standard condition number as defined in equation (5). Since the probability density function of the SCN under $\mathcal{H}_0$ and synchronous traffic only depends on number of received samples N, we can characterize the SCN-based spectrum sensing detector by its detection performance for a fixed false-alarm probability. In order to measure the gap in term of false-alarm probabilities between the classical synchronous and the asynchronous PU traffic sensing, one must use λ^{syn} as a threshold to evaluate the false-alarm probability under the asynchronous scenario. Therefore, we can calculate the false alarm probability under asynchronous scenario as follows:

$$\begin{aligned} P^{Asyn}_{fa} &= \int_{\lambda^{syn}}^{+\infty} f^{Asy}_{\kappa_N^{\mathcal{H}_0}}(t)dt \\ &= \frac{M P^{Syn}_{fa}}{N+M} + \frac{N\sum_{d=1}^{N-1} \Xi^{Asy}(d,N,\rho)}{(N-1)(N+M)} \end{aligned} \tag{39}$$

where $\Xi^{Asy}(d,N,\rho) = \int_{\lambda^{syn}}^{+\infty} f_{\kappa_{N,2\rho(N-d)}^{\mathcal{H}_1}}(t)dt$. It is important to note that the limit of $\Xi^{Asy}(d,N,\rho)$ as ρ approaches zero is P^{Syn}_{fa}.

In Fig. 4, the probability density functions $f_{\kappa_N^{\mathcal{H}_0}}(x)$ and $f^{Asy}_{\kappa_N^{\mathcal{H}_0}}(x)$ are depicted for $N = 128$ and $\alpha = \beta$. Under asynchronous PU traffic the probability density function of the SCN, $f^{Asy}_{\kappa_N^{\mathcal{H}_0}}(x)$, is shown for three different values of the SNR ρ. To quantify the similarity of the distribution of the SCN under the synchronous null hypothesis and the asynchronous one, we use the Hellinger distance [19, Definition 2.3] as a metric to evaluate the difference between the mentioned distributions. The Hellinger distance ranges from 0 to 1, where 1 means that the probability distributions are completely different, however 0 means that two probability distributions are identical. The Hellinger distance values obtained are 0.00255, 0.02144 and 0.07606 corresponding to SNR equal to -6 dB, -3 dB, and 0 dB respectively. Therefore, it is easy to conclude

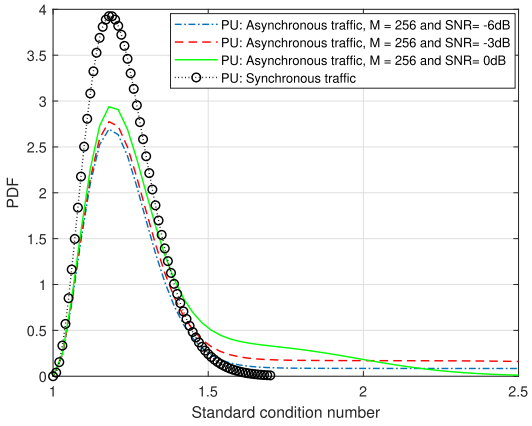


FIGURE 4. PDF under $\mathcal{H}_0$ of the SCN for different values of the SNR and $N = 128$ number of samples.

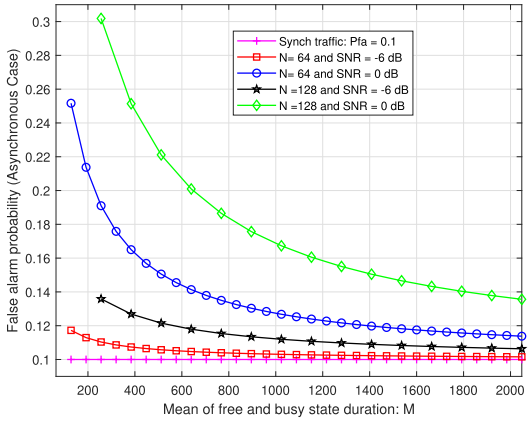


FIGURE 5. False alarm probabilities under Synchronous and Asynchronous PU traffic versus SNR.

that as the SNR tends to minus infinity (no signal), the SCN probability density function $f^{Asy}_{\kappa_N|\mathcal{H}_0}$ converges to $f_{\kappa_N|\mathcal{H}_0}$.

In Fig. 5, we evaluate the effect of the mean length M on the false-alarm probabilities of synchronous and asynchronous scenarios. We can see that the false-alarm probability under asynchronous PU traffic approaches the false-alarm probability under synchronous case as the mean length $M \to \infty$. This can be justified by recalling that the P^{Asyn}_{fa}, derived in (39), is a function of M. As a matter of fact, asymptotically when $M \to +\infty$ and a fixed value of N, the first term tends to P^{Syn}_{fa}, however the second term vanishes and approaches to zero.

Fig. 6 depicts the variation of false-alarm probability under asynchronous PU activity as a function of signal-to-noise ratio in following parameters pairs: $(N = 64, M = 128)$, $(N = 64, M = 256)$, $(N = 128, M = 256)$, and $(N = 128, M = 512)$. Two main observations can be made. First, the false-alarm probability under asynchronous PU traffic converges at low SNR to 0.1, which is the false-alarm probability under synchronous PU traffic. The convergence at low SNR can be explained using (39) by noting that at small SNR value ($i.e$ ρ approaches zero), the function $\Xi^{Asy}(d, N, \rho)$ converges to P^{Syn}_{fa}. Second, depends from parameters N and M, the false-alarm probability under asynchronous PU traffic converges to constant values as the signal-to-noise ratio

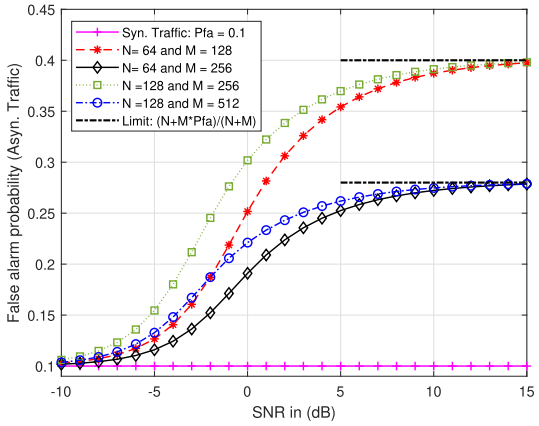


FIGURE 6. False alarm probabilities under Synchronous and Asynchronous PU traffic versus SNR.

increases. Those limits can be deduced from (39) as follows: For high SNR, the function $\Xi^{Asy}(d, N, \rho)$ converges to the unity as $\rho \to +\infty$. Thus, for the high SNR regime, we have the following approximation of P^{Asyn}_{fa} as:

$$P^{Asyn}_{fa,wc} = \lim_{\rho \to \infty} P^{Asyn}_{fa} \approx \frac{N + MP^{Syn}_{fa}}{N + M}, \tag{40}$$

where $P^{Asyn}_{fa,wc}$ denotes the "worst-case" false alarm probability under an asynchronous PU traffic. We can also note that, for fixed SNR and mean length M, the false-alarm probability under asynchronous PU traffic increases as the number of received samples N increases. Intuitively, this can be explained by the fact that, if N increases, the probability that PU is first active at the begin of sensing interval and then becomes idle by the end of the sensing interval or vice-versa, increases.

Fig.7 shows the false-alarm probability under asynchronous PU traffic in worst-case condition, $P^{Asyn}_{fa,wc}$, for different pairs (N,M). The horizontal axis represents the false-alarm probability under synchronous PU traffic P^{Syn}_{fa}. As can be seen in Fig. 7, for a fixed N, as the mean length M increases the false-alarm probability $P^{Asyn}_{fa,wc}$ is improved as the part of the distribution under $\mathcal{H}_0$ to the left of the threshold λ^{syn} increases. Fig. 7 is well-known as probability–probability (P-P) plot tool for the applied statistics. A P-P plot that lies on the first bisector line indicates that the two studied cumulative distributions are identical, whereas a P-P plot lying strictly above the first bisector line, as the case in Fig.7, indicates that $P^{Asyn}_{fa,wc}$ stochastically dominates[1] P^{Syn}_{fa}.

[1]Let X_1 and X_2 two random variables. We say that X_1 first-order stochastically dominates X_2 if and only if $Pr\,(X_1 > c) \geq Pr\,(X_2 > c)$, whatever the value of c.

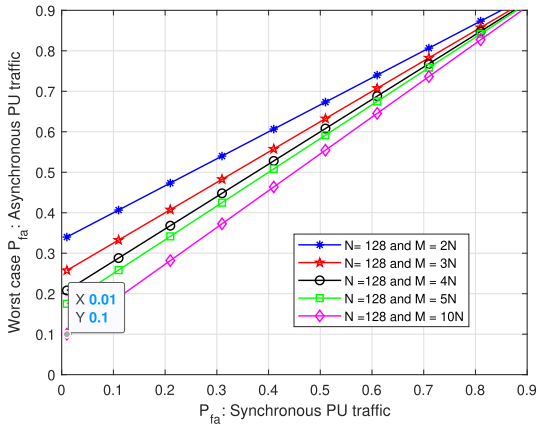


FIGURE 7. Worst case of false-alarm probability under asynchronous PU traffic versus P_{fa}^{Syn}. The graph compares five different pairs (N, M).

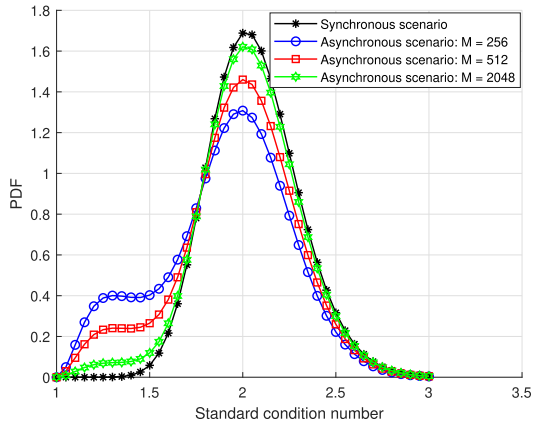


FIGURE 8. Probability density function of the SCN under $\mathcal{H}_1$ as function of M with $N = 128$ and SNR = 0dB.

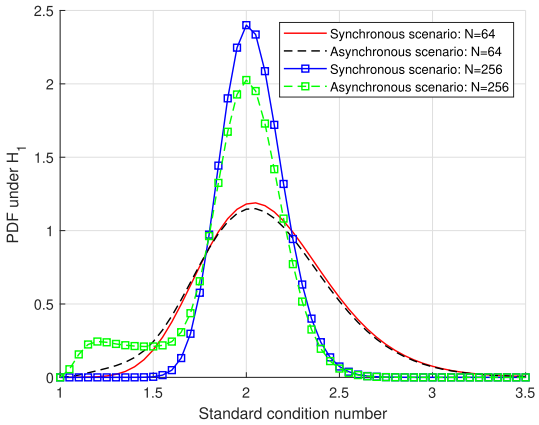


FIGURE 9. Probability density function of the SCN under $\mathcal{H}_1$ as function of N with $M = 1024$ and SNR = 0dB.

B. DETECTION PROBABILITY

In order to guarantee a high reuse probability of the unused spectrum by the secondary user, we adopt the constant false alarm rate (CFAR) based detection criterion. According to the CFAR principle, the probability of false-alarm is fixed to a small value while the detection probability should be maximized. Let us fix the false-alarm probability to a given value γ. Then, the detection probability my be obtained for the two different primary user traffic scenarios. In the case of a synchronous PU traffic, the detection probability is given by

$$P_d^{Syn} = \int_{\lambda^{syn}}^{+\infty} f_{\kappa_{N,\omega_1}}^{\mathcal{H}_1}(x)dx \tag{41}$$

where $\lambda^{syn} = F_{\kappa_N}^{-1\,\mathcal{H}_0}(1-\gamma)$. In case of asynchronous PU traffic, the detection probability can be expressed as

$$P_d^{Asyn} = \int_{\lambda^{Asyn}}^{+\infty} f_{\kappa_N}^{Asy\,\mathcal{H}_1}(x)dx \tag{42}$$

where the decision threshold λ^{Asyn} was derived for the worst-case condition as

$$\lambda^{Asyn} = F_{\kappa_N}^{-1\,\mathcal{H}_0}\left((1-\gamma)(1+\frac{N}{M})\right) \tag{43}$$

From (43), it can be seen that the detection threshold under asynchronous PU traffic depends on the pair (N,M). Moreover, given a fixed number of received samples N, λ^{Asyn} converges to λ^{syn} as the mean length M tends to infinity.

Fig. 8 shows the probability density function of the SCN under synchronous and asynchronous primary user traffic with fixed SNR equal to 0 dB, number of received samples $N = 128$, and different values of the the mean length of idle/busy state M. As same as in Fig. 4, the Hellinger distance is used to quantify the similarity between different probability distributions. Given fixed SNR and $N = 128$, the Hellinger distances between $f_{\kappa_{N,2\rho N}}^{\mathcal{H}_1}$ and $f_{\kappa_N}^{Asy\,\mathcal{H}_1}$ are quite significant, being 0.0751, 0.0405, and 0.0095 for $M = 256$, $M = 512$ and $M = 2048$, respectively. Moreover, analyzing the expression (37), it can be seen that as M approaches infinity, $f_{\kappa_N}^{Asy\,\mathcal{H}_1}$ approaches $f_{\kappa_{N,2\rho N}}^{\mathcal{H}_1}$ due to the fact that the second term tends to zero.

Under both scenarios, the numerically evaluated probability density function of the SCN under $\mathcal{H}_1$ hypothesis, for fixed M and SNR values, as function of the number of received samples N is plotted in Fig. 9. Given fixed SNR and $M = 1024$, the probability density functions curves change slightly between both scenarios when $N = 64$ (i.e. Hellinger distance = 0.0036). However, in the case when $N = 256$, we note that the probability density functions curves are significantly different from each other (i.e. Hellinger distance = 0.0567), which indicates that the detection probabilities values under both scenarios will be very different too. Theoretically, for any N, as the ratio M/N tends to infinity, the probability density function of the SCN under asynchronous scenario converges to $f_{\kappa_{N,2\rho N}}^{\mathcal{H}_1}$.

Fig. 10 shows a comparison between the considered two scenarios, asynchronous and synchronous PU traffic, in term of detection probabilities as function of SNR. To keep the performance comparison consistent, we fixed the false-alarm probability target to the value of 0.1 for all graphs (*i.e.* $P_{fa}^{Syn} = P_{fa,wc}^{Asyn} = 0.1$). It is obvious that under synchronous scenario,

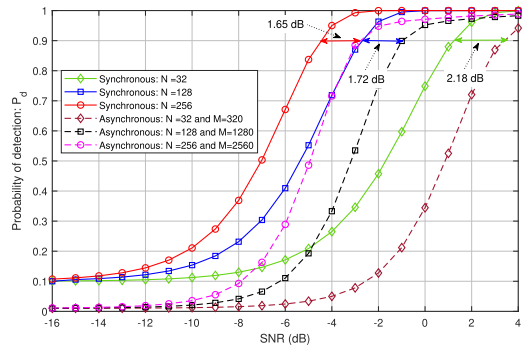


FIGURE 10. Probability of detection versus SNR with target false-alarm probability equal to 0.1, and different (N, M) pairs.

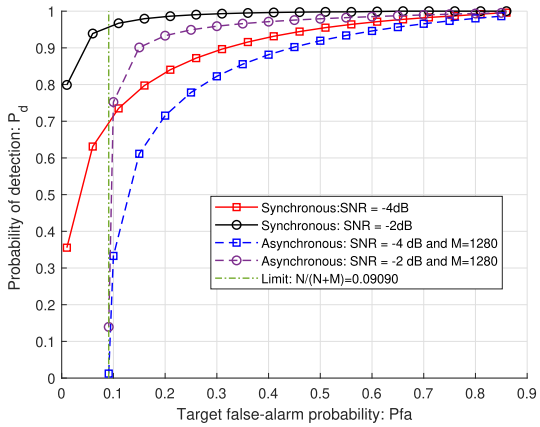


FIGURE 11. The ROC curves for $N = 128$. Solid lines represent different SNR values, -4 dB and -2 dB, under synchronous PU traffic, dashed lines represent different SNR values, -4 dB and -2 dB, under asynchronous PU traffic for fixed $M = 10 \times N$, and dot-dashed line represent the theoretical lower bound of false-alarm probability under asynchronous scenario with $N = 128$ and $M = 1280$.

the SCN-based detector achieves best performances. Moreover, as expected, the detection performance improves with the increase of the number of received samples N during the sensing time. Using (40), we are able to fix the mean duration of primary user busy/idle state M to $10 \times N$ in order to ensure that the target false-alarm probability is at most 0.1 under asynchronous scenario. From Fig. 10, we can see that the detection performance of the SCN-based sensing technique degrades severely under asynchronous PU traffic. Moreover, it can be noted that the SNR gap between synchronous and asynchronous scenarios decreases as M increases. Thus, for a detection probability of 0.9, the SNR gap between the two considered scenarios with $(N = 32, M = 320)$, $(N = 128, M = 1280)$, and $(N = 256, M = 2560)$, is 2.18 dB, 1.72 dB, and 1.65 dB respectively.

C. RECEIVER OPERATING CHARACTERISTIC

The SCN-based detector performance can be represented by the receiver operation characteristic (ROC) curves, in which the detection probabilities and false alarm probabilities are plotted. Thus, the ROC curves indicate the practical regions inside of which the SCN-based detector is capable of providing reliable results. Fig. 11 shows that, for the SCN-based detector over Rayleigh fading channel, ROC curves move to the upper left corner as increasing SNR under any scenarios (synchronous/asynchronous), confirming better overall detection performance. By comparing the results collected in Fig. 11, it is noticeable that the no synchronization between PU and SU during the sensing time decreases the probability of detection for a given false alarm probability. We note that, the main observation is in Fig. 11, which shows the existence of the "P_{fa} *wall*" phenomenon, below which SCN-based detector will fail to be robust under asynchronous scenario. The position of this "P_{fa} *wall*" is determined by N and M, and it can be derived as follows:

$$P_{fa}^{Asy,wall} = \frac{N}{N+M} \tag{44}$$

IV. CONCLUSION

From the view of an opportunistic spectrum access, correct detection of the absence or presence of primary user is a vital component. Traditionally, perfect synchronization between PU and SU is assumed, which is infeasible in the absence of a centralized control unit. In this paper, we discuss the impact of an asynchronous situation, which means SU have no idea about the communication time of PU, on the detection performance of SCN-based technique for spectrum sensing. By assuming that the PU activity follows with an idle/busy Markov chain, we derived the false-alarm and detection probabilities under asynchronous scenario. The existence of the "P_{fa} *wall*" has been established, and its exact expression has been derived. While the present paper considered a single SU and a single primary spectrum band, our main future objective is to extend the presented analytical framework to the case of asynchronous cooperative spectrum sensing where there are no synchronization among the different secondary users, and also between them and the primary user.

ACKNOWLEDGMENT

The author gratefully acknowledges the helpful discussions with Gurvan Priem, an engineer at Biosency company.

REFERENCES

[1] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
[2] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, 1st Quart., 2009.
[3] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: Challenges and solutions," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, pp. 1–15, Dec. 2010.
[4] A. Ali and W. Hamouda, "Advances on spectrum sensing for cognitive radio networks: Theory and applications," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1277–1304, 2nd Quart., 2017.
[5] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 4–17, Feb. 2008.
[6] A. Nafkha and B. Aziz, "Closed-form approximation for the performance of finite sample-based energy detection using correlated receiving antennas," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 577–580, Dec. 2014.
[7] H. Kobeissi, A. Nafkha, Y. Nasser, Y. Louët, and O. Bazzi, "Approximating the standard condition number for cognitive radio spectrum sensing with finite number of sensors," *IET Signal Process.*, vol. 11, no. 2, pp. 145–154, Apr. 2017.
[8] H. Pradhan, S. S. Kalamkar, and A. Banerjee, "Sensing-throughput tradeoff in cognitive radio with random arrivals and departures of multiple primary users," *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 415–418, Mar. 2015.

[9] P. Dhakal, S. K. Sharma, S. Chatzinotas, B. Ottersten, and D. Riviello, "Effect of primary user traffic on largest eigenvalue based spectrum sensing technique," in *Proc. Int. Conf. Cogn. Radio Oriented Wireless Netw. (CrownCom)*, 2016, pp. 67–78.
[10] I. Atef, A. Eltholth, A. S. Ibrahim, and M. S. El-Soudani, "Energy detection of random arrival and departure of primary user signals in cognitive radio systems," in *Proc. IEEE Int. Conf. Comput. as a Tool (EUROCON)*, Sep. 2015, pp. 1–6.
[11] M. Jin, Q. Guo, Y. Li, J. Xi, and Y. Yu, "Energy detection with random arrival and departure of primary signals: New detector and performance analysis," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10092–10101, Nov. 2017.
[12] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of wishart matrices," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1050–1060, Apr. 2009.
[13] D. K. Nagar, S. K. Jain, and A. K. Gupta, "Distribution of LRC for testing sphericity of a complex multivariate Gaussian model," *Int. J. Math. Math. Sci.*, vol. 8, no. 3, pp. 555–562, 1985.
[14] L. T. Tan and L. B. Le, "Distributed MAC protocol design for full-duplex cognitive radio networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Diego, CA, USA, Dec. 2014, pp. 1–6.
[15] C. Fu, Y. Li, Y. He, M. Jin, G. Wang, and P. Lei, "An inter-frame dynamic double-threshold energy detection for spectrum sensing in cognitive radios," *EURASIP J. Wireless Commun. Netw.*, vol. 2017, no. 1, p. 118, Dec. 2017.
[16] M. R. Amini, M. Mahdavi, and M. J. Omidi, "Discrete-time Markov chain analysis of energy efficiency in a CR network regarding primary and secondary traffic with primary user returns," *IEEE Access*, vol. 6, pp. 22305–22323, 2018.
[17] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integral, Series, and Products*, 7th ed. New York, NY, USA: Academic, 2007.
[18] R. J. Muirhead, "Aspects of multivariate statistical theory," in *Wiley Series in Probability and Mathematical Statistics*. New York, NY, USA: Wiley, 1982.
[19] A. B. Tsybakov, *Introduction to Nonparametric Estimation*. New York, NY, USA: Springer, 2009.

AMOR NAFKHA (Senior Member, IEEE) received the B.Sc. (Eng.) degree from the Higher School of Communications (SupCom), Tunis, Tunisia, in 2001, and the Ph.D. degree from the University of South Brittany (UBS), Lorient, France, in 2006, all in information and communications technology. From 2006 to 2007, he was a Postdoctoral Researcher with the Signal, Communication, and Embedded Electronics (SCEE-IETR) Research Group, CentraleSupélec, France. During this time at SCEE, he was actively involved in the reconfigurable hardware platform implementation for software-defined radio, co-authoring several contributions on FPGA dynamic partial reconfiguration. Since January 2008, he has been an Associate Professor at CentraleSupélec. He has published more than 70 papers in international peer-reviewed journals and conferences. His research interests include multiuser and MIMO detection, hardware implementation, information theory, sample rate conversion, and spectrum sensing techniques.

5.6 Summary

To overcome the noise uncertainty problem effect and improve the conventional energy detector performance, we have proposed a hybrid architecture combining the simplicity of the energy detector and the robustness of the cyclostationary feature detection. Using adaptive energy detector threshold, the computational complexity of the new proposed hybrid sensing technique tends to decrease over time and converges slowly to the complexity of an energy detector with reduced noise uncertainty. Moreover, the hybrid sensing structure performances are close to that of the conventional cyclostationary feature detection (CFD) proposed by Giannakis and Dandawate. To reduce the computational complexity of the conventional cyclostationary feature detection, we have exploited the sparse property of the cyclic auto-correlation function (CAF) and we have proposed a new blind spectrum sensing detector by taking advantage of symmetry property that exists in the cyclic frequency domain when the alternative hypothesis is true. Simulation results indicate that the proposed blind sensing scheme offers high sensing performance and lower computational complexity in comparison with the classical CFD detector.

The problem of analyse the performance of eigenvalue-based spectrum sensing detector in a finite-dimensional context has been also studied. We derived exact expressions for the PDF and the CDF of the standard condition number (SCN) of the received covariance matrix using results from finite random matrix theory (RMT). In addition, we derived exact expressions for the moments of the SCN and we proposed a new approximation based on the generalized extreme value distribution. Using results from the asymptotic RMT, we further provided a simple forms for the central moments of the SCN and we end up with a simple and accurate expression for the CDF, PDF, P_{fa}, P_d and the decision threshold that could be computed and hence provide a dynamic SCN detector. In the Massive MIMO context, the exploitation of the large number of antennas for spectrum sensing purpose has also be considered. Two antenna exploitation scenarios are studied: **(i)** Full antenna exploitation and **(ii)** Partial antenna exploitation.

We resume below the combined outputs of the work on spectrum sensing topic described in the present chapter.

Research Outputs of Spectrum Sensing Techniques

- **Publications**: 8 Journals + 15 Conferences + 1 Chapter.
- **Collaborators**: Palicot J., Louet Y., Nasser Y., Bazzi O., Siala M., Abdelkefi F.
- **Postdoctoral researcher**: Aziz B.
- **PhD Students**: Khalaf Z., Kobeissi H., Traore S., Gouldieff V.
- **Master Students**: Ajmi K., Bahamou S., Jellali M., Priem G.
- **Supported by**: MENRT grant, IETR grant, SOFTRF project

BIBLIOGRAPHY

[1] Y.-. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, 2008.

[2] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116–130, 2009.

[3] B. Wang and K. R. Liu, "Advances in cognitive radio networks: A survey," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 5–23, 2011.

[4] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio : State-of-the-art and recent advances," *IEEE Signal Processing Magazine*, vol. 29, no. 3, pp. 101–116, 2012.

[5] A. Ali and W. Hamouda, "Advances on spectrum sensing for cognitive radio networks: Theory and applications," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 1277–1304, 2017.

[6] Y. Arjoune and N. Kaabouch, "A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions," *Sensors*, vol. 19, no. 1, 2019.

[7] G. Aswathy and K. Gopakumar, "Sub-nyquist wideband spectrum sensing techniques for cognitive radio: A review and proposed techniques," *International Journal of Electronics and Communications*, vol. 104, pp. 44–57, 2019.

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

[9] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proceedings of the IEEE*, vol. 55, no. 4, pp. 523–531, 1967.

[10] F. Digham, M.-S. Alouini, and M. Simon, "On the energy detection of unknown signals over fading channels," in *IEEE International Conference on Communications, 2003. ICC '03.*, vol. 5, 2003, pp. 3575–3579 vol.5.

[11] S. P. Herath, N. Rajatheva, and C. Tellambura, "Energy detection of unknown signals in fading and diversity reception," *IEEE Transactions on Communications*, vol. 59, no. 9, pp. 2443–2453, 2011.

[12] R. Tandra and A. Sahai, "Snr walls for signal detection," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 4–17, 2008.

[13] W. Jouini, "Energy detection limits under log-normal approximated noise uncertainty," *IEEE Signal Processing Letters*, vol. 18, no. 7, pp. 423–426, 2011.

[14] L. Shen, H. Wang, W. Zhang, and Z. Zhao, "Blind spectrum sensing for cognitive radio channels with noise uncertainty," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1721–1724, 2011.

[15] Y. Wenshan, R. Pinyi, C. Jun, and S. Zhou, "Performance of energy detector in the presence of noise uncertainty in cognitive radio networks," *Wireless Networks*, vol. 19, p. 629–638, 2013.

[16] S. Shellhammer, "Performance of the power detector with noise uncertainty." IEEE P802.22 Wireless RANs, 2006.

[17] C. D. Hou, "A simple approximation for the distribution of the weighted combination of non-independent or independent probabilities," *Statistics and Probability Letters*, vol. 73, no. 2, pp. 179–187, 2005.

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

[19] W. Gardner, "Signal interception: a unifying theoretical framework for feature detection," *IEEE Transactions on Communications*, vol. 36, no. 8, pp. 897–906, 1988.

[20] ——, "Exploitation of spectral redundancy in cyclostationary signals," *IEEE Signal Processing Magazine*, vol. 8, no. 2, pp. 14–36, 1991.

[21] M. Tsatsanis and G. Giannakis, "Transmitter induced cyclostationarity for blind channel equalization," *IEEE Transactions on Signal Processing*, vol. 45, no. 7, pp. 1785–1794, 1997.

[22] B. Ramkumar, "Automatic modulation classification for cognitive radios using cyclic feature detection," *IEEE Circuits and Systems Magazine*, vol. 9, no. 2, pp. 27–45, 2009.

[23] A. Dandawate and G. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2355–2369, 1994.

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

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

[26] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Prentice-hall Englewood Cliffs, 1999.

[27] W. A. Gardner, *Statistical Spectral Analysis: A Non-Probabilistic Theory.* Prentice-hall, 1988.

[28] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of wishart matrices," *IEEE Transactions on Communications*, vol. 57, no. 4, pp. 1050–1060, 2009.

[29] C. G. Khatri, "Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations," *The Annals of Mathematical Statistics*, vol. 35, no. 4, pp. 1807–1810, 1964.

Chapter 6

Software Defined Radio: Sample Rate Conversion & Advanced Hardware Implementation

"In theory, there is no difference between theory and practice. In practice there is."
—Yogi Berra

The advent of increasingly powerful and reconfigurable integrated circuits has led to the emergence of the software defined radio (SDR) concept, which brought endless possibilities for future communication technologies. The SDR concept refers to the capability to change the feature of a radio transceiver in order to adapt it to various air interfaces without requiring any new hardware. SDR is done using software in order to carry out signal processing tasks which are normally processed by a specific hardware. Thus, under the SDR concept, a single hardware platform can be reconfigured on run-time to support a wide range of wireless communication standards and technologies [1]. The signal processing tasks can be run on general purpose processor (GPP), digital signal processors (DSP), programmable system-on-chip (SoC), field-programmable gate array (FPGA) or other application-specific programmable circuits (ASPC).

In this chapter, I will present my research topics related to the concept of software defined radio on which I have been working in the last several years. This chapter is mainly concerned with generic digital front-end, dynamic partial reconfigurable FPGAs, and spectrum sensing for cognitive radio using SDR platforms. For each topic, I will first present a brief technology overview before specifying my main contributions.

6.1 Reconfigurable Digital Front-End

In order to transmit a modulated data, the signal must be placed on a radio-frequency carrier so it can be sent through the antenna with a reasonable size and power. This operation takes place in the radio frequency (RF) front-end. The RF front-end has been commonly implemented

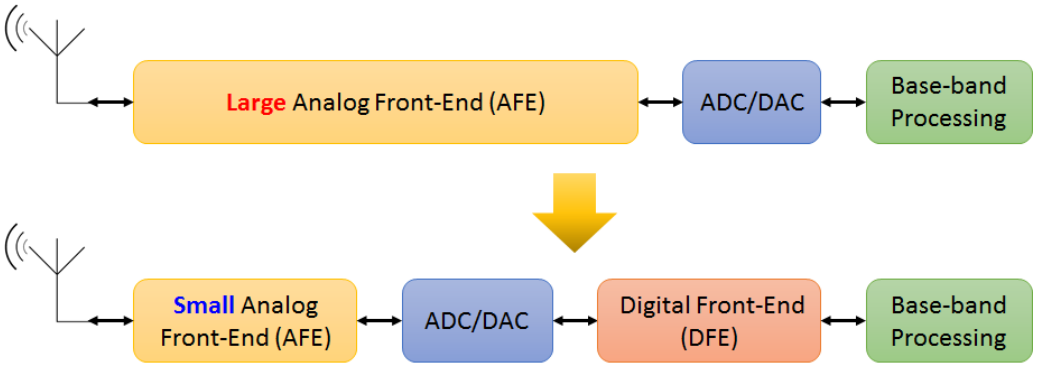


Figure 6.1 – The radio frequency front-end without & with digital front-end

using analog components, such as frequency mixers, local oscillators, and analog filters, all constituting the analog front-end (AFE). However, these implementation of front-ends proved to be bulky, power consuming, vulnerable to many types of interference, and most importantly they lacked the ability to be configurable. In today's systems, and with the rise of the software defined radio (SDR), we can't tolerate the inconveniences of the analog front-end (AFE), where we are dealing with mobile devices with limited battery life, and we need the capability of reconfiguring the front-end to be able to have a cognitive radios that can adapt themselves to their environment. This is what gave rise to the digital front-end (DFE), which aims to reduce the operation of the AFE by placing the digital-to-analog converter (DAC) and the analog-to-digital converter (ADC) as close as possible to the antenna as illustrated in Fig.6.1.

The digital font-end plays multiple roles in the transmission and reception chains, such as digital pre-distortion (DPD) and crest factor reduction (CFR) in the transmission path, and I/Q imbalance compensation and DC offset correction in the reception path. Another main role of the DFE is the sampling rate conversion (SRC), which consists of manipulating the sampling frequency of the signal in order to bring it from base-band to RF, and vice versa. Two types of SRC exist in the digital front-end. The first is the *coarse* SRC, where the sampling rate is increased or decreased by a large integer factor. The other type known as *fine* SRC or arbitrary sample rate conversion (ASRC), which is more complicated to implement than coarse SRC due to the extra required precision.

6.1.1 [C1]: Unified vision of FIR-based SRC

One major drawback of the literature until 2016 is its continued focus on particular SRC structures (Polyphase, Farrow, Newton, *etc.*.) without given any theoretical relationships between them. Addressing this issue was performed during the PhD thesis of Ali Zeineddine in partnership with TDF company and IRT-BCOM. This work provides an up to date coverage of SRC solutions by including the modifications of the generalized Newton structure developed

during Ali's PhD thesis. The practical implementation aspects are also presented for each SRC solution.

As shown in Fig. 6.2, starting from the linear phase FIR filter based sample rate conversion, we proceed by taking more particular filter types, and then by exploiting the characteristics of the impulse response of these filters, we optimize and find the different implementation structures. For the linear phase FIR filters, we present the U-F-D and the polyphase structures. Then we take particular cases of linear phase FIR filters, which are the polynomial based, the B-spline, and the Lagrange interpolation impulse responses. By exploiting the particular properties of each response, we derive the Farrow, Cascaded-integrator-comb (CIC), and Newton structure, respectively. This presentation offers an easy way to understand the relationship between the different SRC structures, clarifying thereby how to select the most appropriate structure for a certain SRC operation, how to compare the different structures, and how they can be combined together to perform any SRC operation. In order to avoid signal distortions, we only considered the linear phase FIR based SRC because the linearity of the phase response will be constant over all frequency components of input signal.

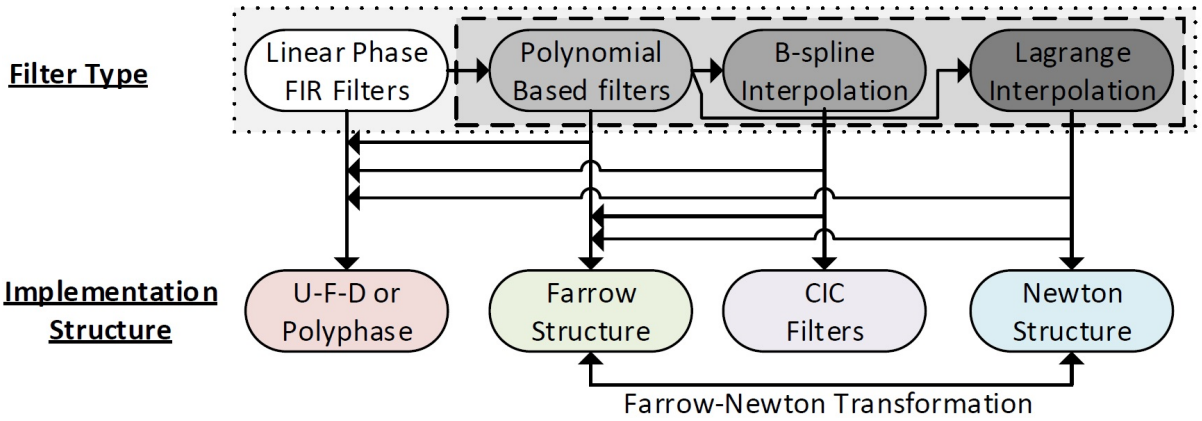


Figure 6.2 – The radio frequency front-end without & with digital front-end

For more details about the unified vision of FIR-based sample rate conversion, the reader is referred to our paper published in Journal of Signal Processing Systems 2020 [J17], and included hereafter.

