Spectral-Energy Efficiency Trade-off for Next-Generation Wireless Communication Systems

Fourat Sami Haider

A thesis submitted in partial fulfilment for the degree of
Doctor of Philosophy
at
Heriot-Watt University
School of Engineering and Physical Sciences

August 2015

Abstract

The data traffic in cellular networks has had and will experience a rapid exponential rise. Therefore, it is essential to innovate a new cellular architecture with advanced wireless technologies that can offer more capacity and enhanced spectral efficiency to manage the exponential data traffic growth. Managing such mass data traffic, however, brings up another challenge of increasing energy consumption. This is because it contributes into a growing fraction of the carbon dioxide (CO_2) emission which is a global concern today due to its negative impact on the environment. This has resulted in creating a new paradigm shift towards both spectral and energy efficient orientated design for the next-generation wireless access networks. Acquiring both improved energy efficiency and spectral efficiency has, nonetheless, shown to be a difficult goal to achieve as it seems improving one is at the detriment to the other. Therefore, the trade-off between the spectral and energy efficiency is of paramount importance to assess the energy consumption in a wireless communication system required to attain a specific spectral efficiency. This thesis looks into this problem. It studies the spectral-energy efficiency trade-off for some of the emerging wireless communication technologies which are seen as potential candidates for the fifth generation (5G) mobile cellular system. The focus is on the orthogonal frequency division multiple access (OFDMA), mobile femtocell (MFemtocell), cognitive radio (CR), and the spatial modulation (SM).

Firstly, the energy-efficient resource allocation scheme for multi-user OFDMA (MU-OFDMA) system is studied. The spectral-energy efficiency trade-off is analysed under the constraint of maintaining the fairness among users. The energy-efficient optimisation problem has been formulated as integer fractional programming. We then apply an iterative method to simplify the problem to an integer linear programming (ILP) problem.

Secondly, the spectral and energy efficiency for a cellular system with MFemtocell deployment is investigated using different resource partitioning schemes. Femtocells are low range, low power base stations (BSs) that improve the coverage inside a home or office building. MFemtocell adopts the femtocell solution to

be deployed in public transport and emergency vehicles. Closed-form expressions for the relationships between the spectral and energy efficiency are derived for a single-user (SU) MFemtocell network. We also study the spectral efficiency for MU-MFemtocells with two opportunistic scheduling schemes.

Thirdly, the spectral-energy efficiency trade-off for CR networks is analysed at both SU and MU CR systems against varying signal-to-noise ratio (SNR) values. CR is an innovative radio device that aims to utilise the spectrum more efficiently by opportunistically exploiting underutilised licensed spectrum. For the SU system, we study the required energy to achieve a specific spectral efficiency for a CR channel under two different types of power constraints in different fading environments. In this scenario, interference constraint at the primary receiver (PR) is also considered to protect the PR from harmful interference. At the system level, we study the spectral and energy efficiency for a CR network that shares the spectrum with an indoor network. Adopting the extreme-value theory, we are able to derive the average spectral efficiency of the CR network.

Finally, we propose two innovative schemes to enhance the capability of (SM). SM is a recently developed technique that is employed for a low complexity multiple-input multiple-output (MIMO) transmission. The first scheme can be applied for SU MIMO (SU-MIMO) to offer more degrees of freedom than SM. Whereas the second scheme introduces a transmission structure by which the SM is adopted into a downlink MU-MIMO system. Unlike SM, both proposed schemes do not involve any restriction into the number of transmit antennas when transmitting signals. The spectral-energy efficiency trade-off for the MU-SM in the massive MIMO system is studied. In this context, we develop an iterative energy-efficient water-filling algorithm to optimises the transmit power and achieve the maximum energy efficiency for a given spectral efficiency.

In summary, the research presented in this thesis reveals mathematical tools to analysis the spectral and energy efficiency for wireless communications technologies. It also offers insight to solve optimisation problems that belong to a class of problems with objectives of enhancing the energy efficiency.

To Miriam, Rose, & my parents

Acknowledgements

This thesis would not have been possible without the support of many people. I would firstly like to express my profound gratitude to my wife Miriam who has been an invaluable support to me, both practically taking time to help me with my thesis, but also always providing me with encouragement, motivation during difficulty and understanding throughout the duration of my adventure. I would like to gratefully acknowledge the distinguished supervision of Professor Cheng-Xiang Wang for all the useful advice, expert knowledge, guidance and patience without which this work would not have materialised. My special thanks goes to Professor Erol Hepsaydir for his enthusiastic support, mentoring and his unfailing belief in me. I would also like to express my gratitude to Professor Harald Haas for his helpful advice and contributions to my work. I am forever indebted to my family and especially my parents overseas for their tireless practical support, strength and help in every way during my studies. I hope I have realised my father's dream and have made him proud with my achievement. Last, but not least, I am deeply thankful to all my colleagues and friends who travelled alongside me on my PhD journey, shared in my struggles, provided laughter when it was most needed, and helped me with the completion of this thesis.

F. S. Haider

August 23, 2015

Contents

List of Figures

List of Tables

Abbreviations

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
AI	Antenna Interface
APD	Avalanche Photo-Diode
AWGN	Additive White Gaussian Noise
B4G	Beyond 4G
BaB	Branch and Bound
BBU	Base Band Unit
BD	Block Diagonalisation
BER	Bit Error Rate
BPSK	Binary Phase-Shift Keying

BS	Base Station
CIR	Channel Impulse Response
CN	Core Network
CO_2	Carbon Dioxide
CPU	Central Processing Unit
CR	Cognitive Radio
CRC	Cyclic Redundancy Check
CSI	Channel State Information
CSIT	Channel State Information at Transmitter
D2D	Device-to-Device
DAS	Distributed Antenna System
dB	Decibel
DD	Direct Detection
EMO	European Mobile Observatory
EPS	Evolved Packet System
FCC	Federal Communications Commission
FDD	Frequency Division Duplexing
FFT	Fast Fourier Transform
flops	Floating Point Operations
FPGA	Field Programmable Gate Arrays
G-SM	Generalised Spatial Modulation
GDP	Gross Domestic Product

GEVD	Generalised Extreme Value Distribution
HetNet	Heterogeneous Network
HSDPA	High-Speed Downlink Packet Access
ICSI	Interference Channel State Information
ICT	Information and Communication Technologies
IFFT	Inverse Fast Fourier Transform
ILP	Integer Linear Programming
IM	Intensity Modulation
IMT-A	International Mobile Telecommunications-Advanced
IoT	Internet of Things
IP	Internet Protocol
IT	Information Technology
ITS	Intelligent Transportation System
LED	Light Emitting Diode
LNA	Low Noise Amplifier
LTE	Long-Term Evolution
LTE-Advanced	Long-Term Evolution Advanced
MA	Margin Adaptive
MA-SM	Multiple Active Antennas Spatial Modulation
MFemtocell	Mobile Femtocell
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood

mm	millimeter
MMSE	Minimum Mean Square Error
MRC	Maximum Ratio Combining
MU	Multi-User
MU-MIMO	Multi-user MIMO
NLoS	Non-Line-of-Sight
Ofcom	Office of Communications
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OPEX	Operational Expenditure
PCM	Power Consumption Model
PD	Photo-Diode
PF	Proportional Fairness
PIN	P-Intrinsic-N
PR	Primary Receiver
PT	Primary Transmitter
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase-Shift Keying
RA	Rate Adaptive
RB	Resource Block
RF	Radio Frequency

RRA	Radio Resource Allocation
RRH	Remote Radio Head
RRU	Remote Radio Unit
SINR	Signal-to-Interference-plus-Noise Ratio
SM	Spatial Modulation
SNR	Signal-to-Noise Ratio
SSK	Space Shift Keying
SSL	Solid-State Lighting
ST	Secondary Transmitter
SU	Single User
SU-MIMO	Single User MIMO
SVD	Singular Value Decomposition
TDD	Time Division Duplexing
TXB	Transceiver Board
V-BLAST	Vertical-Bell Laboratories Layered Space-Time
VLC	Visible Light Communication
VNI	Visual Networking Index
WWRF	Wireless World Research Forum
ZF	Zero Forcing

Symbols

$\subseteq$	a subset
$\sum$	summation
$\prod$	production
$(\cdot)^T$	transpose of a matrix
$(\cdot)^{-1}$	inverse of a matrix
$(\cdot)^{\perp}$	orthogonal complement of a matrix
$\dot{\tilde{C}}(\cdot)$	first derivative
$\ddot{\tilde{C}}(\cdot)$	second derivative
$\acute{Q}(\cdot)$	constellation slicing function
$(\cdot)^H$	Hermitian transpose of a matrix
$(\cdot)^{\dagger}$	pseudoinverse of a matrix
$\|\| \cdot \|\|_F^2$	Frobenius norm of a matrix
$\arg\min\{f\}$	an index at which the vector f has the minimum value
$\arg\max\{f\}$	an index at which the vector f has the maximum value
$[x]^+$	equivalent to $\max(x, 0)$
$\mathtt{tr}[\cdot]$	trace of a matrix
$\mathbb{E}[\cdot]$	statistical expectation operator
$\mathbf{a}$	a vector
$\mathbf{A}$	a Matrix
$\mathbf{D}$	non-negative vector
$\mathbf{P}$	power loading vector

$\mathbf{A}_e$	exclusively matrix
$\mathbf{H}$	MIMO channel matrix
$\mathbf{W}$	weight matrix
$\mathbf{P}^{\perp}$	orthogonal projection
$\mathbf{I}$	identity matrix
$\mathbf{\Lambda}$	diagonal matrix
$\mathbf{S}$	convex set
$\mathbf{n}$	noise vector in MIMO channels
$\mathbb{K}$	Rician factor
$\mathbb{C}^{x\times y}$	the space of $x \times y$ complex matrices
$\mathbb{Z}$	the space of $x \times y$ integer matrices
$\mathbb{R}$	set of real number
$\mathcal{K}$	set of macro/primary users
$\mathcal{J}$	set of MFemtocells
$\mathcal{N}$	set of secondary users
$\mathcal{N}^{\mathrm{d}}$	set of direct transmission users
$\mathcal{K}$	set of primary users
$\mathcal{M}_j$	set of access users within j_{th} MFemtocell
ϕ	fractional of spectrum
α	a ratio ($0 \leq \alpha \leq 1$)
γ_{d}	instantaneous direct user SNR
γ_{a}	instantaneous access user SNR
γ_{pk}	peak transmit power
γ_s	instantaneous transmit power
γ_s^*	optimum power allocation
γ_{avg}	average transmit power
γ_0	water-filling cutoff value
$\acute{\gamma}_d$	direct transmission user SINR
$\acute{\gamma}_a$	access transmission user SINR
σ^2	noise variance
λ_i	eigenvalue of $\mathbf{H}_i\mathbf{H}_i^H$

β	pathloss exponent
δ_c, δ_p, λ_c, λ_p, and $\bar{\beta}$	scale parameters of type II or Fréchet distribution
η_{load}	processing consumption factor
η	power amplifier efficiency
η_{BH}	backhaul consumption factor
μ_Ω	mean of the shadowing
σ_Ω^2	variance of the shadowing
Ω	feasible region
Υ	Euler-Mascheroni constant
Pr	the probability
Q^{-1}	inverse Q-function
Q	peak interference constraints
$f(c_{n,k})$	required received power level to support $c_{k,s}$ bits on the sth sub-channel sent to user k
$c_{n,k}$	the number of loaded bits on the sth sub-channel sent to user k
U	total number of users
V	number of possible modulation schemes
U^{d}	number of direct users
U_j^{a}	number of access users in jth MFemtocell
$\bar{S}$	number of subcarriers within a sub-channel
EE	energy efficiency in Joules/bit/Hz
SE	spectral efficiency in bps/Hz
$\bar{I}$	binary indicator
R	achievable rate
$\bar{R}_n(t)$	average delivered rate in the past for nth direct transmission user
$R^{\min}$	minimum required user rate
$\acute{Q}(\cdot)$	constellation slicing function
$a \in A$	element a in the set A
N_r	number of receive antennas
N	total number secondary users
N_t	number of transmit antennas

N_0	noise power
N_{sys}	number of control signalling bits
$\bar{N}$	MIMO degrees of freedom
M_{mod}	bit per symbol
M	number of transmit antenna groups
K	total number of macro/primary users
S	number of sub-channels (or resource blocks)
S_0	wideband slope in low SNR regime
S_∞	slope of the spectral efficiency in the high SNR regime
G	backhaul link gain
C_b	backhaul spectral efficiency
C_a	access links spectral efficiency
C_{max}	maximum spectral efficiency
h_{d}	complex-value direct link channel gain
h_{b}	complex-value backhaul link channel gain
h_{a}	complex-value access link channel gain
g_c	cognitive channel gain
g_i	interference channel gain
$g_{c(\max)}$	supremum of a random variable g_c
A_{d}	mean power of the direct transmission channel
B	system bandwidth
Z_{ctrl}	size of the control signal
CR	coding rate
I	interference
I_k	aggregate interference at primary receiver
I_p	aggregate interference of primary receivers
I_c	aggregate interference of CR network
$\left(\frac{E_b}{N_0}\right)_{\text{min}}$	minimum energy efficiency
$\left(\frac{E_b}{N_0}\right)_{\text{penalty}}$	horizontal penalty
P_{cpu}	digital unit processes power consumption

P_{CE}	channel estimation power consumption
P_{RRM}	radio resource management power consumption
$P_{\text{BBU-TXB}}$	BBU transceiver board Power consumption
$P_{\text{RRU-TXB}}$	RRU transceiver board Power consumption
$P_{\text{BBU-crt}}$	BBU circuit power
$P_{\text{RRU-crt}}$	RRU circuit power
P_{sta}	static power consumption
P_{LNA}	LNA power consumption
P_{BH}	backhaul power consumption
P_{BS}	base station transmit power
P_{MF}	mobile femtocell transmit power
P_{fix}	fixed backhaul power consumption
P_{fan}	internal fan power consumption
P_{external}	external fan power consumption

Chapter 1

Introduction

1.1 Problem Statement

A world without wireless communication is almost difficult to imagine. We are becoming dependent upon its advantages of immediate transfer of information, the flexibility it affords us to work anywhere at anytime, interpersonal use of messaging, chat and browsing, and intercultural communication to name just a few. The continuous evolution of wireless communication with more sophisticated technologies has had a massive impact in changing how people anywhere and everywhere on the globe can communicate with each other in all aspects of life including business operations, individuals, and society. Not to mention that it has an obvious welcome impact on economic growth. The world of wireless communication has undergone some drastic changes during the past years due to technical and political reasons. In fact, the concept of digital globalisation has also made great impacts on the wireless communication. Today, the use of the Internet has been extended from the wired to the wireless and more precisely the mobile cellular communication, in order to meet the public's demands to Internet access on the move. This has resulted in organisation of a standard that will be practised universally. However as the advances and availability of attractive technology continues, it becomes clear that wireless communication is a victim of its own success. As the public increases its appetite for faster Internet

connection and for more powerful smart phones, tablets and laptops with multimedia capabilities, the resulting uncontrolled explosion of wireless mobile devices has inevitably caused rapid traffic growth. Many organisations have highlighted this fact. For example, the European mobile observatory (EMO) reported that there is a 92% growth in mobile broadband per year since 2006 [1]. Furthermore, it has been predicated by the Wireless World Research Forum (WWRF) that 7 trillion wireless devices will serve 7 billion people by 2017, i.e., the number of network-connected wireless devices will extend to 1000 times the world's population [2]. This places a hitherto unprecedented amount of pressure on the telecom stakeholders using the current generation of wireless communication to cope with such a tremendous traffic growth. It is therefore essential to develop a groundbreaking wireless system that with the network is able to manage such traffic. Globally researchers in the wireless communication technical domain have already started to investigate and identify the potential candidate technologies for the next-generation wireless systems, i.e., fifth generation (5G), with the objective of obtaining superior sustainability. 5G is expected to outperform the previous generation. This includes the ability to provide a much higher network capacity with a better coverage whiles keeping the complexity and cost of the network at a low possible as possible.