Journal of Signal Processing Systems (2021) 93:113–125
https://doi.org/10.1007/s11265-020-01575-6

Comprehensive Survey of FIR-Based Sample Rate Conversion

Ali Zeineddine[1] · Amor Nafkha[2] · Stéphane Paquelet[3] · Christophe Moy[4] · Pierre Yves Jezequel[1]

Received: 25 February 2020 / Revised: 19 June 2020 / Accepted: 24 June 2020 / Published online: 27 July 2020

Abstract

Sample rate conversion (SRC) is ubiquitous and critical function of software defined radio and other signal processing systems (speech coding and synthesis, computer simulation of continuous-time systems, *etc*..). In this paper, we present a survey on linear phase finite impulse response (FIR) based sampling rate conversion. Many different FIR-based SRC solutions exist, such as classical FIR, polyphase, Farrow, cascaded-integrator-comb, and Newton structures. Each one of these solutions is presented differently in the literature, and SRC reference books introducing the subject are often missing hardware implementation aspects. The main objective of this paper is to provide a simple and comprehensive overview of main FIR-based SRC techniques from theoretical to hardware implementation aspects. The state of the art of FIR-based SRC filters is summed-up through a concise derivation of the different solutions from a common root: linear phase FIR filters. Each SRC solution is presented from both theoretical and practical implementation points of view. The paper provides a succinct tutorial that introduces SRC, and helps identifying and implementing the appropriate FIR-based SRC architecture for any given applications.

Keywords Sample rate conversion · FIR filter · Polyphase filter · Farrow structure · Newton structure · CIC filter

1 Introduction

Flexibility and reconfiguration are essential requirements in modern transceivers. To implement these features at the front-end interface, sample rate conversion (SRC) is necessary [1]. However, each transceiver is designed for certain application requirements, that may range from high performance and low latency, down to low cost and low power consumption applications. In each case, the needed SRC implementation differs to comply with the system requirements. This does not apply to transceivers only, but also a wide range of signal processing applications are based on SRC, most notably in the audio processing domain. The importance of SRC made it a ubiquitous signal processing function used in many domains for over four decades [2]. Being a classical topic, the literature of SRC is very huge, has been addressed in a large number of reference books, and many solutions have been developed. However, extracting the needed information from the literature can be complicated as each solution is presented from its own point of view. And when referring to books about the subject, practical implementation aspects are often overlooked.

This article provides a concise tutorial presenting the different available SRC solutions, mainly for engineers in this field, looking for an efficient SRC implementation. This work focuses on finite impulse response (FIR) SRC filters due to their linear phase, that is required in many applications which do not tolerate distortion of the signal's phase. Understanding these solutions facilitates the exploration of other SRC types, and also building new ones.

✉ Amor Nafkha
amor.nafkha@centralesupelec.fr

Ali Zeineddine
Ali.Zeineddine@tdf.fr

Stéphane Paquelet
Stephane.Paquelet@b-com.com

Christophe Moy
christophe.moy@univ-rennes1.fr

Pierre Yves Jezequel
Pierre-Yves.Jezequel@tdf.fr

[1] TDF, 155 bis Avenue Pierre Brossolette, 92120, Montrouge, France

[2] IETR UMR CNRS 6164, SCEE, CentraleSupélec, Avenue de la Boulaie, 35510, Cesson Sévigné, France

[3] IRT-BCOM, 1219 Avenue des Champs Blancs, 35510, Cesson Sévigné, France

[4] CNRS, IETR - UMR 6164, University Rennes, Rue du Thabor, 35000, Rennes, France

In the FIR-based SRC literature, five main filter structures can be identified: classical up-sampling/filtering/down-sampling (U-F-D) [2], polyphase [3], Farrow [4], cascaded-integrator-comb (CIC) [5], and Newton structures [6]. This tutorial emphasizes the unified representation shown in Figure 2. This vision derives the different SRC solutions from a common root, which helps in intuitively understanding the features and limitations of each solution, and its relation to the other structures. This tutorial also provides an up to date coverage of SRC solutions by including the modifications of the generalized Newton structure developed in the last four years [7–10]. Last but not least, the practical implementation aspects are also presented for each SRC solution, providing thereby a guide to select and implement the best adapted SRC solutions for a given desired application.

This article is organized as follows. Section 2 presents the basic SRC concepts, and summarizes the unified vision approach. Then each of the following sections studies one of the main SRC solutions by deriving its structure through the introduced unified vision, and then by discussing its practical implementation aspects. Sections 3, 4, 5, and 6 present the U-F-D/Polyphase, Farrow, CIC, and Newton structures respectively. The generalization of the Newton structure is studied in Section 7. Then, Section 8 compares the different presented solutions, and shows how the proposed unified vision is used to study and select the appropriate sample rate conversion structures for a given application. Finally, Section 9 concludes this tutorial.

2 Linear Phase FIR SRC Filters

Let $x(t)$ be a limited-band continuous time signal with its highest frequency component designated as f_m. The sampling rate conversion (SRC) operation consists of changing the sampling frequency of the signal $x(t)$ without going through reconstruction and re-sampling. The definition of SRC operation and its relation to sampling and reconstruction is shown in Figure 1. Sampling $x(t)$ at a frequency $F_{in} = 1/T_{in}$ and $F_{out} = 1/T_{out}$ results in two series of samples $x(mT_{in})$ and $x(kT_{out})$, respectively. In order to correctly sample a signal, the Nyquist theorem states that F_{in} and F_{out} should be greater than or equal to $2 \times f_m$ [11]. Sampling rate conversion is then defined as the operation that takes the samples $x(mT_{in})$ as input, and outputs $x(kT_{out})$, with the change in the sampling frequency defined by the SRC factor R:

$$R = \frac{F_{out}}{F_{in}} = \frac{U}{D} \quad \text{such that} \quad U, D \in \mathbb{N}. \tag{1}$$

The SRC operation with $R > 1$ is called interpolation, while decimation is used to refer to the case of $R < 1$. The supposition that U and D are integers results from the fact that the SRC operation will be implemented in a digital system, that is quantified and can only handle finite precision. In the SRC jargon, the term "$coarse$" SRC is used to refer to a large sampling rate conversion factor that can be a large integer or simple rational ($e.g.$ $R = 1/128$, $R = 64/10$), while the term "$fine$" SRC is used to refer to a very small modification of the sampling factor and it concerns a fine-tuning factor close to unity ($e.g.$ $R = 160/159$).

To implement an SRC operation of a factor $R = U/D$ using discrete time signal processing, the most straightforward solution is to increase the sampling frequency by U through inserting $U - 1$ zeros between input samples, followed by a decrease of factor D by keeping one sample from each D samples. Low-pass filtering is required between these two steps to find the values of the added zero samples (i.e. Figure 4) and potentially to remove any frequency components that may cause aliasing [2]. This work addresses the FIR based SRC methods due to their multiple advantages: stability, linear-phase, and implementation efficiency.

The frequency response of the SRC filter is either optimized to reject up-sampling images or to protect from aliasing. The filter is then called an anti-imaging (AI) or an anti-aliasing (AA) filter, respectively. In practice, the AI filter is used for interpolation, while the AA filter is used for decimation. However, nothing prevents the two types of filter to be used for either kind of SRC factors if the filtering response is acceptable. The AI and AA filters used for SRC are duals. Duality is the property where one operation performs the complementary of the other, such as modulation and de-modulation. The dual of an SRC system is found using the transposition theorem developed in [12], and also presented in chapter three of [2]. The unified vision in this article is developed for the AI filters version,

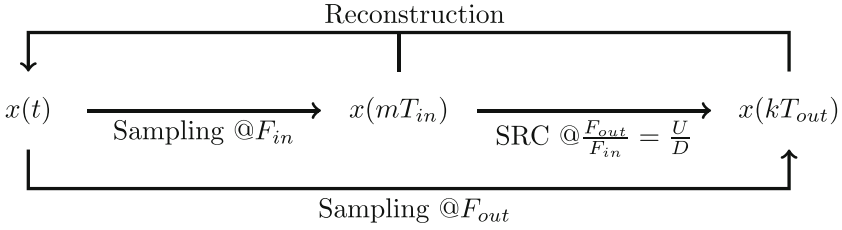


Figure 1 Signal sampling, sampling rate conversion, and signal reconstruction.

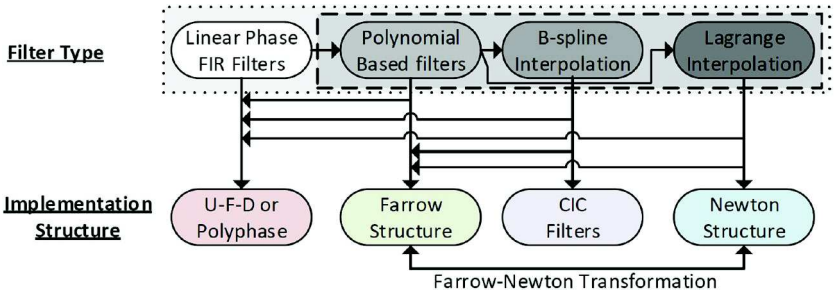


Figure 2 Types of linear phase FIR filters and their corresponding implementation structures for sample rate conversion.

and their corresponding AA duals can be found using the transposition theorem.

All of the FIR based SRC filters can be related to at least one of the five main FIR based SRC structures: the U-F-D, polyphase, Farrow, CIC, and Newton structures. This article offers a unified derivation, starting from one common root, in order to derive these structures and to explain clearly their respective correspondences and differences as shown in Figure 2.

The linear phase FIR filter type is first considered, then more particular filter types are studied, and by exploiting the characteristics of the impulse response of these filters, the different implementation structures are derived. The linear phase FIR filters basic structure is the U-F-D or the polyphase structure. Then particular cases of linear phase FIR filters are considered, which are: polynomial based, B-spline, and Lagrange interpolation impulse responses. By exploiting the particular properties of each response, the Farrow, CIC, and Newton structure are found respectively, and reveal how to generalize the Newton structure in order to implement any polynomial based filter.

3 Fundamental SRC: U-F-D & Polyphase Structures

The U-F-D structure is a direct implementation of the theoretical SRC model presented in last section. It is theoretically capable of implementing any SRC operation using any linear phase FIR filtering response. However, it does not always create a practical efficient implementation [2, 13]. The polyphase structure improves the efficiency while keeping the flexibility of the U-F-D structure [3, 14, 15]. The derivation of these two structures are obtained as presented below.

3.1 Filter Structure Derivation

In the case of the U-F-D structure, the impulse response is directly implemented using an FIR filter structure as shown in the top of Figure 3. Since the input of the filter is the output of the up-sampling of $x[m]$, then we know that for each U input samples to the filter, only one is non-zero. An example of $U = 4$ is shown in Figure 4, where for each output only a specific set of coefficients are used (squares,

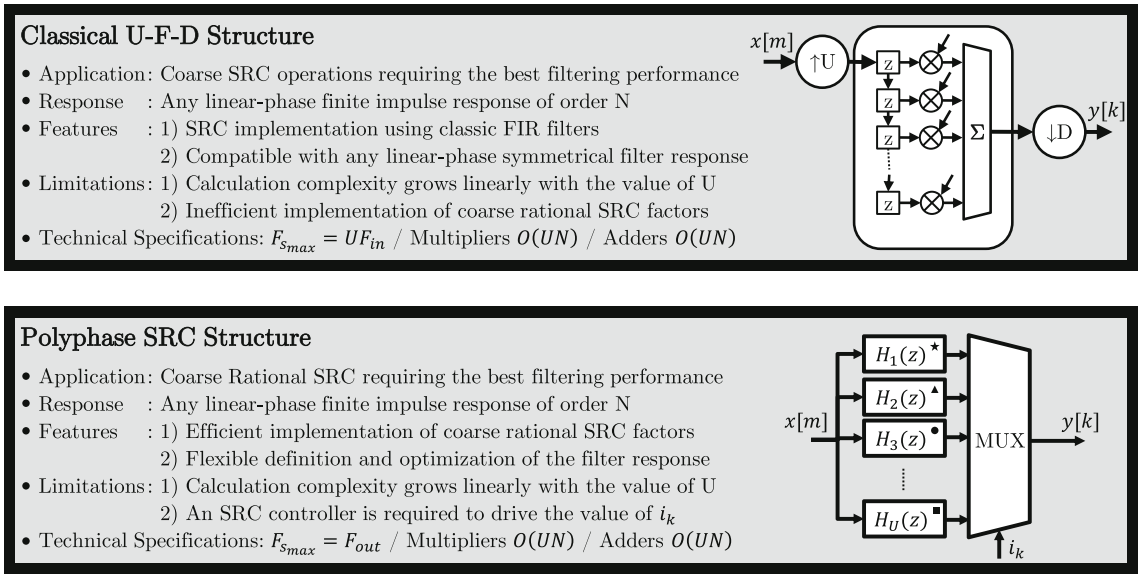


Figure 3 Fundamental SRC: U-F-D and polyphase structures.

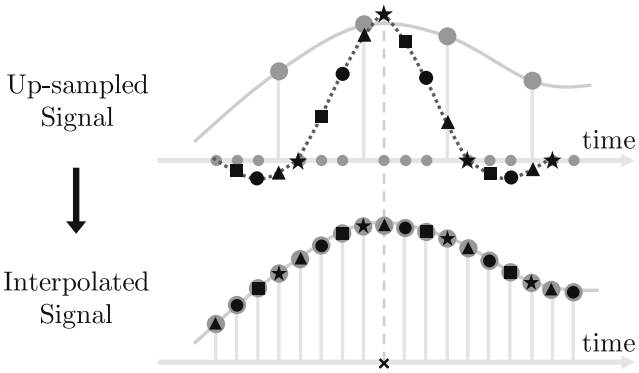


Figure 4 Example order 3 interpolation of an up-sampled signal.

circles, triangles, or stars). Therefore, if a filtering operation of order $N-1$ is required, the number of taps of the FIR filter has to be equal to $N \times U$. The transfer function of this filter is written as:

$$H(z) = \mathcal{Z}\{h[n]\} = \sum_{n=-\infty}^{+\infty} h[n]z^{-n} = \sum_{n=0}^{NU-1} h[n]z^{-n}, \quad (2)$$

with $h[n]$ being the impulse response coefficients. When implementing a rational SRC factor with $D > 1$, a number of the filter output samples is dropped by the following down-sampling operation, meaning that the filter is inefficient since it calculates unneeded samples. This is addressed by the polyphase structure. The classical polyphase structure is deduced from the U-F-D structure by rearranging the filter coefficients, through re-writing $H(z)$ in Eq. 2 as:

$$H(z) = \sum_{i=0}^{U-1} z^{-i} \sum_{j=0}^{N-1} h[i+jU]z^{-jU} = \sum_{i=0}^{U-1} z^{-i} H_i\left(z^U\right), \quad (3)$$

where the FIR filter is split into U filters $H_i(z)$ of order N in parallel, with each filter containing the coefficients used to find a certain output. The matrix representation of $H(z)$ offers a clear visual description of the filter structure:

$$\underset{U\times 1}{H(z)} = \begin{bmatrix} \star_1 & \star_2 & \dots & \star_N \\ \blacktriangle_1 & \blacktriangle_2 & \dots & \blacktriangle_N \\ \vdots & \vdots & \ddots & \vdots \\ \blacksquare_1 & \blacksquare_2 & \dots & \blacksquare_N \end{bmatrix} \times \begin{bmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-N+1} \end{bmatrix}, \quad (4)$$

with each row i of the matrix H containing the coefficients of the sub-filter $H_i(z)$. The polyphase structure is then built using U sub-filters $H_i(z)$ of order N in parallel. For each output $y[k]$, the corresponding sub-filter $H_{i_k}(z)$ is selected by a multiplexer combined with a controller operating at the output sampling frequency F_{out}. This results in the polyphase structure shown in Figure 3, that operates efficiently using the sampling frequencies F_{in} and F_{out} only, without needing the higher UF_{in} sampling domain [16]. However for fine-tuned SRC factors, when the values of U and D are large, the structure still requires having U filters in parallel, which is problematic concerning the structure efficiency. This can be solved using polynomial based filters that are presented in the next section.

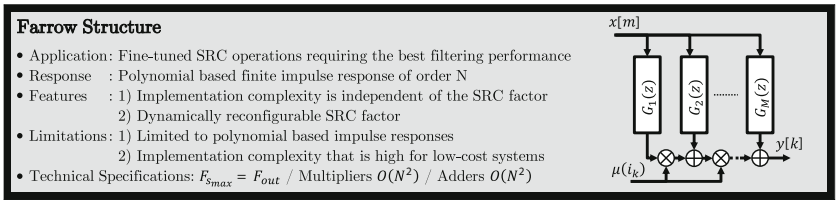
3.2 Practical Implementation

To implement the derived filter structures, a number of practical implementation aspects need to be addressed. In the case of the U-F-D and polyphase structure, these aspects are the following.

Hardware Implementation: For a given filter structure, many hardware implementation strategies exist, with each favoring a certain implementation constraint. Pipeline implementations favor high operation speed at the cost of added complexity [17], while a lower consumption at the cost of reduced speed can be achieved through reusing hardware elements [18]. The second design aspect concerns the quantization of the implementation, that defines the calculation precision. A wide range of quantization methods for both the filter parameters [19–21] and signals [22–24] exist in the literature.

SRC Controller: It is often supposed in the literature that for the SRC implementation, F_{in} and F_{out} are actually available digital clocks. However, practical systems often operate using only one digital clock higher than both sampling frequencies. The multi-rate operation is then made possible through an SRC controller that controls the samples flow. This is required because the number of inputs to the SRC filter is not equal to the number of outputs. In the case of the polyphase structure, for a given output instant k, two indexes are needed to control the filter state. The first being m_k that indicates the input samples used by the filters $H_i(z)$, and the

Figure 5 Efficient arbitrary SRC: Farrow structure.



second being i_k that specifies the branch of the polyphase filter that calculates the output $y[k]$. These two indexes are managed by the SRC controller in order to implement the desired SRC operation defined by U and D. The control algorithm that finds these indexes is shown in Algorithm 1, that describes the AI SRC controller operation.

Algorithm 1 Algorithm of the SRC Controller.

Input: U, D
Output: m_k, i_k, k
1: $m_k = 0\ ,\ i_k = 0\ ,\ k = 0$
2: **while** True **do**
3: $\quad i_k = i_k + D$
4: $\quad$ **while** $i_k \geq U$ **do**
5: $\quad\quad m_k + +$ {Get next input}
6: $\quad\quad i_k = i_k - U$
7: $\quad$ **end while**
8: $\quad k + +$ {Find next output}
9: **end while**

Half-Band filters: When performing SRC of large factors, it is possible to improve the implementation efficiency by breaking the SRC operation into multiple stages as proposed in [13]. This is also more advantageous when it is possible to break the SRC factor into factors of two, that can be implemented efficiently using half-band filters. These filters are a particular case of the U-F-D structure, with a very efficient implementation that uses only 25% of the multipliers [2].

4 Efficient Arbitrary SRC: Farrow Structure

An efficient solution for finely-tuned SRC implementation is the structure developed by C. W. Farrow in [4]. This structure implements polynomial based filters, and is known in the literature under many names, most notably as the Farrow structure [25, 26]. The introduction of this structure enabled implementing high precision finely-tuned SRC far more efficiently than it was possible with the previously existing structures. A modification of the structure was developed in [27] that exploits the symmetry of the impulse response in order to reduce the complexity. The transposed Farrow structure preferably used for SRC decimation was addressed on multiple occasions in [28–31]. The transposed Farrow structures has many other applications beside SRC as discussed in [28, 32, 33].

4.1 Filter Structure Derivation

It was shown in the last section that for each output $y[k]$, only one polyphase branch is selected by the multiplexer. Therefore it would be interesting that instead of having U filters in parallel, to have only one filter with its coefficients being calculated dynamically for each output using the index i_k. This can be made possible by defining the impulse response as being constructed using N polynomial pieces $P_j(\mu)$ with $j \in \{1, 2, \ldots, N\}$. Each polynomial piece is of order $M - 1$ and is defined as:

$$P_j(\mu) = \sum_{l=1}^{M} c_l(j)\mu^{l-1}, \tag{5}$$

then P_j is used to calculate the filter's coefficient $h_{i_k}[j]$ corresponding to the output $y[k]$. It is then possible to rewrite the $H_{i_k}(z)$ filter expression in Eq. 3 as:

$$H_{i_k}(z) = \sum_{j=1}^{N} \left[\sum_{l=1}^{M} c_l(j)\mu(i_k)^{l-1} \right] z^{-j+1}, \tag{6}$$

where $\mu(i_k)$ is used to indicate the position of the coefficients of the filter i_k on the different polynomial pieces. To exploit the symmetry of the impulse response of a linear phase filter [34], the variable $\mu(i_k)$ used to develop the polynomials is defined as:

$$\mu(i_k) = \frac{i_k}{U} - \frac{1}{2} \in [-0.5; 0.5[. \tag{7}$$

Using this type of impulse response, the coefficients of each filter $H_{i_k}(z)$ in Eq. 3 are found as:

$$\begin{aligned} \underset{U\times 1}{H_{i_k}(z)} &= \underset{1\times M}{\mu_{i_k}} \times \underset{M\times N}{P} \times \underset{N\times 1}{Z(z)} \\ &= \begin{bmatrix} 1 \\ \mu_{i_k} \\ \vdots \\ \mu_{i_k}^{M-1} \end{bmatrix}^T \begin{bmatrix} c_1(1) & \ldots & c_N(1) \\ c_1(2) & \ldots & c_N(2) \\ \vdots & \ddots & \vdots \\ c_1(M) & \ldots & c_N(M) \end{bmatrix} \begin{bmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-N+1} \end{bmatrix}. \end{aligned} \tag{8}$$

Through a factorization by μ in Eq. 6, the Farrow structure definition is found as:

$$H_{i_k}(z) = \sum_{l=1}^{M} G_l(z)\mu(i_k)^{l-1} \quad \text{with} \quad G_l(z) = \sum_{j=1}^{N} c_l(j)z^{-j+1}. \tag{9}$$

The implementation of this structure is shown in Figure 5, which consists of M FIR filters $G_l(z)$, each of N taps. The outputs of the G_l filters are then evaluated following the Horner's scheme for polynomial evaluation with $\mu(i_k)$. In this structure, the coefficients of the filters G_l are constant and symmetrical, allowing an efficient implementation by using only half of the multipliers.

4.2 Practical Implementation

The same practical implementation aspects presented in the last section apply the Farrow structure as well. Two other specific implementation aspects to the Farrow structure are: the polynomial coefficients and the fractional delay quantization.

Polynomial Coefficients: The Farrow structure has $M \times N$ degrees of freedom to define the filter response, i.e. the coefficients $c_l(j)$. When the linear-phase filter response is used, the degrees of freedom are divided by two, due to the symmetry of the response [34]. These coefficients can be found generally in two ways: using well known polynomial interpolation methods, or applying filter optimization techniques.

Three types of polynomial interpolation methods are widespread in the literature [35]. Due to their particular mathematical definition, these interpolation types can be implemented more efficiently using other structures than the Farrow structure as it will be developed later in this article. These are B-spline interpolation [36–39], Lagrange interpolation [25, 40, 41], and Hermite interpolation [42, 43].

However, to correctly profit from the Farrow structure filtering performance, the polynomial coefficients should be defined using filter optimization methods. In the literature, different techniques are employed [4, 44–46]. The thesis of M. T. Hunter [47] develops in details the Farrow structure optimization problem, and provides a software implementation of the developed optimization algorithm.

Fractional Delay Quantization: Quantization of the fractional delay $\mu(i_k)$ affects two implementation aspects: the output sampling rate precision and the filtering performance. This work uses the approach developed in [47–49].

5 Low-cost Coarse SRC: CIC Filter

In 1981, E. B. Hogenauer developed the cascaded-integrator-comb (CIC) filters [5], illustrated in Figure 6, which are a class of SRC FIR filters with an implementation that does not require multipliers. These filters are mainly used to implement coarse SRC operation of a large integer SRC factor.

5.1 Filter Structure Derivation

The Farrow structure previously presented offers an SRC solution with a high degree of freedom permitting the implementation of different filter responses. Considering the most simple case of a polynomial filter, where the impulse response consists of only one polynomial piece ($N = 1$) of degree zero ($M = 1$), having a value of one. This corresponds to the Pi or rectangular function normalized to the input sampling period T_{in} noted as $\Pi(t/T_{in})$. Used for interpolation, the Pi function results in a zero-order-hold reconstruction. An interesting property of the Pi function is its frequency response $H_\Pi(f) = sinc(f/F_{in})$, with its zeros located at multiples of F_{in}, which is the optimal zero position for images suppression, since the up-sampling images are located on multiples of F_{in} [2]. The filtering performance of a single Pi function may not be sufficient, however it can be improved by cascading N filters. This results in the impulse response of the B-spline interpolation of order $N - 1$:

$$\beta^N(t) = (\underbrace{\Pi * \Pi * \ldots * \Pi}_{N})(t). \tag{10}$$

The B-spline function $\beta^N(t)$ is piece-wise polynomial with N pieces of degree $N - 1$ each [36]. Therefore, this interpolation can be implemented using the Farrow structure as it was developed in [7, 37]. However, by exploiting the particularity of the impulse response definition as a cascade of rectangular functions, a more efficient implementation structure is found. It was presented previously that NU filter taps are needed for an SRC filter of order $N - 1$. In the case of the rectangular function, we have U coefficients taken from the causal rectangular function, with all coefficients having a value of one, which can be expressed through the Z transform as:

$$H(z) = \sum_{n=0}^{U-1} z^{-1} = \left(1 - z^{-U}\right)\left(\frac{1}{1 - z^{-1}}\right) = C(z^U)I(z). \tag{11}$$

This is the non normalized moving average filter, which can be implemented in a recursive running-sum structure. It consists of two construction blocks: the integrator having the transfer function $I(z)$ and the comb with a delay z^{-U} having the transfer function $C(z^U)$ [50]. Then by cascading N rectangular function filters, the B-spline interpolation of order $N - 1$ is implemented. We know that the comb and the integrator are linear time-invariant systems and, therefore, their order of placement relative to one another in the structure can be exchanged. Then, benefiting from the second Noble identity, the comb blocks can be placed before the up-sampling operation optimizing the implementation

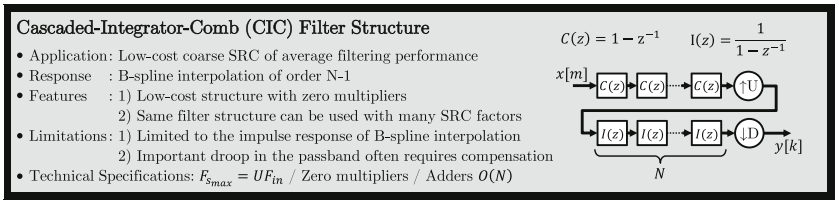


Figure 6 Low-cost coarse SRC: CIC filter.

by reducing the delay line of the comb blocks z^{-U} to only one delay element z^{-1}. This results in the structure presented by Hagenauer in [5], and shown in Figure 6, called the CIC interpolation filter, where a structure having N comb and N integrator blocks implements B-spline interpolation of order $N-1$.

5.2 Practical Implementation

For CIC filters, an SRC controller is not needed since the filter implements a U-F-D structure. This section develops the particular implementation aspects to the CIC filters.

Output Re-scaling: The CIC filter has an inherent gain that can be expressed as $G_{CIC} = U^{N-1}$, which requires adding $I_G = \lceil N \log_2(U) \rceil$ bits to the processed signal in the case of fixed-point quantization. Certain implementations may require that the output of the CIC filter to be normalized to its input, requiring thereby an output re-scaling. This is done in two steps: coarse and fine gain adjustments. The coarse gain adjustment consists of normalizing the integer gain factor, by right shifting the quantization by $s = \lfloor \log_2(G_{CIC}) \rfloor$. Fine gain adjustment then finalizes the normalization by multiplying the samples with the factor $SF = 2^s / G_{CIC}$. The output of this multiplication is then rounded to the desired bit width. The reader may refer to Hogenauer's work [5] for more details about the quantization in the intermediate stages.

Selecting the Filter Order: Three measures are used to select the appropriate CIC filter order for a given SRC operation: pass-band width, image attenuation level, and side lobe rejection levels. To simplify the design process, Hogenauer proposed in the original article tables that sums-up these measures for certain scenarios [5].

CIC Compensation Filter: The frequency response of B-spline interpolation is known to have an important attenuation in its pass-band. A solution to improve the frequency response was proposed in [51] that uses filter sharpening techniques resulting in the sharpened CIC filter structure, with better pass-band and improved filtering. It is also possible to improve the response by using compensation filters that aim at correcting the attenuation of the CIC filter [52–56].

Non-Recursive CIC Filter Structures: Another downside of the classical CIC filter is the recursive structure of the integrator blocks, that limits the maximum speed of the implementation. A non-recursive structure was proposed in [57] offering lower power consumption and higher operating speeds compared to the recursive CIC structure. Further improvements were achieved in [58] through a polyphase non-recursive CIC implementation.

6 Low-cost Arbitrary SRC: Newton Structure

The Newton interpolation structure presented in [6], and illustrated in Figure 7, offers a low-cost arbitrary SRC filter solution based on Lagrange interpolation. The structure modifies what was proposed in [59, 60] to accommodate for a variable fractional delay.

6.1 Newton Structure Derivation

Lagrange interpolation is a widely used interpolation technique due its simplicity and the maximally flat pass-band of its frequency response. The Lagrange interpolation filter is piece-wise polynomial by definition and, therefore, can be implemented using the Farrow structure [25, 26]. In [61], the Lagrange interpolation filter response was developed using the B-spline response, and showed that the zeros of Lagrange interpolation are the same for both interpolations, hence sharing the role of image suppression. The frequency response of Lagrange interpolation focuses on approximating the ideal filter at $f = 0$ Hz the best way possible [62].

The Newton structure is a particular way of implementing the maximally flat FIR filter at $f = 0$ Hz. First, the delay factor μ in the case of the Farrow structure is transformed into the causal version of the delay expressed as $d = [(N-1)/2] - \mu$. Then, the Z-domain transfer function of the ideal filter definition can be written as:

$$H^d(z) = z^{-d} \quad \text{with} \quad z = e^{j2\pi f T_{in}}. \tag{12}$$

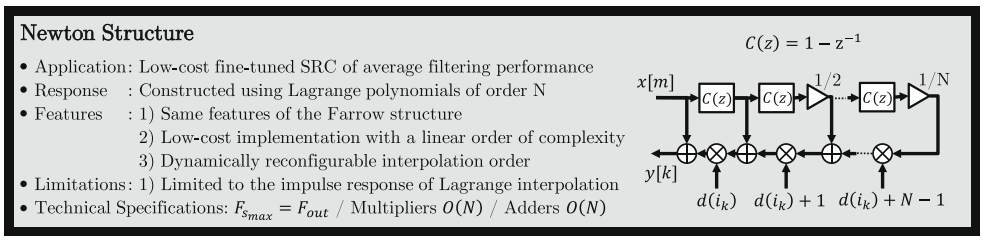


Figure 7 Low-cost arbitrary SRC: Newton structure.

The main mathematical problem of the above expression Eq. 12 is that in the Z-domain, the terms z are used to express delays by an integer number of samples. However, in this case d is a real value and, therefore, evaluating $H^d(z)X(z)$ in this form will result in an ambiguous output expression. This was addressed in [63], where the transfer function $z^{-p/q}$ was approximated as a partial series of terms z^{-i}, with p and q being real numbers, and i being integer. This is done using the generalized binomial theorem, where it is possible to re-write $H^d(z)$ as:

$$\begin{aligned} H^d(z) &= \left[1 + \left(z^{-1} - 1\right)\right]^d \\ &= \sum_{k=0}^{+\infty} \frac{d(d-1)\dots(d-k+1)}{k!} \left(z^{-1} - 1\right)^k. \end{aligned} \tag{13}$$

This series converges for $|z^{-1} - 1| < 1$. Truncating this expression to the first N terms is known to result in the partial sum that satisfies the maximally flat criterion presented previously [60, 64, 65]. And since this criterion has a unique solution, then the partial sum of N terms corresponds to the Lagrange interpolation of order $N-1$ [62, 66]. The partial sum of Eq. 13 of order N corresponds to the Newton backward difference formula of order $N-1$, that implements the Lagrange interpolation [59, 60]. The filter structure is to be modified as it is shown by [6] in order to support variable fractional delays, resulting in the Newton interpolation structure, presented in Figure 7. This structure has a complexity of $O(N)$ compared to the Farrow structure implementing Lagrange interpolation of order N with a complexity of $O(N^2)$. It is important to note that a rigorous demonstration of the convergence of the Newton fractional-delay filter to the ideal fractional delay filter was recently published in [67].

6.2 Practical Implementation

The Newton structure does not have any different practical implementation aspects than those presented earlier in this paper. The same hardware implementation considerations and the same coefficients quantization guidelines presented for the U-F-D/polyphase structures apply for the Newton structure. The quantization of the fractional delay follows the same approach presented for the Farrow structure. Finally, the implementation of the SRC controller is also identical, with the only difference being the added circuit to calculate the value of the causal fractional delay d.

7 Improved Low-Cost Arbitrary SRC: Generalized Newton Structure

To extend the use of the Newton structure, D. Lamb *et al.* developed the transformation matrices that show the relation between the Farrow implementation of the Lagrange interpolation and the Newton interpolation structure [7]. Then they applied these transformations to B-spline interpolation of order 3 using the Farrow structure to get a modified Newton structure implementation. Later on, this transformation was used to modify the Newton structure in different ways to obtain improved low-cost arbitrary SRC solutions [8–10].

7.1 Farrow to Newton Transformation

The matrix representation of the Farrow structure transfer function presented in Eq. 8 uses a monomial basis vector μ_k, and a cumulative delay base vector $Z(z)$. The work developed in [7] shows that it is possible to transform these base vectors to those of the Newton structure, allowing thereby to deduce the corresponding Newton structure implementation of any Farrow structure. The Newton structure is based on differentials between samples and their derivatives. In this case, the polynomial evaluation is done with the Newton polynomial base vector d^T expressed as:

$$\underset{1\times M}{d^T} = \left[1, d, d(d-1), \dots, \Pi_{i=0}^{M-2}(d-i)\right]. \tag{14}$$

For the delay vector, the Newton structure uses a differential delay base vector $\bigtriangledown Z(z)$ expressed as:

$$\underset{N\times 1}{\bigtriangledown Z(z)} = \left[1, (1-z^{-1}), \dots, (1-z^{-1})^{N-1}\right]. \tag{15}$$

Then, the matrix representation of the Newton structure transfer function is written as:

$$H_{i_k}(z) = \underset{1\times M}{d^T} \times \underset{M\times N}{Q} \times \underset{N\times 1}{\bigtriangledown Z(z)}, \tag{16}$$

where the matrix Q containing the coefficients describing the Newton structure implementation. By transforming the base vectors in Eq. 8 to those of the Newton structure in Eq. 16, the Newton structure coefficients matrix Q corresponding to the matrix P of the Farrow structure can be found as follows:

$$\begin{aligned} H(z,\mu) &= \mu^T P\ Z(z) \\ &= \mu^T \left(T_d^T\ T_d^{-T}\right)\ P\ \left(T_z^{-1}\ T_z\right)\ Z(z) \\ &= (T_d\ \mu)^T \left(T_d^{-T}\ P\ T_z^{-1}\right)\ (T_z\ Z(z)) \\ &= d^T\ Q\ \bigtriangledown Z(z), \end{aligned} \tag{17}$$

where $T_d^{-T} = \left(T_d^T\right)^{-1}$ and the matrices T_d and T_z are the ones transforming the vectors μ and $Z(z)$ to d and $\bigtriangledown Z(z)$, respectively [7]. The analytical expressions of these matrices are developed in [8]. This transformation allows implementing any polynomial based response using a modified Newton structure as shown in Figure 8 for a random example. Concrete examples of modified Newton structures are provided in the following paragraphs.

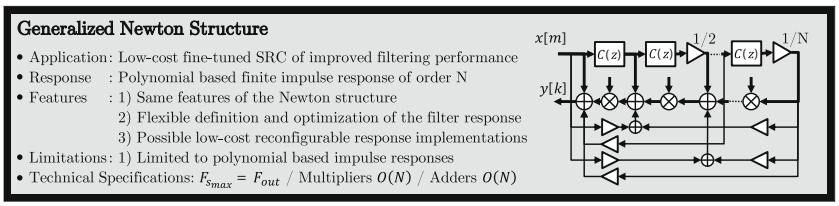


Figure 8 Improved low-cost arbitrary SRC: generalized Newton structure.

7.2 Modified Newton Structures:

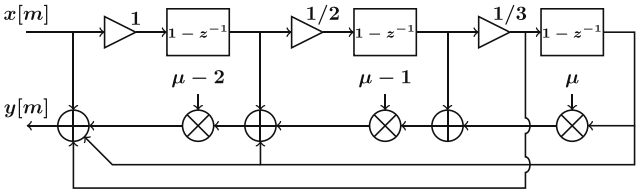
In the most general case, the generalized Newton structure can implement any polynomial based filter response, however, the order of complexity will be comparable to that of the Farrow structure. Nevertheless in certain particular cases, the generalized Newton structure can offer much more efficient implementations than it is possible using the Farrow structure. This section presents some of these cases.

B-spline interpolation: The original work that proposed this transformation applied it to the case of B-spline interpolation of order 3 [7]. This resulted in the modified structure shown in Figure 9-(a), that modifies the classical Newton structure of order 3 using only three additional adders.

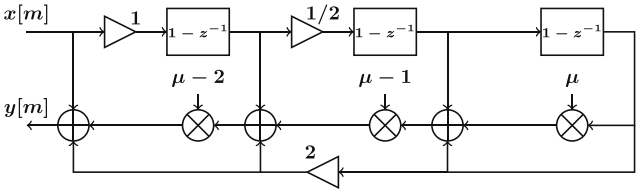
Hermite interpolation: A disadvantage of B-spline interpolation is the important attenuation in its pass-band, limiting its practicality for certain transceiver systems. Hermite interpolation is a more interesting interpolation method that keeps the flat pass-band of Lagrange interpolation while improving the side lobes rejection levels. The modification of the Newton structures to implement this type of interpolation was developed in [9], that resulted in filter structures having lower complexity than the classical Newton structures of the same order, while offering improved filtering performance. An example modified structure of order 3 is shown in Figure 9-(b).

Optimization methods: By using the Farrow to Newton transformation, the filter response relation to the generalized Newton structure coefficients is not explicit. However, this relation can be found using the analytical expressions of this transformation as developed in [8]. This results in the closed-form expression of the generalized Newton structure frequency response. It is then possible to use filter optimization methods to find customized Newton structures with the best filter response possible. An example is shown in Figure 9-(c) for a customized structure of order 3, where it is supposed that the modification introduces feedback lines from the last delay element only. Then by using filter optimization methods, the coefficients c_1 to c_8 that correspond to the optimal approximation of the desired filter response are found. A concrete example is developed in [8] that shows how it is possible to achieve similar filtering performance to the Farrow structure, using a modified Newton structure of a much lower complexity.

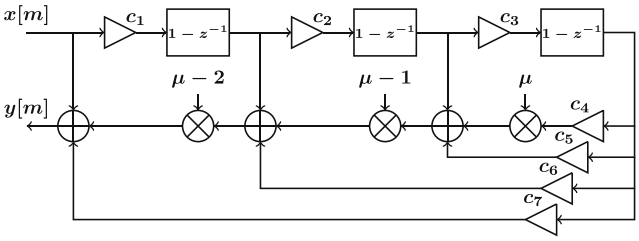
Reconfigurable structures: All the previously presented structures are modifications of the classical Newton structure. By linearly combining the transfer function of the classical structure with any other modified one, it is poss-



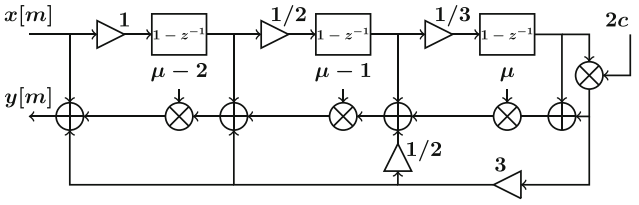
(a) Spline interpolation of order 3 [7]



(b) Hermite interpolation of order 3 [9]



(c) Optimized interpolation of order 3 [8]



(d) Reconfigurable interpolation of order 3 [10]

Figure 9 Modification examples of the generalized Newton structure.

ible to find reconfigurable structures using only one variable parameter c. This is done by expressing the filter response as $H_R(z) = (1-c)H_{Newton}(z) + cH_{Modified}(z)$. The structure implementing $H_R(z)$ is then capable of performing both interpolations (Lagrange for $c = 0$, and the second interpolation for $c = 1$), and also a range of interpolations that combine the characteristics of both interpolation methods by using a value of $c \in [0; 1]$. This design approach is developed in [10], where two examples are provided. The reconfigurable structure implementing both Lagrange and Hermite interpolations is shown in Figure 9-(d).

8 SRC Solutions Comparison

The developed presentation of the FIR based SRC filters, summed-up in Figure 2, offers an easy way to grasp the relations between the different structures, clarifying thereby how to select the most appropriate structure for a certain SRC operation, how to compare the different structures, and how they can be combined together to perform any SRC operation efficiently.

A simple chart to select the best adapted SRC solution for a given application is represented in Table 1. For coarse SRC, the choice has to be made between the U-F-D/polyphase solution and the CIC filter for best performance and low-cost applications respectively. While for fine SRC, the choice has to be made between the Farrow and Newton structures. With the generalization of the Newton structure, it is possible to achieve low-cost implementation with filter response approaching what is possible with the Farrow structure as presented in the last section.

The CIC filter is the reference SRC solution with the lowest complexity. However, it is only suited for coarse SRC operations. For fine SRC, the Newton structure used to be the most efficient implementation before our proposition of the modified Newton structures. We developed in [17] the implementation of low-cost SRC filters on application specific integrated circuits (ASIC). The results, illustrated in Figure 10, show that a Newton structure of order 5 uses only 72% of the hardware resources of an equivalent implementation using the Farrow structure. However, the modified Newton structure for Hermite interpolation requires only half the resources while offering improved filtering performance. This demonstrates the advantage of the modified Newton structures that make possible the implementation of fine SRC with a complexity approaching

Table 1 Choosing the best adapted SRC solution.

	Best Performance	Low-Cost
Coarse SRC	U-F-D / Polyphase	CIC filter
Fine SRC	Farrow structure	Newton structure

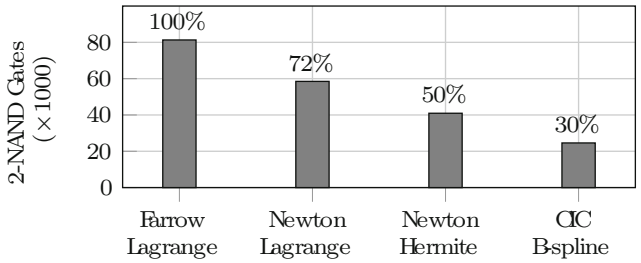


Figure 10 Complexity on ASIC hardware of order 5 implementations.

that of the coarse SRC CIC filters of a similar order. This was not conceivable a decade ago before the propositions in [6, 7, 9].

For an efficient implementation, the SRC operation is implemented as a cascade of multiple SRC filters [1, 2, 13]. In order to partition an SRC operation into a multistage implementation, a simple guide is proposed below:

1. For SRC applications with a large changes in the sampling frequency, extract the first factor R_1 representing an interpolation or decimation operation by an integer factor larger than eight. Then either the CIC filters or the U-F-D/polyphase structures are used depending on the filtering requirements. If the factor R_1 can be broken into multiple factors of 2 or $1/2$, then half-band filters offer improved filtering performance with a very low complexity.
2. After extracting the factor R_1, identify if there exists a lower coarse rational or integer SRC factor R_2. Usually when a CIC filter is used, this second stage combines further SRC with the CIC filter response compensation. The implementation choice of R_2 largely depends on the system, but is often implemented using the U-F-D / polyphase structures.
3. Finally, the fine adjustment of the output sampling frequency is represented by the factor $R_3 \approx 1$. Then depending on the application requirements, a Farrow or Newton structure is used for implementation, offering an average side lobe attenuation of 60-70 dB or 35-45 dB respectively for an order 5 implementation.

9 Conclusion

In this paper, a succinct presentation of the main FIR based SRC solutions was developed. This presentation is based on a deep analysis of the previously proposed solutions for FIR-based SRC filtering, which resulted first in the derivation of a unified vision of this issue, and second in the development of a design methodology guide for SRC designers. Both cases of coarse and fine SRC were considered, and two types of solutions for each case were given: best performance and low-cost solutions. Recently proposed fine

SRC filters based on the Newton structure were also presented, that allow implementing fine SRC with a complexity approaching that of the more simple coarse SRC implementations. This presentation also covered the different practical implementation aspects in hardware, and provided a simple guide to help implementing any SRC operation using a multi-stage architecture. Through this tutorial, any engineer in the field should be able to easily understand and use the basics of the presented SRC solutions.

Author Contributions AZ is the main author of the current paper. SP and AN came up with the original idea. As academic supervisors, CM and AN have proofread the paper several times and provided guidance throughout the whole preparation of the manuscript. SP, AZ, and PYJ analyzed and compared with an existing survey of sample rate converters. All of the authors participated, and they read and approved the final manuscript.

Funding Information This work has been achieved within the Institute of Research and Technology IRT-B<>COM, dedicated to digital technologies. It has been funded by TéléDiffusion de France (TDF) and the French government through the National Research Agency (ANR) under the Investissements d'Avenir program, with reference number ANR-A0-AIRT-07.

Availability of data and materials Data sharing is not applicable to this article as no datasets were generated or analysed during the study.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no competing interests.

Abbreviations SRC, Sample rate conversion; FIR, Finite impulse response; CIC, Cascaded integrator-comb; U-F-D, Up-sampling/ Filtering/ Down-sampling; AI, Anti-imaging; AA, Anti-aliasing.

References

1. Fa-Long, L. (2011). *Digital Front-End in wireless communications and broadcasting: circuits and signal processing*. United Kingdom: Cambridge University Press.
2. Crochiere, R.E., & Rabiner, L.R. (1983). Multirate digital signal processing, prentice hall, englewood cliffs, NJ USA.
3. Bellanger, M., Bonnerot, G., Coudreuse, M. (1976). Digital filtering by polyphase network: Application to sample-rate alteration and filter banks. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *24*(2), 109–114.
4. Farrow, C.W. (1988). A continuously variable digital delay element. In *IEEE International Symposium on Circuits and Systems, Espoo, Finland,* (Vol. 3 pp. 2641–2645).
5. Hogenauer, E. (1981). An economical class of digital filters for decimation and interpolation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *29*(2), 155–162.
6. Lehtinen, V., & Renfors, M. (2009). Structures for interpolation, decimation, and nonuniform sampling based on Newton's interpolation formula, International Conference on Sampling Theory and Applications.
7. Lamb, D., Chamon, L.F.O., Nascimento, V.H. (2016). Efficient filtering structure for spline interpolation and decimation. *Electronics Letters*, *52*(1), 39–41.
8. Zeineddine, A., Paquelet, S., Nafkha, A., Moy, C., Jezequel, P. (2018). Generalization and coefficients optimization of the newton structure. In *IEEE 25th International Conference on Telecommunications, St. Malo, France* (pp. 98–103).
9. Zeineddine, A., Nafkha, A., Moy, C., Paquelet, S., Jezequel, P. (2018). Variable fractional delay filter: a novel architecture based on hermite interpolation. In *IEEE 25th International Conference on Telecommunications, St. Malo, France* (pp. 93–97).
10. Zeineddine, A., Paquelet, S., Kanj, M., Moy, C., Nafkha, A., Jezequel, P.Y. (2018). Reconfigurable Newton structure for sample rate conversion. In *IEEE Global Conference on Signal and Information Processing, Anaheim, CA, USA* (pp. 271–275).
11. Shannon, C.E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, *37*(1), 10–21.
12. Claasen, T.A.C.M., & Mecklenbrauker, W.F.G. (1978). On the transposition of linear time-varying discrete-time networks and its application to multirate digital systems. *Philips Journal Research*, *33*, 78–102.
13. Crochiere, R., & Rabiner, L. (1975). Optimum FIR digital filter implementations for decimation, interpolation, and narrow-band filtering. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *23*(5), 444–456.
14. Hsiao, C.C. (1987). Polyphase filter matrix for rational sampling rate conversions. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, Dallas, TX, USA* (pp. 2173–2176).
15. Bregovic, R., Yu, Y.J., Saramäki, T., Lim, Y.C. (2011). Implementation of linear-phase FIR filters for a rational sampling-rate conversion utilizing the coefficient symmetry. *IEEE Transactions on Circuits and Systems I:, Regular Papers*, *58*(3), 548–561.
16. Harris, F. (1997). Performance and design considerations of the Farrow filter when used for arbitrary resampling of sampled time series. In *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers (Cat. no.97CB36136), Pacific Grove, CA, USA* (Vol. 2 pp. 1745–1749).
17. Zeineddine, A., Paquelet, S., Nafkha, A., Jezequel, P., Moy, C. (2019). Efficient arbitrary sample rate conversion for multi-standard digital front-ends. In *IEEE 17th International New Circuits and Systems Conference, Munich, Germany* (pp. 1–4).
18. Choi, J.I., Jun, H.S., Hwang, S.Y. (1996). Efficient hardware optimisation algorithm for fixed point digital signal processing ASIC design. *Electronics Letters*, *32*(11), 992–994.
19. Crochiere, R. (1975). A new statistical approach to the coefficient word length problem for digital filters. *IEEE Transactions on Circuits and Systems*, *22*(3), 190–196.
20. Wong, P.W. (1991). Quantization and round off noises in fixed-point FIR digital filters. *IEEE Transactions on Signal Processing*, *39*(7), 1552–1563.
21. Qiao, J., Fu, P., Meng, S. (2006). A combined optimization method of finite wordlength FIR filters. In *First International Conference on Innovative Computing, Information and Control - Volume I, Beijing, China* (pp. 103–106).
22. Jackson, L.B. (1970). On the interaction of round off noise and dynamic range in digital filters. *Bell System Technical Journal*, *49*(2), 159–184.
23. Sung, W., & Kum, K.I. (1995). Simulation-based word-length optimization method for fixed-point digital signal processing systems. *IEEE Transactions on Signal Processing*, *43*(12), 3087–3090.
24. Menard, D., Rocher, R., Sentieys, O. (2008). Analytical Fixed-Point Accuracy Evaluation in Linear Time-Invariant Systems. *IEEE Transactions on Circuits and Systems I:, Regular Papers*, *55*(10), 3197–3208.
25. Erup, L., Gardner, F.M., Harris, R.A. (1993). Interpolation in digital modems. II. Implementation and performance. *IEEE Transactions on Communications*, *41*(6), 998–1008.

26. Laakso, T.I., Valimaki, V., Karjalainen, M., Laine, U.K. (1996). Splitting the unit delay [FIR/all pass filters design]. *IEEE Signal Processing Magazine*, *13*(1), 30–60.
27. Vesma, J., & Saramäki, T. (1996). Interpolation filters with arbitrary frequency response for all-digital receivers. In *IEEE International Symposium on Circuits and Systems. Circuits and Systems Connecting the World, Atlanta, GA, USA* (Vol. 2 pp. 568–571).
28. Henker, M., & Fettweis, G. (2000). Combined filter for sample rate conversion, matched filtering and symbol synchronization in software radio terminals. Proceedings of the European Wireless pp. 61–66.
29. Hentschel, T., & Fettweis, G. (2001). Continuous-time digital filters for sample-rate conversion in reconfigurable radio terminals. *Frequenz*, *55*, 185–188.
30. Babic, D., Vesma, J., Saramäki, T., Renfors, M. (2002). Implementation of the transposed Farrow structure. In *IEEE International Symposium on Circuits and Systems. Proceedings (Cat. no.02CH37353), Phoenix-Scottsdale, AZ, USA*.
31. Babic, D., & Renfors, M. (2005). Power efficient structure for conversion between arbitrary sampling rates. *IEEE Signal Processing Letters*, *12*(1), 1–4.
32. Reconstruction of non-uniformly sampled signal using transposed Farrow structure (2004)
33. Li, W., & Tomisawa, M. (2004). Transposed-farrow-structure-based multirate filters for symbol timing synchronization in software defined radio (SDR). In *IEEE 60th Vehicular Technology Conference, Los Angeles, CA* (Vol. 3 pp. 1668–1672).
34. Paquelet, S., & Savaux, V. (2018). On the symmetry of FIR filter with linear phase. *Digital Signal Processing*, *81*, 57–60.
35. Meijering, E. (2002). A chronology of interpolation: from ancient astronomy to modern signal and image processing. *Proceedings of the IEEE*, *90*(3), 319–342.
36. Unser, M. (1999). Splines: a perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, *16*(6), 22–38.
37. Dooley, S.R., Stewart, R.W., Durrani, T.S. (1999). Fast on-line B-spline interpolation. *Electronics Letters*, *35*(14), 1130–1131.
38. Milovanovic, G.V., & Udovicic, Z. (2010). Calculation of coefficients of a cardinal b-spline. *Applied Mathematics Letters*, *23*(11), 1346–1350.
39. Huang, X., Guo, Y.J., Zhang, J.A. (2012). Sample rate conversion using B-Spline interpolation for OFDM based software defined radios. *IEEE Transactions on Communications*, *60*(8), 2113–2122.
40. Välimäki, V. (1995). A new filter implementation strategy for Lagrange interpolation. In *IEEE International Symposium on Circuits and Systems, Seattle, WA, USA* (Vol. 1 pp. 361–364).
41. Deng, T. (2007). Coefficient-symmetries for implementing arbitrary-order lagrange-type variable fractional-delay digital filters. *IEEE Transactions on Signal Processing*, *55*(8), 4078–4090.
42. Soontornwong, P., Chivapreecha, S., Pradabpet, C. (2011). A cubic hermite variable fractional delay filter. In *IEEE International Symposium on Intelligent Signal Processing and Communications Systems, Chiang Mai, Thailande*.
43. Tseng, C., & Lee, S. (2012). Design of fractional delay filter using hermite interpolation method. *IEEE Transactions on Circuits and Systems I:, Regular Papers*, *59*(7), 1458–1471.
44. Vesma, J., & Saramäki, T. (1997). Optimization and efficient implementation of FIR filters with adjustable fractional delay. In *IEEE International Symposium on Circuits and Systems, Hong Kong,* (Vol. 4 pp. 2256–2259).
45. Lu, W.S., & Deng, T.B. (1999). An improved weighted least-squares design for variable fractional delay FIR filters. *IEEE Transactions on Circuits and Systems II:, Analog and Digital Signal Processing*, *46*(8), 1035–1040.
46. Johansson, H., & Lowenborg, P. (2003). On the design of adjustable fractional delay FIR filters. *IEEE Transactions on Circuits and Systems II:, Analog and Digital Signal Processing*, *50*(4), 164–169.
47. Hunter, M.T. (2008). *Design of polynomial-based filters for continuously variable sample rate conversion with applications in synthetic instrumentation and software defined radio, PhD thesis*. Orlando: University of Central Florida.
48. Lopez, F., Vesma, J., Saramäki, T., Renfors, M. (2000). The effects of quantizing the fractional interval in interpolation filters. In *Proceedings of the Nordic Signal Processing Symposium, Sweden*.
49. Lopez, F., Vesma, J., Renfors, M. (2002). Defining the word length of the fractional interval in interpolation filters. In *11th European Signal Processing Conference, Toulouse, France*.
50. Lyons, R. (2005). Understanding cascaded integrator-comb filters. *Embedded Systems Programming*, *18*(4), 14–27.
51. Kwentus, A.Y., Jiang, Z., Willson, A.N. (1997). Application of filter sharpening to cascaded integrator-comb decimation filters. *IEEE Transactions on Signal Processing*, *45*(2), 457–467.
52. Altera (2007). Understanding CIC compensation filters. Application Note 455.
53. Molnar, G., & Vucic, M. (2011). Closed-Form Design of CIC compensators based on maximally flat error criterion. *IEEE Transactions on Circuits and Systems II:, Express Briefs*, *58*(12), 926–930.
54. Fernandez-Vazquez, A., & Dolecek, G.J. (2012). Maximally flat CIC compensation filter: design and multiplierless implementation. *IEEE Transactions on Circuits and Systems II:, Express Briefs*, *59*(2), 113–117.
55. Dolecek, G.J., & Mitra, S.K. (2008). Simple method for compensation of CIC decimation filter. *Electronics Letters*, *44*(19), 1162–1163.
56. Dolecek, G.J. (2009). Simple wideband CIC compensator. *Electronics Letters*, *45*(24), 1270–1272.
57. Gao, Y., Jia, L., Isoaho, J., Tenhunen, H. (2000). A comparison design of comb decimators for sigma-delta analog-to-digital converters. *Analog Integrated Circuits and Signal Processing*, *22*, 51–60.
58. Shahana, T.K., James, R.K., Jose, B.R., Jacob, K.P., Sasi, S. (2007). Polyphase Implementation of Non-recursive Comb Decimators for Sigma-Delta A/D Converters. In *IEEE Conference on Electron Devices and Solid-State Circuits, Tainan* (pp. 825–828).
59. Depalle, P., & Tassart, S. (1996). Fractional delay lines using Lagrange interpolators. In *International Computer Music Conference, Hongkong* (pp. 341–343).
60. Tassart, S., & Depalle, P. (1997). Analytical approximations of fractional delays: lagrange interpolators and allpass filters. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, Munich* (Vol.1, pp. 455–458).
61. Franck, A., & Brandenburg, K. (2009). A Closed-Form description for the continuous frequency response of lagrange interpolators. *IEEE Signal Processing Letters*, *16*(7), 612–615.
62. Hermanowicz, E. (1992). Explicit formulas for weighting coefficients of maximally flat tunable FIR delayers. *Electronics Letters*, *28*(20), 1936–1937.
63. Ramachandran, V., Gargour, C.S., Ahmadi, M. (1990). Generation of digital transfer functions of the FIR type approximating z/sup -p/q/. In *IEEE International Symposium on Circuits and Systems, New Orleans, LA, USA* (Vol.4 pp. 3260–3262).
64. Pei, S.C., & Wang, P.H. (2001). Closed-form design of maximally flat FIR Hilbert transformers, differentiators, and fractional delayers by power series expansion. *IEEE Transactions on Circuits and Systems I:, Fundamental Theory and Applications*, *48*(4), 389–398.

6.1.2 [C2]: FPGA/ASIC implementation of reconfigurable SRC

The main objective of this second contribution is to design and optimize a FIR-based SRC function that is implemented in hardware. We address the hardware implementation aspects of the SRC on both ASIC and FPGA targets. The SRC may be implemented in different strategies depending on the application and the deployment context. Each deployment environment may have different hardware requirements (lower computational complexity, lower consumption, higher operating speeds, *etc*...). Depending on the deployment requirements, certain implementation strategies may be more suited for the given hardware constraints. The comparisons include various combinations of nine different filter responses (CIC/Spline, Hermite, Lagrange, and their combinations) of different orders, five different SRC structures (CIC, Farrow, Newton and generalized Newton), and three implementation strategies (ALU-based, pipelined and time-shared).

For both hardware targets, ASIC and FPGA, the implementation results validated the advantages of the different hardware design strategies. As shown in Fig. 6.3, the implementation of different architectures on ASIC showed a better optimization of the complexity compared to the implementation on FPGA. The power consumption levels were also more proportional to the complexity of the implementation comparing to FPGA. Most importantly, implementation results showed that the different structures are able to operate at very high frequencies, making them useful not only for IoT standards, but also for high performance wireless standards. The results also validated the high efficiency of the proposed Newton structure for Hermite interpolation relatively to the classical Farrow and Newton structures.

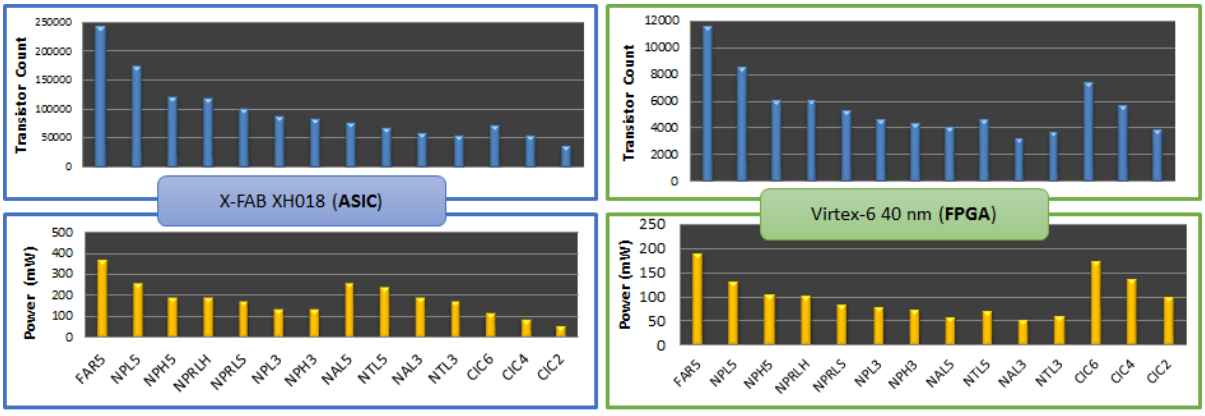


Figure 6.3 – ASIC and FPGA implementation results (Transistors & Power consumption)

This work was published in IEEE International New Circuits and Systems Conference in 2019 [CI53], and included hereafter.

Efficient Arbitrary Sample Rate Conversion for Multi-Standard Digital Front-Ends

Ali Zeineddine*‡, Stéphane Paquelet†, Amor Nafkha‡, Pierre-Yves Jezequel§, Christophe Moy¶

*TDF, Direction Technique Rennes, F-35510, Cesson-Sévigné, France
†Institut de Recherche Technologique b-com, , F-35510, Cesson-Sévigné, France
‡SCEE/IETR UMR CNRS 6164, CentraleSupélec campus de Rennes, F-35510 Cesson-Sévigné, France
§TDF, Centre de mesures d'antennes, F-35340, Liffré, France
¶Univ Rennes, CNRS, IETR - UMR 6164, F-35000, Rennes, France

***Abstract*—In this paper, we study the hardware implementation of arbitrary sample rate conversion (ASRC) using recently proposed variable fractional delay filter (V-FDF) structures. The most commonly used solution to implement V-FDFs has been the Farrow structure for the last three decades. In this work, we develop and compare the implementations of different recently proposed V-FDF options based on the Newton structure. These implementations are done on both ASIC and FPGA targets. The obtained results show that the recently proposed solutions offer similar ASRC performance while using up to 3 times less resources relatively to the classical Farrow structure. The generic nature of these filters make them suited for a large number of standards.**

***Index Terms*—Arbitrary Sample Rate Conversion, Digital Front-End, Newton Structure**

I. Introduction

For many years, the evolution of digital signal processing (DSP) in telecommunication systems has been focused on improving processing power performance. Recently, more focus is increasingly given to improving the system's energy efficiency, and to building lower cost hardware. One of the main motivators of this trend is the evolution of the Internet of Things (IoT) domain, where billions of devices and their corresponding gateways are being deployed. To make this deployment possible at an acceptable cost, both devices and gateways need to be maximally optimized. A main part of every telecommunication system is the interface between radio frequency and baseband domains, implemented digitally in modern systems and known as the digital front-end (DFE) [1]. The DFE has two main roles, sample rate conversion (SRC) and filtering [2]. In this work, we are concerned with SRC. Two types of SRC exist in the DFE. The first is coarse SRC, where the sampling rate is increased or reduced by an important integer factor. The other type is known as fine SRC or arbitrary sample rate conversion (ASRC). Generally, fine SRC is more complicated to implement than coarse SRC due to the extra required precision.

To implement ASRC, a variable fractional delay filter (V-FDF) is the most efficient solution. The Farrow structure [3], dating back to 1988, is the most widely adopted implementation option that is found in most of today's systems. An advantage of this structure is its capability of implementing any ASRC operation, with a reconfigurable conversion factor. However more efficient structures can be found in the literature. In 2009, the Newton structure [4] was adapted to ASRC in [5], however limited to the filtering response of Lagrange interpolation. Later more recent work generalized this structure to Spline and Hermite interpolations [6][7]. Hermite interpolation is preferred for the DFE, due to its wide pass-band and good SRC image rejection performance.

In this paper we aim to investigate the practical hardware complexity of these recently proposed solutions. The rest of the paper is organized as follows. Section II presents an overview of both the most common classical SRC solutions in modern DFE systems, and the recently proposed Newton structures. In Section III, the implementation methodology based on the pipeline approach is developed, and the hardware architecture model is detailed. In Section IV, we implement the developed architectures on both FPGA and ASIC targets. We then compare the Hermite based Newton structures to the modern SRC solutions in both terms of complexity and performance. The conclusion is finally presented in Section V.

II. ASRC Solutions for DFE Systems

One of the main DFE roles in modern transceiver systems is adapting the sampling rate between the radio frequency and baseband domain [1][2]. This conversion is done practically using both coarse and fine-tuned SRC modules. The SRC operation can be used at different locations in the same DFE of a radio transceiver. Moreover, fine SRC have other applications in the DFE, most notably for implementing synchronization functions. Therefore, ASRC modules are a main part of today's transceivers, and their implementation efficiency plays an important role in optimizing the total system.

We usually find the coarse SRC implemented using cascaded-integrator-comb (CIC) filters [8], due to their efficient structure consisting of only registers and adders. However, the CIC filters cannot implement all kind of SRC operations, and are only practical for coarse factor SRC. In the case of fine-tuned SRC, ASRC modules are required, and the Farrow structure is the most common solution [3]. This structure implements a polynomial based V-FDF. Polynomial based means that the filter impulse response is constructed

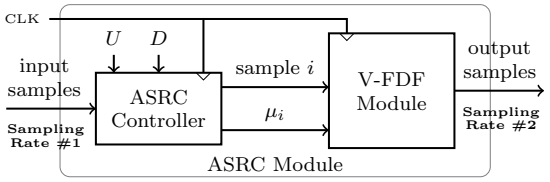


Fig. 1. Simplified high-level architecture of the ASRC module

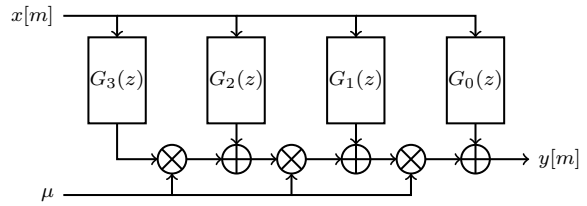
using polynomials pieces. A V-FDF filter interpolates the value of the signal at variable instants defined by a given input μ.

To perform the ASRC operation, the value of μ is updated for every output sample, as explained in [5]. The μ value is then used by the V-FDF to calculate the value of the corresponding output. In practical implementations, there are two important points to consider. First, the number of input and output samples is different, since the sampling rate is modified. Second, the hardware module operates at a single digital clock rate higher than both the input and output sampling rates. Therefore a control module is required alongside the V-FDF to manage the samples stream, and to calculate the required output instant μ for each sample. This control function is configured using the input up-sampling U and down-sampling D parameters representing the SRC factor $R = U/D$. The proposed ASRC module architecture is shown in Figure 1.

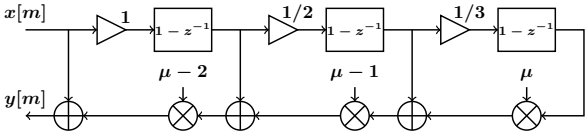
The implementation of the V-FDF using a Farrow structure of order 3 is shown in Figure 2-a. This structure consists of multiple FIR filters $G_i(z)$ that have their outputs multiplied by μ according to the Horner scheme, to find the final output $y[m]$. For a theoretical understanding of the structure, the reader can refer to [3] and [9]. The Farrow structure can implement any polynomial-based filter response, however the most simple and commonly used option is Lagrange interpolation. To achieve side lobes rejection levels of at least 30 dB, an order 5 is required.

A structure developed in [5] modifies the original Newton structure [4] into a V-FDF form compatible with ASRC. The structure of order 3 is shown in Figure 2-b. Compared to the Farrow structure, the Newton structure is designed for Lagrange interpolation only. However structurally, an order N interpolation can be achieved with a complexity of order $O(N)$ through the Newton structure, compared to a quadratic order of complexity $O(N^2)$ for the Farrow structure. Later work extended the Newton structure to implement Spline [6] and Hermite [7] interpolation. Spline interpolation is not very interesting for multi-standard DFE systems due to its very small passband, and its bad scaling with interpolation order. On the other hand, Hermite interpolation keeps the same passband of Lagrange interpolation, while offering superior side lobes rejection levels. At the same time, the Hermite based Newton structures offer lower complexity by eliminating the need for some multipliers. The Hermite based Newton structures of order 3 and 5 are shown in Figure 2-c and Figure 2-d respectively.

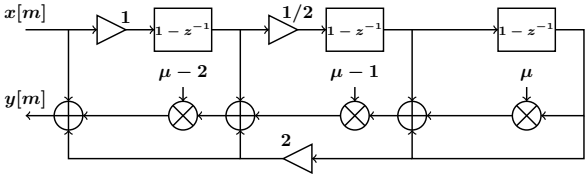
These two architectures are very promising from a structural level point of view. In this paper, we are interested in investi-



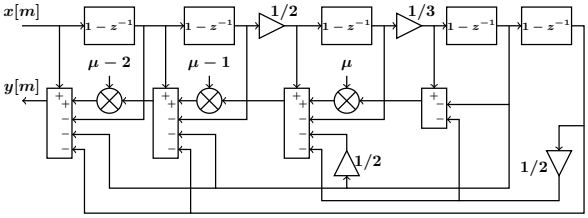
(a) Farrow structure of order 3 [1]



(b) Newton structure for Lagrange order 3 [5]



(c) Newton structure for Hermite order 3 [7]



(d) Newton structure for Hermite order 5 [7]

Fig. 2. V-FDF module implementation options for (a) Farrow order 3, (b) Newton Lagrange order 3, (c) Newton Hermite order 3, (d) Newton Hermite order 5

gating the hardware implementation complexity of these recent structures, and compare them to the modernly used ASRC solutions. In the next section, we develop the methodology we used to implement these structure in order to compare their complexity on both ASIC and FPGA targets.

III. Hardware Implementation Methodology

This section develops the implementation approach used to obtain the results discussed in Section IV. The complete filter is composed of two modules, the controller and the V-FDF filter. As discussed in the last section, the controller is responsible for managing the SRC operation, while the V-FDF filter is only responsible for calculating the output samples.

To implement the ASRC controller, a finite state-machine (FSM) is the most appropriate approach. This FSM has the U and D parameters as inputs. The control FSM has 4 states, starting from the initial IDLE state, the FSM may be in the NEUTRAL state when there is only one sample output corresponding to the current input. The FSM may also be

in the INTERPOLATION and DECIMATION states when more outputs than inpust or the opposite conditions exist respectively. The value of μ is continuously updated to keep track of the output sample time delay.

Since there could be more outputs than inputs or vice versa, a mechanism to block or advance the samples stream at certain instants is needed. In this work we used a handshake protocol based on "ready to send and receive" signals. Each module is responsible of signaling its own status. Between the controller and the V-FDF filter modules, extra control signals are used to signal if a register update or a new calculation is required.

To implement the V-FDF module, we use the pipeline approach that is the most adapted for multi-standard DFE, where processing speed is privileged. The pipeline approach consists of breaking the filter structure into stages, with each stage containing only one calculation operation. This has the objective of minimizing the critical path length, and maximizing thereby the operation frequency. The pipeline architecture for the modified Newton structure for Hermite interpolation of order 3 is shown in Figure 3.

The quantization of this implementation is done using fixed point representation. The signal's quantization parameters are found by developing the analytical expression of the quantization error [10]. This expression is then used to find the signal to quantization noise ratio (SQNR). The optimal quantization parameters are then found using exhaustive research for a given SQNR. Considering for example an input signal quantized on 2 and 16 bits for the integer and fractional parts respectively, Figure 3 shows the quantization parameters for tolerated SQNR degradation of less than 0.6 dB. This quantity is chosen in order to have negligible deterioration of the effective number of bits due to quantization. The signals are quantized relatively to the input, where the number of added bits for integer and fractional parts is shown between braces, e.g. $\{\text{x.y}\} \rightarrow$ (2+x.16+y). For the fractional delay μ quantization on 6 bits, we referred to the work presented in [11]. Finally, the U and D parameters are quantized on 18 unsigned bits for an ASRC precision of 5 ppm.

Using the approach developed above, the hardware implementations of the different V-FDF options shown in Figure 2 are developed with the same quantization performance. An implementation of the CIC filter is also developed to compare the complexity between fine and coarse SRC solutions.

IV. Implementation Results and Discussion

We start the discussion by comparing the filtering performance of the different structures. The objective of an SRC filter is to remove the signal images on the multiples of the input sampling frequency F_s. As it is shown in Figure 4, the first side lobe attenuation for Lagrange interpolation of order 5 is around -33 dB. However for both cases of Hermite interpolation of order 3 and 5, we find a side lobes rejection improvement by around 10 dB. All three responses have a maximally flat passband, and zeros on multiples of F_s. The only major compromise is the Hermite interpolation of order 3 having weaker zeros. However this may be tolerable for certain

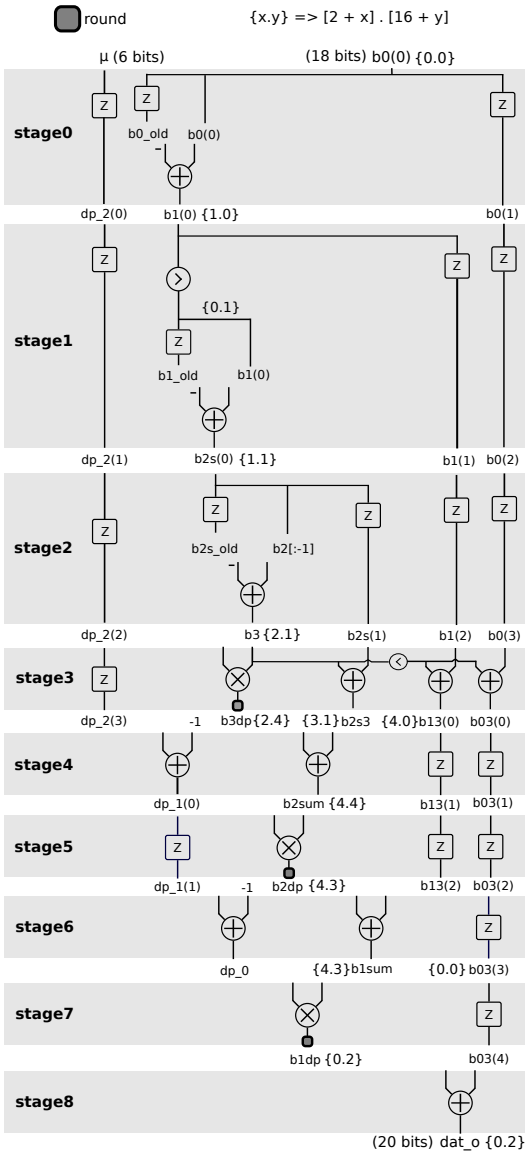


Fig. 3. Pipeline architecture of Newton Structure for Hermite order 3

SRC applications that do not have high image attenuation requirements, which is the case of the SRC modules in the DFE that come after the filtering operations.