1.2 Motivation

Managing the future traffic growth with the deployment of such advanced wireless technologies, however, comes at another cost in a form of high energy consumption. The concept of energy-efficient communication, also known as Green Radio [3], has triggered a recent surge in research. This is because the increase of energy consumption in wireless communication systems has indirectly caused a rise in the emission of carbon dioxide (CO_2), which is considered as a major threat for the environment currently. Moreover, it has been reported by cellular operators that the energy consumption of base stations (BSs) contributes to over

70% of their electricity bill which accounts for approximately 18% of the total operational expenditure (OPEX) [3, 4]. Therefore, reducing energy consumption is indeed essential to minimising the carbon footprint, and reduce the OPEX for the cellular networks. In fact, energy-efficient communication was not one of the initial requirements even in the fourth generation (4G) wireless systems, but was highlighted as an emerging need at a later stage. The previous generation of the wireless communication systems have been evaluated in terms of spectral efficiency, i.e., an indication of how much traffic a limited frequency spectrum can carry. Even so, it neglects to offer any insight on how efficient the energy consumption is, i.e., the energy required to handle the traffic. It is now of paramount importance to adapt to the current climate and create a new paradigm shift towards joint spectral and energy efficient oriented design for the next-generation wireless access networks. However, there is a fundamental trade-off between the spectral and energy efficiency which refers to the idea that a communication system chooses the amount of spectral efficiency to achieve by enhancing the energy consumption. Jointly attaining both enhanced energy efficiency and spectral efficiency is unfortunately a challenging problem to solve. Often, achieving enhancement of one of them means sacrificing the other. Investigating this trade-off for the potential technology candidates of the 5G wireless systems, is of primary significance so one can design a system that achieves the target spectral efficiency with the lowest possible energy consumption cost. Those 5G technology candidates include orthogonal frequency division multiple access (OFDMA), mobile femtocell (MFemtocell), cognitive radio (CR), and the spatial modulation (SM). OFDMA technique is still considered very promising radio access technology because of its flexibility to allocate the radio resources among the users. CR and Femtocell can play key roles to enhance the spectrum utilisation. Femtocell is an indoor solution while CR can be used for both indoor and outdoor scenarios. Adopting the femtocell technology inside vehicles allows the users can enjoy the high data rate services in the high speed train with reduced signalling overhead. Combining SM with massive MIMO systems is expected to be another potential candidates in 5G wireless cellular networks to increase the spectral and energy

efficiency.

1.3 Contributions

The key contributions of the thesis are summarised as follows:

- **Designing an energy-efficient resource allocation algorithm**
  We study the problem of allocating the sub-channel and bit loading among the users with an objective of minimising the energy per transmitted bit for a given spectral efficiency in an OFDMA system. The energy-efficient optimisation problem is formulated first as an integer fractional programming, a problem which is difficult to solve. To relax the complexity of the optimisation problem, we use a parametric approach and based on that propose an iterative algorithm that optimally allocates the radio resources among the users. Using the proposed algorithm, the trade-off between spectral and energy efficiency is analysed under the constraints of maintaining the fairness among users and total transmit power.

- **Analysing the spectral and energy efficiency for cellular systems with MFemtocell deployment**
  We investigate the spectral and energy efficiency for the MFemtocell-assisted network with two resource partitioning schemes, i.e., orthogonal and non-orthogonal schemes. Closed form expressions for the relationships between the energy efficiency and the spectral efficiency are derived in low and high signal-to-noise ratio (SNR) regimes for a single-user (SU) MFemtocell system with these two resource partitioning schemes. We also present a spectral efficiency analysis of multi-user (MU) MFemtocells with OFDMA-based spectrum reuse and opportunistic scheduling schemes.

- **Analysing the spectral-energy efficiency trade-off for CR network**
  We compare the spectral-energy efficiency trade-off for an SU CR network in the low and high SNR regimes when transmitting a signal under average

power constraint and transmitting a signal under peak power constraint, whilst keeping the interference on the primary receiver (PR) below an acceptable level for both. We also propose a cellular architecture which is based on the CR network. Here, we derive the achievable spectral efficiency for the proposed network using extreme value theory. A general analytical framework to evaluate the energy-spectral efficiency trade-off of the CR based cellular network is established for all SNR values using peak-power interference constraint.

- **Analysing the spectral-energy efficiency trade-off for enhanced SM systems**

  We propose two transmission schemes that overcome the limitations of the SM. The first scheme, pre-coded SM, outperforms other single user MIMO (SU-MIMO) schemes. In this scheme, the transmitter pre-codes the transmit signals by a random vector. The receiver adopts matrix projection method in order to decode the information. The second scheme enables the SM scheme to be adopted in MU-MIMO systems. We apply block diagonalisation (BD) linear precoding technique to mitigate the inter-antenna interference. The spectral and energy efficiency for the MU-SM in the presence of massive MIMO system is then studied.

1.4 Publications

The work presented in this thesis has led to the following publications:

Journals

1. C.-X. Wang, **F. Haider**, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122-130, Feb. 2014.

2. **F. Haider**, C.-X. Wang, H. Haas, E. Hepsaydir, X. Ge, and D. Yuan, "Spectral and energy efficiency analysis for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 2969–2980, Feb. 2015.

3. **F. Haider**, C.-X. Wang, B. Ai, H. Haas, and E. Hepsaydir, "Spectral-energy efficiency trade-off of cellular systems with mobile femtocell deployment," *IEEE Trans. Veh. Technol.*, accepted for publication.

4. **F. Haider**, C.-X. Wang, H. Haas, E. Hepsaydir, M. Shu, C. Wang, and C. Jiang, "Spectral-energy efficiency trade-off for multi-user spatial modulation in massive MIMO networks," *IEEE Trans. Wireless Commun.*, submitted for publication.

Conferences

1. **F. Haider**, C.-X. Wang, H. Haas, D. Yuan, H. Wang, X. Gao, X.-H. You, and E. Hepsaydir, "Spectral efficiency analysis of mobile femtocell based cellular systems," in *Proc. IEEE ICCT'11*, Jinan, China, Sept. 2011.–Best Paper Award

2. **F. Haider**, C.-X. Wang, X. Hong, H. Haas, D. Yuan, and E. Hepsaydir, "Spectral-energy efficiency trade-off in cognitive radio networks with peak interference power constraints," **invited paper**, in *Proc. IEEE ICCT'11*, Jinan, China, Sept. 2011.–Best Paper Award

3. **F. Haider**, C.-X. Wang, H. Haas, E. Hepsaydir, and X. Ge, "Energy-efficient subcarrier-and-bit allocation in multi-user OFDMA systems," in *Proc. IEEE VTC'12-Spring*, Yokohama, Japan, May 2012.

1.5 Thesis Organisation

The remainder of this thesis is organised as follows:

Chapter 2 gives some essential background information for the research work

presented in this thesis. It begins by discussing the importance of wireless communication and how the cellular system evolved especially in reaction to the phenomenon of mobile data traffic growth. Next, it goes on to explore the vision of 5G mobile cellular systems and the potential candidate technologies that can deliver this goal. It then enters into the discussion of green communication and why it is necessary. This discussion includes an overview about energy efficiency metrics and the fundamental trade-off between spectral and energy efficiency are both given. This is followed by laying out the mathematical framework based on the fractional programing that suits the energy efficiency optimisation problems and how it can solve these kind of problems. At the end, a modern power consumption model (PCM) is introduced.

Chapter 3 studies the energy efficient resource allocation scheme for an MU-OFDMA system. It first gives a review on a family of radio resource allocation (RRA) techniques applicable to OFDMA systems. Next, the system model is presented whereby the employed network setup and transmission protocol are explained. Then, we confine our attention to energy-efficiency resource allocation of which the optimisation problem of subcarrier allocation and bit loading, subject to some constraints, is formulated. Following that, the proposed energy efficiency algorithm, together with its mathematical formulation, which takes into consideration the required spectral efficiency, the number of users, and the minimum user rate, are then presented.

Chapter 4 analyses the spectral and energy efficiency for a cellular system with MFemtocell deployment. It begins by introducing the concepts of the MFemtocell, highlighting its benefits and the challenges faced. It then outlines the possible resource partitioning schemes that can be deployed by the MFemtocell. Following that, the relationship between energy efficiency and spectral efficiency is analysed within the SU-MFemtocell with two different resource partitioning schemes. Finally, it investigates the spectral efficiency of an MU MFemtocell network with MU and opportunistic scheduling schemes.

Chapter 5 investigates the spectral and energy efficiency for CR networks. The

concepts of the CR is first introduced. Next, the analysis on spectral-energy efficiency trade-off for an SUCR fading channel in low and high SNR regimes is given. This is followed by introducing the proposed cellular architecture of which a CR network shares a spectrum that belongs to the indoor primary network each with multiple users. Here, it explores mathematical preliminaries in extreme value theory which is used to model the achievable spectral efficiency of CR network in the proposed system. This is followed by presenting a mathematical framework that can be used to study the spectral-energy system for the proposed system.

Chapter 6 analyses the spectral and energy efficiency for single and MU-SM modulation systems in massive MIMO networks. It begins with essential background information on SM, underlining its pros and cons. Following that, the proposed pre-coded SM, together with its performance evaluation, are then presented. Then the discussion takes us to the MU-SM whereby the employed system model along with the transmission protocol are given. Through this discussion, the formulation for the optimisation problem and proposed energy-efficient water-filling algorithm are described. Analysis on the achievable spectral and energy efficiency are then presented.

Finally, Chapter 7 concludes the thesis and suggests some future research topics.

Chapter 2

Background

2.1 Information Communication Technology

Information and communication technology (ICT) is the comprehensive integration of information technology (IT) with telecommunications serving as an umbrella to broadcast media, all variations of audio and video processing and transmission, television, cellular phones, computer and network hardware and software, satellite systems, network based control and monitoring functions, and so on. It is anticipated to intrinsically improve the quality of human life for most of the world. The ramifications of ICT in the world arena has already been recognised by local and international governments. For the first time in the history of civilisation, individuals and communities are able to immediately and affordably communicate and interact through data streaming and advanced communication. This transformation to the globalisation of communication far outweighs any potential gains of the matching technological advances to ICT. The use of ICT and digital technology is fast becoming a priority for governmental bodies, business establishments, life style choices of individual, industries, education institutions, and the environment. The innovative ICT has had and will have a significant positive impact upon the economy in the world [5]. In addition to that, ICT has provided society with a vast array of new telecommunication capabilities in the last few decades. Increasing capacity of ICT has further been empowered by

the growth of a global network of connected computers or the Internet. Today, the Internet has grown into the foundation of our information society. E-mails, web shopping and video broadcasting have become indispensable components of our everyday life. The present and future importance of this economically and developmentally is so critically significant that governments have actually set up specialised departments and organisations solely for the promotion of ICT, to ensure that even previously technologically disadvantaged communities, particularly in developing nations, will experience equal opportunity to information access, technological growth and communication services in order to be able to fully engage with the digital globalisation phenomenon. It would not then be an exaggeration to claim that ICT should be ranked alongside the discovery of Penicillin, and the industrial revolution when we consider it's place in human history.

2.1.1 Importance of Wireless Communication

The wireless communication system is a, perhaps the, critical element in the global ICT strategy, underpinning many other industries. It is one of the fastest growing and most dynamic sectors in the world. The impact of wireless communication cannot be minimised. It's advantages have been felt in all spheres of life including personal, commerce, healthcare, news and media reporting and access, education, government and has provided instant connection indiscriminately between populations throughout the world. Historically it's position is unrivalled, remote communities and populations now have the ability to engage instantly with each other, propelling forward intercultural evolution at an unparallelled rate. Indeed, wireless communication has even facilitated civil revolution and its future effects can only be imagined. As technology develops, wireless communication in general and particularly the cellular communication has a welcome impact on the economic growth. Many reports are optimistic that usage of wireless communication is steadily on the increase, a trend that will impact positively upon the economic growth and development of those countries. For instance,

the EMO reported, that the mobile communication sector has its own total revenue of €174 billion in 2010, thereby bypassing the aerospace and pharmaceutical sectors [1]. In addition to that, a report from the GSMA association has shown that cellular communication contributes to a growth rate of 3.4 % gross domestic product (GDP) per capita in a period from 2008 until 2011 [6]. The report also highlighted that a doubling of mobile data use, leads to a growth in the GDP per capita growth rate of 0.5 % points each year.

2.1.2 The Evolution of Cellular Systems

Unlike the fixed communication networks, wireless mobile networks can be built quickly. As a result of that, the wireless mobile communication has evolved dramatically over the last three decades. The phenomenal success of wireless mobile communications is mirrored by a rapid pace of technology innovation. From the early analogue mobile first generation (1G), deployed in Norway in 1981, followed by the second generation (2G) mobile communication system which debuted in 1991 to the third generation 3G system first launched in 2001, the wireless mobile network has transformed from a pure telephony system to a network that can transport rich multimedia contents. A notable milestone in the cellular revolution occurred when high-speed downlink packet access (HSDPA) enhanced 3G protocol was rolled out to address the increasing public demands of accessing the Internet on the move. The HSDPA has evolved to HSPA+ later on. HSDPA can support downlink speeds of up to 21Mbps. HSPA+ offers further speed increases, providing speeds of up to 42Mbps. The next step was the introduction of long-term evolution (LTE) proposed by third generation partnership project (3GPP) to meet the increasing demands for higher mobile broadband data rates and quality of service (QoS). LTE adopts OFDMA as the technique for resource sharing among multiple users [7]. LTE is able to offer peak rates up to 160Mbps within 20MHz channel bandwidth. Unlike the previous generation, LTE was designed from the beginning with the goal of developing the radio access technology (RAT) under the assumption that all the services can be based

on Internet protocol (IP). Advanced technologies like MIMO and link adaptation have been selected for the LTE system in order to boost up the performance. In addition to that, LTE has emerged with an evolution of the non-radio aspects of the overall system. The work on specifying the core network (CN) in LTE is commonly known as evolved packet system (EPS). Afterwards, international mobile telecommunications-Advanced (IMT-A) has specified the requirements of the 4G wireless systems. These were that a wireless system which should have the ability to support data rates of up to 1 Gbps and 100 Mbps in low mobility and high mobility environments, respectively, and also specified the use of IP for all services [9]. The LTE was applied as the baseline for further enhancements in order to meet IMT requirements. Hence, 3GPP proposed the LTE Advanced (LTE-Advanced) system as a candidate to be a practical 4G system [10]. Both LTE and LTE-Advanced share the same CN, however, LTE-Advanced extended the radio capabilities of LTE by introducing new features such as carrier aggregation, high orders MIMO, smart beamforming. etc.. Fig. 2.1 illustrates how the mobile broadband speed has evolved with different cellular generations.