To develop the ASIC implementation, we used the Cadence Encounter tools with the X-FAB XH018 technology. The results are resumed in Table I. The implementations were optimized to operate at 180 MHz. It is seen that implementing Lagrange interpolation using the Newton structure only requires 70% the resources used by the reference Farrow structure. However, the Newton Hermite structure of order 5 requires less than half of the resources while offering a superior filtering response. Moreover, when using the Hermite order 3 response is tolerable, it is possible to implement fine SRC operations at a similar cost of coarse SRC CIC filters, with a reduction of complexity by a factor of three compared to the Farrow reference. The dynamic power draw of the

TABLE I
ASIC IMPLEMENTATION RESULTS USING THE X-FAB XH018 TECHNOLOGY

Module	Gate Count	Core Surface	Transistors Count	Complexity %	Max. Frequency	Core Power
Farrow Lagrange 5	27328	$950 \times 950\ \mu m^2$	243937	100%	180 MHz	362 mW
Newton Lagrange 5	19824	$800 \times 800\ \mu m^2$	175475	70.9%	179 MHz	251 mW
Newton Hermite 5	14803	$660 \times 660\ \mu m^2$	122838	48.3%	180 MHz	187 mW
Newton Hermite 3	10903	$540 \times 540\ \mu m^2$	84520	32.3%	180 MHz	134 mW
CIC Spline 5	8299	$510 \times 510\ \mu m^2$	73768	28.8%	180 MHz	113 mW

TABLE II
FPGA IMPLEMENTATION RESULTS USING THE XILINX VIRTEX-6 XC6VLX365T

Module	Registers	Look-Up Tables	Total Logic	Complexity %	Max. Frequency	Core Power
Farrow Lagrange 5	5497	6126	11623	100%	172 MHz	100 mW
Newton Lagrange 5	3744	4811	8555	73.6%	162 MHz	84 mW
Newton Hermite 5	2621	3470	6091	52.4%	172 MHz	71 mW
Newton Hermite 3	1577	2838	4415	38.0%	164 MHz	63 mW
CIC Spline 5	4173	3275	7448	64.0%	232 MHz	77 mW

ASIC cores is estimated using Encounter RTL Compiler for an operating frequency of 180 MHz, a signal toggle rate of 20 MHz, and 1.8V core voltage.

For the FPGA implementations, we used the Xilinx Virtex-6, that gave the results shown in Table II. The dynamic power draw of the SRC cores in this case are estimated using Xilinx Power Estimator (XPE) for an operating frequency of 160 MHz, a signal toggle rate of 20 MHz, and 1.0V core voltage. The objective is not to compare the ASIC and FPGA implementations, but rather to study the implementation complexity on FPGA. The order of complexity between the Newton and Farrow structures stay the same, however the implementation of a CIC filter using look-up tables on an FPGA is not optimal, where the relative complexity is much larger on FPGA than on ASIC. Regardless, the results clearly show that the Hermite based Newton structures offer an advantageous replacement of the widely used Farrow structure, by offering improved performance at a lower complexity cost.

V. CONCLUSION

In this paper, we developed the hardware implementation of different ASRC modules, including the recently proposed modified Newton structures for Hermite interpolation. The quantized hardware architectures were developed using a pipeline approach, and the implementation was then done on both FPGA and ASIC. The results showed that the different structures are able to operate at very high frequencies, making them useful not only for IoT standards, but also for high performance wireless standards. The results also validated the high efficiency of the recently proposed Newton structure for Hermite interpolation relatively to the classical Farrow and Newton structures.

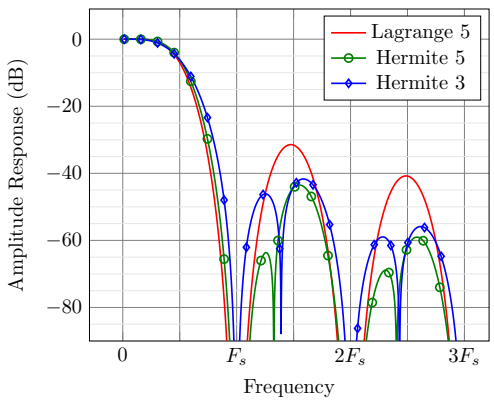


Fig. 4. Frequency responses of Lagrange and Hermite interpolations

VI. REFERENCES

[1] L. Fa-Long, ed. *Digital Front-End in Wireless Communications and Broadcasting: Circuits and Signal Processing*. United Kingdom: Cambridge University Press, 2011.

[2] G. Fettweis and T. Hentschel. "The Digital Front End: Bridge Between RF and Baseband Processing". In: *Software Defined Radio*. Wiley-Blackwell, 2002. Chap. 6, pp. 151–198. ISBN: 9780470846001.

[3] C. W. Farrow. "A continuously variable digital delay element". In: *Circuits and Systems, 1988., IEEE International Symposium on*. 1988, 2641–2645 vol.3.

[4] S. Tassart and P. Depalle. "Fractional delay lines using Lagrange interpolation". In: *Proceedings of the 1996 International Computer Music Conference*. The International Computer Music Association. 1996, pp. 341–343.

[5] V. Lehtinen and M. Renfors. "Structures for interpolation, decimation, and nonuniform sampling based on Newton's interpolation formula". In: *SAMPTA'09*. 2009, Special–session.

[6] D Lamb, L. Chamon, and V. Nascimento. "Efficient filtering structure for spline interpolation and decimation". In: *Electronics Letters* 52.1 (2016), pp. 39–41.

[7] A. Zeineddine, A. Nafkha, C. Moy, S. Paquelet, and P.-Y. Jezequel. "Variable fractional delay filter: A novel architecture based on hermite interpolation". In: *25th International Conference on Telecommunications, ICT 2018* (2018).

[8] E. Hogenauer. "An economical class of digital filters for decimation and interpolation". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 29.2 (1981), pp. 155–162. ISSN: 0096-3518.

[9] F. Harris. "Performance and design considerations of the Farrow filter when used for arbitrary resampling of sampled time series". In: *Signals, Systems amp; Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on*. Vol. 2. 1997, 1745–1749 vol.2.

[10] D. Menard, R. Rocher, and O. Sentieys. "Analytical fixed-point accuracy evaluation in linear time-invariant systems." In: *IEEE Trans. on Circuits and Systems* 55.10 (2008), pp. 3197–3208.

[11] J. Vesma, F. Lopez, T. Saramäki, and M. Renfors. "The effects of quantizing the fractional interval in interpolation filters". In: *Proc. IEEE Nordic Signal Processing Symp*. 2000, pp. 215–218.

6.2 Dynamic Partial Reconfiguration of FPGAs

For traditional field programmable gate array (FPGA) reconfiguration computing, one major drawback is the lack of flexibility, since the whole FPGA is required to be reconfigured even when the modification is only required for a small part of FPGA. As dipected in Fig. 6.4, dynamic partial reconfiguration (DPR) allows a given region (*i.e.* reconfigurable region) of the FPGA fabric to be reconfigured while the remainder of the fabric (*i.e.* static region) can continue to operate unaffected [2, 3, 4, 5]. DPR allows the implementation of more power-efficient designs by using hardware on demand, that is, only instantiate the logic that is necessary at a given time and remove unused logic. Moreover, DPR has the benefit of reduced reconfiguration time since this time is directly proportional to the area of the reconfigurable region.

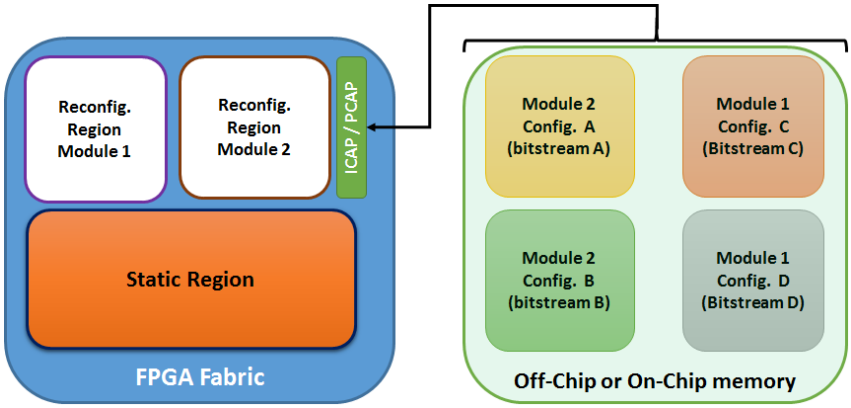


Figure 6.4 – Dynamic Partial Reconfiguration scheme.

Configuration overheads (time, power) are major performance bottlenecks of the dynamic partial partial reconfiguration and they can seriously limits the usefulness of this technique. The performed work, in the context of the improvement of the reconfiguration speed, was done during my postdoctoral fellowship (Apr. 2006 - Dec. 2007) and was supported by E2R and IDROMEL projects. In collaboration with my colleague Pr. Pierre Leray, we have proposed the use of direct memory access techniques to directly transfer partial configuration bitstreams to the internal configuration access port (ICAP) primitive present in the Xilinx FPGA fabric without using HWICAP core provided by Xilinx. Our publication in IEEE International Conference on Reconfigurable Computing and FPGAs confirm that the DMA-based technique provides a significant reconfiguration time overhead reduction [CI11].

As a natural evolution, in 2016, I have investigated experimentally the problem of the power consumption overhead during the dynamic partial reconfiguration process. The Xilinx ML550 board has been used to perform different power consumption measurements. This board contains a series of shunt 10 $m\Omega \pm 1\%$ current sense resistors on each voltage regulator lines (V_{CCint}, V_{CCo}, and V_{CCaux}). Since the sensitivity of the voltage regulator sensors are pretty low (*i.e* 0.5 mV - 2

mV) and the partial reconfiguration time is very small (*i.e* a few tens of microseconds), voltage amplifiers and a very sensitive oscilloscope were needed to take an accurate measurements. To this end, analog devices' instrumentation amplifiers AD620 and TDS2024C oscilloscope (200 MHZ & 2 GSPS) from Tektronix are chosen to accurately measure different FPGA voltages.

Once the FPGA fabric is powered on and configured, power consumption takes two basic forms: dynamic and static. The dynamic power is the power consumed by the device during the switching activity of the core or in the input/output pins. The static power is the power consumed by the device due to leakage currents when there is no activity (*i.e.* circuit is powered up, configured and doing nothing). Currently, the static power consumption represents a significant percentage of the total FPGA power dissipation. In fact, below 65 nm technology node, the increase in power consumption is mainly related to the leakage current. It's well known that the current leakage is dramatically influenced by junction temperature of the integrated circuit. This is why it is important to evaluate how we can use dynamic partial reconfiguration to manage the thermal profile of the FPGA fabric and by consequence the static power consumption. Still targeting the benefits of dynamic partial reconfiguration technique, I have investigated leakage power dissipation in FPGAs and the possibility to reduce it by means of effective management of FPGA thermal profile through dynamic partial reconfiguration.

6.2.1 [C3]: Timing Overhead Reduction

The dynamic partial reconfiguration of FPGAs provides good advantages to software defined radio systems to load the desired standard in runtime manner. Indeed, the ability of modifying a specific hardware block inside FPGA device in runtime, while all other blocks are working normally, provides an opportunity to create an extremely flexible design. Thus, DPR can be used to achieve high hardware flexibility and an efficient utilization of hardware resources. However, the major issue with the utilization of DPR is the configuration speed because FPGA reprogramming process induces power and performance overheads. Thus, by maximizing the reconfiguration speed, these overheads can be minimized.

For Xilinx FPGA fabric, it is possible to perform DPR using the internal configuration access port (ICAP) primitive that allows us to access to the FPGA configuration memory both in read and write mode. As depicted in Fig. 6.5, the ICAP data interface can be set to one of three data widths: 8, 16, or 32 bits. CSB is the active-low interface select signal, RDWRB is the read/write select signal. BUSY is valid only for read operations and remains low for write operations. The maximum recommended frequency of operation for the ICAP is 100 MHz and the theoretical maximum performance is 400 MB/s when the input data width is 32 bits. In practical, the reconfiguration process is controlled by a processor (MicroBlaze, PowerPC or Cortex-A9, ...) via a vendor-provided controller such as OPB_HWICAP and XPS_HWICAP [6], connected as a slave peripheral to the soft/hard processor bus. Using these controllers leads to a relatively

low throughput in the range of 6.0 MB/s - 14.6 MB/s as published in [7].

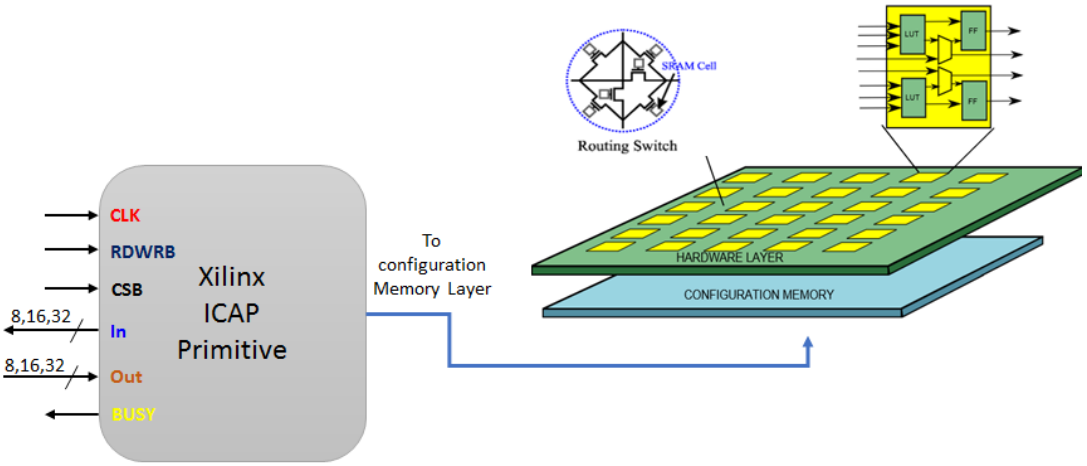


Figure 6.5 – The ICAP primitive on Xilinx FPGAs.

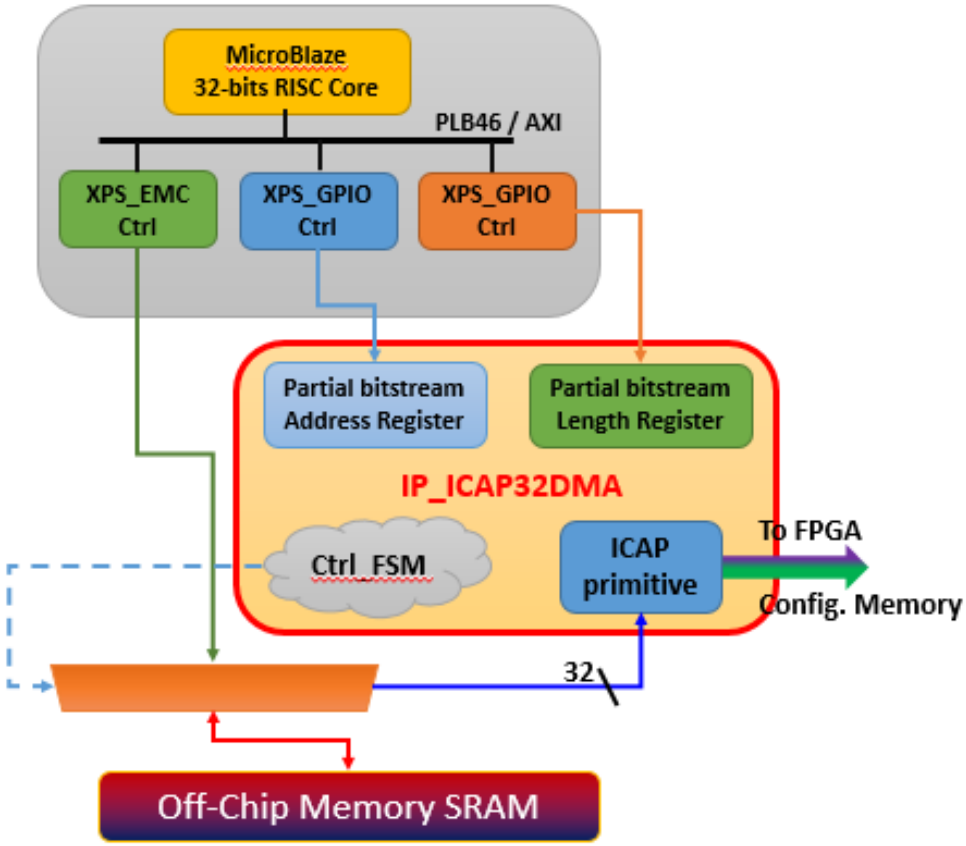


Figure 6.6 – IP_ICAP32DMA block diagram and the structure of the test system.

In order to enhance the reconfiguration speed, Hübner *et* al. proposed to connect the ICAP

controller to the fast simplex link (FSL) bus of the embedded processor in order to maximize the reconfiguration throughput [8]. The maximum speed they have attained is 28.28 MB/s with FSL-ICAP (32-bit, 100MHz) on PowerPC hard processor and 25.89 MB/s on MicroBlaze soft processor. For Xilinx Zynq platform which integrates an ARM processor with a FPGA, Vipin *et* al. developed a custom ICAP controller, called ZyCAP, that significantly improves reconfiguration throughput.

In 2008, in collaboration with Pr. Pierre Leray, we have proposed a high reconfiguration speed ICAP controller, called IP_ICAP32DMA, which is very close to theoretical performance (*i.e* 400 MB/s). The main idea is to modify the ICAP controller by using DMA to transfer partial bitstreams from external memory to ICAP controller. A block diagram of the IP_ICAP32DMA is shown in Fig. 6.6. In Table 6.1, the ICAP reconfiguration throughput speed reported and presented by different authors previously in the literature, is depicted and compared with the results obtained with our approach

In 2011, researchers have tried overclocking the ICAP controller in order to achieve higher reconfiguration speed. However, following Xilinx recommendations, the maximum clock frequency at which ICAP controller should operate is 100 MHz. In [9], authors reported the possibility to overclock the ICAP interface up to 550 MHz without any failure. Thus, the theoretical maximum reconfiguration speed can reach 2200 MB/s.

Source	**Device**	**Bit Width**	**Throughput**
2008 [6]	Virtex-4	32	295 MB/s
2008 [10]	Virtex-4	32	350 MB/s
2009 [7]	Virtex-4	32	378 MB/s
2014 [11]: ZyCAP	Zynq-7000	32	382 MB/s
2009 [12]: ICAP_DMA	Virtex-4	32	392 MB/s
2012 [13]: ICAP_DMA	Virtex-4	32	399 MB/s
2014 [14]: DyRACT	Virtex-6 / Virtex-7	32	364 MB/s
2015 [15]: AC_ICAP	Virtex-5 / Kintex7	32	381 MB/s
2015 [16]: MiCAP-Pro	Zynq-7000	32	272 MB/s
2017 [17]: RT-ICAP	Zynq-7000	32	382 MB/s
Proposed 2008 [CI11]	Virtex-4/Virtex-5	8/32	380 MB/s
Proposed 2011 [CI21] + Assembly code	Virtex-4/Virtex-5	8/32	385 MB/s

Table 6.1 – Reconfiguration throughput speed of the ICAP port clocked et 100 MHz.

Table 6.1 summarises the maximum measured throughput speed of different ICAP controllers that are proposed in literature. We can notice that our proposed ICAP controller based on DMA access has a relatively high throughput compared to existing implementation approaches.

6.2.2 [C4]: Power Consumption Overhead Analysis

Always on the FPGA dynamic partial reconfiguration technique, another contribution concerns the evaluation of the power consumption during the reconfiguration process. The analysis of the power consumption overheads has been the target of many research works [18, 19, 20, 21]. In [19], authors proposed model to estimate the power consumption during the reconfiguration process. Moreover, in order to verify the accuracy of their model, they carry out experimental measurements of the FPGA power consumption using the shunt resistor method and a high-speed digital oscilloscope.

In our publication [CI46] included hereafter, we measured the power consumption overhead when applying partial reconfiguration on Virtex 5 FPGAs. The MicroBlaze processor core and the IP_ICAP32DMA controller were used to apply partial reconfiguration and results showed that the additional power consumption does not exceed 160 mw during the reconfiguration process as depicted in Table 6.2.

Reconf. (Mode)	Reconf. time	FPGA core power	Power Overhead
FPGA Power-on	ND	340 mW	ND
Total (JTAG)	3.8 s	220 mW	-120 mW
Partial (JTAG)	208.9 ms	ND	ND
Partial (IP_ICAP32DMA)	324 μs	500 mW	160 mW

Table 6.2 – Time and power overheads during dynamic partial reconfiguration of Virtex-5 device.

Fig. 6.7 presents the setup used to measure the power consumption during dynamic partial reconfiguration process. We have used a ML550 development board from Xilinx, which provides us with five power rails (core, IOs, peripherals) and current sense resistors. The board hosts a two header connector which provides test points for power regulators.

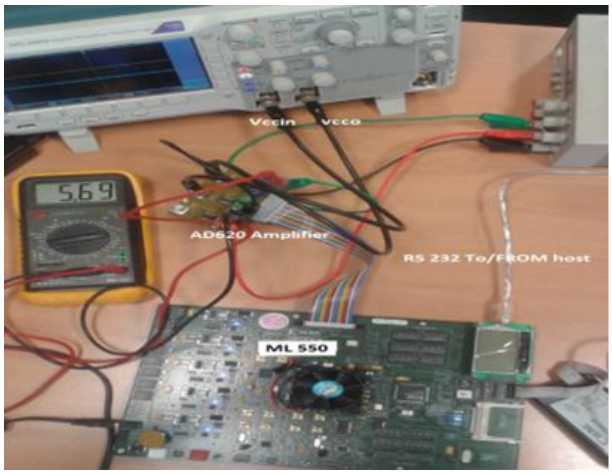


Figure 6.7 – Power measurement setup.

Accurate Measurement of Power Consumption Overhead During FPGA Dynamic Partial Reconfiguration

Amor Nafkha, and Yves Louet
CentraleSupélec/IETR, Campus de Rennes
Avenue de la Boulaie, CS 47601, 35576 Cesson Sévigné Cedex, France.
Email: {amor.nafkha,yves.louet}@centralesupelec.fr

***Abstract*—In the context of embedded systems design, two important challenges are still under investigation. First, improve real-time data processing, reconfigurability, scalability, and self-adjusting capabilities of hardware components. Second, reduce power consumption through low-power design techniques as clock gating, logic gating, and dynamic partial reconfiguration (DPR) capabilities. Today, several application, *e.g.*, cryptography, Software-defined radio or aerospace missions exploit the benefits of DPR of programmable logic devices. The DPR allows well defined reconfigurable FPGA region to be modified during runtime. However, it introduces an overhead in term of power consumption and time during the reconfiguration phase. In this paper, we present an investigation of power consumption overhead of the DPR process using a high-speed digital oscilloscope and the shunt resistor method. Results in terms of reconfiguration time and power consumption overhead for Virtex 5 FPGAs are shown.**

I. Introduction

Software Defined Radio (SDR) refers to reconfigurable or re-programmable radios that can perform different functionality with the same hardware [1], [2]. The main objective is to produce communication devices which can support any wireless standards and services. The reconfigurability poses a set of new challenges in terms of real-time processing, power consumption, flexibility, performance. Consequently, many heterogeneous platforms composed of digital signal processors, application specific integrated circuits, and FPGAs have been developed to achieve sufficient processing power and flexibility. To cope with flexibility and computational issues, the concept of FPGA Dynamic Partial Reconfiguration (DPR) can be used [3]–[8].

The DPR is achieved by loading the partial bitstream of a new design into the SRAM-based FPGA configuration memory and overwriting the current one. Thus the reconfigurable portion will change its behavior according to the newly loaded configuration. In this procedure, the reconfiguration time, $i.e.$ reconfiguration throughput, represents an important and critical design parameter which determines the switching time switching time between the two configuration. This factor must be taken into account in many cases where performance critical applications require fast switching of IP cores. To achieve runtime DPR, authors in [4] provide a high-speed interface that transfers the DPR bitstreams into the Internal Configuration Port (ICAP). They measured a maximum DPR speed of 400 MB/s that can be achieved by using Direct Memory Access (DMA) for data transfers between external memory and the ICAP interface (clock speed of 100 MHz). In [9], authors use an over-clocking ICAP interface (clock speed of 133 MHz) and provide a maximum DPR measured speed of 418.5 MB/s.

To the best of our knowledge, there is a few work which address the experimental measurement and/or the power consumption estimation during the DPR process. In [10], authors presents a detailed investigation of power consumption of the DPR process. However, they use underclocked ICAP interface in order to catch power consumption during DPR given the fact that the partial configuration time is very small. Recently, Texas Instruments provides the Fusion Digital Power Designer software package [11] with the USB Interface Adapter EVM [12] to monitor real-time power consumption values. However, this method did not provide high temporal resolution to measure the FPGA power consumption during DPR process. In the present paper, we will present a experimental approach using high-speed digital oscilloscope and the shunt resistor method to provide power consumption during DPR process. One of the main results of this work is that the power consumption overhead, during the FPGA dynamic partial reconfiguration, is low and doesn't exceed 160 mW.

The paper is structured as follows. Section II provides introduction to different actors providing the utilization of Dynamic Partial reconfiguration concept. In order to make experimental results more representative, variable digital filters are introduced and implemented in Section III. Hardware architecture is provided in section IV. The experimental setup is introduced in Section V. The results of the experiments are presented and discussed in Section VI. Finally, Section VII provides a conclusion.

II. Partial reconfiguration

Xilinx company proposes two main FPGA family: Spartan and Virtex devices. Both families can support partial reconfiguration through ICAP component inside the FPGA. The section gives different actors providing the utilization of dynamic partial reconfiguration design approach

A. The ICAP primitive

The ICAP allows to access to configuration data and it has the same signaling interface as SelectMAP and can be configured in 8 or 32 bits mode, depending on the target device use. The ICAP primitive is the hardwired FPGA logic by which the bitstream can be dynamically loaded into the configuration memory. For partial reconfiguration management, we use an embedded microprocessor (either MicroBlaze or PowerPC) to transmit the partial reconfiguration bitstream from storage devices to ICAP in order to accomplish the reconfiguration process. The DPR feature allows to partially reconfigure the FPGA internal configuration logic at the runtime. As a consequence, a specific design flow has to be used so as to define the static areas (which will not change at runtime) and the dynamic areas (which can be changed at runtime). This flow allows to create partial bitstreams for the dynamic and static part. These bitstreams are merged to give the global bitstream which is mandatory for the initial global programming of the FPGA. Then the dynamic part can be reconfigured using the partial bitstreams sent through the ICAP component at runtime.

B. Partial reconfiguration manager

To manage partial reconfiguration, we use an embedded processor soft core, called MicroBlaze, which is a reduced instruction set computer processor optimized for implementation on any Xilinx FPGA. Moreover, Xilinx provides through Embedded Development Kit software an environment tool to connect peripherals to the MicroBlaze and develop application program to drive it. Several communication ports (OPB bus, FSL links, PLB, LMB) are available. As mentioned before, partial reconfiguration is realized through ICAP component inside the FPGA. This reconfiguration is possible by sending partial bitstream to the ICAP. As a consequence, the different bitstreams required have to be stored in external memory like SRAM or internal FPGA memory like BRAM in the order to reduce writing and reading transfer during communication mechanism to ICAP. In this paper, we use the Xilinx ML550 board to make power and time overhead measurements of the dynamic partial reconfiguration process. We have made the choice to store the bitstreams in an internal dual-RAM memory.

C. Partial reconfiguration modes

FPGA dynamic partial reconfiguration can be performed externally, through the JTAG or SelectMap port, or internally through ICAP component. These two modes takes the same configuration files to reconfigure the desired area inside FPGA. A restriction has to be made for the internal reconfiguration, where only partial reconfiguration can be made. Indeed, the management of the process of reconfiguration has to be monitored by a specific interface. For the external configuration, this management could be realized by different components like DSP or GPP. For the internal configuration, this could be managed by a hardware or a software processor as a MicroBlaze or a specific computing unit. As a consequence, for the internal case a global reconfiguration is impossible, after an initialization phase, due to the lost of the control of reconfiguration process it self.

D. Enhancement ICAP architecture

Considering the inefficiency when the processor moves data to ICAP, we propose a design approach based on an appropriate architecture in charge of the partial bitstreams management through the ICAP. We use the MicroBlaze to catch the order of reconfiguration. Partial bitstreams are stored in a RAM and are accessible by the MicroBlaze through the OPBEMC controller. Then bitstream dataflow is sent to the ICAP component using one of the ICAP32 DMA IP. The ICAP32 DMA is an IP that we have developed in our team in order to take advantage of new 32 bit mode of Virtex 5. The goal is to get closer to the theoretical throughput of the ICAP capacity, 400MByte/s (ICAP running at a maximum frequency of 100MHz). The Microblaze is only used for the setup phase, after that, the IP works alone and acts like a DMA (Direct Memory Access) for bitstreams transfers between BRAM and ICAP component. This IP is connected to the MicroBlaze via the GPIO, and the ICAP component is configured in 32 bits mode. Specific registers store the start addresses and the offset of the physical memory where partial bitstreams are stored in RAM. The MicroBlaze sends the order of reconfiguration to the IP for the selected configuration, and the IP makes the memory transfer to the ICAP itself using a SRAM controller included in ICAP32 DMA component. So data writing phases are realized by words of 32 bits at each FPGA clock cycle.

III. Variable Digital Filters

This section provides a detailed description of the implemented design in order to make different measurements

A. APT-VDF filter

All-pass transformation based variable digital filter was initially proposed in [7]. Consider a prototype FIR filter with impulse response $h(n)$ and its z-transform $H(z)$. The APT-VDF version of $H(z)$ is obtained by replacing each unit delay z with the first order all-pass structure, $A(z)$, defined as

$$A(z) = \frac{-\alpha + z^{-1}}{1 - \alpha z^{-1}},\ |\alpha| < 1 \qquad (1)$$

where α is the first order warping coefficient and its value determines the cut-off frequency of the APT-VDF filter. This filter can provide variable frequency response with unabridged control over cut-off frequency. The advantage of the APT-VDF filter is that it allows fine control over cut-off frequency without updating the filter coefficients or structure. The hardware architecture of the APT-VDF filter is depicted in figure 1.

In Fig. 1, the signals sel_f_1 and sel_f_2 are used to select which filter type will be created:

- Lowpass filter when $sel\_f_1 = 0$ and $sel\_f_2 = 0$,
- Bandpass filter when $sel\_f_1 = 1$ and $sel\_f_2 = 0$,
- Highpass filter when $sel\_f_1 = 0$ and $sel\_f_2 = 1$,
- Bandstop filter when $sel\_f_1 = 1$ and $sel\_f_2 = 1$

Consider a N^{th} order lowpass prototype filter with cut-off frequency f_{co}, if $0 < \alpha < 1$, then the resultant cut-off

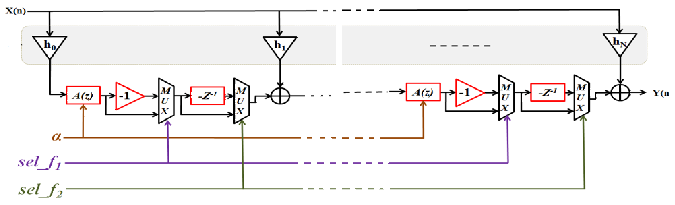


Fig. 1. APT-based VDF filter architecture

frequencies are lower than f_{co}. For $\alpha = 0$, the transfer function of the APT-VDF filter is equal to the original transfer function of the prototype filter. If $-1 < \alpha < 0$, then the resultant cut-off frequencies of the APT-CDF filter are higher than f_{co}. The parameter α can be expressed as

$$\alpha = \frac{sin[(f_{co} - f_c)\pi/2]}{sin[(f_{co} + f_c)\pi/2]} \quad (2)$$

where f_c is the cut-off frequency of the desired APT-VDF filter. Given a desired frequency response specifications, Matlab filter design tool can be used to obtain the prototype filter coefficients.

B. Floating to Fixed point conversion

The APT-VDF architecture has been implemented in a Virtex5 XC5VLX50T FPGA using fixed-point representation. Herein, we consider the generalized fixed-point number representation $[w_l, f_l]$, where w_l and f_l correspond to the word length and the fractional length of the number, respectively. The difference $i_l = w_l - f_l$ is referred to as the integer length of the number. In this paper, the minimum integer word length is calculated under large amounts of simulation data, and the fractional word length are chosen through extensive simulation and tolerable error limit. The root mean square error (RMSE) of output data between floating-point filter and fixed-point filter is evaluated and it is considered as an optimization criteria. Due to the limitation of space, this study will not be presented in this paper. The input, output, and internal signals were represented as an integer and their word length was varied in order to observe the effects of finite precision. Simulations using a serial configuration of the APT-VDF filter at different precisions were made. The results are shown that 12-bit word length and 10-bit fractional precision has significant root mean square error ($RMSE \approx -32dB$), whereas 16-bit word length and 14-bit fractional precision or higher precision results in very low error ($RMSE \approx -44dB$).

C. VHDL Implementation

The APT-VDF filter has to be implemented in such a way that the design has a static region (SR) and a reconfigurable region (RR) which depends on parameters sel_f_1 and sel_f_2. In this work, we assume that the parameter α is known and fixed. The static region of the APT-VDF filter contains the prototype filter structure and the all-pass filter given the fact that the parameter α is known. Both filters coefficients are given as constants in the hardware architecture. The reconfigurable region of the APT-VDF filter contains the minimum

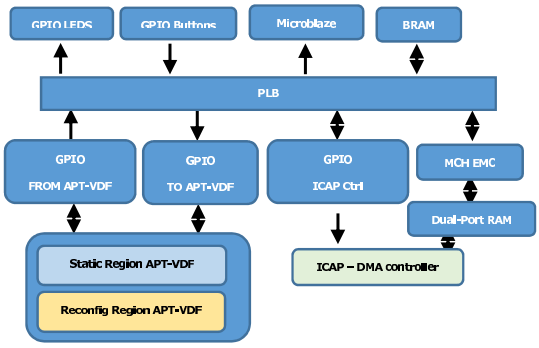


Fig. 2. Hardware architecture of the designed system

logic that is necessary to each combination of sel_f_1 and sel_f_2 parameters values. The VHDL code of the APT-VDF filter is written in such a way that it is reusable. By changing the filter's order N, word length w_l, and fractional length f_l, an APT-VDF filter of any order and word length can be generated. The ModelSim tool is used to simulate and validate the designed APT-VDF filter.

IV. Hardware Architecture

The hardware architecture of the implemented APT-VDF filter is shown in Fig. 2. This architecture was built on Xilinx ML550 development board and designed using Xilinx ISE 14.7, XPS 14.7, and PlanAhead 14.7 environments. The architecture contains the following components: Microblaze processor, processor local bus (PLB), memory BRAM, local memory bus (LMB), universal asynchronous receiver transmitter (UART RS232), multi-channel external memory controller (MCH EMC), and five general purpose input/output (GPIOs), true dual-port RAM, internal configuration access port controller, the SR and RR regions of the APT-VDF filter. The static region communicates with Microblaze via GPIO. The UART is integrated to allow the communication between the Microblaze and the RS232 interface of the board. All the peripherals communicate with the Microblaze via the PLB bus. The memory space of the BRAM is configured to be 16 KBytes. The on-chip dual-port RAM is configured to 128 Kbytes and it is used to store the partial bitstream files. The host PC is responsible for transferring those bitstream files to the Microblaze through the UART. The ICAP32 DMA IP is used to take advantage of the 32-bit mode of virtex5. The goal is to get closer to the theoretical throughput of the ICAP capacity, 400MByte/s (ICAP running at a maximum frequency of 100MHz). The Microblaze is only used for the setup phase, after that, the ICAP32 DMA IP controller operates autonomously, and it acts as direct memory access for partial bitstream files transfer between the dual-port RAM and the ICAP component. The ICAP32 DMA IP is connected to the MicroBlaze through a general purpose input/output interface. Two specific registers are used to store the start addresses and the offset of the physical memory where partial bitstream files are stored.

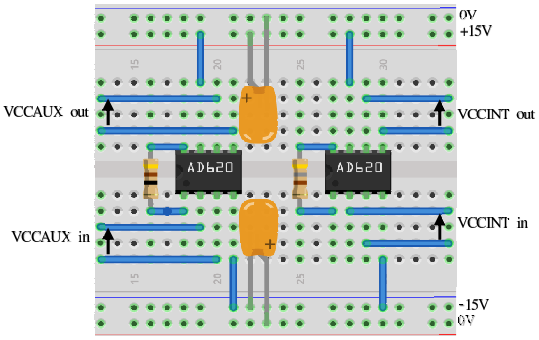


Fig. 3. AD620 amplification circuit.

V. Experimental Setup

The ML550 evaluation boards have been used in this study. It is well suited to power consumption measurement for several reasons. This board provides us with 5 power rails (core, IOs, peripherals) and current sense resistors which would simplify the experimental measurement. The recommended Virtex5 core voltage, designated VCCINT, is 1.0 ± 10%V. Depending on the input/output standard being implemented, the Virtex5 I/O voltage supply, designated VCCO, can vary from 1.2 V to 3.3 V. Moreover, Xilinx defines an auxiliary voltage, VCCAUX, which is recommended to operate at 2.5 ± 10%V to supply FPGA clock resources. The board hosts a two header connector which provides test points for the ML550 power regulators. Moreover, to measure different currents drained by the FPGA, the ML550 board contains a series of shunt 10 $m\Omega$ ± 1% 3W Kelvin current sense resistors on each voltage regulator lines. Thus, the current will be the voltage across the shunt resistor divided by the resistance value itself. Since the sensitivity of the VCCINT, VCCO, and VCCAUX sensors are pretty low (0.5-2mV), voltage amplifiers are needed. To this end, we have used integrated instrumentation amplifiers AD620, from Analog Devices. The AD620 has a feature to increase gain between 1-10.000 times with an external resistor. Those amplifiers allow to regulate the gain only by changing a single resistance, called RG. We wanted to reach a voltage of about 100-400 mV, at the output of the voltage amplifier then, according to the component's availability of the laboratory, was established gains of about 100 and 1000 to amplify the VCCINT and the VCCO voltages using resistances of 483 Ω, and 49 Ω. Fig. 3 shows the developed circuit board to amplify VCCINT and VCCO voltages. The input voltages are connected to a 2 x 13 0.1-inch male header connector which provides test points for the ML550 power regulators. The Tektronix TDS2024C oscilloscope is being used to display the voltages output from the two AD620 amplifiers.

VI. Experimental results

In this section, the implementation results and power consumption measurements of the APT-VDF filter are given with the help of a suitable design example. Herein, all mentioned frequencies are normalized with respect to half of the sampling frequency. Consider a prototype lowpass filter where the cut-off frequency, the transition bandwidth, the passband, and the stopband ripple specifications are 0.08, 0.14, 0.8 dB, and -40 dB, respectively. Using the filter design and analysis tool and direct-form FIR structure, the order of generated prototype filter is equal to $N = 21$.

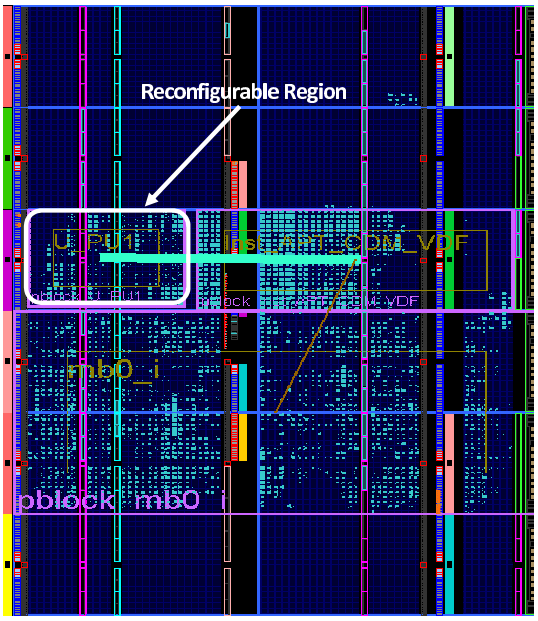


Fig. 4. Placement of top-level modules.

TABLE I
FPGA physical resources of the reconfigurable region

Resources	Highpass	Bandpass
LUT/FD	315/0	336/336
SliceL/SliceM	67/23	77/26
Frames	16	16
Frame Region	2	2
Bitstream	94464 Bytes	94464 Bytes

A. DPR area

The partial reconfiguration module has been implemented using the Xilinx PlanAhead tool. The Design view shows the placement of each instance and the post-place information, as shown in fig. 4.

Table I summarizes the key (post-place-and-route) implementation results of the reconfigurable region when the couple of parameters (sel_f_1, sel_f_2) is equal to $(0, 1)$ and $(1, 1)$, respectively. The results are coherent with our expectation given the fact that when $sel\_f_1 = 0$ and $sel\_f_2 = 1$, the generated APT-VDF filter is equivalent to highpass filter. However, when $sel\_f_1 = 1$ and $sel\_f_2 = 1$, the generated APT-VDF filter is equivalent to a bandstop filter, which requires to insert Nw_l delay units.

B. DPR timing overhead

The FPGA devices are composed of columns and rows. The biggest advantage brought by the recent Virtex family is that the reconfiguration can be realized by frames, contrary to Virtex II which imposed to reconfigure the whole columns. This improvement gives more flexibility on the modularity of the design flow during the placement phase of the dynamic and static part on the floorplan of the device. Frame by frame reconfiguration can significantly reduce the size of the generated partial reconfiguration files, and hence the configuration time overhead.

Table II presents the experimental FPGA reconfiguration time needed to perform APT-VDF filter reconfiguration. Three reconfiguration approaches are used:

- Full FPGA reconfiguration: we use JTAG connector to configure the FPGA with a total pre-made bitstream.
- External partial FPGA reconfiguration: we use JTAG connector to configure just the reconfigurable region of the APT-VDF filter with a pre-made partial bitstream
- Internal partial FPGA reconfiguration: we use the ICAP32 DMA IP to configure just the reconfigurable region of the APT-VDF filter with a pre-made partial bitstream.

TABLE II
EXPERIMENTAL RESULTS ON RECONFIGURATION TIME OVERHEAD

Reconf. type	Bitstream	Reconf. time
Full (JTAG)	1716 KB	3.80 s
Partial (JTAG)	94.464KB	208.9 ms
Partial (ICAP32 DMA)	94.464KB	324 μs

MicroBlaze is only used to configure specific registers which give the beginning address and the offset in RAM where partial bitstreams are stored. As a consequence, we achieve the maximum theoretical input throughput of the ICAP32 DMA which equal to 400 MB/s. We send a word each clock cycle, but due to registers loading and experimental uncertainty, we have an constant overhead of around 80 μs.

C. DPR power consumption overhead

The overall resource required to implement the designed system is given in table III. Moreover, the static and dynamic power consumption of the design mapped on FPGA can be estimated using the Xilinx XPower analyser tool.

Fig. 5 and Fig. 6 show the power consumption of the FPGA core during total reconfiguration and partial reconfiguration processes, respectively. The power consumption overhead is given by VCCINT voltage value. During total reconfiguration, the power consumption of the FPGA core is stable around 220 mW, otherwise it is around 360 mW. During the dynamic partial reconfiguration phase of the APT-VDF filter, the FPGA power consumption overhead overhead doesn't exceed 160 mW (500 mW - 340 mW) during 324 μs.

TABLE III
TOTAL FPGA PHYSICAL RESOURCES USAGE AND POWER CONSUMPTION ESTIMATION UNDER TWO CONFIGURATIONS

Resources		
(sel_f_1, sel_f_2)	(0,1)	(1,1)
Slices	2364/7200	2427/7200
Registers	3843/28800	4179/28800
LUTS	4850/28800	4998/28800
St. power	450 mW	450 mW
Dy. power	256 mW	270 mW
Total power	707 mW	720 mW

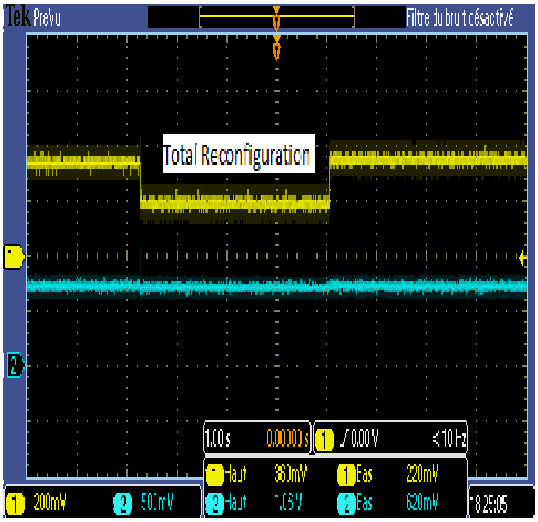


Fig. 5. FPGA core power consumption during total reconfiguration

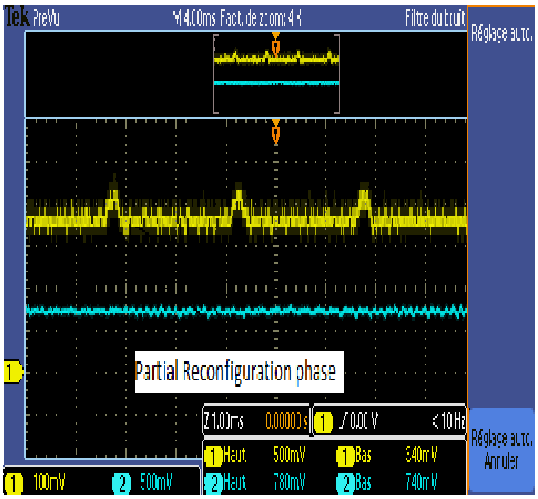


Fig. 6. FPGA core power consumption during partial reconfiguration

VII. CONCLUSIONS

Based on experiment measurement, this paper presents the FPGA core power consumption overhead during the dynamic partial reconfiguration. Three main conclusions can be drawn from our analysis. First, the FPGA dynamic partial recon-

figuration is an efficient design technique which permits to change a part of the devices while the rest of an FPGA is still running. Moreover, this technique can reduce dynamic and static power consumption in comparison with the well-known parametrization design approach. Second, the reconfiguration time is very small (around 324 μs for a partial bitstream size equal to 95Kbytes). Third, as we know, the DPR will introduce power consumption overhead. This work shows that this overhead is small.

VIII. Acknowledgment

The present work was carried out within the framework of Celtic-Plus SHARING project (number C2012/1-8).

References

[1] W. H. W. Tuttlebee, "Software-defined radio: facets of a developing technology," *in IEEE Personal Communications*, vol. 6, no. 2, pp. 38-44, Apr 1999.

[2] R. G. Machado and A. M. Wyglinski, "Software-Defined Radio: Bridging the AnalogDigital Divide," *in Proceedings of the IEEE*, vol. 103, no. 3, pp. 409-423, March 2015.

[3] J. P. Delahaye, J. Palicot, C. Moy and P. Leray, "Partial Reconfiguration of FPGAs for Dynamical Reconfiguration of a Software Radio Platform," *16th IST Mobile and Wireless Communications Summit, 2007 IEEE*, Budapest, hungary, pp. 1-5.

[4] J. Delorme, A. Nafkha, P. Leray and C. Moy, "New OPBHWICAP Interface for Real-time Partial Reconfiguration of FPGA," *International Conference on Reconfigurable Computing and FPGAs, 2009 IEEE*, Quintana Roo, 2009, pp. 386-391.

[5] E. J. McDonald,"Runtime FPGA Partial Reconfiguration," *Aerospace Conference, 2008 IEEE*, Big Sky, MT, 2008, pp. 1-7.

[6] Xilinx Inc. "Partial Reconfiguration User Guide (UG702)". 2013.

[7] Xilinx Inc. "Virtex-5 FPGA Configuration User Guide (UG191)". 2012.

[8] J. Becker, M. Hubner, G. Hettich, R. Constapel, J. Eisenmann and J. Luka, "Dynamic and Partial FPGA Exploitation," *in Proceedings of the IEEE*, vol. 95, no. 2, pp. 438-452, Feb. 2007.

[9] J. C. Hoffman and M. S. Pattichis, "A High-Speed Dynamic Partial Reconfiguration Controller Using Direct Memory Access Through a Multiport Memory Controller and Overclocking with Active Feedback," *International Journal of Reconfigurable Computing*, vol. 2011, Article ID 439072, 10 pages, 2011.

[10] R. Bonamy, D. Chillet, S. Bilavarn and O. Sentieys, "Power consumption model for partial and dynamic reconfiguration," *International Conference on Reconfigurable Computing and FPGAs, IEEE 2012*, Cancun, 2012, pp. 1-8.

[11] Information available at: http://www.ti.com/tool/fusion_digital_power_designer

[12] Information available at: http://www.ti.com/tool/usb-to-gpio

6.2.3 [C5]: Static Power Reduction

This contribution deals with electronics power consumption in CMOS technology, particularly for FPGA integrated circuits. Traditionally, the power consumption of CMOS circuit was largely due to dynamic power consumption. However, as technology advances to shrink the feature size and increase the transistor density, the static (*i.e* leakage) power consumption, which caused by leakage currents in the absence of any switching activity, becomes more critical in the overall power consumption of the FPGA circuits.

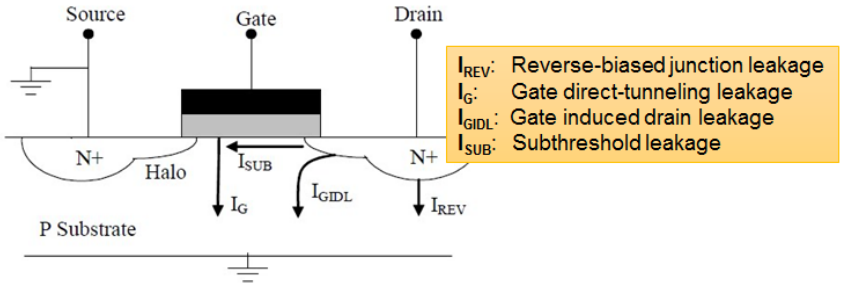


Figure 6.8 – Main CMOS leakage currents.

As threshold voltage, channel length and gate oxide thickness are scaled down (*i.e.* 65 nm, 40 nm, 32 nm and 28 nm), the inherent CMOS leakage currents, as shown in Fig. 6.8, have become the main source of power consumption in CMOS circuits [22, 23, 24, 25, 26, 27]. The static power dissipation in CMOS can be approximated by the following formula [28, 29]:

$$P_{static} = V_{dd} I_{sub} \tag{6.1}$$

where V_{dd} is the supply voltage and I_{sub} is the subthreshold leakage current which occurs when a CMOS gate is not turned completely off and its value is given by [30, 31]:

$$I_{sub} = \frac{\alpha \mu_0 C_{ox} W}{L} \times V_{th}^2 \times e^{\frac{(V_{gs}-V_t)}{nV_{th}}} \times \left(1 - e^{\frac{-V_{ds}}{V_t}}\right) \tag{6.2}$$

where α is a constant, μ_0 is the zero bias carrier mobility, C_{ox} denotes the gate-oxide capacitance, W and L are respectively the width and the channel length of the transistor, n is the process parameter, V_{gs} represents the gate to the source voltage, V_t is the threshold voltage, V_{ds} denotes the drain to source voltage and V_{th} is the thermal voltage given by kT/q ($\sim$ 26 mV at 273 °K). Equation (6.2) tells us that subthreshold current increases exponentially with the difference $V_{gs} - V_t$ and the transistor temperature (T). The reader can refer to [32, Fig.2], as this figure represents temperature versus leakage current characteristics. As a result, the static power also increases exponentially with the increase in die temperature and an accurate and efficient measurement of the thermal properties of an FPGA die is needed to avoid violating the

thermal constraints and ensure reliability during operation.

Given the relationship between static power consumption and temperature, it's no surprise that, an accurate thermal map of the die is highly required in order to manage and reduce static power consumption in a circuit. To this end, researchers have paid a particular attention to soft on-chip thermal sensor network based on ring oscillator (RO) circuit. ROs sensors are mainly based on the relationship between the delay time and the temperature of the die as depicted in Fig. 6.9. Thus, if we can obtain the frequency of the ring oscillator then we are able to measure the temperature accordingly. In the FPGA fabric, these smart temperature sensors are reconfigurable and can be modified, inserted or removed at run-time using dynamic partial reconfiguration technique.

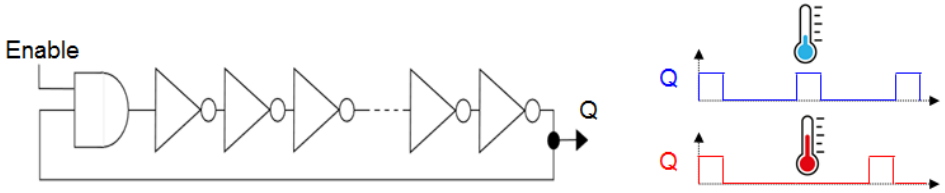


Figure 6.9 – Ring oscillator with an odd number of inverters.

To reduce static power consumption via an efficient management of circuit's temperature, in 2012, we have investigated the possibility to migrate self-heating modules, which cause hot spots, to the lowest temperature reconfigurable regions in the FPGA fabric using dynamic partial reconfiguration technique as shown in Fig. 6.10. The main conclusion of this contribution is that the maximum temperature difference between warmer and cooler areas inside FPGA does not exceed 16 °C. Based on the above result, the temperature difference is not sufficient to save static power since dynamic partial configuration process also wastes power. For more details, the reader is referred to our work published in book chapter [BC4] and in IEEE international symposium on wireless communication systems 2012 [CI27] (included hereafter).

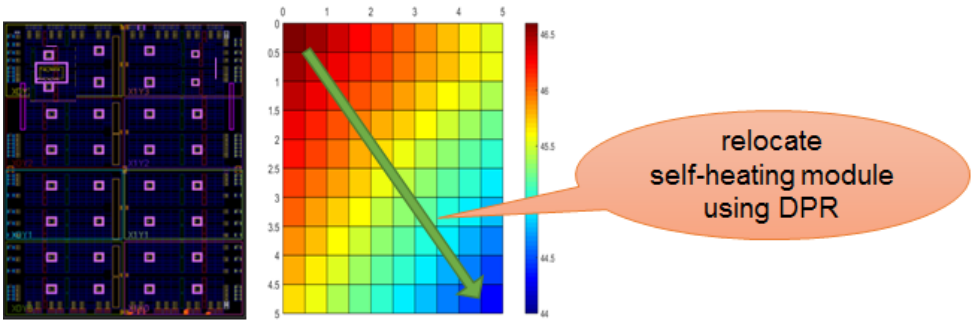


Figure 6.10 – Thermal map of Virtex-5 (65 nm) obtained using an uniform interpolation of RO sensors responses.

Leakage Power Consumption In FPGAs: Thermal Analysis

(Invited Paper)

Amor Nafkha*, Jacques Palicot*, Pierre Leray* and Yves Louet*
* Signal, Communication & Embedded Electronics / IETR
SUPELEC, Avenue de la Boulaie, CS 47601
35576 Cesson Sévigné Cedex, France
Email: {amor.nafkha,jacques.palicot,pierre.leray,yves.louet}@supelec.fr

***Abstract*—Current power saving techniques have been focused on reducing the dynamic power consumption induced by switching activity in CMOS digital circuits. Among these techniques, we can cite the clock gating, dynamic voltage frequency scaling, adaptive voltage scaling, and multiple voltage thresholds. Recently, as transistor sizes get smaller, the leakage power consumption has become a non-negligible and dominating part of the total power consumption. Thus, it is essential to take care of this new constraint in order to design low-power embedded wireless communication systems. It's well known that there are a strong relationship between leakage power and die-temperature, so that the larger leakage power consumption is associated with the higher temperature. In this paper we give a detailed analysis of FPGA leakage power consumption based on die temperature measurement. Also, after discretizing the FPGA area into several rectangular regions, we investigate whether dynamic partial reconfiguration technique can be performed over different regions to decrease average die-temperature, and consequently the leakage power consumption.**

I. Introduction

Saving energy to mitigate power consumption and allow green communication has become one of the most challenging research fields especially in advanced CMOS technology nodes, below 65nm. As a result, many power management techniques have been developed throughout the design flow (algorithmic level, architecture, and circuit layout) in order to provide both low power consumption and chips reliability. Traditionally, the most important integrated circuits power consumption was due to dynamic power consumption which arises due to switching activity of the circuit and supply voltages. Different design strategies has been developed and proven to reduce this latter power such as clock gating, dynamic frequency scaling, Multi-Voltage, dynamic voltage scaling, multi-threshold designs. Recently, as technology moved down to 65nm and 28nm, we reaching the point where leakage power is the same as dynamic power. The leakage (static) power is due essentially to the leakage current increase and supply voltages. A lot of research papers are dealing with leakage power consumption itself within the framework of nanoscale technology [1]. If, in the past, FPGAs were mainly concerned with dynamic power consumption [2], it remains no more true with new technologies [3] and with the large number of transistors inside FPGA circuits. The leakage power consumption becomes more and more important both in the active and non active part of the component.

Real measurements in [4] show that dynamic power consumption may be decreased with the dynamic and partial reconfiguration (DPR), and this saved power consumption depends on the application and the designs architecture. Dynamic activation and deactivation of the FPGAs fabric clock network can be implemented to decrease dynamic power consumption. However, neglecting the reconfiguration process power consumption makes only sense if the reconfiguration time is very low compared to the application execution time's constraints. The study in [5] indicates that the polarity of the inputs in FPGA hardware structures may significantly impact leakage power consumption (average reduction of roughly 25%). In [6], it has been shown that the total power consumption can be decreased by exploiting the dynamic partial reconfiguration capability of FPGAs using classical reduction techniques such as clock gating, hardware deactivation and removal. To the best of our knowledge, the benefit of dynamic partial reconfiguration to reduce the die temperature and then the static power consumption has not been yet evaluated.

The paper is organized as follows. The next section gives a brief overview of power consumption origins and traditional power reduction techniques. For thermal computations, the FPGA is discretized into an N regions, where each of them contains a digital thermal sensor described in Section III. The relationship between chip temperature and the leakage power consumption will be given. Section IV gives both the hardware[1] and software's set-up. In this section, two different PEs used to increase the temperature are described. Experimental results concerning the temperature difference between hot and cool regions inside the FPGA are provided and analyzed in section V. Finally, some discussion and conclusions are given in the last section.

[1]We have performed experiments on the Virtex-5 FPGA to study the FPGA power consumptions behaviors. The ML550 development board is well suited to power consumption's measurement. This board provides us with five power rails (core, IOs, peripherals) and current sense resistors which could simplify the experimental measurement.

978-1-4673-0762-8/12/$31.00 ©2012 IEEE

II. BACKGROUND

In this section, we give some useful definitions related to power consumption as well as present classical power reduction techniques used in FPGA design.

A. Power consumption sources

The total power consumption in the FPGA can be defined as: $P_t = P_{le} + P_{dy}$, where P_{le} defines the leakage power that a circuit consumes even when it is in the standby mode and P_{dy} defines the dynamic power consumption that occurs when the output signal of a CMOS logic cell makes a transition. Each component of the power consumption is discussed in more detail in the following subsections.

1) Dynamic power: Dynamic power is the power consumed through device during switching events in the core or in the input/output pins. The main variables affecting this power are capacitance charging and discharging, toggling frequency, and supply voltage. Transistors are used for logic and programmable interconnects so in general a large portion of the dynamic power is due to the routing fabric. The formula for dynamic power is given by:

$$P_{dy} = nCV^2 f \tag{1}$$

where, n is the number of toggling nodes, C is the capacitance, V and f are the supply voltage and the toggle frequency respectively. The core FPGA supply voltage and capacitance are reduced with each new process node. A very large number of papers have considered the distribution of dynamic power consumption in FPGAs [2].

2) Leakage power: Leakage power is the power consumed by the device due to leakage currents when there is no activity or switching in the design. Using finer semiconductor technologies has caused an increase of this total power part. In fact, as transistor size decreases and lower voltages are utilized, a greater sub-threshold leakage current occurs in the transistor channel when the transistor is in the off state. The sub-threshold leakage current depend on temperature and the threshold voltage. So, the leakage power can be expressed as

$$P_{le} = I_{le} V \tag{2}$$

where I_{le} is the leakage current, which consists of both the sub-threshold leakage current and the reverse bias junction current in the CMOS circuit. In the current generation of 65nm, 40nm, and 28nm FPGAs, static power is becoming a more dominant power component.

B. Power reduction techniques

In the following, a number of techniques for reducing power consumption are outlined. These include use of clock gating, dynamic voltage and frequency scaling and multi voltage threshold.

1) Clock gating: Clock gating allows the ability to turn off different parts of the FPGA when they are not being used and is normally done by disconnecting the clock using the gating control logic [7]. For example, a set of registers could have their clock source turned off so that the gates do not switch, there is no charging or discharging and thus no wasted power. The used AND gate will dissipate some power, but this is much less than the register or block of registers. This technique is used to minimize dynamic power.

2) Dynamic Voltage and Frequency Scaling: Power may also be reduced using DVFS which adjust the supply voltage and clock rate dynamically according to circuit parameters. The energy efficiency of this technique is highly dependent on the slack of the circuit. However, according to [8], larger voltage ranges does not improve power efficiency. They showed that reducing both the dynamic and leakage power consumption simultaneously is critical for an overall energy-efficient design.

3) Multi Voltage Threshold: Multiple threshold voltage techniques use both low V_t and high V_t cells. Lower threshold gates are used on critical path and high threshold gates everywhere else in the design. With this technique, the energy consumption is decreased without affecting the overall frequency [8]. This scheme is used to limit the leakage current during sleep mode.

III. FPGA'S TEMPERATURE SENSOR

On-chip measurements of local temperature present an opportunity to incorporate temperature management techniques and performance optimization into the FPGA fabric. This section presents the design of a sensors system that continually measures the temperature values at various locations on the FPGA. The mean idea behind this work is to perform a smart power management when an high local die temperature is detected. In the following, we suppose that the FPGA is discretized into several rectangular regions, each of them incorporated a simple and accurate digital thermal sensor, capable of detecting a wide range of temperature.

A. Ring-oscillator based thermal sensor

An originally work in [9] proposes a simple way to measure chip heating using a ring oscillator associated to a frequency counter [10]. A ring oscillator consists of a feedback loop that includes an odd number of inverters needed to produce the phase shifting that maintains the oscillation. In [11], Lopez-Buedo et al. proposed the minimization of the operation frequency of ring oscillator in order to minimize the problems related to self-heating, power consumption, and counter size. Due to the lack of knowledge of hot spot locations inside the die, all thermal sensors nodes are spread into different regions across the FPGA according to the work published in [13].

Exhibiting a linear dependence of oscillation frequency on chip temperature, at a fixed value of V_{ccint} supply voltage, the ring oscillator circuit can be embedded inside any FPGA fabric and in conjunction with a counter it paves way for efficient and highly accurate temperature measurement. The

ring oscillator is characterised by its reduced area and energy overhead which allowing it to be replicated as many times as needed to create a FPGA thermal map. The thermal sensor can be divided into three major parts. The first entity is composed of a ring oscillator controlled by an external enable signal, the second entity is 12-bit cyclic counter and the last entity is an 14-bit reference time counter which indicate the number of rising edge reference 50Mhz clock between two events. During the process, the first 12-bit counter is clocked by the output of the ring oscillator, and generates a boolean event to the second 14-bit reference time counter. This boolean is equal to true when the 12-bit counter value is equal to 2^{12}, otherwise the enable signals value equals false. The reference 14-bit counter determines the number of rising edge reference clock that have been counted between two true events. The reference counter value is sent back to the host via the RS232 port. The FPGA temperature is gathered from an on-chip hardware sensor called System Monitor which is located at the center of the FPGAs core and developed around a 10 bits analog-to-digital converter. The ring oscillators inverters are implemented in look up table, and their relative location constraints were used to equalize the physical distances between inverters.

B. ring oscillator's frequency vs temperature

To have a good estimation of the ring oscillator frequency as function of an FPGA temperature, we have placed four sensors inside FPGA at the following locations: North East (NE), South East (SE), South West (SW) and North West (NW) (as indicated in figure 1).

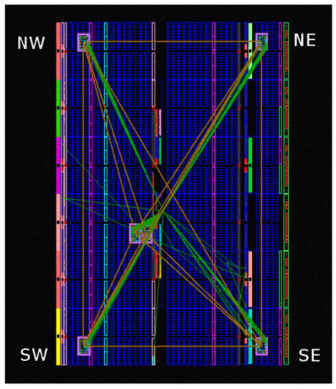


Fig. 1. Thermal sensors' position inside the FPGA.

Figure 2 shows the frequencies of different ring oscillators as a function of the FPGA temperature given by the System Monitor. Notice that the experimental results are given when V_{ccint} supply voltage is equal to $0.998V$. The variation of frequency between different sensor's locations is due to the propagation delay variation between ring oscillator's inverters, which is related place-and-route phase. Experimental results show that there is indeed a high correlation and proportionality between propagation delay inside ring oscillator and the FPGA temperature. Linear approximation seems to be a near fit of the recorded curve. Below is the equation that expresses average ring oscillator frequency value to FPGA temperature conversion using the classical linear first-order fitting equation.

$$freq = 0.0094T + K \tag{3}$$

Where $freq$ is the ring oscillator frequency in Megahertz, T is the temperature in degree Celsius($°C$), and K is a calibration constant which can be easily calculated for a given initial temperature and ring oscillator frequency. The frequency value varies linearly with increase in temperature. The slope of the line, equal to $9.4Khz/°C$, gives us the effective change in ring oscillator frequency corresponding an increase or decrease of temperature.

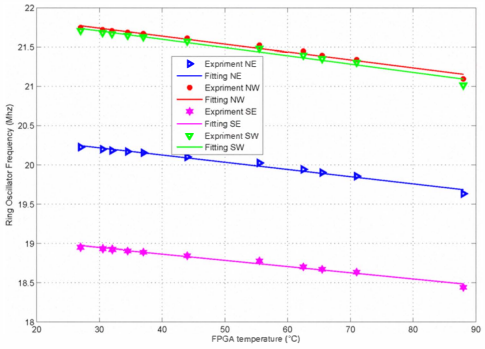


Fig. 2. Measured ring oscillators frequency in MHz versus FPGA's temperature in degree Celsius.

IV. Processing Elements Design

In order to increase temperature locally inside FPGA, we described in this section the design of two processing elements used for different experiments.

A. PE number 1: Random Number Generator

We have designed a parametric processing unit based on the random number generator. A 512-bit random number is being generated with the help of the exclusive-or (XOR) gate. The most significant digit and the least significant digit are both processed through a XOR gate and give the first bit of the random number and then this equation runs inside the loop 2^9 times so as to give a 2^9 bits random number. So the generated number contains 2^9 random bits per clock cycle. To significantly increase the temperature, the used clock's frequency can be doubled by using the digital content managers components inside the FPGA.

B. PE number 2: Microblaze processor

To increase the temperature of a predefined region inside FPGA, a microblaze embedded soft-core processor has been implemented. The advantage of using this type of soft processors is that they are highly reconfigurable and can be customized according to our needs. Here, the microblaze runs at 125MHz, the floating point unit being enabled, and we chose a 64×64 standard matrix multiplication as an application target.

V. Experimental results

As said previously all results are based on the measurements made at a nominal V_{ccint} level of $0.998V$. The clock frequency was kept fixed at 100 MHz. The relationship between temperature, leakage power consumption and the on-chip temperature has been explored and the results are presented in the following subsections.

A. Temperature vs leakage current consumption

In order to measure leakage current for different die-temperatures, experimental and analytical results are carried out using Agilent multimeter and Xilinx XPower analyzer tool (XPA), respectively. The XPA is part of the Xilinx ISE design suite and provides an accurate way to analyze the power profile of post place&route designs. Figure 3 shows the measured leakage current for the various die-temperature. At that level we are considering current leakage rather than power leakage in order to be in accordance with the Xilinx general curve from. Data points were obtained via the running of each experiment until temperature stabilization. This takes several minutes.

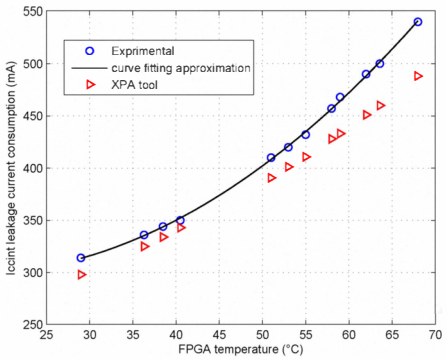


Fig. 3. Leakage currents consumption versus FPGAs fabric temperature.

Measuring voltages across different sense resistors enables the computation of the FPGA I_{ccint} leakage current consumption at any time. In Figure 3, the continuous line presents the best-fitting quadratic function. As regard the temperature T parameter, and the data of the present experiment. the leakage current consumption can be defined as:

$$I_{ccint}[mA] = 0.0899T^2 - 2.889T + 321.724, \qquad (4)$$

where T is the junction temperature in degree Celsius (°C). An important conclusion that can be drawn here is that for low-temperature range (30-40°C), ten-degree temperature variation can cause a small increase of the leakage current (roughly 33 mA). By contrast, the same temperature range (60-70°C) in a hot FPGA chip significantly affects the leakage current (roughly 90 mA).

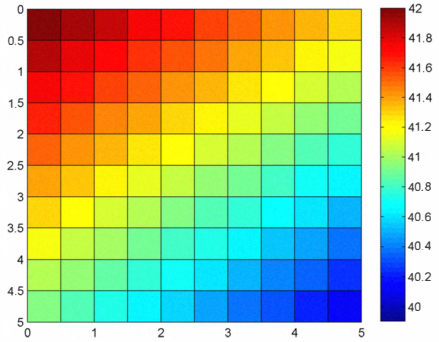


Fig. 4. Thermograph obtained by an uniform interpolation of sensors responses: Processing element number 1 at 100 Mhz

B. Temperature map analysis

Figure 4 and figure 5 show experimental temperature maps of the FPGA in case of use of the processing element number one and two respectively. Data intensive processing element PE1 has been inserted into the NW location. The core idea here is that we are able to increase locally the temperature by running the processing elements. A linear temperature gradient can be assumed inside the FPGA. Consequently, we are able to draw the thermograph. Furthermore, the temperature values are obtained after a long time, this means we have not considered in this study the temperature peaks, which have very short duration of less than 1 ms. The steady states temperature prior to the running of the PE number 1 is equal to 32°C. The maximum temperature's increase is 5°C. The difference between the hottest and coolest regions after several seconds is roughly 2°C. Figure 5 presents the same type of results obtained with PE number 2. The steady states temperature prior to the running of the PE is equal to 40°C. The maximum temperatures increase is 6.5°C. The difference between the hottest and coolest regions after several seconds is roughly 2.5°C.

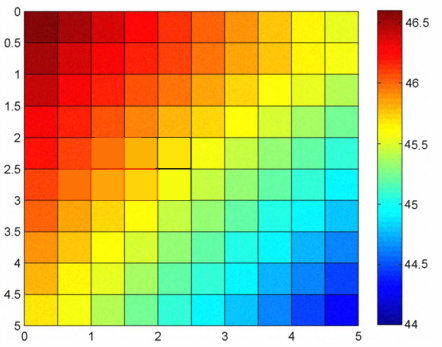


Fig. 5. Thermograph obtained by an uniform interpolation of sensors responses: Processing element number 2 at 125 Mhz

To get a better information of the impact of the filling rate of the FPGA on the temperatures difference between the hot

and cooler FPGA regions, figure 6 illustrates the temperature's difference between hottest and coldest region as a function of the FPGA's occupancy. The occupancy is performed with several PEs number one running at 100 MHz. As expected the difference increases vs the occupancy rate to a maximum when the rate is equal to 50 percent. Then, as this can be easily understood, the difference decreases, since the temperature of the coolest region in this zone increases, when PEs are implemented in the free zone.

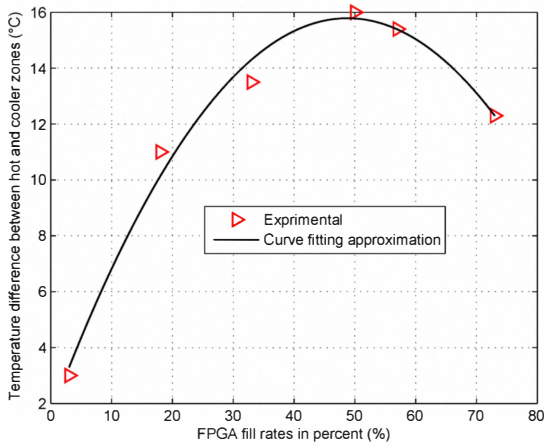


Fig. 6. Temperatures difference between hottest and coolest spots vs FPGA occupancy.