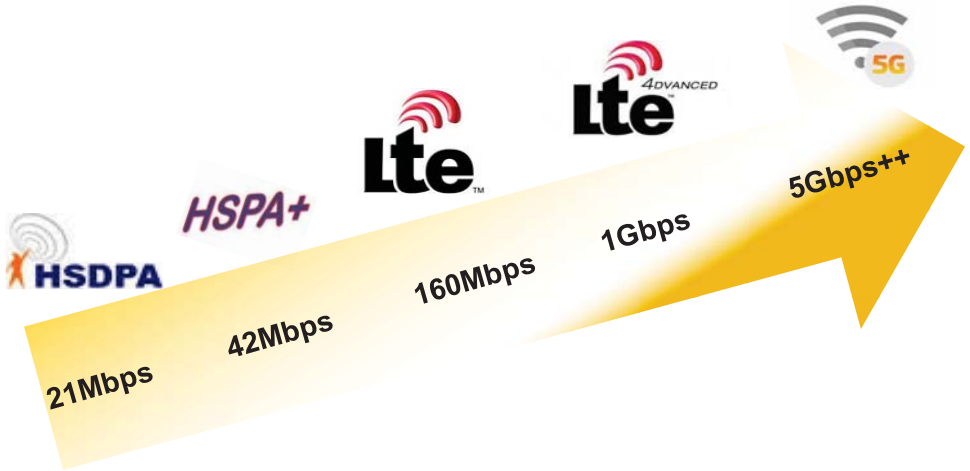

Figure 2.1: Evolution of Mobile Cellular Networks [8].

2.1.3 Mobile Data Traffic Growth

During the evolution of cellular systems, the number of wireless devices were in continuous growth, which was driven by the innovation of "smartness" of mobile devices. Each year several new devices in different shapes, with advanced features, and enhanced capabilities and intelligence, are being brought into the market. Smartphones together with large screens and different software applications, from browsers to video and audio streaming applications are projected to rise dramatically within the next few years. Cisco has reported that an estimated of 7 billion plus number of wireless devices were accessing the network in 2013, 77% more than in 2012 [11]. Furthermore, the introduction of Internet of Things (IoT) raised the challenges not simply in terms of network capacity, but also on the number of devices connected to the network. From Fig. 2.2 we can see the increase in number of connected wireless devices in each year until 2018. Cisco expects that the number of mobile devices will increase by more than 56% in 2018. It is noticeable that there will be a significant increase in the number of smart devices, i.e., smartphones and tablets, from 26% in 2013 to 42% by 2018.

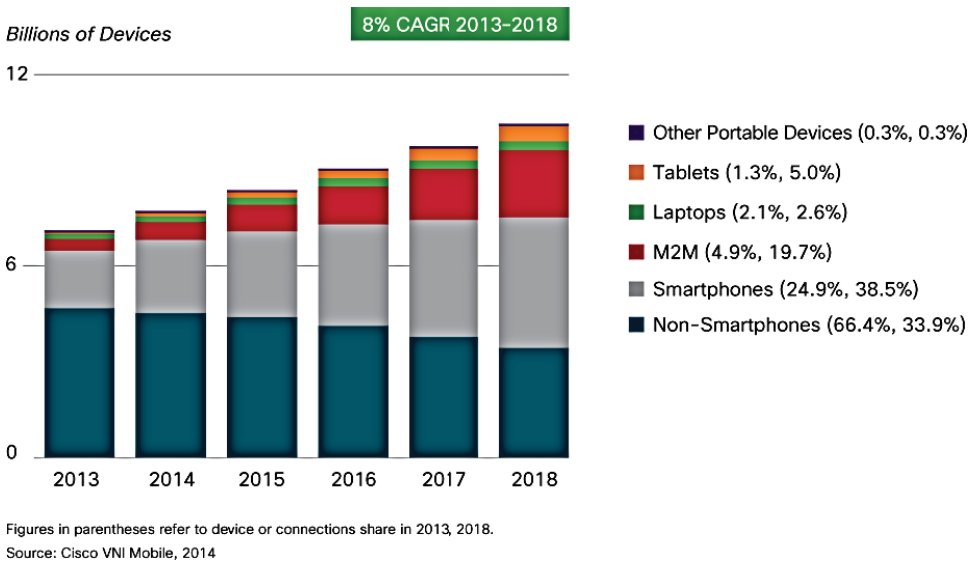

Figure 2.2: Cisco forecasts on mobile devices and connections growth [11].

Thus, the easy-to-use smart devices becomes more and more central in peoples lives and gradually essential in doing every-day activities. This massive number of wireless devices is one of the primary contributors to global mobile traffic growth in the next few years. Each year, Cisco releases its updated visual networking index (VNI) survey, which includes forecasts on the volume of mobile traffic. In its latest report, Cisco highlighted that the global mobile data traffic grew by 81% in 2013 where it reached 1.5 exabytes per month at the end of 2013, compared to 820 petabytes per month at the end of 2012 [11]. Hutchinson 3G, a prestigious mobile operator in the UK, has revealed that their data volume in 2014 has increased by 24 times more than that in 2007, as shown in Fig. 2.3. The reality, however, is that the potential for higher growth is still to come. Cisco's forecasts predict that the overall mobile data traffic will grow to 15.9 exabytes per month by 2018, nearly an 11-fold increase compared to 2013, the fact illustrated in Fig. 2.4. Moreover, mobile data traffic will have a growth rate considerably larger than the growth rate of traffic within the fixed network,

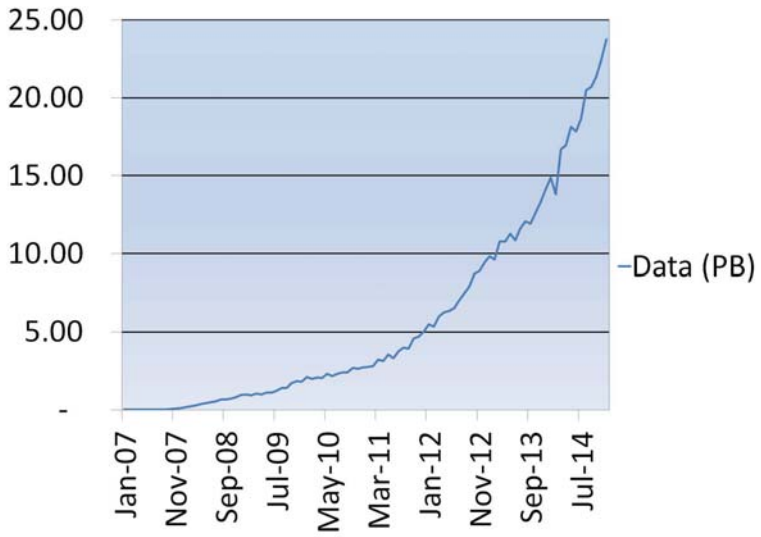

Figure 2.3: Hutchinson 3G traffic growth portfolio from 2007 until 2014. Reproduced by permission of © Hutchison 3G, U.K.

reported in [12]. The traffic has been driven mostly by the demand for video-related services, such as web video or downloading. It is understandable that, from a consumer point of view, as the broadband access speed increases, so will the desire for higher quality video be stepped up. A potential consequence of this situation, is that if the capacity of 4G cellular system increases at a level less than that required to support such the traffic growth, then this will contribute to congested networks, poorer user experience, and reduced innovation.

2.2 5G Mobile Cellular System

The continuing growth of traffic motivates the technologists and researchers to look ahead at how networks can be forged to meet the future anticipated extreme capacity demand, high quality performance, and a better mobile broadband

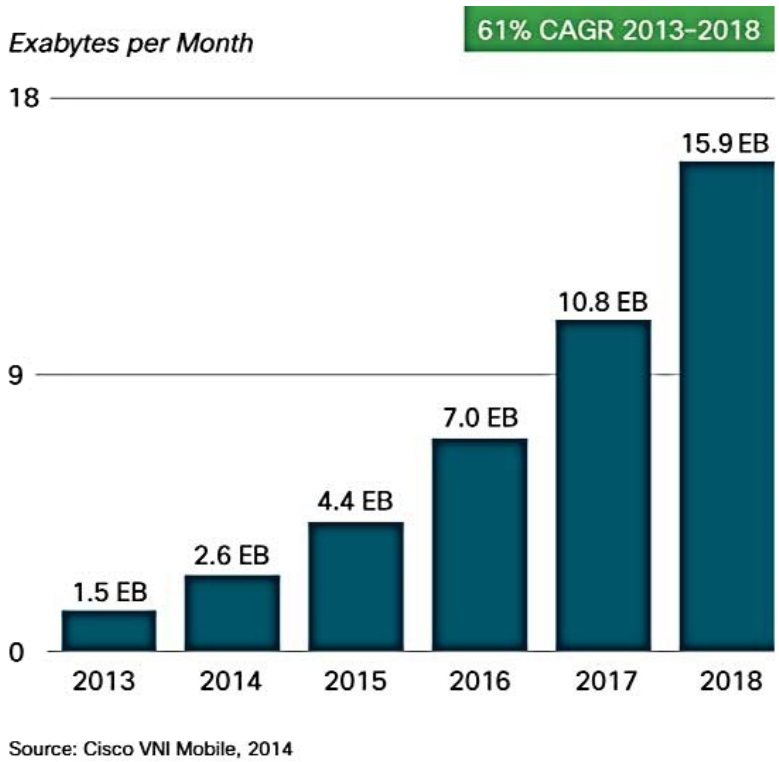

Figure 2.4: Cisco forecasts on mobile data traffic until 2018 [11].

experience. The 4G system may be able to cope with the traffic demands for a short period of time, but it cannot hope to survive in the long term. 4G systems presents challenges that have contributed to the inability in keeping up with such a growth demand in the traffic. One of the most crucial challenges is the physical scarcity of RF spectrums allocated for cellular communications. Hence, cellular frequencies use ultra high frequency bands for cellular phone use, normally ranging from several hundred MHz to several GHz. These frequency spectrums have been mostly allocated, making it difficult for operators to acquire more. Other backlogged challenges are for example, high-speed mobility, seamless coverage, diverse QoS requirements, and energy efficiency. To this end, we need cutting-edge cellular architecture to address these challenges from a fresh perspective in order to be a step ahead from the increasing traffic demands. Initial research has already been encouraged by various stakeholders in the wireless community to investigate beyond 4G (B4G) or the 5G wireless standards. The project entitled "UK-China Science Bridges: (B)4G Wireless Mobile Communications" (http://www.ukchinab4g.ac.uk) is perhaps one of the first projects in the world to start B4G research, where some potential B4G technologies were identified. Nokia Siemens Networks described how the underlying radio access technologies can be further developed to support up to 1000 times higher traffic volumes compared to 2010 travel levels over the next 10 years [13]. Samsung demonstrated a wireless system, using advanced technologies which can also be a potential candidate for 5G wireless communication networks with data rates faster than 1 Gbps [14].

What will the 5G mobile cellular system, which will be standardised around 2020, look like? It is now too early to define for sure. However, it is widely agreed that compared to the 4G, the 5G mobile cellular system should achieve 1000 times the capacity, 10 times the spectral efficiency, 25 times the cell average throughput, better energy efficiency, and higher data rate for both low and high mobility environments. The 5G mobile cellular system should also be able to support communications for some special scenarios not supported by 4G networks, e.g., for high speed train users. The aspiration is to connect the entire world, truly creating a

smaller global village, and accomplish seamless and ubiquitous communications between Anybody (people to people), Anything (people to machine, machine to machine), Anywhere (wherever they are), Anytime (whenever they need), and Anyhow (by whatever electronic devices, services, or networks). Achieving this is a challenge and requires a dramatic change in the design of cellular architecture.

2.2.1 A Potential 5G Mobile Cellular Architecture

Fig. 2.5 depicts a vision for 5G mobile cellular architecture. It comprises some promising key wireless technologies that can enable 5G mobile cellular system to fulfil performance requirements. It has been observed that mobile users stay indoors for about 80% of the time, whilst they only stay outdoors for about 20% of the time [15]. In reality, the distribution of users is not as uniform as this, as they tend to concentrate in certain areas, i.e., hotspots. In this case, a small cell will be able to offer high data throughput to the hotspot's users. In the meantime, users on the macrocell will also experience a more pronounced increase in throughput because a higher proportion of traffic has been offloaded from the macro cell into the small cell. Using this knowledge, one of the focus areas in designing the 5G cellular architecture is to separate outdoor and indoor scenarios. The 5G cellular architecture should then be a heterogeneous one, with macro-cells, micro-cells, small cells, and relays. This enables the macro base station (BS) to pay the attention to outdoor traffic while allowing the indoor traffic to be managed by indoor technologies. In this context, many emerging technologies can be employed indoors, which are suited for short-range communications to provide high capacity. Some examples include femtocell, ultra wideband (UWB), mm wave communications (3–300 GHz) [14], and visible light communication (VLC) (400–490 THz) [16]. The mm wave and VLC technologies use higher frequencies not traditionally used for cellular communications. These high-frequency waves do not penetrate solid materials very well and can be readily absorbed or scattered by gases, rain, and foliage. Therefore, it is hard to use these waves for outdoor and long distance applications.

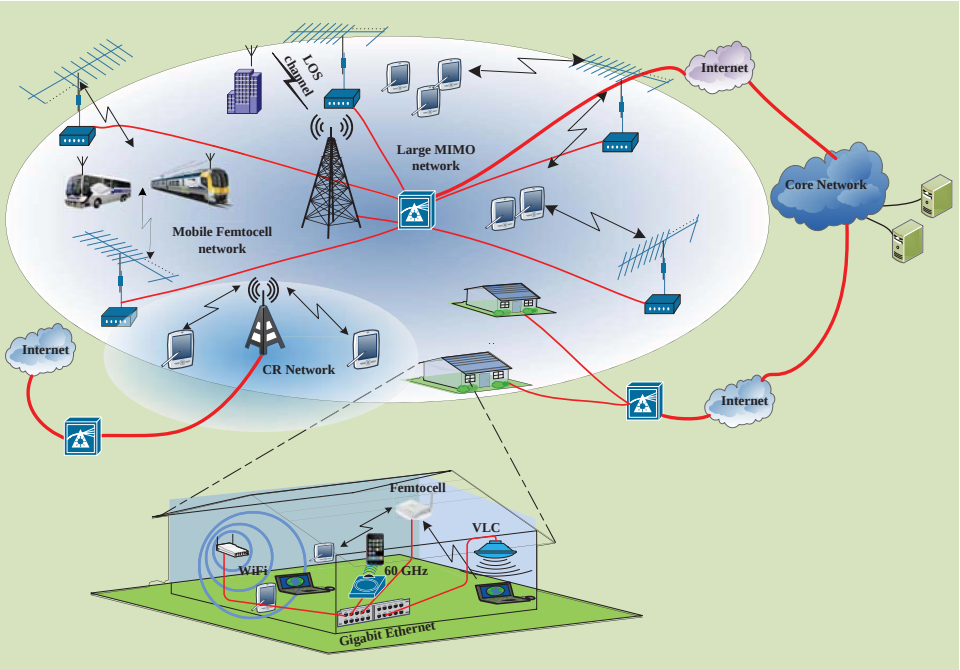

Figure 2.5: A vision 5G heterogeneous mobile cellular architecture [17].