VI. Discussion

As stated above, the static power is mostly due to sub-threshold leakage current. This current, although is small for low FPGA's temperature, increases exponentially with temperature. In order to reduce power consumption by means of reducing hotspot's temperature, we have studied in depth the temperature's difference between hot and cooler spots inside the FPGA. On the basis of the results given in the above section and the original work published by S.Liu et al in [12] a detailed analysis has brought us to the following interpretation:

- The temperature's differences between the hotter and colder FPGA regions do not exceed 16°C, so moving hotspots to colder regions can not save more than 100 mW when the die's temperature range is (35-50°C). This optimistic gain is made under the assumption than moving a PE will result in a decreasing temperature on the spot, which is not so obvious as we experimented it.
- On the basis of the original work of S.Liu et al in [12] dealing with partial reconfiguration to reduce Virtex4 power consumption, we are able to have a good estimation of the chip's power consumption during the partial reconfiguration's processing time ($\approx$ 600 mW). This power consumption overhead is larger than the saved one given in the above item. Therefore, with current available technologies (mainly DPR), there is no gain in static power consumption by moving PEs from a hot to a cool spot.

VII. Conclusion

The power reduction is an important design issue especially in energy constrained applications and green cognitive radio. In this paper, we have shown that the temperatures obtained from the digital sensors correlate well with temperatures given by System Monitor. This study provide answers to the questions raised in the introduction: Using run-time partial reconfiguration, is't possible to reduce the total power consumption by moving hotspots from hotter FPGA region to another colder region? The experimental results demonstrate that due to thermal propagation speed on a FPGA, using partial reconfiguration can not lead to total power consumption reduction. FPGA architectures offer the possibility to perform dynamic and partial hardware reconfiguration. By exploiting this feature, smaller chip sizes can be used which also leads to decreased static power consumption. In future work, we plan to investigate the power saving in the case of a multi-FPGA architecture. The main idea behind this future study is that instead of using one high density FPGA to implement cognitive radio, it will be more interesting to use multiple low density FPGAs. Exploiting reconfiguration can allow using a smaller chip size and this way achieve lower static power dissipation. In this case, we are able to maintain a high temperature difference between adjacent FPGAs.

References

[1] J. Kao, S. Narendra, A. Chandrakasan, A. Ch, and I. I. S, "Subthreshold leakage modeling and reduction techniques," in In Proc. ICCAD, 2002, pp. 141148.

[2] L. Shang, A. Kaviani, and K. Bathala, "Dynamic power consumption in the Virtex-II FPGA family," in International Symposium on Field-Programmable Gate Arrays, 2002, pp. 157164.

[3] I. Xilinx, Virtex-5 Family Overview: Product Specification

[4] J. Becker, M. Hubner, and M. Ullmann, "Real-time dynamically runtime reconfiguration for power/cost optimized Virtex FPGA realizations," in VLSI-SOC, 2003, pp. 283288.

[5] J. H. Anderson, F. N. Najm, and T. Tuan, "Active leakage power optimization for FPGAs," in Proc. ACM/SIGDA Int. Symp. Field Programmable Gate Arrays, Monterey, CA, 2004, p. 3341.

[6] M. G. Lorenz, L. Mengibar, M. Garca-Valderas, and L. Entrena, "Power consumption reduction through dynamic reconfiguration." in FPL, ser. Lecture Notes in Computer Science, J. Becker, M. Platzner, and S. Vernalde, Eds., vol. 3203. Springer, 2004, pp. 751760.

[7] S. Huda, M. Mallick, J. Anderson, "Clock gating architectures for FPGA power reduction," In IEEE International Conference on Field-Programmable Logic and Applications, Prague, Czech Republic 2009, P: 112118.

[8] P. Girard, N. Nicolici, X. Wen, "Power-Aware Testing and Test Strategies for Low Power Devices, " Electrical engineering, Springer 2009.

[9] N. P. G. Quenot and B. Zavidovique, " A temperature and voltage measurement cell for VLSI circuits," E. A. Conf., Ed. Piscataway, N.J.: IEEE Press, pp. 334338.

[10] R.O. E. Boemo, and S. Lopez-buedo, "Thermal monitoring on FPGAs using, in Proc. FPL 1997 Workshop, Lecture Notes in Computer Science 1304. Springer-Verlag, 1997, pp. 6978.

[11] S. Lopez-Buedo, J. Garrido, and E. I. Boemo, "Thermal testing on reconfigurable computers, " IEEE Design & Test of Computers, vol. 17, no. 1, pp. 8491, 2000.

[12] S. Liu, R. N. Pittman, and A. Forin, "Energy reduction with run-time partial reconfiguration," in Proceedings of the 18th annual ACM/SIGDA international symposium on Field programmable gate arrays, FPGA 10.

[13] S. Lopez-Buedo and E. Boemo, "Making visible the thermal behaviour of embedded microprocessors on FPGAs: a progress report," in Proceedings of the 2004 ACM/SIGDA 12th international symposium on Field programmable gate arrays, 2004, pp. 7986.

6.3 Experimental Evaluation of Spectrum Sensing Techniques

Software defined radio enables a flexible approach to support a wide range of wireless communication systems by updating the software and the reconfigurable logic without making any changes to the hardware platform. As a result, a software defined radio shifts the transition from the digital to the analog domain and vice verso close to the radio frequency front-end while keeping the main part of the communications processing operations in software. In general, the hardware aspects of a software defined radio platform consist of programmable radio frequency (RF) front-end elements, some baseband signal processing (*e.g.* sample rate conversion (SRC), digital up converter (DUC), digital down converter (DDC), *etc.*.) and communications link to the FPGA-based, DSP-based, GPP-based signal processing system [33, 34, 35, 36]. In the sequel of the present section, we mainly focus on the implementation of spectrum sensing techniques on a general purpose processor (GPP) based SDR platforms.

As depicted through Fig. 6.11, in GPP-based SDR platform, a GPP is used to handle a wide variety of generic tasks. It is designed to implement and perform complex arithmetic functions such as filtering, modulation/demodulation and encoding/decoding using fixed or floating point operations. Modern general purpose processors provide very high flexibility and easy development environments. Some commonly used GPP architectures are ARM, X86 and AMD64.

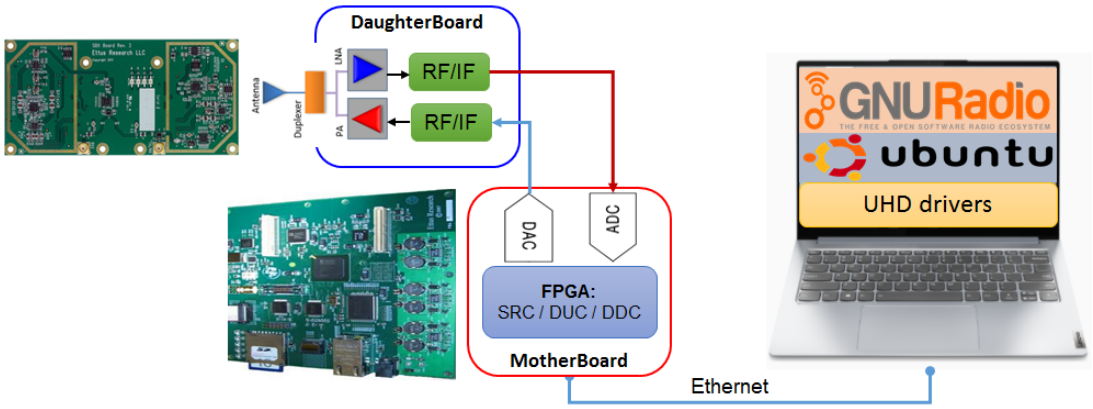


Figure 6.11 – GPP-based SDR platform (Ettus USRP N210 board & GNU Radio framework).

For the experimental contributions reported in this section, we adopted the universal software radio peripheral (USRP) hardware platforms, provided by Ettus Research, which operate by digitizing the baseband signal processing and performing the required digital signal processing on computers. Particularly, we investigate the performance of some spectrum sensing techniques presented in chapter 5 using the USRP N210 SDR kit which consists of a motherboard and a daughterboard. The motherboard is powered by a Xilinx FPGA chip which contains the USRP hardware driver (UHD), allowing control and synchronization between the GPP and the USRP board. From the FPGA, the signal at intermediate frequency is sent and received

to and from daughterboard through analog-to-digital converters (ADCs) and digital-to-analog converters (DACs). The daughterboard, inside which resides the radio-frequency (RF) front-end, is plugged onto a socket of the motherboard. Different types of daughterboards can be used in conjunction with the USRP N210 motherboard, in this work, the wideband WBX daughterboard with the transceiver of 50MHz - 2.2GHz frequency range is used.

The USRP N210 SDR platform is aimed to work with GNU Radio open source software environment [37, 38], however, it is also compatible with LabView software tool [39] and Matlab-Simulink software environment using communications system toolbox support package for USRP radio [40]. For experimental spectrum sensing measurements, we have opted to use GNU Radio which has a large community of professionals, experts, scientists and hobbyists that is constantly pushing the tool forward.

GNU Radio uses Python to connect different signal processing blocks which are written in C++ software development language, so adding custom block to GNU Radio is simple [41]. Based on signal-processing packages and GNU Radio Companion (GRC) interface, users can graphically place and connect blocks in order to construct their own radio. A top-down description of the combined GRC and USRP SDR hardware platform is provided in fig 6.12.

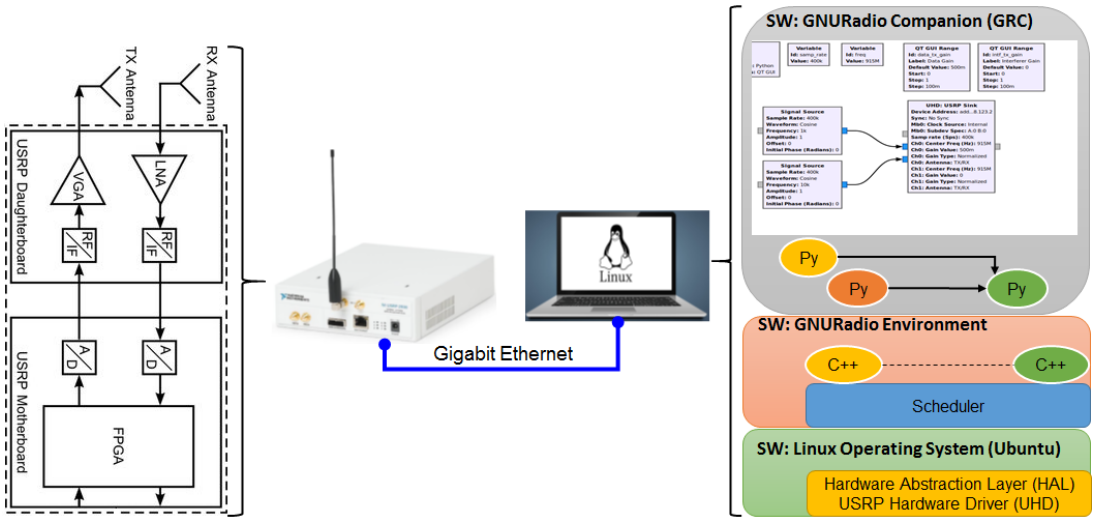


Figure 6.12 – USRP hardware platform and GNU Radio software architecture.

The following two contributions concern the experimental validation of symmetry-based cyclc autocorrelation function (SPCAF) detector and two eigenvalue-based spectrum sensing detectors: maximum-minimum eigenvalue (MME) detector and energy with minimum eigenvalue (EME) detector, readers can refer to Chapter 5 more algorithmic details. All the experiments are conducted in real world environment using Ettus USRP N210 hardware platform and GNU Radio companion tool.

6.3.1 [C6]: Experimental Evaluation of SPCAF Detector

In order to evaluate the real performance of the SPCAF detector, we need to design new out-of-tree custom block, which can be easily drag and dropped to GRC flowgraph and connected to other inbuilt predefined signal processing blocks. For computational efficiency, we implemented the SPCAF module in C++ and we used the simplified wrapper and interface generator (SWIG) to generate the necessary component to make it accessible from Python. The developed module has been tested using a USRP N210 as a primary user transmitter and a second USRP N210 as a secondary user equipment as shown in Fig 6.13.

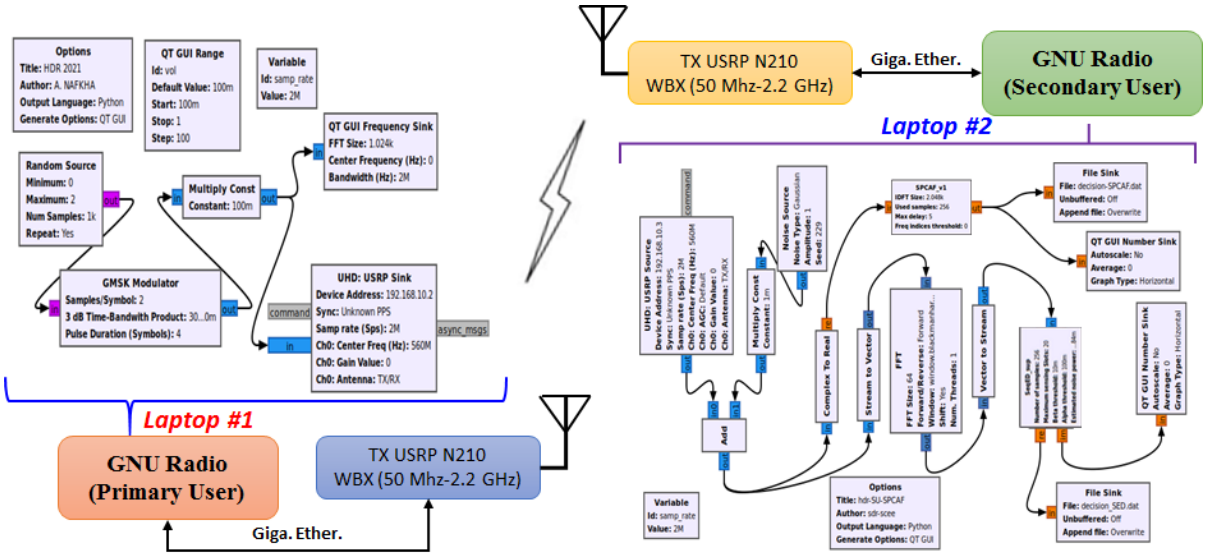


Figure 6.13 – A simplified block diagram of sensing performance measurements with GRC flowgraphs associated to PU and SU equipments.

Without loss of generality, the OFDM signals are used to illustrate the sensing performances of the studied SPCAF approach. Signal processing of transmitter (*i.e.* primary user, laptop #1) consists of random source data generator, OFDM modulator, amplitude adjusting, and USRP sink as shown in Fig 6.13. The receiver (*i.e* secondary user, laptop #2) does the inverse signal processing, including USRP source, artificial noise injection to mitigate DC offset and I/Q imbalance issues, converts the time-domain signal to the frequency-domain signal using FFT block, and SPCAF module as depicted in Fig 6.13. The output of FFT block is used to perform spectrum sensing using a simple sequential energy detector (SED) method published in [42]. For both USRP N210 devices, the carrier frequency has been set to 560 MHz, as this fragment of the TV band is free of DVB-T transmission in the location where the experiments have been conducted, and the complex sampling frequency has been defined to be equal to 2 Msps. Fig. 6.14 shows the laboratory configuration used for the conducted experiments. The host computers used in this real world experimentation run on Linux Ubuntu 16.04 LTS, with

GNU Radio software version 3.7.7 installed on them.

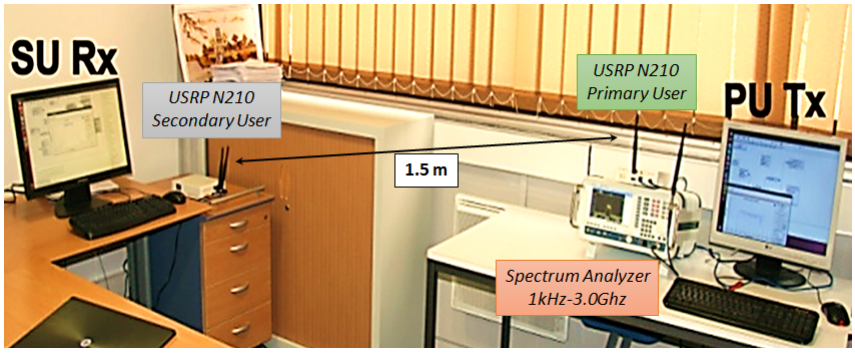


Figure 6.14 – Experimental setup for spectrum sensing using GNU Radio and USRP devices.

For all experiments, the decisions about the presence or absence of the PU signal have been made within GNU Radio blocks. However, the whole post-processing required for graphical visualization of the results has been done using Octave software. Comparison of the spectrum sensing performance between the SPCAF, SED and an hybrid architecture combining SED and SPCAF, under assumption that the noise power is unknown, is presented in Fig. 6.15.

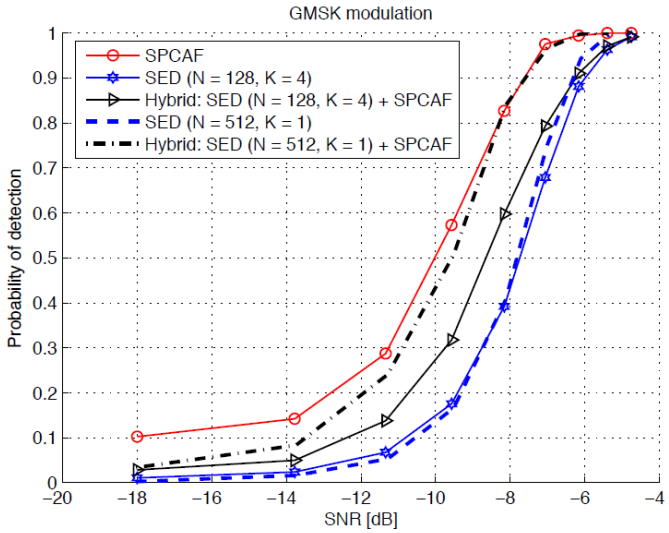


Figure 6.15 – Performance comparison between SPCAF, SED and hybrid architecture.

As shown in Fig. 6.15, the SPCAF detector demonstrates the better performance in comparison with other above mentioned detectors. For example, the SPCAF achieves 2.0 dB SNR gain in comparison with the SED at the detection probability equal to 0.7; the SNR gain is equal approximately to 1.2 dB in favor of the SPCAF comparing with hybrid architecture (SED followed by SPCAF). Interested readers are referred to our journal paper [J6] and conferences [CI33], [CI34] (**included hereafter**), [CI37] in which much more details are given.

2014 9th International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM)

Experimental Spectrum Sensing Measurements using USRP Software Radio Platform and GNU-Radio

(Invited Paper)

Amor Nafkha, Malek Naoues
SCEE/IETR, SUPELEC, Avenue de la Boulaie
CS 47601, 35576 Cesson Sévigné Cedex, France
Email: {amor.nafkha, malek.naoues}@supelec.fr

Krzysztof Cichon, Adrian Kliks
Poznan University of Technology
Polanka 3, 60-965 Poznan, Poland
Email: {krzysztof.z.cichon@doctorate, akliks@et}.put.poznan.pl

***Abstract*—In cognitive radio, the secondary users are able to sense the spectral environment and use this information to opportunistically access the licensed spectrum in the absence of the primary users. In this paper, we present an experimental study that evaluates the performance of two different spectrum sensing techniques to detect primary user signals in real environment. The considered spectrum sensing techniques are: sequential energy and cyclosationary feature based detectors. An Universal Software Radio Peripheral platform with GNU-Radio is employed for implementation purpose. We analyzed the performances of both spectrum sensing methods by measuring the detection probabilities as a function of SNR for a given false alarm probability. As predicted theoretically, experimental measurements show that the cyclostationnary feature detector performs better than the sequential energy detector. However sequential energy detector can be used for reduction of sensing time in the presence of strong signals.**

I. Introduction

Cognitive Radio (CR) is an emerging concept to increase spectrum usage efficiency by allowing a secondary user (SU) to access some licensed spectrum bands temporarily unoccupied by the primary user (PU). Two basic approaches to spectrum sharing have been considered: spectrum *overlay* and spectrum *underlay*. According to the spectrum overlay approach, the secondary users sense and identify unused frequency bands and use them for communication purposes. Thus, the secondary users (SU) are responsible for detecting the unused bands and they should vacate the spectrum as soon as the primary user begins its [1] activities. The underlay approach imposes constraints on the secondary users' transmission power level so that it operates below the noise floor of primary users. Here, we focus on the implementation aspects of the overlay spectrum sharing. To determine the absence or presence of the primary user signals, several spectrum sensing techniques have been developed [2], [3], [4]. These techniques can be classified into three categories: (**i**) methods requiring both primary user signal and noise variance information, (**ii**) methods requiring only noise variance information (also called *semi-blind* methods), and (**ii**) methods not requiring any information on primary user signal or noise variance (also called *blind* methods). Examples of blind spectrum sensing methods would be wavelet based detection [5], eigenvalue based detection [6], second order statistical based detection [7], and symmetry property of cyclic autocorrelation function based detection [8].
In the practical implementation, the simplest spectrum sensing method capable of detecting the presence of a PU's signal is based on energy detection which is a semi-blind method. Several papers address experimental results of the energy detector and outline the impact of noise uncertainty on the performance of detection [9], [10]. Due to this shortcoming, there is an signal to noise ratio wall, SNR_{wall}, in which energy detector can not guarantee a detection performance.
The main aim of the conducted experiment is to sense the spectrum in a given frequency range and make as reliable decision as possible on the potential presence of the primary user (PU) signal in the observed spectrum fragment. In order to achieve this goal selected algorithms for spectrum sensing have been implemented in hardware.
This paper is organized as follows. In section II, the system model is presented as well as sequential energy and the cyclostationary feature-based detectors. In section III, we describe our experimental setup for both detectors using two different modulations for the PU. Section IV presents the experimental evaluation results of the considered algorithms. Finally, conclusions and future works are presented in section V.

II. System Model and Sensing Techniques

Thus, spectrum sensing and detecting the presence of a radio in the environment can be treated as a classical detection problem [11] [12]. Two binary hypotheses H_0 and H_1 can be defined to indicate the absence or the presence of the PU in the environment. The received signal at the SU, $r(t)$, can be expressed as:

$$r(t) = \begin{cases} n(t) + i(t) = \hat{n}(t) & \longrightarrow H_0 \\ h(t) \cdot s(t) + n(t) + i(t) = s(t) + \hat{n}(t) & \longrightarrow H_1 \end{cases} \quad (1)$$

where $s(t)$ and $h(t)$ stand for the PU signal and channel impulse response, respectively, $n(t)$ is an additive white Gaussian noise (AWGN), and $i(t)$ represents other sources of distortions such as ambient noise or interferences; the equivalent noise observed at the antenna input can be then represented as $\hat{n}(t)$. The objective of the spectrum sensing operation is to decide between H_0 et H_1 based on the observation of the received signal $r(t)$; one can find in the literature the papers dealing with interference mitigation or reduction during the spectrum sensing process. The detection performance is characterized by two probabilities: probability of detection, P_{d}, where the decision is H_1, while H_1 is true; and probability of false alarm, P_{fa}, which corresponds to the case where the decision is H_1 while H_0 is true.

978-1-63190-003-7 © 2014 ICST
DOI 10.4108/icst.crowncom.2014.255415

In our experiments, two algorithms have been tested i.e. sequential version of energy detection, and a method for cyclostationarity-based detection called Symmetry Property of Cyclic Autocorrelation Function (SPCAF). Let us briefly summarize the theoretical basis of these solutions.

A. Traditional Energy-based spectrum sensing

One of the simplest way for primary user detection is to calculate the amount of received power in the considered frequency subband and compare this value with the noise variance. In the case that the received power is greater than the previously approximated power of noise the algorithm will make a decision on the spectrum occupancy by the primary user signal. In turn, the channel will be assumed to be vacant if the computed noise power will be close to the noise variance at the certain level of certainty. There are several parameters that influences the reliability of any spectrum sensing algorithm. In case of traditional energy detection, the crucial role is played by properly defined decision threshold, and in consequence, by the correctness of noise variance, and the duration of sensing time (expressed in seconds or - for discrete signals - in terms of number of gathered samples). Following [12], for the given values of probability of false alarm, P_{fa}, number of collected samples N, and the (equivalent-)noise variance $\hat{\sigma}_n^2$, the decision threshold can be defined as following:

$$\gamma_{\text{thr}} = \hat{\sigma}_n^2 \cdot Q^{-1}(P_{\text{fa}}) \cdot \left(\sqrt{2N} + N\right), \tag{2}$$

where $Q()$ represents the Q-function. Having in mind that the total power of N collected samples in the given frequency band can be represented as the random variable $P_N = \sum_{k=0}^{N-1} |r[n]|^2$, then based on (1) the generic decision rule D_N can be then modified to the considered case:

$$D_N = \begin{cases} P_N \leq \gamma_{\text{thr}} & \longrightarrow H_0 \\ P_N > \gamma_{\text{thr}} & \longrightarrow H_1 \end{cases} \tag{3}$$

It is worth noticing that the reliability of energy-detectors strongly depends on the received power and on the accuracy of approximated variance noise $\hat{\sigma}_n^2$. The latter can be improved by increasing the number of collected samples N. In practice, however, the sensing time will be fixed, thus it is feasible that for low signal-to-noise ratios the performance of the traditional energy detection algorithms will be relatively low (i.e. SNR_{wall}) [13] [14].

B. Sequential Energy detector

The behaviour of the traditional energy detector can be improved in various ways, e.g. by application of the adaptively modified threshold. In our experiment we have selected the double-threshold, sequential energy detector which possess the same reliability as the traditional one, but its application could reduce the sensing time. The main concept is based on the assumption that for very strong PU signal, or - contrarily - in the presence of noise only, the number of samples that should be collected for decision making can be reduced. If this is a case, the sensing time is minimized which increases the time devoted for data transmission and reduces the energy consumption of the observation phase. In order to achieve this goal two decision thresholds have to be applied, γ_{HI} and γ_{LO}, which will be used for decision making if the signal is or is not present in the observed frequency fragment. In a nutshell, the procedure can be realized in the iterative way. The energy detector collects the signal samples in the shorter period $N_s < N$ and tries to made the decision. If the amount of power is greater than γ_{HI}, the decision of the PU signal presence can be made; if the received power is lower than γ_{LO} one can state that the considered channel is vacant. If the calculated value falls between these thresholds, the sequential energy detector collects next block of samples and repeats the procedure. When the total number of gathered samples reaches the maximally allowed value (i.e. the maximum sensing time will be finished), the decision will be made as for traditional algorithms. The decision rule for i-th iteration can be written as follows:

$$D_N^i = \begin{cases} P_N^i \leq \gamma_{\text{LO}} & \longrightarrow H_0 \\ P_N^i \in (\gamma_{\text{LO}}, \gamma_{\text{HI}}) & continue \\ P_N^i \geq \gamma_{\text{HI}} & \longrightarrow H_1 \end{cases} \tag{4}$$

C. Cyclostationary feature based spectrum sensing

In wireless communications, the transmitted signals show very strong cyclostationary features [15]. Therefore, identifying a unique set a features of a particular radio signal can be used to detect its presence based on its cyclostationary features. In the context of spectrum sensing many works have been conducted in using the cyclostationary features to detect the presence of PU in the radio environement [4]. In general, this method can perform better than the energy based detector. However its main drawbacks are the complexity associated with the detection technique and needs of some a-priori knowledge of the PU signal.
The cyclostationary detector can be realized by analyzing the Cyclic Autocorrelation Function (CAF) of a received signal $r(k)$. The CAF of a received signal $r(k)$ at the SU can be expressed as illustrated in (5).

$$R_r(k,\tau) = \sum_{\alpha} R_r^{\alpha}(\tau) \mathrm{e}^{2\pi j \alpha k} \tag{5}$$

where τ is lag associated to the autocorrelation function, α the cyclic frequency and $R_r^{\alpha}(\tau)$ is given by (6).

$$R_r^{\alpha}(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} R_r(k,\tau) \mathrm{e}^{-2\pi j \alpha k} \tag{6}$$

1) Classical Cyclostationary feature based detector: The classical approach to realize the cyclostationary detector is based on the Cyclic Spectrum Density (CSD) or the spectral correlation function of the received signal $r(k)$.

$$S_r^{\alpha}(f) = \frac{1}{N} \sum_{k=0}^{N-1} R_r^{\alpha}(\tau) \mathrm{e}^{-2j\pi f \tau} \tag{7}$$

The CSD function presented in (7), exhibits peaks when the cyclic frequency α equals the the fundamental frequencies of $s(k)$ the transmitted signal. Under the H_0 hypothesis, the CSD function does not have peaks since the noise is generally non-cyclostationary.
Using this technique, it is possible to distinguish even weak PU signals from the noise at very low SNR, where the energy detector is not applicable.

2) SPCAF detector: The discrete-time consistent and unbiased estimation of the CAF of a random process is given as:

$$\tilde{R}^{\alpha}_{rr^*}(\tau) = \frac{1}{M}\sum_{k=0}^{M-1} r(k)r^*(k+\tau)\mathrm{e}^{-2j\pi\alpha k} \tag{8}$$

For a given lag parameter $\tau \in \{1, 2, \ldots, L\}$, the cyclic autocorrelation function (CAF) can be seen as Fourier transform of $V = [r(0)r^*(0+\tau), r(1)r^*(1+\tau), \ldots, r(M-1)r^*(M-1+\tau)]$, where M is FFT size. As shown in the work of Khalaf *et al.* [8], the CAF is an M-dimensional sparse vector in cyclic frequency domain for a fixed lag parameter τ. Moreover, it presents a symmetry property as illustrated in (9).

$$||\tilde{R}^{\alpha}_{rr^*}(\tau)||_2 = ||\tilde{R}^{-\alpha}_{rr^*}(\tau)||_2 \tag{9}$$

Using a compressed sensing (CS) recovery technique like the Orthogonal Matching Pursuit (OMP) algorithm [16], we can accurately estimate the CAF using a limited and small number of received samples $N << M$. If the obtained CAF verifies the property (9) then H_1 is true otherwise H_0 is true. Its important to note that even under H_0 the obtained CAF verifies the symmetry property. However, when using a small number of samples, the probability to obtain a symmetrical CAF under H_0 is very small [8]. This SPCAF technique, can perform with a limited number of samples and consequently with lower complexity and shorter observation time compared to the classical cyclostationary feature detector.

III. Spectrum Sensing Experimental Setup

A. Hardware/Software overview

The performance of the selected spectrum sensing algorithms has been verified in conducted experiments realized by means of Universal Software Radio Peripheral (USRP) boards by Ettus Research. USRP platforms, as the low-cost and high-quality realization of the software-defined-radio (SDR) concept, delivers to the users various functionalities allowing for efficient, real-time realization of even very complicated wireless systems that operate in the radio-frequency (RF) band. The main role of the URSP platform is to convert the digital base-band signal delivered from the computer to analogue signal in the RF band. This process is realized in two-steps. In the first step the digital signal is converted to the digital intermediate-frequency (IF) domain; this phase is realized in the so-called mother-board, being the basis of the USRP platform. After that the signal is processed in the dedicated daughter-board, which is responsible for transforming of the digital IF signal to its analogue form in RF band. Finally, the signal is radiated by means of the mounted RF aerial. The variety of available daughter-boards creates big opportunities to the user, since these are designed to convert the IF signal to different part of the RF spectrum. Being the realization of the SDR concept, USRP are steered from the software level, i.e. the whole data processing in the base-band is realized on the computer side. Various software platforms can be applied for that purposes, including commercial and open-source solutions.

In our experiments two USRP boards have been utilized: the PU signal has been generated by means of the first board,

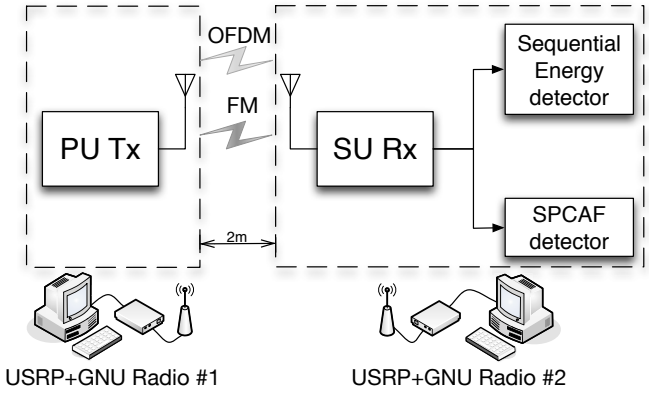


Fig. 1. Schematic system diagram

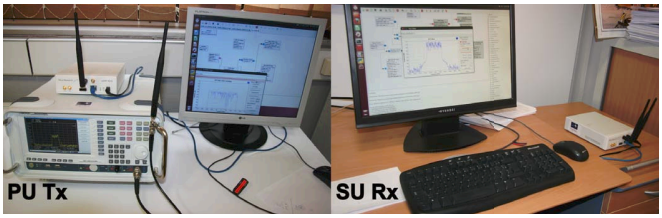


Fig. 2. PU transmitter and SU receivers realized by means of USRP board and personal computers

whereas the second one has been used for spectrum sensing purposes and acted as the secondary user. The whole software processing has been realized in the open-source GNU-Radio environment [17]. This set of libraries together with the appropriate drivers for manipulating the USRP boards and graphical programming environment allowed for efficient and accurate implementation of the selected spectrum sensing algorithms. In our experiment two sensing scenarios have been considered: first, where the narrow-band frequency-modulated radio signal, and second, where the multicarrier signal should be detected. The schematic diagram and the dedicated photographs of the experimentation setup are shown in Fig. 1 and 2, respectively. Finally, the whole system is presented in Fig. 3.

B. Transmitter side

At the transmitter side two types of signals were generated, the narrowband FM signal, and wideband mutlicarrier signal based on orthogonal frequency division multiplexing. In the former case the composite radio signal has been created and frequency modulated before sending to the USRP board via

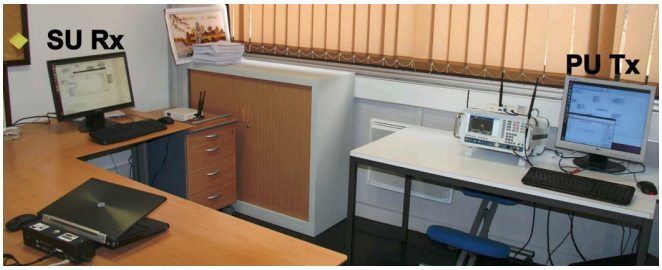


Fig. 3. The whole experimentation setup

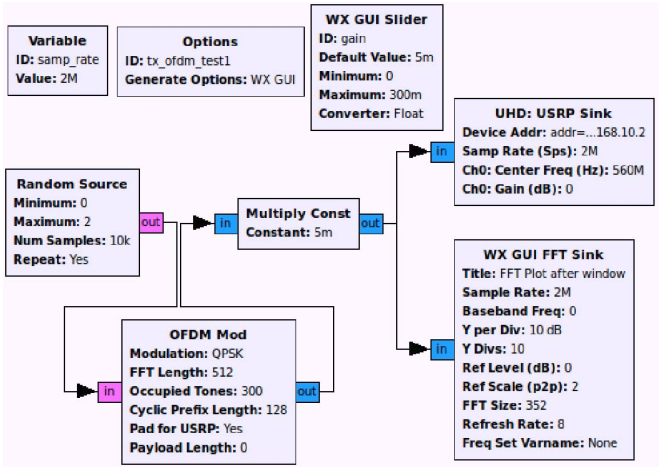


Fig. 4. Diagram of the PU OFDM transmitter realized in the GNU radio (Screenshot from GRC)

Ethernet cable. It means that assumed frequency deviation ($\pm$ 75 kHz deviation from the assisted center frequency) the bandwidth of the spectrum occupied by the FM signal is narrow (144 kHz). On the contrary, the spectrum of the multicarrier signal is assumed to be wider - the OFDM symbol with $N_{\mathrm{OFDM}} = 512$ subcarriers of the width 1.2 MHz has been used. As it has already been mentioned, the whole baseband processing has been realized on the PC computer in the GNU Radio environment, and in particular in the graphical tool called GNU Radio Companion (GRC), where the whole system is built from blocks. The screen-shot from the GRC illustrating the OFDM transmitter side is shown in Fig. 4. One can observe the presence of the signal source block (*Random Source*) that generates repeatedly random data, which are mapped to QPSK symbols and then are subject to OFDM modulation (realized in *OFDM Mod* block). Only 300 subcarriers from 512 available has been occupied, and the cyclic prefix of the size equal to one-quarter of that of IFFT was used. Finally, after proper power adjustment, the signal was sent to the local spectrum analyzer (*FFT plot*) and to the USRP block (*USRP Sink*), responsible for sending data to the USRP platform. It can be noticed that the complex sampling frequency has been set to 2 MHz, and the center frequency was set to 560 MHz. This frequency band has been chosen intentionally - it is within the TV band and is not occupied in the physical location where the experiment was conducted (i.e. no interference from the distance digital-television station could be observed).