However, with large bandwidths available, mm wave and VLC technologies can greatly increase the transmission data rate for indoor scenarios. Furthermore, to accommodate traffic that's generated from indoor high-mobility vehicles, we introduce the concept of MFemtocell. MFemtocell uses the femtocell technology with mobile wireless backhauling.

The use of advanced MIMO techniques such as massive MIMO [18–20] and SM [21] can provide a potential solution to accommodate prominent portion of the outdoor traffic. While most current MIMO systems utilise 2–4 antennas, the goal of massive MIMO systems is to exploit the potentially large capacity gains that can arise in larger arrays of antennas. Another scenario is that BSs will be equipped with large antenna arrays with some antenna elements (also large antenna arrays) distributed around the cell and connected to the BS via optical fibres, i.e., benefiting from both the distributed antenna system (DAS) and massive MIMO technology [20]. Outdoor mobile users are normally equipped with limited numbers

of antenna elements but they can collaborate with each other to form a virtual large antenna array, which together with BS antenna arrays will construct virtual massive MIMO links. Large antenna arrays will also be installed outside of every building to communicate with outdoor BSs or distributed antenna elements of BSs, possibly with line-of-sight (LoS) components. Large antenna arrays have cables connected to the wireless access points inside the building communicating with indoor users. This will certainly increase the infrastructure cost in the short term while can significantly improve the cell average throughput, spectral efficiency, and data rate of the cellular system in the long run. To reduce the complexity of the decoding algorithm, the SM concept can be integrated with massive MIMO systems [21].

To solve the spectrum scarcity problem, besides finding new spectrum not traditionally used for mobile wireless services, e.g., millimeter (mm) wave communications and the visible spectrum, we can also try to improve the spectrum utilisation of the existing radio spectrums using the concept of CR networks [22].

2.3 Promising Key 5G Wireless Technologies

In this section, based on the above proposed heterogeneous cellular architecture, we will discuss some key wireless technologies that can considered as promising candidates for 5G mobile cellular system. The purpose of developing these technologies is to enable a dramatic capacity increase of the 5G mobile cellular system with efficient utilisation of all possible resources.

2.3.1 Massive MIMO

MIMO systems consist of multiple antennas at both the transmitter and receiver. By adding multiple antennas, a greater degree of freedom (in addition to time and frequency dimensions) in wireless channels can be offered to accommodate more information data. Hence, a significant performance improvement can be obtained

in terms of the reliability and spectral efficiency. In massive MIMO systems, the transmitter and/or receiver are equipped with a large number of antenna elements (typically tens or even hundreds). Note that the transmit antennas can be co-located or distributed (i.e., DAS system) in different applications. Also, the enormous number of receive antennas can be possessed by one device or distributed to many devices. By properly using multi-user MIMO (MU-MIMO) in massive MIMO systems, the multiple access control (MAC) layer design can be simplified by avoiding complicated scheduling algorithms [23]. With MU-MIMO, the BS can send separate signals to individual users using the same frequency at the same time. Consequently, these main advantages enable the massive MIMO system to be a promising candidate for 5G mobile cellular system.

2.3.2 Spatial Modulation

In SM, blocks of the number of data bits are represented in the constellation point in the signal domain and a constellation point in the spatial domain (transmit antenna index). At each time instant, only one active antenna is active. The receiver needs first to estimate the activated antenna to use it as a basis to decode the information. SM can mitigate three major problems in conventional MIMO systems: inter-channel interference, inter-antenna synchronisation, and multiple RF chains. Moreover, low-complexity receivers in SM systems can be designed and configured for any numbers of transmit and receive antennas even for unbalanced MIMO systems. We have to mention that the transmission rate in SM increases logarithmically with the increase of the number of transmit antennas, whilst it increases linearly in the MIMO systems. Hence, the low implementation complexity comes at the expense of sacrificing some degrees of freedom. Therefore, it is worth investigating methods that enhance the transmission rate of SM whilst keeping the complexity as its low level. In this thesis we consider two schemes that improve the capability of SM. Both schemes can be another potential candidates for 5G mobile cellular system.

2.3.3 Cognitive Radio Networks

CR network is an innovative software defined radio technique which has been considered as one of the promising technologies to improve the utilisation of the congested RF spectrum. Adopting CR is motivated by the fact that a large portion of the radio spectrum is underutilised most of the time as has been declared by spectrum authority regulators. In CR networks, a secondary system can share spectrum bands with the licensed primary system, either on an interference-free basis or on an interference-tolerant basis [22]. The CR network should be aware of the surrounding radio environment and regulate its transmission accordingly. In interference-free CR networks, CR users are allowed to borrow spectrum resources only when licensed users do not use them. A key to enable interference-free CR networks is how to detect the spectrum holes (white space) that spread out in wideband frequency spectrums. CR receivers should first monitor and allocate the unused spectrums via spectrum sensing (or combining with geo-location databases) and feed back this information to the CR transmitter. A coordinating mechanism is required in multiple CR networks that try to access the same spectrum to prevent the users from colliding when accessing the matching spectrum holes. In interference-tolerant CR networks, CR users can share the spectrum resource with a licensed system whilst keeping the interference below a threshold. In comparison with interference-free CR networks, interference-tolerant CR networks can achieve enhanced spectrum utilisation, by opportunistically sharing the radio spectrum resources with licensed users, as well as better spectral and energy efficiency. In [24], hybrid CR networks have been proposed to be adopted in cellular networks to explore additional bands and expand the capacity.

2.3.4 Mobile Femtocell

MFemtocell is a new concept that has been proposed recently to be a potential candidate technology in the next generation intelligent transportation systems. It combines the mobile relay concept (moving network) with femtocell technol-

ogy. MFemtocells are located inside vehicles to communicate with users within the vehicle, while large antenna arrays are located outside the vehicle to communicate with outdoor BSs. A MFemtocell and its associated users are all viewed as a single unit to the BS. From the user point of view, a MFemtocell is seen as a regular BS. This is very similar to the above idea of separating indoor (inside the vehicle) and outdoor scenarios. Using this technology, the handover activities can be reduced for the users within the MFemtocell. This makes the deployment of MFemtocells suitable for high-mobility environments. In addition, the energy consumption of users inside an MFemtocell can be reduced due to relatively shorter communication range and low signaling overhead.

2.3.5 Visible Light Communication

VLC uses off-the-shelf white light emitting diodes (LEDs) used for solid-state lighting (SSL) as signal transmitters and off-the-shelf p-intrinsic-n (PIN) photo-diodes (PDs) or avalanche photo-diodes (APDs) as signal receivers [16]. This means that VLC enables systems that illuminate and at the same time provide broadband wireless data connectivity. If illumination is not desired in the uplink, infrared LEDs or indeed RF would be viable solutions. In VLC, the information is carried by the intensity (power) of the light. As a result, the information-carrying signal has to be real valued and strictly positive. Traditional digital modulation schemes for RF communication use complex valued and bi-polar signals. Modifications are therefore necessary and there is a rich body of knowledge on modified multi-carrier modulation techniques for intensity modulation (IM) and direct detection (DD). Data rates of 3.5 Gbps have been reported from a single LED. It has to be noted that VLC is not subject to fast fading effects as the wavelength is significantly smaller than the detector area. While the link-level demonstrations are important steps to prove that VLC is a viable technique to help mitigate spectrum bottlenecks in RF communications, it is essential to show that fully-fledged optical wireless networks can be developed by using existing lighting infrastructures. This includes MU access techniques, interference

coordination, etc.

In this thesis, the focus would be on MFemtocell, CR and SM with massive MIMO systems.

2.4 Green Radio Communications

Global warming is primarily a problem of too much CO_2 in the atmosphere. It refers to a century-scale rise in the average temperature of the earth which leads to an unpredictable climate system that has harmful effects on our communities and our health. The energy consumptions that are involved in different economic sectors, e.g., electrical supply, transport, industry, etc., contribute to global warming. This also includes the energy use in the operations of wireless communication systems. It has been stated that, there have been approximately 24 billion tonnes of total CO_2 emissions sine 2011 alone [25]. The ICT sector contributes to around 2% of that total, of which 0.74% domes from the wireless industry [26, 27]. Although the low percentage contributed by the mobile industries does not cause concern at present. However, with the continuous growth of traffic volume and with the efforts of other economic sectors to reduce their energy use, the percentage of wireless energy consumption is expected to increase correspondingly. Adopting innovative technologies with appropriate wireless architectures to reduce the energy consumption, is therefore needed in order to achieve greener radio communication [28, 29]. This contributes to a reduction of the fraction of CO_2 emissions. Wireless stakeholders allover the world should be encouraged to achieve certain energy consumption levels. As for the cellular operators interest, reducing energy consumption has a welcome impact to its OPEX and consequently into their total profit. Unlike the previous wireless generations, the design of 5G mobile cellular system should take into account the assessment of the energy consumption (i.e., energy efficiency) as a main objective. The indoor communication technologies are promising deployment strategies to achieve greater energy efficiency. This is because of the favourable channel conditions

that they can offer between the transmitters and receivers. It was demonstrated in [30–32] that the energy efficiency of the system can be improved significantly by using femtocells deployment within the existing macrocells. Moreover, by separating the indoor traffic from the outdoor traffic, the marcocell BS will have less pressure in allocating radio resources and can transmit with low power resulting in a significant reduction in the energy consumption. SM with massive MIMO is another potential application to achieve green wireless communications [33]. This is due to the fact that SM-MIMO contributes to a significant reduction into the circuit power by using a single-RF chain. VLC and mm wave technologies can also be considered as energy efficient wireless communication solutions to be deployed in 5G mobile cellular system. For example, in VLC systems the consumed energy in one bulb is much less than that in its RF-based equivalents for transmitting the same high density data.

2.4.1 Energy Efficiency Metrics

Energy efficiency is defined in several ways [34]. One way is to take the ratio of the transmitted bit rate per unit time (or spectral efficiency) to the transmitted power (plus circuit power), measured in bits/Joule (or bits/Hz/Joule). The objective of this definition is to increase the number of transmitted information bits per unit energy which has been widely used in various publications [35, 36]. A different way to define the energy efficiency is to use the energy consumption per bit, i.e., the ratio of consumed power per achievable rate (or spectral efficiency), as a metric measured in Joules/bit [37]. The energy consumption increases linearly with transmitted power. The objective here is to minimise the total energy consumption for a given data rate. Another widely used metric in the literature is energy-per-bit to noise power spectral density ratio i.e., $\left(\frac{E_b}{N_0}\right)$ [38–46]. This metric is useful to evaluate the spectral and energy efficiency from the transmit power perspective. In this thesis, the bits/Hz/Joule and $\left(\frac{E_b}{N_0}\right)$ have been considered as energy efficiency metrics.

2.4.2 Spectral-Energy Efficiency Fundamental Trade-off

The spectral-energy efficiency trade-off is becoming a important tool for bringing insight on how to design an energy efficient communication system. The spectral-energy efficiency trade-off refers to the amount of consumed energy needed to reach a specific spectral efficiency or, put simply, how to express energy efficiency as a function of spectral efficiency. The concept of spectral-energy efficiency trade-off was first introduced in [35] where the authors found the relationship between the energy channel and the channel capacity for the additive white Gaussian noise (AWGN) channel. Two analytical tools used to analyse the spectral-energy efficiency trade-off for any given wireless networks were proposed in [38] and [39] in low and high SNR regimes, respectively. In both tools, the energy efficiency, $\left(\frac{E_b}{N_0}\right)$, is defined as the required energy per bit (in joules/bit) normalised to the background noise power N_0 for reliable communications. The spectral efficiency, C, refers to the number of bits per second transmitted over a given bandwidth (in bps/Hz).

- **Low SNR regime**

  Using the approximation tool of [38] in the low SNR regime, $\left(\frac{E_b}{N_0}\right)$ can be approximated as an affine function with respect to the spectral efficiency and can be expressed by

  $$\left(\frac{E_b}{N_0}\right)\Bigg|_{\text{dB}} = \left(\frac{E_b}{N_0}\right)_{\text{min}} + \frac{3}{S_0}C \tag{2.1}$$

  where $\left(\frac{E_b}{N_0}\right)_{\text{min}}$ is the minimum energy efficiency required for transmitting information reliably over a channel and it is given by

  $$\left(\frac{E_b}{N_0}\right)_{\text{min}} = \lim_{\text{SNR}\to 0} \frac{\text{SNR}}{\bar{C}(\text{SNR})}. \tag{2.2}$$

  Here, $\bar{C}(\text{SNR})$ is the spectral efficiency as a function of SNR. In (2.1), S_0 is the wideband slope of the spectral efficiency, defined as the increase of bits per second per hertz per 3 dB (bps/Hz/(3 dB)), and can be expressed by [38]

  $$S_0 = \frac{2\dot{\bar{C}}(0)}{-\ddot{\bar{C}}(0)} \tag{2.3}$$

where $\dot{\bar{C}}(0)$ and $\ddot{\bar{C}}(0)$ are the first and second derivative, respectively, when SNR=0.

- **High SNR regime**

  In the high SNR regime, the required energy efficiency to obtain a specific spectral efficiency can be expressed by [39]

$$\left(\frac{E_b}{N_0}\right)\Bigg|_{\text{dB}} \approx \frac{C}{S_\infty} 10\log_{10}2 - 10\log(C) + \left(\frac{E_b}{N_0}\right)_{\text{penalty}} 10\log_{10}2 \tag{2.4}$$

  where S_∞ is the slope of the spectral efficiency in the high SNR regime in bps/Hz/(3 dB) and is given by [39]

$$S_\infty = \lim_{\text{SNR}\to\infty} \text{SNR}\dot{\bar{C}}(\text{SNR}). \tag{2.5}$$

  In (2.4), $\left(\frac{E_b}{N_0}\right)_{\text{penalty}}$ is the horizontal penalty which represents the power offset, in dB units, with respect to a reference channel having the same high SNR regime slope but with an unfaded channel (i.e., AWGN) and it is calculated by [39]

$$\left(\frac{E_b}{N_0}\right)_{\text{penalty}} = \lim_{\text{SNR}\to\infty} \left(\log_2(\text{SNR}) - \frac{\bar{C}(\text{SNR})}{S_\infty}\right). \tag{2.6}$$

Both tools were used to analyse the energy efficiency as a function to spectral efficiency for different network scenarios [40–46]. Hence, using the low-SNR tool, the interplay of the spectral and energy efficiency was studied for SU-MIMO channels [40], SU relay channels [41,42], and MU relay channels [44]. The authors of [45, 46] used the high-SNR tool to analyse the energy efficiency of MIMO channels. These tools consider only the transmit power and do not take into the account the power consumption of other parts such as circuit and processing power consumption. However, these tools are still useful for scenarios where the transmit power dominates other parts of power consumption contributors. Practically, for a communication system where the transmitter and receiver has a single antenna.

2.5 Power Consumption Model

Tools (2.1) and (2.4) can be applied for a system with a single transmit antenna as radiated power is the dominate part of power consumption. However, when it comes to the MIMO system, the problem becomes more complicated. In this case, obtaining an accurate PCM is of chief importance to finding a realistic gaudiness for studying the energy consumption especially for MIMO systems. Most of the previous work on modeling the PCM system assumed that the total power consumption is the sum of the radiated transmit power and a constant that accounts for the circuit power and cooling [47–49]. This can be applicable for conventional MIMO where the number of transmit antennas and users are relatively small. However, when it comes to the massive MIMO system, this model can be very deceptive since it leads to unbounded energy efficiency when the number of transmit antennas is high [50]. The reason behind this is that PCM does not take into account the power consumption for the MIMO-related digital signalling processing such as beamforming and channel estimation as well as the power consumption by backhauling.

Building on prior works of [47, 50–57], we introduce a modified and more realistic PCM that fits the next-generation MU-MIMO systems. To have a realistic PCM we first need to understand the structure of communications and break down the power consumption in each component of the communication system. Currently, the BS functionality is subdivided between a base band unit (BBU) processing and a remote radio unit (RRU) or sometimes it is known by a remote radio head (RRH) [58]. The BBU is responsible for digital communication through the physical interface, whilst RRU implements RF (RF) front-end processing for wireless communication. Each unit holds a different power requirement. Moreover, in recent years, the cellular communication system has gradually evolved into a distributed architecture system in which the BBU and the RRU of a BS are physically separated from each other. The BBU is connected with many RRUs via fibre, cable , or microwave transmissions. This can indeed increase the power consumption due to the operational use of the transmission links between

BBU and RRU. This kind of power consumption should be included within the PCM. Aside from these components, the power supply and the cooling system also contribute to the total power use.

2.5.1 BBU Power Consumption

The BBU is responsible for digital signal processing, coding and decoding, mapping and de-mapping of physical channels and transport channels, channel estimation, and radio resource management of the entire BS system. It further provides ports for data communication between the RRU and the CN. The BBU is manufactured by the combination of microprocessors, field-programmable gate arrays (FPGAs), and other digital and analogue circuits. The BBU further comprises a clock module and a power module, which are used for providing clock synchronisation and power for the BBU, respectively. Therefore, the silicon technology has a great impact on the total BBU power consumption. The advent of the next-generation cellular networks, however, will increase the computing burden within the BS, making floating-point operations (flops) has an impact on PCM. The power consumption in the BBU depends on the complexity of the unit design and the algorithms to handle the various load levels or/and the number of active end users. Moreover, the complexity of the unit design increases with the number of active antennas which also means additional power consumption. The power consumption of the BBU happens at three different terms.

- Central processing unit (CPU) power consumption: The total CPU power consumption P_{cpu} is the amount of power consumed by different digital unit processes used for the radio-related functions. This includes the consumption on the radio channel estimation P_{CE}, generating MIMO precoding vectors P_{PV}, and the radio resource management P_{RRM}. Each term depends linearly or non-linearly on the main MIMO system parameters, i.e., the number of transmit antenna N_t, the number of user K, and the number of the receive antenna N_r. Modeling P_{CE}, P_{PC}, and P_{RRM} can be achieved

by calculating the number of flops, e.g. multiplication, addition, division, and so on, that a CPU performs in order to determine each of its related functions [59]. P_{CE} should be included in the PCM because the BS is required to estimate complete channel vectors by using a mutual orthogonal pilot sequence that transmits from each user in a dedicated uplink training phase and so to generate the optimum precoding matrix for the downlink transmission, assuming time division duplex (TTD). Using minimum mean square error (MMSE) detection, the estimation is then performed by multiplying the uplink observation with each user-specific pilot sequence. This operation requires $2\tau K^2 N_t$ flops [50]. Therefore, the consumed power P_{CE} for estimating a pilot sequence of length of τK is equal to $\frac{2\tau K^2 N_t}{\eta_{\text{cpu}}}$ Watt, with τ is relative pilot length. η_{cpu} refers to a computational performance per watt (or flops per watt) [53] and is equivalent to the number of flops or instructions executed per joule of energy over a period of time. Once completed channel coefficients are estimated, the BS can generate a precoding matrix that can be used on the downlink transmission. For such an operation, the cost in Watt depends on which scheme and how many flops that are required to generate a complete precoding matrix. The impact of P_{RRM} on the total processing power consumption depends on the complexity of the used algorithm for allocating the radio resources among users. This includes an additional cost of selecting the optimal (or sub-optimal) users in the case of $K > N_t$.

- Transceiver board Power consumption: This type of power consumption $P_{\text{BBU-TXB}}$ accounts for the power required by all circuit components in one transmit transceiver board (TXB) chain in the BBU such as memory, cyclic redundancy check (CRC) checker, modulator, fast Fourier transform (FFT) and inverse FFT (IFFT), encoder/decoder circuits, etc. it should also include the circuit power of the receiver transceiver chain to perform channel estimation. $P_{\text{BBU-TXB}}$ can be modelled by

$$P_{\text{BBU-TXB}} = N_t P_{\text{BBU-crt}} + P_{\text{sta}} \tag{2.7}$$

where $P_{\text{BBU-crt}}$ is the circuit power that accumulates all mentioned power components in the BBU, while P_{sta} is the static power consumption. Both $P_{\text{BBU-crt}}$ and P_{sta} are load-independent factors and can be reasonably assumed to be constants.

- Bit processing power consumption: This kind of power consumption P_{BP} includes all consumption that are required for procession K sequence of information symbols, such as bit interleaving, scrambling, and modulation. This type of consumption is proportionally linear to the achievable rate R and it can be modelled by

$$P_{BP} = \frac{BR}{\eta_{\text{load}}} \tag{2.8}$$

where η_{load} is the consumption factor in $\frac{\text{bit}}{\text{second}}$ per watt and B is the system bandwidth.

2.5.2 RRU Power Consumption

The RRU contains the BS RF circuitry plus analog-to-digital/digital-to-analog converters and up/down converters all for converting data between a RF transceiver module and the BBU switching module, and vice versa. It uses a backhaul interface (e.g., optical) to connect to the BBU processing part and a cellular air interface to communicate with the user. The main function of RRU is power amplification of baseband signal before forwarding it to the antenna interface (AI). The power consumption resides in the following components.

- Power amplifier consumption: The most distinct component of RRU is the power amplifier (PA). The PA is characterised by its efficiency $\eta_{\text{PA}}(\eta_{\text{PA}} < 1)$ which refers to the ratio of the radiated power P_r to the total direct DC input power P_{PA}. The power consumption of a PA is proportionate to the radiated power of BS, i.e.,

$$P_{\text{PA}} = \frac{P_r}{\eta_{\text{PA}}}. \tag{2.9}$$

Typically, the most efficient PA should have an operating point that is close to the maximum output power. Due to non-linearity effects most of PAs yet operate in a more linear region, thus 7 to 12 dB below the saturation [54]. Consequently, this operating back-off gives rise to poor power efficiency, and that is translated to a high power consumption.

- RF transceiver board power consumption: Similar to the BBU, the RRU consists of multiple TXBs. A TXB comprise AI, digital-to-analog converters, filters, BBU interfere, PA interface, DC-DC power supply, frequency AC-DC power supply unit, mixer, time synchronisation interface, low noise amplifier (LNA), and an optional backhaul interface in the case where the BBU is remotely connected to the RRU. The power consumption of AI includes the amount of losses that comes from feeder, filters, duplexers, and type and length of cable. There are four important factors influence the power consumption of LNA unit, i.e., power gain η_{LNA}, operation bandwidth B, noise figure N_F, and figure-of-merit FOM. Hence, the power consumption of LNA P_{LNA} is modelled by [51]

$$P_{\text{LNA}} = \frac{\eta_{\text{LNA}} B N_0}{(N_F - 1)\text{FOM}}. \tag{2.10}$$

  For typical LNA, the FOM is in the range of 10^{-7} to 10^{-9} [51]. The total power consumption $P_{\text{RRU-TXB}}$ can be modelled by

$$P_{\text{RRU-TXB}} = M\left(P_{\text{RRU}-\text{crt}} + P_{\text{LNA}}\right) \tag{2.11}$$

where $P_{\text{RRU}-\text{crt}}$ is the power consumption on a circuit board which includes all aforementioned components except the LNA part.

2.5.3 Backhauling Power Consumption

The backhaul links are expected to play a significant role on the total network power consumption for the next-generation cellular systems in order to carry the continued growing traffic. The increase that is required backhauling power is

related to the increase in required information exchange between the two units. This means that the fraction of the power consumption is also dependent on achievable rate at the radio side. In this sense, backhaul powers must be taken into account in a realistic PCM especially when BBU is physically separated from RRU. However, power consumption levels depend on the used backhaul solution. In general, many operators are now aligning their backhauling strategy toward optical technology because of its high capacity and low power consumption [55, 56]. The power consumption in the backhaul link P_{BH} can be modelled using the following formula

$$P_{\mathrm{BH}} = P_{\mathrm{fix}} + \frac{R}{\eta_{\mathrm{BH}}} \tag{2.12}$$

where P_{fix} is the fixed part of the power consumed independently from the traffic load, and η_{BH} is ratio between the maximum capacity offered by a backhaul link and maximum power available for the backhaul transmission.

2.5.4 Cooling Power Consumption

In order to keep the BSs working properly, they must be kept cool. Therefore, the power consumption should not just include the requisite power for BS operation, but in addition the power consumption of cooling unit. For example, cooling helps to quickly restore the microprocessors back to ambient temperature which consequently contributes to a greater extent of power consumption. Hence, it was highlighted in [57] that the cooling power consumption that is demanded for the data processing is equivalent to almost half the total power use of data processing. The type of cooling varies, depending on the local weather and environment conditions. This includes external air conditioning, FAN installed inside the cabinet, or internal small FANs. Air conditioning needs a lot of power, but it is essential in hot countries at summer time. Alternative direct air cooling uses small fans to conduct heat away from the BS cabinet, and this requires 90% less energy than traditional air conditioning. The power consumption of external air condition is a constant. However, the power consumption of internal fan P_{fan}

depends on its speed S and, indirectly, on load level. Thus, P_{fan} can be modelled by $P_{\text{fan}} = \kappa S^3$ [57], where κ is a constant. The total power consumption for cooling can be modelled by

$$P_{\text{cooling}} = \bar{T} P_{\text{fan}} + P_{\text{external}} \tag{2.13}$$

where $\bar{T}$ is the number of internal FAN and P_{external} represents the power consumption at the external air conditioning.

2.6 Chapter Summary

In this chapter, we have first highlighted the importance of mobile wireless communication and how cellular communication has evolved over time. We have also introduced a potential key technologies that can be deployed in 5G mobile cellular system to satisfy the expected traffic growth. The 5G term should be an integration of radio frequency (RF) wireless communications with optical wireless communication. The design of the 5G mobile cellular system should take into the account the energy-efficiency as one main targets to deliver greener communication. Secondly, two analytical methods to analyse the spectral-energy efficiency tradeoff for wireless systems were presented. The tools consider only the transmit power as an input, thus they suits only the scenarios where transmitter has few number of transmit antennas. Thirdly, the concept of fractional programming that fits into the class of energy efficiency optimisation programs has been presented and an approach for computing the solution has been highlighted. Finally, in this chapter, more realistic PCM that suits the structure of the next-generation cellular systems has been introduced. The PCM includes the power consumption of BBU, RRU, backhauling, and cooling components.

Chapter 3

Energy-Efficient Radio Resource Allocation for Multi-User OFDMA Systems

3.1 Introduction

A rapid evolution has occurred in the cellular system in the last few years. The cellular industries have started to pay more attention to high speed mobile broadband systems which has led to the 4G wireless systems. Meanwhile the 3GPP has been working toward the development and maintenance of the global cellular system, including evolved radio access technologies in order to achieve the requirements of the 4G wireless system, resulting in the introduction of LTE-Advanced [10]. Orthogonal frequency division multiplexing (OFDM) has become the multicarrier transmission technique of choice for the 4G wireless systems. It is also expected that OFDM can play a role in the 5G wireless system [60]. OFDM is one of the promising technologies that can provide improved spectral efficiency and the flexibility of the allocation of radio resources among users. Instead of transmitting symbols sequentially over a single wideband channel for the 3G wireless systems, OFDM divides the wideband broadband channel into a number of

narrowband subchannels, each having a bandwidth much smaller than the coherence bandwidth of the channel [61]. Hence, each sub-channel appears to be as a flat fading channel. In an MU-OFDM system, there is a need for a multiple access scheme to allocate a subset of subcarriers with certain power to different users. This is because particular subcarriers at different instants of time may appear differently among the users due to the diversity nature of wireless channels. This offers an opportunity to allocate certain subcarriers to users who can utilise them in the most beneficial way at that particular moment. This mechanism is the principle of OFDMA scheme. Using OFDMA, a subset of subcarriers is assigned exclusively to a user at any given time. Moreover, other resources such as power and modulation format (bit loading) also need to be allocated to each assigned subcarrier. There will be more freedom in allocating subcarriers, transmit power and modulation, when there are greater number of users, i.e., MU diversity. Therefore, it is essential to form a resource allocation scheme that adapts to users' varying channel conditions on a temporal basis. In the last few decades, there were extensive researches which had sought to study the idea of adaptively allocating radio resources to users in OFDMA systems in order to optimise a certain metric of interest such as achievable sum rate, transmission power, or any utility, subject to certain constraints [62, 63].

3.1.1 Radio Resource Allocation for OFDMA Systems

RRA in OFDMA systems has been extensively studied. Radio resources, which in this context we mean orthogonal subcarriers and power, are allocated to different users to increase the performance of the system by the use of different scheduling algorithms. It has been shown that the cell capacity improves as the number of simultaneous users increases in a cell due to MU diversity gain. Previous research on resource allocation in OFDMA-based systems are mainly focused on two approaches, either rate adaptive (RA) or margin adaptive (MA). The objective of the first approach is to maximise the total throughput with the constraints on the total transmit power as well as the users' data rates [64–66]. Whereas the

objective of the second approach minimise the total transmit power with the constraint on users' data rates [67–69]. It has been shown that the RRA problems in OFDMA systems are integer linear programming (ILP) in nature, thereby allowing the optimal subcarrier and bit allocation to be achieved [70, 71]. Converting the non-linear optimisation into the integer programming is performed by inserting a new indicator variable that refers to the subcarrier assignment variable and the number of bits loaded into each subcarrier. The number of loaded bits is taken from a particular set. The ILP optimisation problem can be solved by any standard package, for example, the branch and bound (BaB) method [71].
Nevertheless, neither RA nor MA provides an energy-efficient solution because, for both approaches, the overall performance of the algorithms are evaluated with the aim of enhancing spectral efficiency and user fairness. Recently, much effort was spent in developing energy-efficient resource allocation solutions for link-level OFDM systems [72–75]. The energy-efficient link level performance was studied for flat fading [72] and frequency-selective fading channels [73, 74]. The authors of [75] developed an OFDM water-filling power allocation scheme that maximises energy efficiency. Furthermore, some works considered a muti-user scenario [76, 77]. In [76], an energy-efficient power allocation scheme was proposed for interference-limited OFDMA systems. An algorithm was further developed in [77] that reduces the complexity with the iterative search technique. None of the aforementioned works considered the use of the ILP approach to solve the RRA problems with energy-efficiency objectives. The ILP approach, nevertheless, can not be applied directly to the energy-efficient optimisation problem because the optimisation problem has a fractional objective function. Therefore, in this chapter, we reformulate the energy-efficient optimisation problem in order to attain a solution by ILP approach and find the optimum solution. The contribution of this chapter can be summarised as follows.