C. Receiver side

As indicated in Fig. 1, two spectrum sensing algorithms have been implemented: the one based on energy detection, and the second that analyzes the cyclostationarity features of the received data. Analogously to the transmitter side, the whole base-band processing - that will be performed by the SU wireless terminal - has been realized in the computer side using the GNU Radio environment. The schematic diagram of the receiver is shown in Fig. 5. One can observe the presence of the *USRP Source* block responsible for delivering data from RF spectrum to the computer; it operates at the center frequency equal to 560 MHz and covers the band of 1MHz (what corresponds to complex sampling frequency equal to 1 Msps, as well). In order to evaluate the influence of noise on the performance of selected spectrum sensing algorithms, additional block for noise generation has been used and the noise-signal of appropriate power has been added to the signal produced by the *USRP Source* block. After, the signal is split into parallel chains: one dedicated for energy detection, and one for cyclostationariy feature-based algorithm. In such a configuration both algorithms operates on the same received samples making the comparison fair. One can also observe the presence of the *FFT plot* block used for displaying the received signal on the computer screen. Let us focus on the sequential energy detection algorithm (lower processing chain in the analyzed figure). The signal disturbed by the additive white Gaussian noise is then transformed to the frequency domain by means of *FFT* block. The algorithm for sequential energy detection, implemented in C language and assigned to the *DTED* block, make the decision on the occupancy of each frequency bin separately. In other words for the presented case 256 decisions will be made. The decisions are then transferred to the graphical sink. In the upper processing chain, devoted for cyclostationary feature spectrum sensing algorithm, the signal is converted from complex to real type and such modified signals are subject to processing in the *SPCAF_v1* block, realizing the functionality of the SPCAF algorithm described in the previous sections. All of the decisions are stored into the files.

IV. Experimental results

In order to compare the performance of the selected spectrum sensing algorithms let us analyze the results obtained during the conducted experiments.

A. FM signal as PU

Here, we compare performance of the SPCAF based blind detector with the sequential energy detector. A frequency modulated signal is used as primary user's signal. In the experiments, the central carrier frequency is set to 560 MHz. We compute the good detection probability (P_d) for the two detectors at different values of estimated SNR. In order to estimate the SNR, the noise power σ^2 is estimated at the receiver with no transmitted signal. Then, the transmitter is switched on and its transmission power is varied to obtain different signal-to-noise ratios (SNRs) at the receiver. Fig. 6 shows the detection probability of the two detectors obtained through experiments as function of SNRs. As concluded from the measurements, the probability of false alarm (P_{fa}) is approximately equal to 0.08 for both detectors. Furthermore, for the SPCAF detector, the maximum value of the lag parameter is $L = 5$ and the FFT size is $M = 2048$. It is clear from Fig. 6 that the performance of the SPCAF is better than the sequential ED. Another important point to note is that the number of received samples used by SPCAF is $N = 256$. However, the sequential ED requires at least N samples for detection.

B. OFDM signal as PU

In this part, the primary user signal is an orthogonal frequency division multiplexing (OFDM) signal. Fig. 7 shows

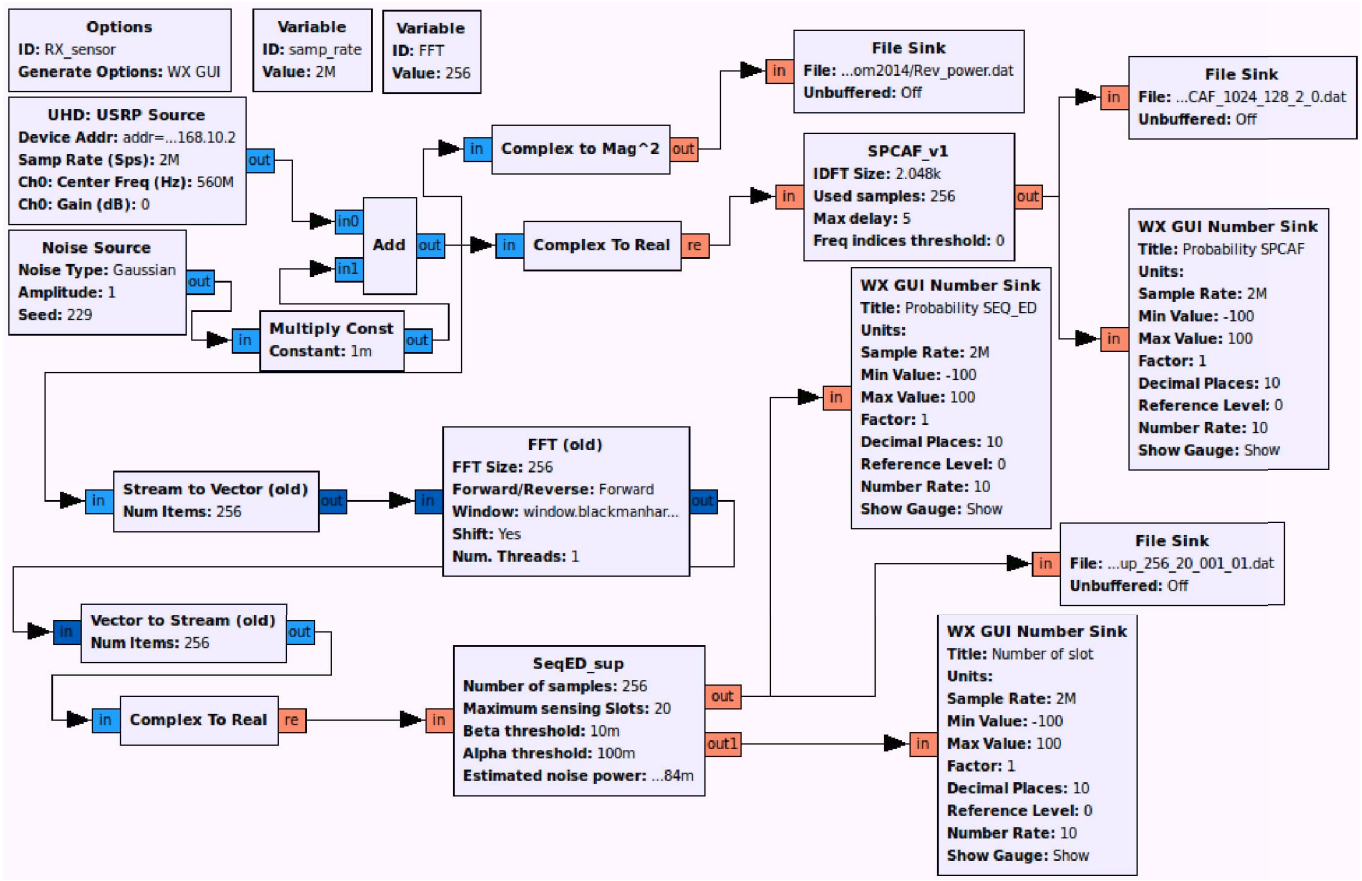


Fig. 5. Diagram of the SU receiver realized in the GNU radio (Screenshot from GRC)

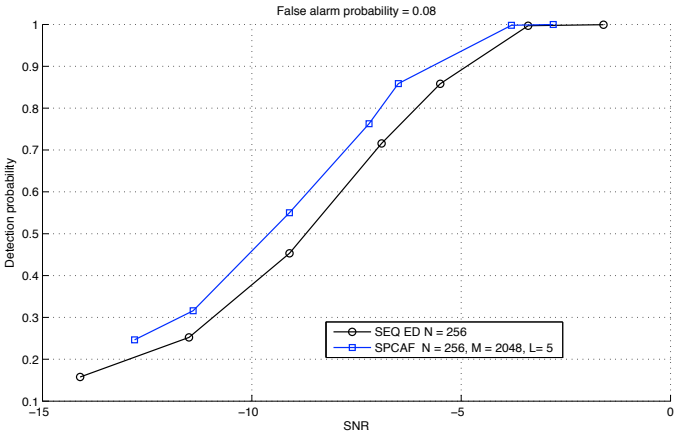


Fig. 6. Probability of detection for both algorithms when using FM signal as PU ($P_{fa} = 0.08$)

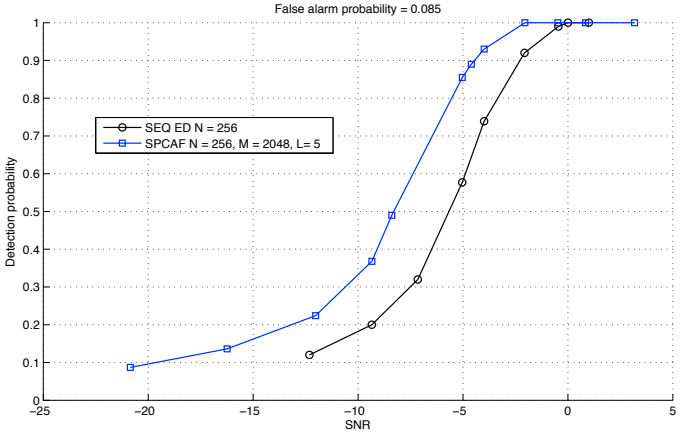


Fig. 7. Probability of detection for both algorithms when using OFDM signal as PU ($P_{fa} = 0.085$)

the detection probability achieved by the secondary user using SPCAF and sequential ED, while maintaining the false alarm probability below 0.08. Based on the Fig. 7, it can be concluded that the SPCAF gives better results compared to the sequential Energy detection method at low signal-to-noise ratios (SNRs).

V. Conclusions and Future Work

One can observe that for high SNRs both algorithms behave similar achieving high detection efficiency. However, it is not so challenging to detect strong signal, since the influence of noise in such a case will be minimal. Thus it is more important to focus on the low-SNR regions for both: narrowband and wideband signals. As it could be expected, the energy detector

achieves much poorer results comparing to the cyclostationary feature-based detector. It is due to the fact that the latter are not so sensitive to the signal imperfections and channel influence. However, let us remind that the reason for application of sequential energy-detection algorithm was to reduce the sensing time in the case when the reliable decision could be made before collecting the maximum allowed number of samples for spectrum sensing. The obtained results confirm that for high SNRs values the performance of sequential algorithms is as good as the performance of more advanced ones but the price paid for it - understood as the computational complexity - is much less. This bring us to the concept of the hybrid structure of spectrum sensing algorithm. In such a case, the low-complex double-threshold algorithm should be applied in the first phase, followed by the cyclostationarity-based one. When the signal of the PU will be strong enough or the observed signal variance will be close to the noise variance, the sequential energy detection algorithm will made reliable decision in the very short time, and the application of the more complicated algorithms will be not necessary. On the other hand, if the energy-detection procedure will not finish after collecting of N signal samples, the cyclostationary based algorithm shall be applied for final decision. Such a scenario will be investigated in the future. However, beside measurements of the sensing-time reduction obtained in the hybrid approach as well as of its overall performance, the whole system will be implemented on the FPGA chips. It will allow for detailed analysis of the energy consumed in each phase of the hybrid algorithm. Thus, the conclusions on the real energy-efficiency of the proposed hybrid approach could be drawn.

ACKNOWLEDGMENT

This work has been supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM# (FP7 Contract Number: 318306).

REFERENCES

[1] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 849–877, 2009.

[2] 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*, vol. 2010, p. 2, 2010.

[3] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *Signal Processing Magazine, IEEE*, vol. 29, no. 3, pp. 101–116, 2012.

[4] L. Lu, X. Zhou, U. Onunkwo, and G. Y. Li, "Ten years of research in spectrum sensing and sharing in cognitive radio," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–16, 2012.

[5] Z. Tian and G. B. Giannakis, "A wavelet approach to wideband spectrum sensing for cognitive radios," in *Cognitive Radio Oriented Wireless Networks and Communications, 2006. 1st International Conference on*. IEEE, 2006, pp. 1–5.

[6] Y. Zeng and Y.-C. Liang, "Maximum-minimum eigenvalue detection for cognitive radio," in *Proc. IEEE PIMRC*, vol. 7, 2007, pp. 1–5.

[7] P. Cheraghi, Y. Ma, and R. Tafazolli, "A novel blind spectrum sensing approach for cognitive radios," in *PGNET 2010 Conference*, 2010.

[8] Z. Khalaf, A. Nafkha, and J. Palicot, "Blind spectrum detector for cognitive radio using compressed sensing and symmetry property of the second order cyclic autocorrelation," in *Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), 2012 7th International ICST Conference on*. IEEE, 2012, pp. 291–296.

[9] P. Alvarez, N. Pratas, A. Rodrigues, N. R. Prasad, and R. Prasad, "energy detection and eigenvalue based detection: an experimental study using gnu radio," in *Wireless Personal Multimedia Communications (WPMC), 2011 14th International Symposium on*. IEEE, 2011, pp. 1–5.

[10] M. A. Sarijari, A. Marwanto, N. Fisal, S. K. S. Yusof, R. A. Rashid, and M. H. Satria, "Energy detection sensing based on gnu radio and usrp: An analysis study," in *Communications (MICC), 2009 IEEE 9th Malaysia International Conference on*. IEEE, 2009, pp. 338–342.

[11] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 1, pp. 116–130, First 2009.

[12] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proceedings of the IEEE*, vol. 55, no. 4, pp. 523–531, April 1967.

[13] R. Tandra and A. Sahai, "Snr walls for signal detection," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, no. 1, pp. 4–17, 2008.

[14] S. Bahamou, A. Nafkha *et al.*, "Noise uncertainty analysis of energy detector: Bounded and unbounded approximation relationship," in *Proceedings of the 21st European Signal Processing Conference*, 2013.

[15] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *Signal Processing Magazine, IEEE*, vol. 8, no. 2, pp. 14–36, 1991.

[16] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *Information Theory, IEEE Transactions on*, vol. 57, no. 7, pp. 4680–4688, 2011.

[17] Gnu radio - the free and open software radio ecosystem. [Online]. Available: http://gnuradio.org/

6.3.2 [C7]: Experimental Evaluation of MME and EME Detectors

Eigenvalue-based detector (EBD) has been proposed as an efficient way for spectrum sensing in cognitive radio network [43, 44, 45], as it does not need any prior knowledge about the noise power or signal to noise ratio. Due to this blindness property, EBD technique has been shown to overcome the conventional energy detector technique. The EBD method is based on the eigenvalues of the received signal covariance matrix and it uses results from random matrix theory (RMT) [46]. It detect the presence/absence of the PU by exploiting receiver diversity and includes the largest eigenvalue (LE) detector proposed in [45], the scaled largest eigenvalue (SLE) detector [44], the maximum-minimum eigenvalue (MME) detector [43], and the energy to minimum eigenvalue (EME) detector [45]. In the remaining of the present contribution, we consider only MME and EME for experimental evaluations using USRP N210 platform and GNU Radio software. For a constant false alarm probability (CFAP), P_{fa}, both blind spectrum sensing detectors (*i.e.* MME and EME) could be summarized in Algorithm 3 and Algorithm 4, respectively. In the sequel, N_s will denote the number of received samples per antenna and M the number of received antennas either distributed or not.

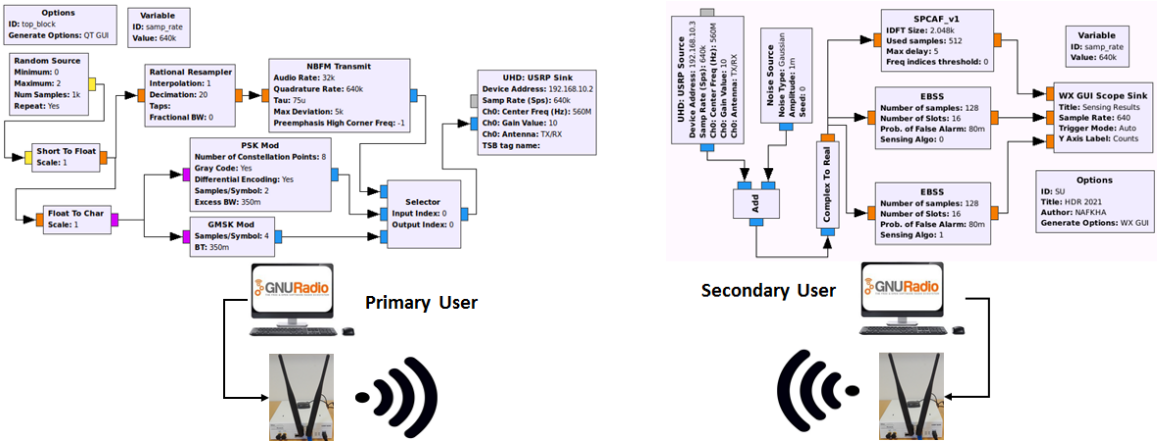
Algorithm 3: Maximum-Minimum Eigenvalue (MME) spectrum sensing detector

Data: Create the received signal matrix $\mathbf{Y} \in \mathbb{C}^{M \times N_s}$ and set the P_{fa} value
Result: Accept the null or alternative hypothesis
if *Sensing == true* **then**
(1) Compute $\mathbf{W} = \mathbf{Y}\mathbf{Y}^{\dagger}$;
(2) Compute the maximum (*i.e* λ_1) and the minimum (*i.e* λ_M) eigenvalues of $\mathbf{W}$;
(3) Evaluate the standard condition number (SCN) of the matrix $\mathbf{W}$ as: $\mathcal{T}_{A1} = \frac{\lambda_1}{\lambda_M}$;
(4) Decide in favor of null hypothesis if and only if $\mathcal{T}_{A1} \leq \frac{\alpha^2}{\gamma^2}\left(1 + \frac{\alpha^{-2/3} f_{\beta}^{-1}(1-P_{fa})}{(MN_s^2)^{1/6}}\right)$;
where $\alpha = \sqrt{N_s} + \sqrt{MN_s}$, $\gamma = \sqrt{N_s} - \sqrt{MN_s}$, and f_{β} is the cumulative distribution function of the Tracy-Widom distribution of order β;
end

Algorithm 4: Energy-Minimum Eigenvalue (EME) spectrum sensing detector

Data: Create the received signal matrix $\mathbf{Y} \in \mathbb{C}^{M \times N_s}$ and set the P_{fa} value
Result: Accept the null or alternative hypothesis
if *Sensing == true* **then**
(1) Compute $\mathbf{W} = \mathbf{Y}\mathbf{Y}^{\dagger}$;
(2) Compute the minimum (*i.e* λ_M) eigenvalues of $\mathbf{W}$;
(3) Compute the average power of the received signal as $P = \frac{1}{MN_s}\mathrm{Tr}(\mathbf{W})$;
(4) Evaluate the ratio between P and λ_M as: $\mathcal{T}_{A2} = \frac{P}{\lambda_M}$;
(5) Decide in favor of null hypothesis if and only if $\mathcal{T}_{A2} \leq \left(1 + \sqrt{\frac{2}{MN_s}} Q^{-1}(P_{fa})\right)\frac{N_s}{\gamma^2}$;
where $Q(.)$ is the Gaussian Q-function and $\gamma = \sqrt{N_s} - \sqrt{MN_s}$;
end

In order to compare the detection performance of the MME and the EME spectrum sensing algorithms, we have conducted a series of experiments in real radio environments using SDR platforms. In our experiments, two USRP N210 boards were used as depicted in Fig. 6.16, the primary user signal was a generated signal for the first board, whereas the second was used for spectrum sensing purposes, and acted as the secondary user.



(a) PU flowgraph in GNU Radio

(b) SU flowgraph in GNU Radio

Figure 6.16 – An overview of the experimental testbed.

At the secondary user side (Fig. 6.16b), the USRP source block, responsible for delivering data from RF spectrum to the computer, operates at the center frequency equal to 560 MHz and covers the band of 640KHz. An artificial Gaussian noise generation has been added to the signal produced by the USRP block in order to mitigate the effects of direct current (DC) offset.Then, the signal is split into three parallel paths each one of them will be dedicated to one of the tested sensing method algorithm (SPCAF, MME, and EME). Using this configuration, all sensing algorithms operate on the same received samples making a fair performance comparison. In the EBSS processing block, the *Sensing_Algo* parameter selects the spectrum sensing detection method: **0** for MME and **1** for EME.

In this contribution, we have conducted many in-lab experiments which showed that eigenvalue-based detectors (*i.e.* MME and EME) are more sensitive to the radio frequency (RF) impairments than the cyclostationarity-based detectors (*e.g.* SPCAF).

The paper entitled "Cyclostationarity-Based Versus Eigenvalues-Based Algorithms for Spectrum Sensing in Cognitive Radio Systems: Experimental Evaluation Using GNU Radio and USRP" dedicated to the experimental evaluation of SPCAF, MME and EME detectors was published in IEEE Eight International Workshop on Selected Topics in Mobile and Wireless Computing and is included hereafter [CI38].

Cyclostationarity-Based Versus Eigenvalues-Based Algorithms for Spectrum Sensing in Cognitive Radio Systems: Experimental Evaluation Using GNU Radio and USRP

Amor Nafkha*, Babar Aziz †, Malek Naoues* and Adrian Kliks ‡
*CentraleSupélec/IETR, Avenue de la Boulaie, 35576 Cesson Sévigné, France, email: {amor.nafkha malek.naoues}@centralesupelec.fr
† IFSTTAR,LEOST, F-59650 Villeneuve d'Ascq, France, email: babaraziz11@gmail.com
‡ Poznan University of Technology, Polanka 3, 60-965 Poznan, Poland, email: akliks@et.put.poznan.pl

***Abstract*—Spectrum sensing is a fundamental problem in cognitive radio systems. Its main objective is to reliably detect signals from licensed primary users to avoid harmful interference. As a first step toward building a large-scale cognitive radio network testbed, we propose to investigate experimentally the performance of three blind spectrum sensing algorithms. Using random matrix theory to the covariance matrix of signals received at the secondary users, the first two sensing algorithms base their decision statistics on the maximum to minimum eigenvalue ratio and the sum of the eigenvalues to minimum eigenvalue ratio, respectively. However, the third algorithm is based on cyclostationary feature detection and it uses the symmetry property of cyclic autocorrelation function as a decision policy. These spectrum sensing algorithms are blind in the sense that no knowledge of the received signals is available. Moreover, they are robust against noise uncertainty. In this paper, we implement spectrum sensing in real environment and the performance of these three algorithms is conducted using the GNU-Radio framework and the universal software radio peripheral (USRP) platforms. The results of the evaluation reveal that cyclostationary feature detector is effective in finite sample-size settings, and the gain in terms of the SNR with respect to eigenvalues-based detectors to achieve P_{fa} (probability of false alarm) = 0.08 is at least 4 dB.**

***Index Terms*—Cognitive radio, spectrum sensing, random matrix theory, compressive sensing, GNU-Radio, USRP platform.**

I. Introduction

Spectrum sensing to detect the presence of primary user transmissions is a crucial task for a cognitive radio system, which opportunistically accesses the spectrum once an empty subband is detected. Two basic approaches to spectrum sharing have been considered [1]: spectrum *overlay* and spectrum *underlay*. According to the spectrum overlay approach, the secondary users sense and identify unused frequency bands and use them for communication purposes. Thus, the secondary users (SU) are responsible for detecting the unused bands and they should vacate the spectrum as soon as the primary user begins its activities [2]. The underlay approach imposes constraints on the secondary users' transmission power level so that it can operates below the noise floor of primary users. In this paper we focus on the experimental evaluation of the overlay spectrum sharing.

This research is supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM# (FP7 Contract Number: 318306)

The detection of spectral occupancy can be viewed as a binary hypothesis testing problem. The null hypothesis $\mathcal{H}_0$ corresponds to the case where only additive noise is present, whereas the alternative hypothesis $\mathcal{H}_1$ refers to the case when the PU signal is present along with noise. Based on the latter hypothesis testing model, several spectrum sensing techniques have been proposed so far in literature [3], [4], [5]. These techniques are mainly categorized in two family three categories:

- *non-blind* techniques which require both primary user signal and noise variance information,
- *semi-blind* techniques which need just a few parameters, such as the additive noise variance and/or the fundamental cyclic frequency,
- *blind* techniques which exploits only the received signals without any a priori informations about noise or primary user signal,

In this paper, we will conduct an experimental evaluation by implementing three of the most known blind algorithms in the literature and compare their performance in real environments. The choice of the blind algorithms is motivated by the fact that the main aspect of of a typical cognitive radio is related to autonomously exploiting locally unused spectrum. There have been several blind spectrum sensing techniques proposed in literature. They include wavelet-based detection [4], eigenvalue-based detection [6], second order statistical based detection [7], and symmetry property of cyclic autocorrelation function based detection [8].The purpose of the conducted experiment is to sense the spectrum in a given frequency range and to make a reliable decision on the potential presence of primary user signal within frequency subband. In order to achieve this goal, three blind techniques for spectrum sensing have been implemented using USRP platforms and GNU Radio framework. We focus on sensing algorithms based on eigenvalues of received covariance matrix [6] and cyclostationary feature-

978-1-4673-7701-0/15/$31.00 ©2015 IEEE

based detectors [8].

The rest of the paper is organized as follows: in section II, the system model is presented. In section III and section IV, we present the eigenvalue based detection and the cyclostationary feature-based detection methods, respectively. In section V, we describe our testbed implemented to carry out the experiments. Section VI presents the experimental results of the considered algorithms. Finally, conclusions are presented in section VII.

II. System Model

We consider the problem of detecting the presence of a primary user signal at a specific frequency band based on the signal observed by the secondary user. Detecting the presence of a primary signal can be treated as a binary hypothesis testing problem [9]. Assuming a frequency-flat fading channel between the primary and secondary users, the sampled signal $r(t)$ received by the secondary users, defined as $r[n] = r(nT_s)$ with $1/T_s$ being the sampling rate, is expressed as

$$r[n] = \begin{cases} w[n] & \Rightarrow \mathcal{H}_0 \\ (h * s)[n] + w[n] & \Rightarrow \mathcal{H}_1 \end{cases} \quad (1)$$

where $*$ is the convolution operator. $s[n]$ and $h[n]$ stand is the digitally modulated signal of the primary user drawn from a certain modulation and the channel between the primary and secondary users, respectively, $w[n]$ is an additive white Gaussian noise with zero mean and variance σ^2. The objective of the spectrum sensing operation is to decide between null hypothesis $\mathcal{H}_0$ and alternative hypothesis $\mathcal{H}_1$ based on the observation of the received signal $r[n]$. The detection performance is characterized by two probabilities: probability of detection, P_{d}, where the decision is $\mathcal{H}_1$, while $\mathcal{H}_1$ is true; and probability of false alarm, P_{fa}, which corresponds to the case where the decision is $\mathcal{H}_1$ while $\mathcal{H}_0$ is true.

In this paper, three algorithms have been tested and compared through experimental analysis, namely eigenvalue based maximum-minimum eigenvalue detection method [6], energy with minimum eigenvalue detection [10], and a method based on cyclostationarity detection called Symmetry Property of Cyclic Autocorrelation Function [8].

III. Eigenvalue based detection algorithms

In this section we study two spectrum sensing algorithms based on the distribution of eigenvalues in large dimensional random matrix theory [11]. The discrete-time domain received signal $r[m]$ under $\mathcal{H}_1$ can be given as

$$r[m] = \sum_{k=0}^{N_h} h[k]s[m-k] + w[m] \quad (2)$$

where N_h is the channel filter length. At the cognitive radio's receiver, the received samples are split into M vectors each of length N_s. Let us consider the following $M \times N_s$ matrix consisting of the stacking of the M vectors.

$$\mathbf{Y} = \begin{bmatrix} r_1[1] & r_1[2] & \cdots & r_1[N_s] \\ r_2[1] & r_2[2] & \cdots & r_2[N_s] \\ \vdots & \vdots & \ddots & \vdots \\ r_M[1] & r_M[2] & \cdots & r_M[N_s] \end{bmatrix}$$

In the absence of primary user signal ($\mathcal{H}_0$: alternative hypothesis) all the received samples are uncorrelated whatever fading channel model. Moreover, the non-diagonal element of the received covariance matrix is theoretically zero, whereas the diagonal elements contain the noise variance. Hence, for a fixed M and $N_s \to \infty$, the sample covariance matrix $\frac{1}{N_s}\mathbf{YY}^*$ converges to the true covariance matrix $\sigma^2\mathbf{I_M}$.

In this paper, we assume that the noise is additive white Gaussian noise and, furthermore, the noise and the transmitted signal are uncorrelated. Then if the number of received samples N_s are large enough, it can be shown that

$$\mathbf{R_r}(\mathbf{N_s}) \approx \mathbf{R_r} = \mathbf{E}\left[\frac{\mathbf{1}}{\mathbf{N_s}}\mathbf{YY}^*\right] = \mathbf{HR_sH}^* + \sigma^2\mathbf{I_M} \quad (3)$$

where $\mathbf{R_r}$ and $\mathbf{R_s}$ matrices represent the covariance matrices of the received and transmitted signals, respectively. $\mathbf{I_M}$ is the identity matrix of size.

Let λ_{max} and λ_{min} represent the maximum and minimum eigenvalues of $\mathbf{R}_r$, respectively. Now suppose that ρ_{max} and ρ_{min} are the maximum and minimum eigenvalues of the matrix $\mathbf{HR_sH}^*$ then

$$\begin{aligned} \lambda_{\max} &= \rho_{\max} + \sigma^2 \\ \lambda_{\min} &= \rho_{\min} + \sigma^2 \end{aligned}$$

If $\mathbf{Z} = \mathbf{HR_sH}^* = \delta\mathbf{I_M}$ then $\rho_{min} = \rho_{max}$ where δ is a positive integer. In practice it is highly unlikely that the matrix $\mathbf{Z}$ will be equal to $\delta\mathbf{I}_M$ as mentioned in [6]. Hence, if there is no signal present then $\lambda_{max}/\lambda_{min} = 1$ otherwise $\lambda_{max}/\lambda_{min} > 1$. Therefore, this ratio can be used to detect the presence or absence of the signal. Based on the eigenvalues of the received covariance matrix, the following two methods are proposed in literature.

A. Maximum-Minimum Eigenvalue (MME) detection method

The steps of the MME algorithm are stated below

- Compute the received sample covariance matrix

$$\mathbf{R_r}(\mathbf{N_s}) = \frac{\mathbf{1}}{\mathbf{N_s}}\mathbf{YY}^* \quad (4)$$

- Compute the Maximum and Minimum Eigenvalues of the matrix $\mathbf{R_r}(\mathbf{N_s})$ (*i.e.* λ_{max} and λ_{min}).
- Sensing decision is given by:

$$\mathcal{D} = \begin{cases} \mathcal{H}_1 & if \\ \mathcal{H}_0 & otherwise \end{cases} \quad \frac{\lambda_{max}}{\lambda_{min}} \geq \gamma_{MME}$$

Note that γ_{MME} represents the threshold for the MME method. It is shown that the covariance matrix of the received signal in the absence of any signal at the receiver approximates to one class of random matrices called Wishart random matrix. The probability density function of Wishart random matrix has no marginal defined expression, finite dimensional case, and also present complex mathematics. Some studies on the spectral distribution of eigenvalues found the limiting values for maximum and minimum eigenvalues. Based on the proves given in [6], the distribution of the (properly rescaled) largest eigenvalue of the complex (real) Wishart matrix converges to

the Tracy-Widom law as M,N_s tend to $+\infty$ in some ratio $M/N_s > 0$. As a result, the threshold γ_{MME} given by:

$$\gamma_{MME} = \frac{\chi^2}{\delta^2}\left(1 + \frac{\chi^{-2/3}}{(MN_s^2)^{1/6}} F_\beta^{-1}(1 - P_{fa})\right)$$

Where $\chi = \sqrt{N_s} + \sqrt{MN_s}$, $\delta = \sqrt{N_s} - \sqrt{MN_s}$, and F_β is the cumulative distribution function (CDF) of the Tracy-Widom distribution of order β (*i.e.* $\beta = 1$ for real signal,$\beta = 2$ for complex signal).

B. Energy with Minimum Eigenvalue (EME) detection method

The steps of the EME algorithm are stated below

- Compute the covariance matrix $\mathbf{R}_r$ as given in (4).
- Compute the average power of the received signal

$$\Gamma(N_s) = \frac{1}{MN_s}\sum_{i=1}^{M}\sum_{n=0}^{N_s-1}|r_i[n]|^2 \tag{5}$$

- Then the eigenvalues of $\mathbf{R_r}$ are obtained.
- Decision: if

$$\frac{\Gamma(N_s)}{\lambda_{min}} > \gamma_{EME} \tag{6}$$

where γ_{EME} represents the threshold for EME method defined by:

$$\gamma_{EME} = \left(\sqrt{\frac{2}{MN_s}}Q^{-1}(P_{fa}) + 1\right)\frac{N_s}{\delta^2}$$

where $Q(.)$ is the Gaussian Q-function and P_{fa} is the probability of false alarm.

In the above discussion, we can notice that the threshold are not depending on the noise property but rather on the probability of false alarm, the number of segments M and the number of samples N_s per segment.

IV. Cyclostationary based detection: SPCAF Algorithm

In wireless communications, the transmitted signals show very strong cyclostationary features [12]. In the context of spectrum sensing many works have been conducted in using the cyclostationary features to detect the presence of PU in the radio environment [4]. In general, this method can perform better than the eigenvalue based detectors. However its main drawbacks are the complexity associated with the detection technique and needs of some a-priori knowledge of the PU signal.

The cyclostationary detector can be realized by analysing the Cyclic Autocorrelation Function (CAF) of a received signal $r(k)$. The CAF of a received signal $r(k)$ at the SU can be expressed as illustrated in (7).

$$R_r(k,\tau) = \sum_\alpha R_r^\alpha(\tau)\mathrm{e}^{2\pi j\alpha k} \tag{7}$$

where τ is lag associated to the autocorrelation function, α the cyclic frequency and $R_r^\alpha(\tau)$ is given by (8).

$$R_r^\alpha(\tau) = \lim_{N_s\to\infty}\frac{1}{N_s}\sum_{k=0}^{N_s-1}R_r(k,\tau)\mathrm{e}^{-2\pi j\alpha k} \tag{8}$$

The discrete-time consistent and unbiased estimation of the CAF of a random process is given as:

$$\tilde{R}_{rr^*}^\alpha(\tau) = \frac{1}{N_{FFT}}\sum_{k=0}^{N_{FFT}-1} r(k)r^*(k+\tau)\mathrm{e}^{-2j\pi\alpha k} \tag{9}$$

For a given lag parameter $\tau \in \{1, 2, \ldots, L\}$, the cyclic autocorrelation function (CAF) can be seen as Fourier transform of $V = [r(0)r^*(0+\tau), r(1)r^*(1+\tau), \ldots, r(N_{FFT}-1)r^*(N_{FFT}-1+\tau)]$, where N_{FFT} is FFT size. As shown in the work of Khalaf *et al.* [8], the CAF is an N_{FFT}-dimensional sparse vector in cyclic frequency domain for a fixed lag parameter τ. Moreover, it presents a symmetry property as illustrated in (10).

$$||\tilde{R}_{rr^*}^\alpha(\tau)||_2 = ||\tilde{R}_{rr^*}^{-\alpha}(\tau)||_2 \tag{10}$$

Using a compressed sensing (CS) recovery technique like the Orthogonal Matching Pursuit (OMP) algorithm [13], we can accurately estimate the CAF using a limited and small number of received samples $N_s << N_{FFT}$. If the obtained CAF verifies the property (10) then $\mathcal{H}_1$ is true otherwise $\mathcal{H}_0$ is true. Its important to note that even under $\mathcal{H}_0$ the obtained CAF verifies the symmetry property. However, when using a small number of samples, the probability to obtain a symmetrical CAF under $\mathcal{H}_0$ is very small [8]. This SPCAF technique, can perform with a limited number of samples and consequently with lower complexity and shorter observation time compared to the classical cyclostationary feature detector.

V. Spectrum sensing experimental setup

The performance of the previously presented spectrum sensing algorithms has been verified by conducting experiments realized by means of Universal Software Radio Peripheral (USRP) N210 board by Ettus Research. Being the realization of the SDR concept, USRP are steered from the software level, i.e. the whole data processing in the base-band is realized on the computer side. In our experiments two USRP boards have been utilized: the PU signal has been generated by means of the first board, whereas the second one has been used for spectrum sensing purposes and acted as the secondary user. The whole software processing has been realized in the open-source GNU-Radio environment. This set of libraries together with the appropriate drivers for manipulating the USRP N210 boards and graphical programming environment allowed for implementing the selected spectrum sensing algorithms. The schematic diagram of the experimentation setup is presented in Fig. 1 and the testbed environment is shown in Fig. 2.

A. Transmitter side (PU)

At the transmitter side, three types of signals were generated, the narrow-band FM, 8QPSK and GMSK signals. As it has already been mentioned, the whole baseband processing has been realized on the computer in the GNU-Radio environment where the whole system is built from blocks. After proper power adjustment, the signal was sent to the USRP block (*USRP Sink*), responsible for sending data to the USRP

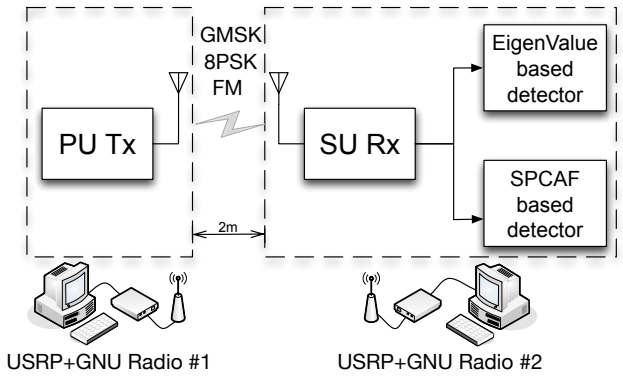


Fig. 1. Schematic system diagram

Fig. 2. The experimentation setup

platform. The center frequency was set to 560 MHz. This frequency band has been chosen intentionally - it is within the TV band and is not occupied in the physical location where the experiment was conducted (i.e. no interference from the distance digital-television station could be observed). The spectrum bandwidth occupied by the FM, 8PSK and GMSK signals are 144 kHz, 600 kHz and 800 kHz, respectively.