- We present a mathematical framework based on fractional programming for energy efficiency optimised algorithm design. We represent the objective function of the optimisation problem by associated parametric single

function in order to simplify the problem.

- We propose a sub-channel and bit allocation algorithm to minimise the energy per transmitted bit for MU-OFDMA system. The algorithm ensure that the minimum energy efficiency is achieved for a given spectral efficiency and user's QoS, providing that this is dose not break the total power constraint.

- We analyse the spectral-energy efficiency trade-off under the constraint of maintaining the fairness among users and total power.

The rest of this chapter is organised as follows. Section 3.2 describes the system model adopted in this work. Section 3.3 formulates the resource allocation optimisation problem and introduce the energy-efficient iterative algorithm. Section 3.4 presents simulation results with detailed analysis. Finally, Section 3.5 summarise the chapter.

3.2 System Model

A downlink OFDMA system is considered, as shown in Fig. 3.1, where a single BS is transmitting data towards a number of users, utilising a number of orthogonal sub-channels. Each sub-channel comprises of a group of subcarriers which are assumed to have correlated channel gains in frequency domain. The number of users and sub-channels are denoted by $k = 1, ..., K$ and $s = 1, ..., S$, respectively. The BS and users are all assumed to be equipped with a single antenna. It has been assumed that each sub-channel which belongs to a particular user is under flat fading and corrupted by AWGN. The perfect channel state information (CSI) is sent to the BS over a feedback channel from each user. The CSI that represents a sub-channel is assumed to be equal to the linear average of all subcarriers' gain within that sub-channel. Based on this information, the BS allocates a set of sub-channels to each user and decides on the number of bits in each sub-channel. It is assumed that each sub-channel is exclusively allocated to one user, i.e., sharing

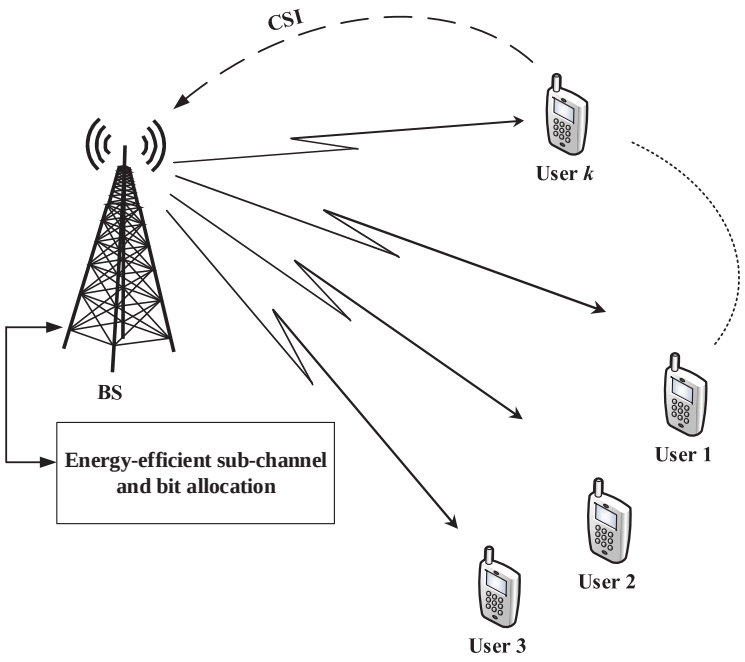

Figure 3.1: System model of an OFDMA-based system where BS serves multiple users.

a sub-channel by two or more users is not allowed at any given time.

Adaptive modulation schemes with binary phase-shift keying (BPSK), quadrature phase-shift keying (QPSK), 16 quadrature amplitude modulation (16-QAM), and 64-QAM are adopted to achieve a target bit error rate (BER). The information regarding subcarrier assignment and bit loading is sent to users over a control channel. The transmit power for user k on sub-channel s is decided upon according to the target BER and the received power level. In other words, the transmit power, normalised to background noise, can be expressed as

$$P_{k,s}(BER, c) = \frac{f(BER, c_{k,s})}{|h_{k,s}|^2} \tag{3.1}$$

where $h_{k,s}$ is the complex-valued channel gain that the uth user experiences on the sub-channel s. $f(c_{k,s})$ [1] is the required received power level that depends on the modulation scheme, and $c_{k,s}$ is defined as the number of loaded bits on the sth sub-channel sent to user k over one OFDM symbol. $c_{k,s}$ takes a value from the an integer non-negative vector $D = \bar{S} \times \{0, b_1 ..., b_v\} \in \mathbb{Z}^{V \times 1}$. Here, $\bar{S}$ is

[1] For notational convenience, the BER index will be omitted in $f(.)$ and $P_{k,s}(.)$ through the rest of the chapter since the value of BER is fixed in this work.

the number of subcarriers within each sub-channel, V is the number of possible modulation schemes, and b_v is the number of bits that equivalents to modulation order. Indeed, V depends on the maximum supported modulation scheme that the BS can adapt. For BPSK and QPSK, $f(c_{k,s})$ equals to $\frac{\bar{S}}{2}(Q^{-1}(\mathtt{BER}))^2$ and $\bar{S}(Q^{-1}(\mathtt{BER}))^2$, respectively [78], where Q^{-1} is the inverse Q-function. For the higher modulation schemes, i.e., 16-QAM and 64-QAM, the required received power can be calculated by [78]

$$f(c_{k,s}) = \frac{\bar{S} \times 2^b}{3} \left(Q^{-1} \left(\frac{b \times \mathtt{BER}}{4} \right) \right)^2. \tag{3.2}$$

3.3 Energy-Efficient Resource Allocation

In this section, the problem of energy-efficient subcarrier and bit allocation, subject to some constraints, will be formulated. Hereafter, an iterative approach will be adopted to achieve the optimum solution.

3.3.1 Problem Formulation

We define the EE as the energy required to transmit information data reliably within one OFDM symbol, measured in Joules/bit/Hz and SE is the spectral efficiency, i.e., the total transmit bits in one OFDM symbol measured in bits/OFDM symbol (or equivalently bps/Hz). Let's define $\bar{I}_{k,s}$ as a binary indicator variable which is equal to one, if the sub-channel s is allocated to user k, or zero if other-

wise. The energy-efficient resource allocation problem can be expressed as

$$(\mathcal{P}) \quad \min_{c_{k,s} \in D} EE = \frac{P_{\mathsf{C}} + \frac{\sum_{k=1}^{K}\sum_{s=1}^{S} P_{k,s}\bar{I}_{k,s}}{\eta}}{\sum_{k=1}^{K}\sum_{s=1}^{S} c_{k,s}\bar{I}_{k,s}} \tag{3.3}$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\sum_{s=1}^{S} c_{k,s} d_{k,s} \geq SE \tag{3.4}$$

$$\sum_{s=1}^{S} c_{k,s}\bar{I}_{k,s} \geq R_k^{\min} \tag{3.5}$$

$$\sum_{k=1}^{K}\sum_{s=1}^{S} P_{k,s}\bar{I}_{k,s} \leq P_{\mathrm{T}} \tag{3.6}$$

$$\sum_{k=1}^{K} \bar{I}_{k,s} = 1 \tag{3.7}$$

$$\bar{I}_{k,s} \in \{0,1\} \tag{3.8}$$

where P_{C} is a circuit power dissipation that is assumed to be constant in this work while η represents the efficiency of a power amplifier. The objective of problem $(\mathcal{P})$ is to minimise the transmit energy per bit for a given spectral efficiency whilst satisfying each user's QoS requirement as well. The minimum required spectral efficiency in this problem is an input under the constraint (3.4) that the solver must satisfy when allocating the radio resources. In (3.3), $D = \bar{S} \times \{0, b_1 ..., b_v\}$ represents the set of all possible values of modulation schemes that can be used by a subcarrier $c_{k,s}$. Thus, adaptive modulation is performed at in each sub-channel. Constraint (3.5) ensures each user obtains the minimum rate to satisfy the QoS requirement. (3.6) is the constraint for the total transmission power at BS, with $P_{k,s}$ being given by (3.1) and P_{T} being the maximal total transmission power. Constraint (3.7) means that each sub-channel is allocated to one user only.

Let's define a variable binary vector $\mathbf{X}$ as $[\mathbf{x}_1\ \mathbf{x}_2\ \ldots\quad \mathbf{x}_S]^T \in \{1,0\}^{SKV\times 1}$, where $\mathbf{x}_s = [\mathbf{x}_1^s\ \mathbf{x}_2^s\ \ldots\quad \mathbf{x}_K^s]^T \in \{1,0\}^{KV\times 1}$ is the allocation vector that allocates a sub-channel s to a user k. Now, $\mathbf{x}_k^s = [x_1^{s,k}\ x_2^{s,k}\ \ldots\quad x_V^{s,k}]^T \in \{1,0\}^{V\times 1}$ is the modulation vector that decides which modulation index will be in sub-channel s that is assigned to user k. Also, let's define a power loading vector $\mathbf{P}$ as $[\mathbf{p}_1\ \mathbf{p}_2 \ldots\quad \mathbf{p}_S] \in \mathbb{R}^{1\times SKV}$, where $\mathbf{p}_s = [\mathbf{p}_1^s\ \mathbf{p}_2^s\ \ldots\quad \mathbf{p}_K^s] \in \mathbb{R}^{1\times KV}$. $\mathbf{p}_k^s = [\frac{f(1)}{h_{k,s}^2}\ \frac{f(2)}{h_{k,s}^2}\ \ldots$

$\frac{f(V)}{h_{k,s}^2}] \in \mathbb{R}^{1\times V}$ represents the possible transmit power. Likewise, a vector $\mathbf{C}$ which denotes the transmit bits in each sub-channel is defined as $\mathbf{C}$= $[\mathbf{c}_1\ \mathbf{c}_2\ \ldots\quad \mathbf{c}_S]$ $\in \mathbb{Z}^{1\times SKV}$, where, $\mathbf{c}_s = [\mathbf{D}_1\ \mathbf{D}_2\ \ldots\quad \mathbf{D}_K] \in \mathbb{Z}^{1\times KV}$. Furthermore, we relax the constraint (3.7) to make the summation equal to either 0 or 1 because. This is because for some channel realisations when they are in a deep fading situation, it is more energy-efficient not to assign the sub-channel s to any particular user. By using these vector definitions, the optimisation problem (3.3)–(3.8) can be reformulated to

$$\min_x EE = \frac{(P_{\mathrm{c}} + \frac{\mathbf{PX}}{\eta})}{\mathbf{CX}} \tag{3.9}$$

$$\text{s.t.} \quad -\mathbf{CX} \leq -SE \tag{3.10}$$

$$-\mathbf{A}_u \cdot \mathbf{X} \leq -\mathbf{R}^{\min} \tag{3.11}$$

$$\mathbf{PX} \leq P_{\mathrm{T}} \tag{3.12}$$

$$\mathbf{A}_e \cdot \mathbf{X} \leq \mathbf{1}^T \tag{3.13}$$

$$\mathbf{X} \in \Omega \subseteq \{0,1\} \tag{3.14}$$

where $\mathbf{A}_u = [\mathbf{a}^1\mathbf{a}^2 \ldots\quad \mathbf{a}^K]^T \in \mathbb{Z}^{K\times SKV}, \mathbf{a}^k = [\mathbf{a}_1^k\mathbf{a}_2^k \ldots\quad \mathbf{a}_S^k] \in \mathbb{Z}^{1\times SKV}, \mathbf{a}_s^k =$ $[\mathbf{0}_{s,1}^k\mathbf{a}_{s,k}^k \ldots\quad \mathbf{0}_{s,K}^k] \in \mathbb{Z}^{1\times KV}$, and $\mathbf{a}_{s,k}^k = \mathbf{D} \in \mathbb{Z}^{1\times V}$. Note that

$$\mathbf{A}_e = \begin{bmatrix} \mathbf{1}_{1,1} & \mathbf{0}_{1,2} & \ldots & \mathbf{0}_{1,S} \\ \mathbf{0}_{2,1} & \mathbf{1}_{2,2} & \ldots & \mathbf{0}_{2,S} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{S,1} & \mathbf{0}_{S,2} & \ldots & \mathbf{1}_{S,S} \end{bmatrix}$$

$\in \{1,0\}^{S\times SKV}$ is to ensure that sub-channel k is either allocated exclusively to one user or excluded from the allocation. Here $\mathbf{1}_{s,s} \in \{1\}^{1\times KV}$ and $\mathbf{0}_{s,s} \in \{0\}^{1\times KV}$. $\mathbf{R}^{\min} = [R_1^{\min}R_2^{\min} \ldots\quad R_K^{\min}]^T$ is the minimum rate vector to ensure each user's QoS and Ω refers to the feasible region for the described problem (3.9)–(3.14). The objective function (3.9) is the ratio of two linear integer functions to be minimised. This type of optimisation problem is commonly known as linear integer fractional programming [79]. The optimisation problems of this kind can be more tractable by adopting Dinkledbach's parametric approach [80] as we will see in the following sub-section.

3.3.2 Problem Transformation and Parametric Algorithm

The energy efficiency optimisation problem is formulated by minimising the proportion between the achievable rate to the entire power consumption (or maximising the proportion between the consumed power to achievable rate). The energy efficiency problems leads to a fractional programming problem [81]. The objective function in a fractional program is a ratio of two functions that are in general nonlinear. A general nonlinear fractional program has the following form

$$(Q) \qquad \min_{x \in \mathbf{S}} \frac{f(x)}{g(x)} \tag{3.15}$$

$$\text{s.t.} \qquad h(x) \geq 0; \tag{3.16}$$

where $f(x), g(x)$, and $h(x)$ are real-valued functions, $\mathbf{S} \subseteq \mathbb{R}$ is a convex set. Problem Q usually comes with some constraints such as constraints on the transmit power and required spectral efficiency. Additional sum rate constraints can represent QoS for a traffic or user. If $f(x)$ is convex and $g(x)$ is concave, the problem Q is called concave-convex fractional programing, providing that $\mathbf{S}$ is a convex set [82]. However, such a problem can not be solved by convex programing algorithm. One way to solve such problem is by representing the optimisation problem by a parametric program [83].

Theorem 1. *Solving problem* $\min \frac{f(x)}{g(x)}$ *is equivalent to obtaining* q^* *that makes* $F(q) = \min\{f(x) - qg(x)\} = 0$*, and thus* $q^* = \frac{f(x^*)}{g(x^*)}$ *is the optimum solution.*

Proof: See Appendix A □

Therefore, according **Theorem** 1 problem Q can be related to

$$(\bar{Q}) \qquad \min_{x \in \mathbf{S}} f(x) - qg(x) \tag{3.17}$$

$$\text{s.t.} \qquad h(x) \geq 0; \tag{3.18}$$

One can see x^* is optimal for problem Q if and only if it is optimal for $\bar{Q}$ with q^* being the only zero of $F(q) = \max\{f(x) - qg(x), x \in \mathbf{S}\}$. It was shown in [80] that $F(q)$ is convex, continuous and strictly decreasing in q. There are various iterative

algorithms available for ascertaining the root of $F(q)$. Dinkelbach in [80] proposed an iterative procedure that solves the equivalent parametric program i.e., $\bar{Q}$, with a sequence that converges to the optimal point with a fast convergence rate [84]. The algorithm terminates once the objective value of the problem $\bar{Q}$ becomes zero. The procedure of Dinkelbach is illustrated in **Algorithm 1**. It is worth mentioning that Dinkelbach's method is still valid for fractional problems with objectives being maximised.

Algorithm 1 Dinkelbach algorithm

Compute $q_1 = \frac{f(x_0)}{g(x_0)}, x_0 \in \mathbf{S}, i = 1$.

Find x_i by solving problem $\bar{Q}$

if $F(q_i) < \delta$ **then**

$q^* \Leftarrow q_i$, $x^* \Leftarrow x_i$, **Stop**.

end if

$q^* = \frac{f(x_i)}{g(x_i)}$, $i = i + 1$, **go to Step 2.**

3.3.3 Energy-Efficient Algorithm

Form (3.9) one can see that $f(x) \equiv (P_{\mathtt{C}} + \frac{\mathbf{PX}}{\eta})$ is convex because it is an affine function whilst $g(x) \equiv \mathbf{CX}$ is concave [70]. Applying the Dinkelbach's approach to the energy-efficient optimisation problem, EE in this case will be equivalent to q^*. This inspires us to propose an energy-efficient allocation algorithm which is summarised in Algorithm 3. Step 2 represents the optimisation problem which is convex and equivalent to $\bar{Q}$. Hence, any binary LP-based solver, e.g., BaB method, can be used in Step 2 to solve the optimisation problem $(\bar{Q})$ for a given q. The complexity of the algorithm depends completely on the optimisation problem solver in Step 2 and how many iterations it takes to converge to q^*.

Algorithm 2 Energy-efficient subcarrier and bit allocation

Require: $q^* \Leftarrow q^{(0)}, j \Leftarrow 0$.

1: **while** $q^* \neq \infty$ **do**
2: **Find** $\mathbf{X}^*$ by solving the problem $\min\{(P_{\mathsf{C}} + \frac{\mathbf{PX}}{\eta}) - q^* \times \mathbf{CX}\}$ problem
3: $Z = (P_{\mathsf{C}} + \frac{\mathbf{PX}^*}{\eta}) - q^* \times \mathbf{CX}^*$, $j = j + 1$
4: **if** $|Z| < \delta$ **then**
5: **Return** q^* & $\mathbf{X}^*$.
6: **else if** $j >$ MAX **then**
7: $\mathbf{X}^* \Leftarrow \mathbf{0}^T$, $q^* \Leftarrow \infty$ {consider this solution infeasible}
8: **else**
9: $q^* = \frac{(P_{\mathsf{C}} + \frac{\mathbf{PX}^*}{\eta})}{\mathbf{CX}^*}$
10: **end if**
11: **end while**
12: **Allocate** the resources according to $\mathbf{X}^*$.
13: $EE = q^*$.

3.4 Simulation Results and Discussions

This section evaluates the proposed energy-efficient algorithm and presents the simulation results. The simulator has a single cell with a radius of 1 km. LTE-Advanced frame structure, as standardised by 3GPP Release 8, is adopted with blocks of 12 contiguous subcarriers in frequency domain and 14 OFDM symbols in the time domain [85]. Hence, the data transmissions are scheduled every 1 ms using 10MHz bandwidth. The simulation results presented in this paper are averaged over 1000 different channel conditions. The bandwidth of each sub-channel is equal to 180 kHz. The channels are subject to path-loss, log-normal shadowing with standard deviation of 8 dB, and a frequency-selective fading with 6 taps. The path loss model is calculated by [86],

$$L[\text{dB}] = 136 + 37.6 \times \log_{10}(d) \tag{3.19}$$

where d is the distance between a user and the BS, in km. The BER is fixed to 10^{-4} in all examples. The circuit power P_{C} is 23 Watt and $\eta = 0.38$.

Fig. 3.2 illustrate how the energy consumption increase with the increases in the spectral efficiency, against different user's QoS requirements. The number of sub-channels is 15 and shared by 4 users. The users are distributed uniformly at the cell edge. The total transmit power is equal to 43dBm. As it can be shown, for a given required QoS, the increase in the spectral efficiency has no impact on the energy consumption until reaching a specific spectral efficiency point of which the energy efficiency starts increasing. This is because, below this point, the energy consumption is stable to satisfy the QoS requirement of each user and that does not break (3.4). However, beyond this certain spectral efficiency point, the QoS is more than adequately compensated by the spectral efficiency, and we can see an increase to energy consumption. This is also the reason why the gaps between energy consumption tend to be small at high spectral efficiency for different QoS requirements because the constraint on the spectral efficiency is dominant and can satisfy the user's QoS requirement.

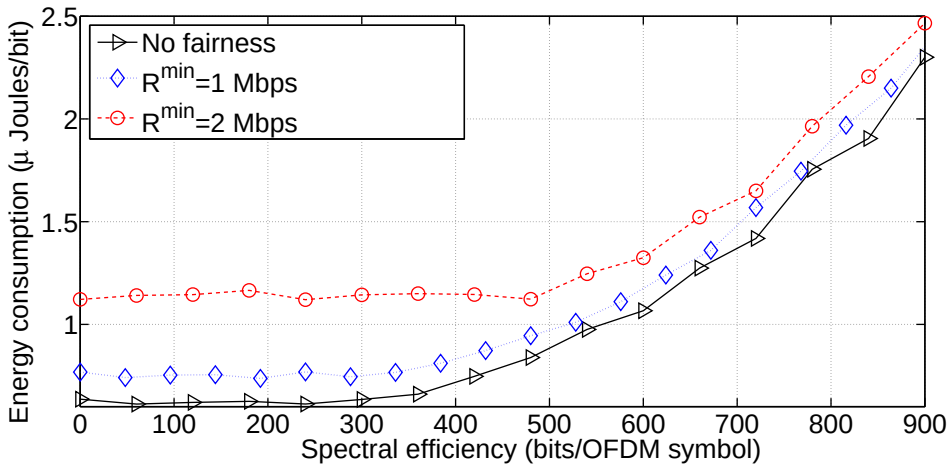

Figure 3.2: System spectral-energy efficiency trade-off against different user's QoS (S=15, K=4, and P_{T}=43dBm).

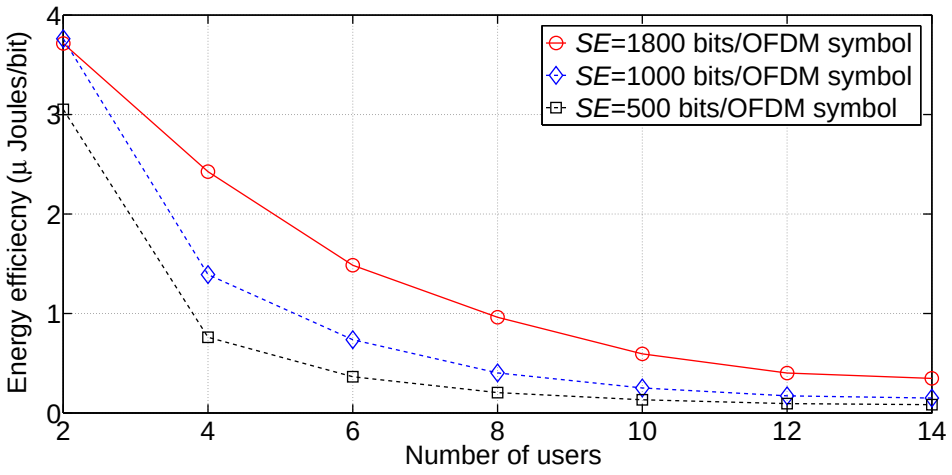

Figure 3.3: Energy efficiency vs. the number of users (S=32, $R^{\min}$=1 Mbps and P_{T}=43dBm).

Fig. 3.3 depicts the energy efficiency as a function of the number of active users in the cell. The number of sub-channels is assumed to be 32 and $R^{\min}$ is equal 1 Mbps. Here, the users are being distributed uniformly in each snapshot. Also, it has been assumed that the BS has an advanced CPU with high computational performance per watt. Thus, the CPU power consumption has a minor impact on the total power consumption for a moderate number of users. The power consumption of channel feedback overhead has not been considered in this analysis. As it is clearly shown, increasing the number of users results in decreasing the consumed energy per bit due to multi user diversity gain. It can also be noticed that with a greater number of users, a smaller amount of energy is required to double the spectral efficiency.

Fig. 3.4 compares the energy efficiency as a function of the distance between the BS and users for two spectral efficiency and different user's QoS requirements. The number of users and the number of sub-channels are assumed to be 5 and 20 respectively. The energy efficiency are approximately the same for the two given spectral efficiency when the users are near the BS. The difference between energy efficiency for the two spectral efficiency increases when the users are far from the BS full power is transmitted in the case of the highest spectral efficiency.

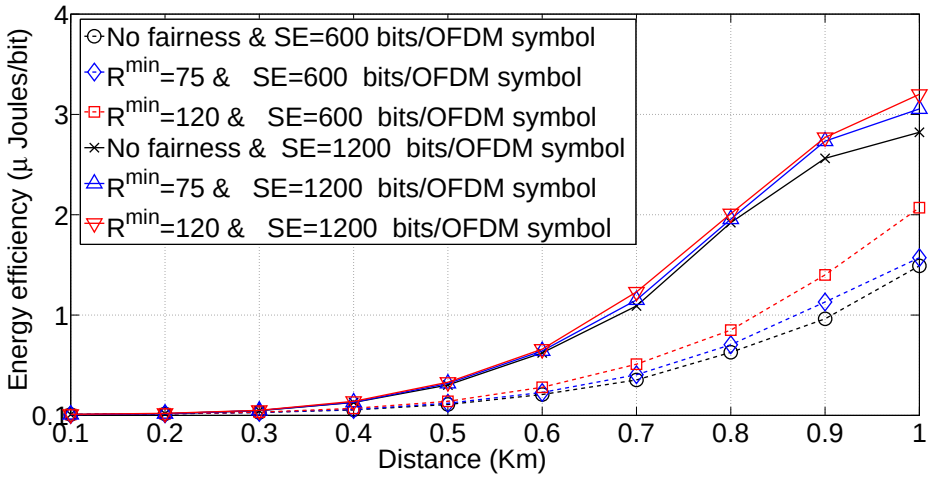

Figure 3.4: Energy efficiency vs. the distance between the BS and users (S=20, K=5, and P_{T}=43dBm).

3.5 Chapter Summary

This chapter has studied the sub-channel and bit allocation scheme to minimise the energy per transmit bit for MU in an OFDMA system. It has explained how the optimisation problem with the energy efficiency objective function can be reformulated to an non-fractional function. It also has proposed an iterative algorithm to find the optimum resource allocation that gives the minimum energy efficiency for given spectral efficiency whilst maintaining the user's QoS and without breaking the total power constraint.

Chapter 4

Spectral-Energy Efficiency Trade-off of Cellular System with Mobile Femtocell Deployment

4.1 Introduction

Heterogeneous network (HetNet) deployments are seen as a promising solution for cellular operators to improve indoor coverage and increase the spectral efficiency with low OPEX [87]. In fact, HetNet is considered crucial for mobile data offloading and will become an essential part of the next-generation cellular systems [88]. One key component in HetNet is femtocell technology. Femtocells are low range, low power mobile BSs that improve the coverage inside a home or office building [89]. They use broadband networks as backhaul to carry the traffic to the operator's data centre. Femtocell technology has proven its excellent capability of grabbing the indoor traffic from the macrocell layer and thus enables the latter to focus its attention toward outdoor traffic. Moreover, femtocell technology is considered as an energy-efficient solution due to its ability to achieve high data rate with low transmitted power due to the short distances toward its users [90]. Public vehicles, e.g., trains and buses, are moving hotspots with many people potentially requesting diverse data services, e.g., web browsing, video streaming,

and gaming. Users inside a moving vehicle may execute multiple handovers at the same time causing a significant increase in signalling load and high connection failures in the network. Furthermore, the vehicle may have high penetration loss through its metallic enclosure, resulting in a poor network connection. Therefore, it is important to minimise the signalling load and drop calls within a fast moving vehicle whilst providing a better internet experience on the move. To this end, adopting femtocell technology inside a vehicle have inspired us to propose a new concept called MFemtocell [91].

4.1.1 Mobile Femtocell

MFemtocell is a mobile small cell that can dynamically change its connection to the operator's CN. The MFemtocell can be deployed on public transport buses, trains, and even private cars. The implementation of MFemtocells can potentially benefit cellular networks. For instance, they can contribute to signalling overhead reduction in the network and improve the system performance [92]. The MFemtocell can also perform handover on behalf of all its associated users. This would reduce the number of handover attempts as the users move between cells in the network, especially in high speed mobility environment. Because of the short communication range with the MFemtocell, the battery life of associated users would increase significantly. Moreover, since the MFemtocells are located inside vehicles with antennas located outside the vehicle, this setup improves the signal quality inside the vehicle by avoiding the penetration loss. In addition to that, a larger number of antennas can be utilised at the MFemtocell with significant diversity/multiplexing gain. Last but not least, the MFemtocell is multi-operator friendly, meaning that the MFemtocell is able to tunnel the traffic of users, which may belong to different operators, through dedicated backhaul.

Various research efforts were made to study and appraise the operation of MFemtocell deployments [93–99]. In [93], the authors studied the performance advantage of using MFemtocell by communicating with the macro BS to improve and extend coverage for users. The potential advantages of using moving cells to

boost performance for UEs in transit vehicles were highlighted in [94]. The authors of [95] proposed the deploying mobile small cells to improve the uplink performance for users within vehicles. The work of [96] investigated the outage performance in the presence of MFemtocells, and discovered that the outage can also be improved in a high speed mobility environment by using MFemtocells. The work of [98] introduced a scheduling algorithm that aims to provide seamless multimedia service within high-speed trains by using small cells onboard. The work of [99, pp. 53–73] investigated an appropriate precoder at the MFemtocell in the vehicle to overcome the degraded performance of the received signal in outdoor wireless links. The proposed precoded helps in extracting the underlying rich multi-path Doppler diversity inherited in this type of double-selective fading link. To utilise the MFemtocell to its full advantage, one must overcome challenges such as finding the most reliable backhaul to carry the traffic between the CN and MFemtocells' users, developing a strategy for sharing the spectrum between the macro BS and MFemtocells, and optimising spectral and energy efficiency. One way to achieve this is to use the standardised cellular air interface as a backhaul link, thus creating a similar situation to the concept of the mobile relay [100]. In this case, an MFemtocell behaves as a femtocell when serving its users and as a regular user when communicating with the macro BS. Furthermore, the traffic of users within an MFetmocell can be treated as an SU traffic when sending to or receiving from the MFemtocell over the air interface. MFemtocell is also able to adopt other backhaul methods for carrying the traffic such as using a different RAT (e.g., Wi-Fi [97] or satellite) on a different spectrum. These options will enable more reliability in case that the serving macro BS is broken down due to failures and/or high congestion. Under these failure situations, a group of MFemtocells within close proximity will be able to create a new network layer and by adopting different backhaul methods the system can maintain the connectivity between the CN and the users. This offers a solution to the challenge that is posed during an emergency situation where the standard networks are jammed.