B. Receiver side (SU)

As indicated in Fig. 1, the three spectrum sensing algorithms explained in Section II have been implemented. Analogously to the transmitter side, the whole baseband processing - that will be performed by the SU - has been realized in the computer side using the GNU-Radio environment. The schematic diagram of the receiver is shown in Fig. 3. One can observe the presence of the *USRP Source* block responsible for delivering data from RF spectrum to the computer; it operates at the center frequency equal to 560 MHz and covers the band of 1 MHz (what corresponds to complex sampling frequency equal to 1 Msps). In order to evaluate the influence of noise on the performance of selected spectrum sensing algorithms, additional block for noise generation has been used and the noise-signal of appropriate power has been added to the signal produced by the *USRP Source* block. After, the signal is split into three parallel chains: one dedicated for each sensing method algorithm. In such a configuration the three algorithms operate on the same received samples making the comparison fair. One can also observe the presence of the *FFT plot* block used for displaying the received signal on the computer screen. In the upper processing chain, devoted for cyclostationary feature spectrum sensing algorithm, the signal is converted from complex to real type and such modified signals are subject to processing in the *Symmetry* block, realizing the functionality of the SPCAF algorithm described in the previous sections. The lower processing chains, devoted for eigenvalue based algorithm realized by the *EBSS_algo* block. In this block, the parameter *Sensing Algo* selects the detection method: "0" for MME and "1" for EME.

VI. Experimental results

In order to compare the performance of the selected spectrum sensing algorithms let us analyse the results obtained during the conducted experiments.

A. FM signal as PU

Here, we compare performance of the SPCAF based blind detector with the MME and EME detectors. A frequency modulated signal is used as primary user's signal. In the experiments, the central carrier frequency is set to 560 MHz. We compute the correct detection probability (P_d) for the three detectors at different values of estimated SNR. In order to estimate the SNR, the noise power σ^2 is estimated at the receiver with no transmitted signal. Then, the transmitter is switched on and its transmission power is varied to obtain different signal-to-noise ratios (SNRs) at the receiver. Fig. 4 shows the detection probability of the three detectors obtained through experiments as function of SNRs. As concluded from the measurements, the probability of false alarm (P_{fa}) is approximately equal to 0.08 for both detectors. Furthermore, for the SPCAF detector, the maximum value of the lag parameter is $\tau = 5$ and the FFT size is $N_{FFT} = 2048$. It is clear from Fig. 4 that the performance of the SPCAF is significantly better than the MME and EME algorithm. MME and EME methods show similar performance. Another important point to note is that the number of received samples used by SPCAF is $N_s = 512$.

B. 8PSK signal as PU

In this section we test the performance of the three algorithm for a 8PSK signal. The central carrier frequency is set to 560 MHz. We compute the correct detection probability (P_d) for the three detectors at different values of estimated SNR. The probability of false alarm (P_{fa}) is approximately equal to 0.08 for both detectors. Furthermore, for the SPCAF detector, the maximum value of the lag parameter is $\tau = 5$ and the FFT size is $N_{FFT} = 2048$. Fig. 5 shows that the performance of the SPCAF is significantly better than the MME and EME algorithm. An important point to note here is that the performance of MME method is better than the EME method. The number of received samples used by SPCAF is $N_s = 512$.

C. GMSK signal as PU

In this part, the primary user signal is a GMSK (Gaussian Minimum Shift Keying) signal. Fig. 6 shows the detection

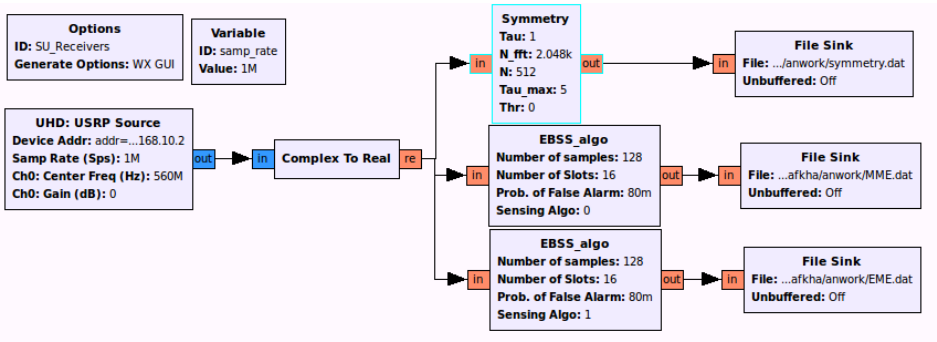


Fig. 3. Diagram of the SU receiver realized in the GNU radio (Screenshot from GRC)

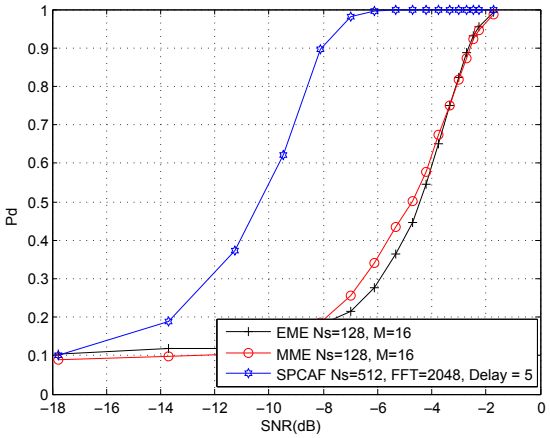


Fig. 4. Probability of detection for the three algorithms when using FM signal as PU ($P_{fa} = 0.08$).

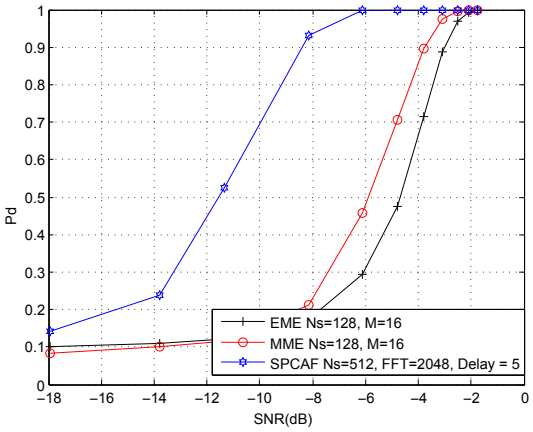


Fig. 5. Probability of detection for the three algorithms when using 8PSK signal as PU ($P_{fa} = 0.08$).

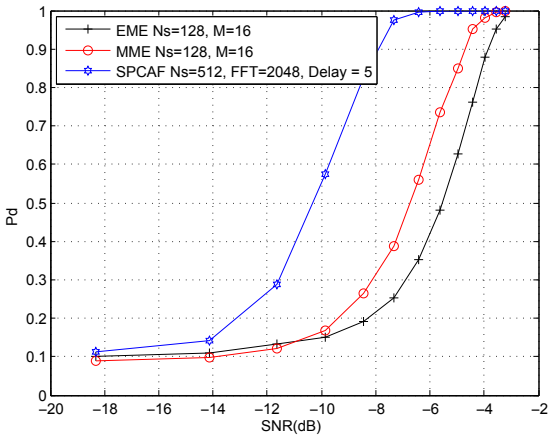


Fig. 6. Probability of detection for both algorithms when using GMSK signal as PU ($P_{fa} = 0.08$)

probability achieved by the secondary user using SPCAF, MME and EME, while maintaining the false alarm probability below 0.08. Based on the results in Fig. 6, it can be concluded that the SPCAF outperforms the eigenvalue based MME and EME methods at low signal-to-noise ratios (SNRs).

VII. Conclusion

In this paper, two classes of sensing methods are presented i.e. eigenvalue based detection methods (MME and EME) and SPCAF method which is based on cyclostationary feature detection. We analysed the performance of these spectrum sensing methods by measuring the detection probabilities as a function of SNR for a given false alarm probability. The testbed setup for the comparison of the three methods is based on two USRP N210 boards and GNU-Radio development toolkit using three types of modulations i.e. FM, 8PSK and GMSK. The cyclostationary feature-based detector achieves results far superior than the eigenvalue based methods in the case of finite received samples. The SPCAF gain in terms of the SNR with respect to eigenvalues-based detectors to achieve $P_{fa} = 0.08$ is at least 4 dB. This is due to the fact that

eigenvalue based techniques are more sensitive to the number of received samples.

ACKNOWLEDGEMENTS

This work has been supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM# (FP7 Contract Number: 318306).

REFERENCES

[1] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks*, vol. 47, no. 4, pp. 445 – 487, 2005.

[2] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 849–877, 2009.

[3] 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*, vol. 2010, p. 2, 2010.

[4] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *Signal Processing Magazine, IEEE*, vol. 29, no. 3, pp. 101–116, 2012.

[5] L. Lu, X. Zhou, U. Onunkwo, and G. Y. Li, "Ten years of research in spectrum sensing and sharing in cognitive radio," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–16, 2012.

[6] Y. Zeng and Y.-C. Liang, "Maximum-minimum eigenvalue detection for cognitive radio," in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, Sept 2007, pp. 1–5.

[7] P. Cheraghi, Y. Ma, and R. Tafazolli, "A novel blind spectrum sensing approach for cognitive radios," in *PGNET 2010 Conference*, 2010.

[8] Z. Khalaf, A. Nafkha, and J. Palicot, "Blind spectrum detector for cognitive radio using compressed sensing and symmetry property of the second order cyclic autocorrelation," in *Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), 2012 7th International ICST Conference on*. IEEE, 2012, pp. 291–296.

[9] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proceedings of the IEEE*, vol. 55, no. 4, pp. 523–531, April 1967.

[10] 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, June 2009.

[11] L. Cardoso, M. Debbah, P. Bianchi, and J. Najim, "Cooperative spectrum sensing using random matrix theory," in *Wireless Pervasive Computing, 2008. ISWPC 2008. 3rd International Symposium on*, May 2008, pp. 334–338.

[12] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *Signal Processing Magazine, IEEE*, vol. 8, no. 2, pp. 14–36, 1991.

[13] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *Information Theory, IEEE Transactions on*, vol. 57, no. 7, pp. 4680–4688, 2011.

6.4 Concluding remarks

In addition to the work performed during my own post-doctoral fellowship, the studies on software defined radio topic were performed in the context of 3 Masters, 3 PhDs (2 visitors) and 2 post-doctoral fellowships. They were supported by a total of 2 EU projects, 1 EU Celtic project, 3 ANR projects, 1 CominLabs project and 1 PHC RILA project. The results were presented in 3 journal papers, 25 conference articles, 2 invited talks and 1 Best Booth Award in CrownCom 2016.

In this chapter, I have covered three of the most important aspects related to software defined radio concept. The first aspect of our work was focused on theoretical study and hardware implementations of arbitrary sample rate conversion for software radio platforms. The second aspect was related to the design of flexible and high speed internal configuration access port controller, which is cornerstone of the FPGA dynamic partial reconfiguration technique. The reconfiguration time and power overheads have been theoretically derived and experimentally verified. The third aspect was more focused on the experimental investigations of various spectrum sensing techniques, developed in chapter 5, using USRP-based software defined radio platform and GNURadio open source software tool.

We resume below the combined outputs of the work on sample rate conversion, FPGA dynamic partial reconfiguration and USRP-based spectrum sensing.

Research Outputs of software defined radio

- **Publications**: 3 Journals + 25 Conferences + 4 Chapter.
- **Collaborators**: Moy C., Louet Y., Palicot J., Paquelet S., Leray P.
- **Postdoctoral researcher**: Babar A., Naoues M., Dark S., Delorme J.
- **PhD Students**: Zeineddine A., Chobanova Z. (visitor), Vásquez J. (visitor)
- **Master Students**: Karmani N., Maghzaoui M., Trotobas B.
- **Supported by**: IoTrust, UniThing, NewCom++, NewCom#, Sharing, WONG5, PHC RILA, GREAT

BIBLIOGRAPHY

[1] J. Mitola, "The software radio architecture," *IEEE Communications Magazine*, vol. 33, no. 5, pp. 26–38, 1995.

[2] E. J. McDonald, "Runtime fpga partial reconfiguration," in *2008 IEEE Aerospace Conference*, 2008, pp. 1–7.

[3] C. H. Huang and P. A. Hsiung, "Hardware resource virtualization for dynamically partially reconfigurable systems," *IEEE Embedded Systems Letters*, vol. 1, no. 1, pp. 19–23, 2009.

[4] P. Kyprianos, D. Apostolos, and H. Scott, "Performance of partial reconfiguration in fpga systems: A survey and a cost model," *ACM Trans. Reconfigurable Technol. Syst.*, vol. 4, no. 4, Dec. 2011.

[5] V. Kizheppatt and A. F. Suhaib, "Fpga dynamic and partial reconfiguration: A survey of architectures, methods, and applications," *ACM Comput. Surv.*, vol. 51, no. 4, Jul. 2018.

[6] C. Claus, B. Zhang, W. Stechele, L. Braun, M. Hubner, and J. Becker, "A multi-platform controller allowing for maximum dynamic partial reconfiguration throughput," in *2008 International Conference on Field Programmable Logic and Applications*, 2008, pp. 535–538.

[7] M. Liu, W. Kuehn, Z. Lu, and A. Jantsch, "Run-time partial reconfiguration speed investigation and architectural design space exploration," in *2009 International Conference on Field Programmable Logic and Applications*, 2009, pp. 498–502.

[8] M. Hübner, D. Göhringer, J. Noguera, and J. Becker, "Fast dynamic and partial reconfiguration data path with low hardware overhead on xilinx fpgas," in *2010 IEEE International Symposium on Parallel Distributed Processing, Workshops and Phd Forum (IPDPSW)*, 2010, pp. 1–8.

[9] S. G. Hansen, D. Koch, and J. Torresen, "High speed partial run-time reconfiguration using enhanced icap hard macro," in *2011 IEEE International Symposium on Parallel and Distributed Processing Workshops and Phd Forum*, 2011, pp. 174–180.

[10] P. Manet, D. Maufroid, L. Tosi, G. Gailliard, O. Mulertt, M. Di Ciano, J.-D. Legat, D. Aulagnier, C. Gamrat, R. Liberati, V. La Barba, P. Cuvelier, B. Rousseau, and P. Gelineau, "An evaluation of dynamic partial reconfiguration for signal and image processing in

professional electronics applications," *EURASIP Journal on Embedded Systems*, no. ID 367860, 2008.

[11] V. Kizheppatt and S. A. Fahmy, "Zycap: Efficient partial reconfiguration management on the xilinx zynq," *IEEE Embedded Systems Letters*, vol. 6, no. 3, pp. 41–44, 2014.

[12] S. Liu, N. Pittman, and A. Forin, "Minimizing partial reconfiguration overhead with fully streaming dma engines and intelligent icap controller," Tech. Rep. MSR-TR-2009-150, September 2009.

[13] V. Kizheppatt and S. A. Fahmy, "A high speed open source controller for fpga partial reconfiguration," in *2012 International Conference on Field-Programmable Technology*, 2012, pp. 61–66.

[14] ——, "Dyract: A partial reconfiguration enabled accelerator and test platform," in *2014 24th International Conference on Field Programmable Logic and Applications (FPL)*, 2014, pp. 1–7.

[15] L. A. Cardona and C. Ferrer, "Ac_icap: A flexible high speed icap controller," *International Journal of Reconfigurable Computing*, no. ID 314358, 2015.

[16] A. Kulkarni, V. Kizheppatt, and D. Stroobandt, "Micap: a custom reconfiguration controller for dynamic circuit specialization," in *2015 International Conference on ReConFigurable Computing and FPGAs (ReConFig)*, 2015, pp. 1–6.

[17] L. Pezzarossa, M. Schoeberl, and J. Sparso, "A controller for dynamic partial reconfiguration in fpga-based real-time systems," in *2017 IEEE 20th International Symposium on Real-Time Distributed Computing (ISORC)*, 2017, pp. 92–100.

[18] R. Bonamy, D. Chillet, O. Sentieys, and S. Bilavarn, "Towards a power and energy efficient use of partial dynamic reconfiguration," in *6th International Workshop on Reconfigurable Communication-Centric Systems-on-Chip (ReCoSoC)*, 2011, pp. 1–4.

[19] R. Bonamy, D. Chillet, S. Bilavarn, and O. Sentieys, "Power consumption model for partial and dynamic reconfiguration," in *2012 International Conference on Reconfigurable Computing and FPGAs*, 2012, pp. 1–8.

[20] A. Kulkarni, R. Bonamy, and D. Stroobandt, "Power measurements and analysis for dynamic circuit specialization," in *2015 International Conference on ReConFigurable Computing and FPGAs (ReConFig)*, 2015, pp. 1–6.

[21] M. A. Rihani, F. Nouvel, J.-C. Prévotet, M. Mroue, J. Lorandel, and Y. Mohanna, "Dynamic and partial reconfiguration power consumption runtime measurements analysis for

zynq soc devices," in *2016 International Symposium on Wireless Communication Systems (ISWCS)*, 2016, pp. 592–596.

[22] N. Kim, T. Austin, D. Baauw, T. Mudge, K. Flautner, J. Hu, M. Irwin, M. Kandemir, and V. Narayanan, "Leakage current: Moore's law meets static power," *Computer*, vol. 36, no. 12, pp. 68–75, 2003.

[23] S. Sharp, "Conquering the three challenges of power consumption," *Xcell Journal Online*, no. April, 2005.

[24] J. Bruce, W. N. Spencer, and T. W. David, "Chapter 29 - power and leakage," in *Memory Systems*. San Francisco: Morgan Kaufmann, 2008, pp. 847–864.

[25] Y. Nasser, J. Lorandel, J.-C. Prévotet, and M. Hélard, "Rtl to transistor level power modeling and estimation techniques for fpga and asic: A survey," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 40, no. 3, pp. 479–493, 2021.

[26] Z. Seifoori, H. Asadi, and M. Stojilović, "Shrinking fpga static power via machine learning-based power gating and enhanced routing," *IEEE Access*, vol. 9, pp. 115 599–115 619, 2021.

[27] M. B. Taylor, "A landscape of the new dark silicon design regime," *IEEE Micro*, vol. 33, no. 5, pp. 8–19, 2013.

[28] J. Rabaey, "Digital integrated circuits: A design perspective," 1995.

[29] D. Fitrio, A. Stojcevski, and J. Singh, *Subthreshold leakage current reduction techniques for static random access memory*. SPIE, 2005, vol. 5649, pp. 673 – 683.

[30] N. Weste and D. Harris, *CMOS VLSI design: a circuits and systems perspective*, 4th ed. United States: Pearson, 2011.

[31] T. Gupta and K. Khare, "A new technique for leakage reduction in 65 nm footerless domino circuits," *Circuits and Systems*, vol. 4, no. 2, pp. 209–216, 2013.

[32] Y. Im, E. S. Cho, K. Choi, and S. Kang, "Development of junction temperature decision (jtd) map for thermal design of nano-scale devices considering leakage power," in *Twenty-Third Annual IEEE Semiconductor Thermal Measurement and Management Symposium*, 2007, pp. 63–67.

[33] G. Hueber, L. Maurer, G. Strasser, K. Chabrak, R. Stuhlberger, and R. Hagelauer, "Sdr compliant multi-mode digital-front-end design concepts for cellular terminals," *WSEAS Transactions on Electronics*, vol. 2, no. 1, pp. 23–27, 2005.

[34] M. Dardaillon, K. Marquet, T. Risset, and A. Scherrer, "Software defined radio architecture survey for cognitive testbeds," in *2012 8th International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2012, pp. 189–194.

[35] E. Grayver, *Implementing software defined radio*. Springer Science & Business Media, 2012.

[36] Z. Geng, X. Wei, H. Liu, R. Xu, and K. Zheng, "Performance analysis and comparison of gpp-based sdr systems," in *2017 7th IEEE International Symposium on Microwave, Antenna, Propagation, and EMC Technologies*. IEEE, 2017, pp. 124–129.

[37] E. Blossom, "Gnuradio: tools for exploring the radio frequency spectrum," *Linux journal*, vol. 2004, no. 122, p. 4, 2004.

[38] O. Holland, H. Bogucka, and A. Medeisis, *On the GNURadio Ecosystem*, 2015, pp. 25–48.

[39] T. B. Welch and S. Shearman, "Teaching software defined radio using the usrp and labview," in *2012 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2012, pp. 2789–2792.

[40] B. Drozdenko, R. Subramanian, K. Chowdhury, and M. Leeser, "Implementing a matlab-based self-configurable software defined radio transceiver," in *International Conference on Cognitive Radio Oriented Wireless Networks*. Springer, 2015, pp. 164–175.

[41] "Out-of-tree oot modules," https://wiki.gnuradio.org/index.php/OutOfTreeModules, accessed: 2021-09-20.

[42] N. Kundargi and A. Tewfik, "A performance study of novel sequential energy detection methods for spectrum sensing," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 3090–3093.

[43] L. S. Cardoso, M. Debbah, P. Bianchi, and J. Najim, "Cooperative spectrum sensing using random matrix theory," in *2008 3rd International Symposium on Wireless Pervasive Computing*, 2008, pp. 334–338.

[44] P. Bianchi, J. Najim, G. Alfano, and M. Debbah, "Asymptotics of eigenbased collaborative sensing," in *2009 IEEE Information Theory Workshop*, 2009, pp. 515–519.

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

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

Chapter 7

Open Issues and Future perspectives

In this chapter, I will conclude my HDR thesis with the description of my current research topics. Then, I will give my research perspectives in the next three years (mid-term) and at a horizon of six years (long-term).

7.1 Short-term: Secure mixed-signal FPGA-based SoCs

Cybersecurity is a major concern in most computing systems, from edge to cloud. Different types of system vulnerabilities are exposed to attackers. Side Channel Attacks (SCAs) constitute one of the most critical vulnerabilities in computing systems, leading to a risk of information leaks. These information can either be sensitive and unencrypted, known as "**red**" information (*e.g* AES Cypher Key), or non-sensitive or encrypted, and known as "**black**" information. SCAs are non-invasive techniques that retrieve **red** information from a system by exploiting indirect sources of information leaked from the device through so-called side channels. An attacker may thus exploit any computation related physical phenomenon and recover **red** information from its correlation to captured signals. SCAs correlate **red** information with unintentional physical leakages like power consumption, electromagnetic radiation (EM) and timing. In general, EM-based SCA require expensive capture setups and a short range access to the target device. However, recent works of Camurati *et al*. [1, 2] have shown and demonstrated that EM side channels can be remotely exploitable in cost-effective manners, due to the discovery of new channels in mixed-signal edge devices that allow data to leak through Analog/RF part of the die. In [1, 2], authors have considered only CPU-based mixed signal chips (*e.g* nRF52832), which feature fixed hardware. However, recent large mixed-signal FPGA-based SoCs chips, known as RFSoCs, combine CPU cores, programmable logic, and programmable analog/Radio Frequency modules in the same die. Therefore, RFSoCs are potential targets for malicious sensor circuits and screaming channel attacks which makes highly probable the presence of unexpected new leakage channels to be discovered and exploited. The planned contributions can be summarized as follows:

- Reproduce screaming channels and leaky noise on RFSoc devices and try to find the the specificities that set RFSoC apart from simpler and fixed hardware IoT devices considered in Camurati's works.
- Introduce malicious sensors such as ring oscillator and tapped delay lines in order to increase and amplify the information leakage. These sensors can allow us gaining insights into the RFSoC device and with what is really going on inside through observing the dynamics of the power distribution network.
- Apply advanced signal processing and machine learning techniques in order to reduce the number of captured traces at the attacker side or to increase the screaming channels attack range.

This research work has already started in the framework of the PhD thesis of Jeremy Guillaume (Jan. 2021 - Dec. 2023).

7.2 Mid-term: Wireless physical layer Authentication

Wireless physical-layer (WPL) authentication is emerging technique to provide a high security and low complexity solution utilizing the unique properties at the physical-layer. Existing wireless physical-layer authentication schemes can be classified into two categories: channel/location based fingerprinting or radio-frequency hardware based fingerprinting. The channel/location based fingerprinting can be used to counter identity spoofing attacks based on both channel state information (CSI) and received signal strength (RSS) [3, 4, 5, 6, 7]. The radio frequency hardware based fingerprinting explores the imperfect features extracted from hardware or modulation to form an radio-frequency fingerprint that can identify a given device [8, 9, 10]. The radio-frequency fingerprints can origin anywhere along the transmitter chain like oscillators, synthesizer for up conversion, clock jitter, I/Q offset due to imbalance between I and Q branch, DAC sampling process and non-linearity of the amplifiers.

Application scenario: Wireless MAC spoofing attack

In infrastructure mode wireless local area networks, all devices are controlled and centrally coordinated through an access point (AP) which serves as the gateway to access other networks (*e.g.* Internet). Media access control (MAC) addresses are used to identify nodes (*i.e.* devices and AP) and to transmit data packets between them. Due to the shared nature of wireless medium, information transmitted over the radio channel is vulnerable to security attacks. Because of MAC headers of management and control frames are usually not encrypted while encryption is only applied to the payload, an adversary with a packet sniffer software can easily intercept packets and extract MAC addresses of the access point or other devices. Then, the malicious attacker can launch a large number of attacks such as denial of service (DoS), man-in-the-middle (MiTM)

[11]. Various strategies have been proposed in literature to detect MAC spoofing attack using wireless physical layer authentication [12, 13, 14, 15]. Recently, we conducted several laboratory experiments to verify the efficiency of frequency-offset based technique to detect unauthorized devices performing MAC address spoofing in the context of IEEE 802.11a/g WLAN standards.

Objectives

Given the variety of physical-layer fingerprinting features that can be used for device authentication. The first work to be carried out will be an extensive literature review on all physical layer features and try to find other interesting features which are not or not enough investigated in literature. The main goal here, is to propose a novel categorization of wireless physical-layer authentication. An important attention will be given to the granularity of various features in order to develop an efficient fingerprinting algorithms to identify wireless devices and detect malicious users. It is very challenging to find the features that are truly discriminative because, in general, the received signal is a mixture mixture of a plurality of components (TX signal, hardware impairments,...). The second study will address this problem by taking advantage from the recent expansion of machine learning methods. As a proof of concept, different developed methods will be evaluated and compared experimentally under different operating conditions.

7.3 Long-term: CR-IoT security

Recently, several research works considered the use of cognitive radio (CR) principle in the conventional IoT system in order to overcome the spectrum scarcity problem [16, 17, 18, 19]. In CR-based IoT networks, the main idea is to increase the spectrum utilisation through the use of dynamic spectrum access (DSA) technology. The DSA provides spectral resource sharing between primary users and secondary users. The secondary users can cooperate, in centralized (through a fusion center, FC) or distributed (consensus-based, relay-assisted) manners, share their spectrum sensing results, and take advantage of their spatial diversity in order to improved detection performance and relaxed sensitivity requirement.

The open nature of the CR-IoT technology can lead to major security concerns about the data being sent and how safe it is. Moreover, the opportunistic spectrum access functionality is power-intensive and this raises a major concern for battery constrained sensors. Indeed, finding the unused spectrum, providing effective transmission, and switching between frequency channels will significantly increase the power consumption of the connected device in comparison with the conventional IoT technology. Thus, more contending CR-based IoT devices bring in high collision probability, large sensing delay and huge power consumption. Recently, the issues in asynchronous spectrum sensing are investigated in the case of asynchronous behavior between primary and secondary users. Another issue should be considered as one of the biggest

challenges of the CR-IoT is to make a decision on the availability of the monitored spectrum with asynchronous sensing report from different connected devices. Indeed, it well known that the secondary user with high signal-to-noise ratio (SNR) finishes the detection earlier than the one with low SNR, and the fusion center makes the final decision depending on the earliest local decision.

Application scenario

There are many applications of CR-based IoT, three of them are:

- Military applications [20] where the main idea is to protect own connected devices, detect and identify the enemies devices in unknown battlefield environment. The connected devices will be flexible and adaptive to rapidly changing environments (battlefield situations and missions).
- Smart traffic applications [21] where the task is to predict public transportation volume by exploiting inter-communication among vehicles and infrastructures. The IEEE 802.11p standard is designed for intelligent transportation systems (vehicle-to-vehicle and vehicle-to/from-infrastructure communications). It typically uses seven channels of 10 MHz bandwidth in the 5.9 GHz bands: six service channels for general purpose messages, and one control channel which carries safety and coordination messages. As vehicular communication becomes a widespread phenomenon, there will be an increase in spectrum scarcity. The cognitive-radio based internet of things technology can be used to overcome this problem.
- Agricultural industry applications [22] where the task is to enable high number of smart farms (*i.e.* using IoT devices) to cooperate between them. This enables smart farms to enhance the management of water resources and irrigation schedules. Since a large number of connected devices in smart farms, denial of service attacks (*i.e.* frequency bands are overcrowded) can be perform in order to disable remote sensors. This type of attack would have deep human and financial consequences.

Objectives

The core of this long-term research work is the development and implementation of new technical solutions to address following issues on CR-based internet of things systems:

- ☐ Designing new power-efficient dynamic spectrum access schemes due to the limited battery power of CR-IoT connected devices. Approximate computing techniques will be extensively used.
- ☐ Protecting asynchronous cooperation spectrum sensing techniques against malicious users attacks (*e.g* primary user emulation, jamming,..). Machine Learning techniques can be leveraged to detect attacks in CR-IoT networks.

- ☐ Preserving the data integrity of CR-IoT devices with a particular focus on spectrum sensing data falsification attack.
- ☐ Conducting wireless experiments using SDR platforms (proof-of-concept of different research findings)

BIBLIOGRAPHY

[1] G. Camurati, S. Poeplau, M. Muench, T. Hayes, and A. Francillon, "Screaming channels: When electromagnetic side channels meet radio transceivers," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, ser. Proceedings of the ACM Conference on Computer and Communications Security. Association for Computing Machinery, oct 2018, pp. 163–177.

[2] G. Camurati, A. Francillon, and F. X. Standaert, "Understanding screaming channels: From a detailed analysis to improved attacks," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, pp. 358–401, 2020.

[3] N. Patwari and S. K. Kasera, "Robust location distinction using temporal link signatures," in *Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, 2007, pp. 111–122.

[4] K. Zeng, K. Govindan, and P. Mohapatra, "Non-cryptographic authentication and identification in wireless networks [security and privacy in emerging wireless networks]," *IEEE Wireless Communications*, vol. 17, no. 5, pp. 56–62, 2010.

[5] J. Yang, Y. Chen, W. Trappe, and J. Cheng, "Detection and localization of multiple spoofing attackers in wireless networks," *IEEE Transactions on Parallel and Distributed systems*, vol. 24, no. 1, pp. 44–58, 2012.

[6] L. Xiao, Y. Li, G. Han, G. Liu, and W. Zhuang, "Phy-layer spoofing detection with reinforcement learning in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10 037–10 047, 2016.

[7] Y. Tian, N. Zheng, X. Chen, and L. Gao, "Wasserstein metric-based location spoofing attack detection in wifi positioning systems," *Security and Communication Networks*, vol. 2021, 2021.

[8] G. Verma, P. Yu, and B. M. Sadler, "Physical layer authentication via fingerprint embedding using software-defined radios," *IEEE Access*, vol. 3, pp. 81–88, 2015.

[9] Y. Tu, Z. Zhang, Y. Li, C. Wang, and Y. Xiao, "Research on the internet of things device recognition based on rf-fingerprinting," *IEEE Access*, vol. 7, pp. 37 426–37 431, 2019.

[10] Y. Li, X. Chen, Y. Lin, G. Srivastava, and S. Liu, "Wireless transmitter identification based on device imperfections," *IEEE Access*, vol. 8, pp. 59 305–59 314, 2020.

[11] H. Alipour, Y. Al-Nashif, P. Satam, and S. Hariri, "Wireless anomaly detection based on ieee 802.11 behavior analysis," *IEEE transactions on information forensics and security*, vol. 10, no. 10, pp. 2158–2170, 2015.

[12] Y. Sheng, K. Tan, G. Chen, D. Kotz, and A. Campbell, "Detecting 802.11 mac layer spoofing using received signal strength," in *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*. IEEE, 2008, pp. 1768–1776.

[13] S. Liu, "Mac spoofing attack detection based on physical layer characteristics in wireless networks," in *2019 IEEE International Conference on Computational Electromagnetics (ICCEM)*, 2019, pp. 1–3.

[14] A. Vaidya, S. Jaiswal, and M. Motghare, "A review paper on spoofing detection methods in wireless lan," in *2016 10th International Conference on Intelligent Systems and Control (ISCO)*. IEEE, 2016, pp. 1–5.

[15] P. Madani and N. Vlajic, "Rssi-based mac-layer spoofing detection: Deep learning approach," *Journal of Cybersecurity and Privacy*, vol. 1, no. 3, pp. 453–469, 2021.

[16] A. A. Khan, M. H. Rehmani, and A. Rachedi, "When cognitive radio meets the internet of things?" in *2016 international wireless communications and mobile computing conference (IWCMC)*. IEEE, 2016, pp. 469–474.

[17] ——, "Cognitive-radio-based internet of things: Applications, architectures, spectrum related functionalities, and future research directions," *IEEE wireless communications*, vol. 24, no. 3, pp. 17–25, 2017.

[18] M. Gummineni and T. R. Polipalli, "Preliminary step for implementing cognitive internet of things through software-defined radio," *SN Computer Science*, vol. 2, no. 4, pp. 1–9, 2021.

[19] R. Ahmed, Y. Chen, B. Hassan, and L. Du, "Cr-iotnet: Machine learning based joint spectrum sensing and allocation for cognitive radio enabled iot cellular networks," *Ad Hoc Networks*, vol. 112, p. 102390, 2021.

[20] A. Kott, A. Swami, and B. J. West, "The internet of battle things," *Computer*, vol. 49, no. 12, pp. 70–75, 2016.

[21] A. Karimi, A. Taherpour, and D. Cabric, "Smart traffic-aware primary user emulation attack and its impact on secondary user throughput under rayleigh flat fading channel," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 66–80, 2020.

[22] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E.-H. M. Aggoune, "Internet-of-things (iot)-based smart agriculture: Toward making the fields talk," *IEEE Access*, vol. 7, pp. 129 551–129 583, 2019.

DOCTORAT BRETAGNE LOIRE / MATHSTIC

UNIVERSITÉ DE RENNES 1

Titre : Radio logicielle et traitement du signal MIMO : De la théorie à la pratique

Mot clés : Radio Logicielle, Radio intelligente, MIMO, FPGA, Détection des bandes libres

Résumé : Les recherches et les contributions présentées portent sur la radio intelligente et les techniques de traitement du signal appliquées aux systèmes MIMO. La thèse d'HDR s'articule autour des points suivants : (1) capacité ergodique du canal MIMO (sans-fil, optique), (2) analyse et amélioration des techniques sous-optimales pour la détection MIMO par le biais de l'interprétation géométrique du problème de détection et des méthodes bio-inspirées, (3) techniques de détection des bandes de spectre libres dans le contexte de la radio intelligente utilisant la cyclostationnarité ou les valeurs propres de la matrice de covariance des signaux reçus, (4) analyse, développement et implémentation matérielle des techniques de changement de rythme (SRC), (5) analyse et optimisation de la technique de reconfiguration dynamique partielle de FPGA, (6) quelques résultats d'études expérimentales de techniques de détection des bandes libres en utilisant les plateformes USRPs.

L'auteur présente également ses futures perspectives de recherches liées à la sécurisation des circuits électroniques (SOC) de type mixed-signal (exp. RFSOC) contre les attaques par canaux auxiliaires, identification des émetteurs via leurs empreintes radiofréquence, et la sécurisation de futures réseaux d'internet des objets basés sur la radio intelligente (CR-IoT). Les thèmes développés ont été décrits dans une vingtaine de revues internationales et plus de soixante articles dans des conférences internationales.

Title: Software Defined Radio & MIMO Signal Processing: From Theoretical to Practical Results

Keywords: Software defined radio, Cognitive radio, MIMO, FPGA, Spectrum sensing

Abstract: The research and contributions presented focus on intelligent radio and signal processing techniques applied to MIMO systems. The HDR thesis focuses on the following points: (1) ergodic capacity of the MIMO channel (wireless, optical), (2) analysis and improvement of suboptimal techniques for MIMO detection through geometric interpretation of the detection problem and bio-inspired methods, (3) free spectrum band sensing techniques in the context of smart radio using cyclostationarity or eigenvalues of the covariance matrix of received signals, (4) analysis, development, and hardware implementation of sample rate conversion techniques (SRC), (5) analysis and optimization of the FPGA partial dynamic reconfiguration technique, (6) some results of experimental studies of free band detection techniques using USRPs platforms.

The author also presents his future research perspectives related to securing mixed-signal electronic circuits (SOCs) (e.g., RFSOCs) against screaming channels, identification of transmitters through their radio frequency fingerprints, and securing future smart radio-based Internet of Things (CR-IoT) networks. The developed topics have been described in about 20 international journals and more than 60 papers in international conferences.