Regarding the second challenge of how to share the spectrum between the macro

BS and MFemtocell, we can examine the research that have studied femtocell-based cellular systems. In this context, the femtocell can share the same spectrum with macro BS (i.e., non-orthogonal transmission scheme) or utilise a dedicated spectrum (i.e., orthogonal transmission scheme) [101–104]. Using the same spectrum would improve spectrum utilisation but contribute to additional interference for indoor and outdoor users. In an orthogonal transmission scheme, the femtocell uses a dedicated spectrum band that is not used by macro network. This mode eliminates the mutual interference between femtocell and macro BS, i.e., intra-cell interference. However, additional spectrum resources are required and this may have a negative impact on the spectrum utilisation.

To the best of our knowledge, no existing work has investigated the spectral and energy efficiency in MFemtocell networks. Therefore, our main contributions of this Chapter are summarised as follows:

1) The investigation of the spectral-energy efficiency trade-off for MFemtocell networks. Closed form expressions for the relationships between the energy efficiency and the spectral efficiency are derived in low and high SNR regimes for an SU MFemtocell with two different resource partitioning (i.e., orthogonal and non-orthogonal) schemes.

2) We also present a spectral efficiency analysis of MU MFemtocells with OFDMA-based spectrum reuse and opportunistic scheduling schemes.

The rest of this chapter is organised as follows. Section 4.2 describes the MFemtocell system model and explain two resource partitioning schemes. In Section 4.3, the relationship between the energy efficiency and spectral efficiency is analysed for an SU MFemtocell network with two different resource partitioning schemes. Section 4.4 investigates the spectral efficiency of an MU MFemtocell network with MUs and opportunistic scheduling schemes. Finally, Section 4.5 concludes the chapter.

4.2 Mobile Femtocell System Model

Let us consider an MFemtocell-assisted cellular network shown Fig. 4.1 with a single BS, multiple MFemtocells, and multiple users. The MFemtocell set is denoted by $\mathcal{J} = \{1, \cdots, J\}$. The total of K users, given by the set $\mathcal{K} = \{1, \cdots, K\}$, are divided into two categories: direct transmission and access users. Let $\mathcal{N}^{\mathrm{d}} = \{1, \cdots, U^{\mathrm{d}}\}$ denote the set of users that communicate directly with the BS. Furthermore, there is the set $\mathcal{M}_j = \{1, \cdots, U_j^{\mathrm{a}}\}$ of access users which communicate with MFemtocell j $\{j \in \mathcal{J}\}$. Both U_j^{a} and U^{d} can be variables, however, $\sum_{j=1}^{J} U_j^{\mathrm{a}} + U^{\mathrm{d}} = K$. The index of the BS is assumed to be zeros and therefore omitted from the analysis. We assume a relaying-based backhauling in which the MFemtocells are using the same standardised air interface to carry the traffic from the BS. The terms backhaul link, access link, and direct link are used to denote BS-MFemtocell, MFemtocell-user, and BS-user links, respectively.

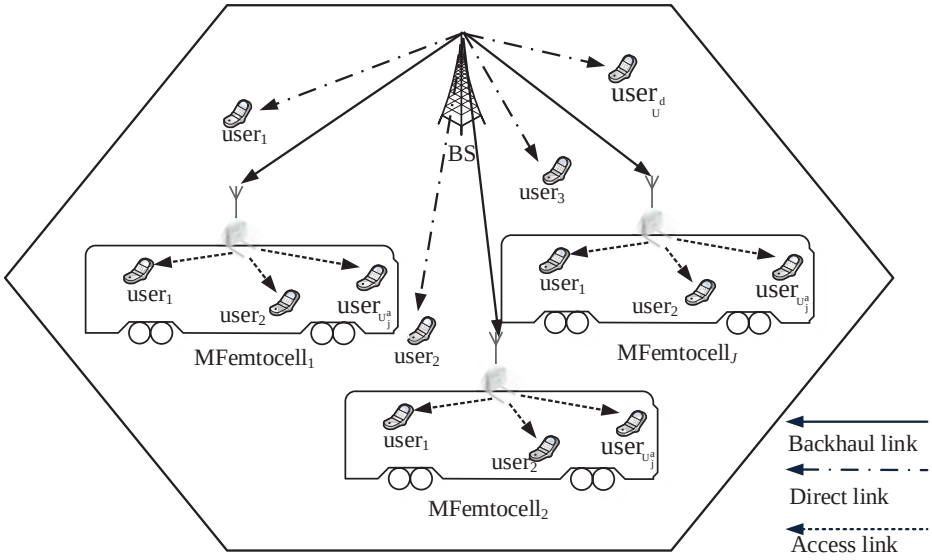

Figure 4.1: System model: a single cell with multiple MFemtocells and users.

We assume that the backhaul, access, and direct links all experience non-line-of-sight (NLoS) Rayleigh block fading channels, which are kept constant within a sub-frame and change independently in the following sub-frame. We also assume that the backhaul link has a gain G over the direct link. This gain can be achieved

by using a highly directional antenna pattern as well as pointing MFemtocell's antenna toward the BS.

4.2.1 Resource Partitioning Schemes

By adopting the MFemtocell, the spectrum has to be allocated (or reused) among different links, i.e., the backhaul, direct, and access links. It is essential to design an efficient resource partitioning policy in the MFemtocell-enhanced system to improve the performance of the whole system. We assume a time division relaying-based backhualing scheme, in which the transmission to end users occurs in two time periods. Each time period contains a specific number of time slots. The BS will transmit traffic to MFemtocells over backhaul links in the first time period. In the second time period, the BS and MFemtocell are communicating simultaneously to the direct transmission and access users, respectively, using either orthogonal or non-orthogonal transmission scheme. The two resource partition policies are explained as follows:

- Orthogonal resource partitioning scheme In this scheme, the radio resources allocated to the backhaul, direct, and access links are all orthogonal either in the time or frequency domain and hence there is no intra-cell interference from the BS to MFemtocell users and vice verse. In this scheme, a fraction ϕ ($0 < \phi < 1$) of the spectrum is allocated exclusively for direct transmissions whilst the rest of the spectrum is allocated to the access transmission, as presented in Fig. 4.2 (a). The interference from an MFemtocell to access users of other MFemtocells can be negligible or considered as background noise. This is because signals that come from an MFemtocell should travel through at least two metallic enclosures to reach the other MFemtocell users. The instantaneous faded SNR for the direct transmission user n $\{n \in \mathcal{N}^{\mathrm{d}}\}$ and access user m_j $\{m_j \in \mathcal{M}_j\}$ can be calculated by

$$\gamma_{\mathrm{d}}(n) = \frac{|h_{\mathrm{d}}(n)|^2 P_{\mathrm{BS}}}{BN_0}, \qquad n \in \mathcal{N}^{\mathrm{d}} \tag{4.1}$$

$$\gamma_{\mathrm{a}}(j, m_j) = \frac{|h_{\mathrm{a}}(j, m_j)|^2 P_{\mathrm{MF}}}{BN_0}, \qquad m_j \in \mathcal{M}_j \tag{4.2}$$

respectively, where $h_{\mathrm{d}}(n)$ and $h_{\mathrm{a}}(j, m_j)$ are complex-valued channel gains over the direct link and access link, respectively, P_{BS} and P_{MF} are the BS and MFemtocell transmit powers, respectively, B is the system bandwidth, and N_0 is noise spectral density.

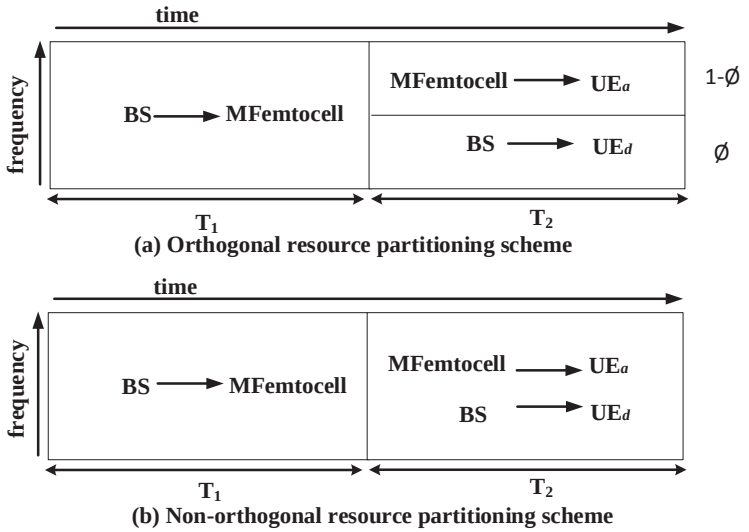

Figure 4.2: Resource partitioning schemes.

- Non-orthogonal resource partitioning scheme In this scheme, the radio resources are reused by the direct and access links, as it is shown in Fig. 4.2(b). However, the radio resources are still orthogonally allocated between backhaul and direct links and between backhaul and access links. Non-orthogonal mode means that there will be intra-cell interference to the access and direct transmission users due to the simultaneous transmissions from MFemtocell and the BS on the same spectrum. The advantage of this scheme is the improvement in resource utilisation compared to the orthogonal scheme. In addition, this scheme gives the flexibility to implement radio resource management (RRM) at the BS and the MFemtocell independently. The instantaneous received signal to interference plus noise ratio (SINR) for

a direct transmission user, denoted by $\acute{\gamma}_d(n)$, can be calculated by

$$\acute{\gamma}_{\mathrm{d}}(n) = \frac{|h_{\mathrm{d}}(n)|^2 P_{\mathrm{BS}}}{I + BN_0}, \qquad n \in \mathcal{N}^{\mathrm{d}} \tag{4.3}$$

where I is the intra-cell interference generated from all the MFemtocells. This type of interference can be scaled down significantly by the indoor penetration loss as well as by constraining the transmit power within the MFemtocell using a directive antenna. On the other hand, the received SINR for an access link user is given by

$$\begin{aligned} \acute{\gamma}_{\mathrm{a}}(j, m_j) &= \frac{|h_{\mathrm{a}}(j, m_j)|^2 P_{\mathrm{MF}}}{|h_{\mathrm{d}}(m_j)|^2 P_{\mathrm{BS}} + BN_0} \\ &= \frac{\gamma_{\mathrm{a}}(j, m_j)}{\gamma_{\mathrm{d}}(j, m_j) + 1}, \qquad m_j \in \mathcal{M}_j. \end{aligned} \tag{4.4}$$

So, $\acute{\gamma}_{\mathrm{a}}(j, m_j)$ can be characterised by the instantaneous SNR received from an MFemtocell, $\gamma_{\mathrm{a}}(j, m_j)$ in (4.2), and the achieved SNR if the same user is served by the BS instead, i.e., $\gamma_{\mathrm{d}}(j, m_j)$. Again, the interference between MFemtocells is negligible or considered as background noise. Now, if we assume that the distance between an MFemtocell and the BS and the distance between the MFemtocell's users and the BS are approximately the same, then (4.4) can be re-written as

$$\acute{\gamma}_{\mathrm{a}}(j, m_j) \approx \frac{\gamma_{\mathrm{a}}(j, m_j)}{\frac{\gamma_{\mathrm{b}}(j)}{G} + 1}, \qquad m \in \mathcal{M}_j \tag{4.5}$$

where G is the backhaul gain over the direct transmission link. In (4.5), $\gamma_{\mathrm{b}}(j)$ is the SNR for a backhaul channel of MFemtocell j and can be calculated by

$$\gamma_b(j) = \frac{|h_{\mathrm{b}}(j)|^2 P_{\mathrm{BS}}}{BN_0}, \qquad j \in \mathcal{J} \tag{4.6}$$

where $h_{\mathrm{b}}(j)$ denotes the complexed-value channel gain of the backhaul link for MFemtocell j.

4.3 Spectral-Energy Efficiency Trade-off for an SU MFemtocell

In this section, the spectral-energy efficiency trade-off is studied for an SU MFemtocell network in low and high SNR regimes. The system we are interested in here consists of the BS, a single MFemtcell and a user. For notational convenience, the user and MFemtocell indexes will be omitted as well since we are dealing with SU performance. The channel gains will be denoted by h_{d}, h_{b}, and h_{a} for direct, backhaul, and access links, respectively. The energy efficiency here is defined as the required energy per bit (in joule per bit) normalised to the background noise spectral level, $\left(\frac{E_b}{N_0}\right)$, for reliable communication. The spectral efficiency refers to the number of bits per second transmitted over a given bandwidth (in bps/Hz). The MFemtocell will receive the data in the T_1 and then transmit the data again to user in the T_2. For simplicity and without lost of generality, both T_1 and T_2 are assumed to be equal to one time slot. In the interest of making a fair comparison, the total bandwidth and the transmit power are set to be the same for cases with or without MFemtocell deployment. Thus, the total transmit power P over the two time slots is shared between the BS and MFemtocell. The transmit power from the BS is then given by

$$P_{\mathrm{BS}} = \alpha P \tag{4.7}$$

and the remainder power, P_{MF}, will be allocated to the MFemtocell and is given by

$$P_{\mathrm{MF}} = (1-\alpha)P \tag{4.8}$$

with α is a ratio, i.e., $(0 \leq \alpha \leq 1)$. Therefore, in the case where no MFemtocell deployment, the BS able to serve users with full transmit power, i.e., $\alpha = 1$. In addition to that, it has been assumed that the same size of bandwidth is used in T_1 and T_2.

4.3.1 Low SNR regime

- Direct transmission scheme:

Let us first analyse the required normalised energy per bit for direct transmission scheme that will be used later as a reference. Here we assume that BS is communicating directly with a user without the assistance of MFemtocell. For direct transmission scheme with a Rayleigh fading channel, the relationship between $\left(\frac{E_b}{N_0}\right)$ and the direct spectral efficiency $\mathrm{C}^{\mathrm{direct}}$ is given by

$$\left(\frac{E_b}{N_0}\right)^{\mathrm{direct}}_{\mathrm{dB}} \approx -1.59 + 10\log_{10}(A_{\mathrm{d}}) + 3\mathrm{C}^{\mathrm{direct}} \tag{4.9}$$

where A_{d} denotes the mean power of the direct transmission channel.

Proof. See Appendix B. □

- Orthogonal scheme:

In this scenario, we assume that a user is communicating with the MFemtocell rather the BS. However, the MFemtocell will receive and buffer the data in T_1 from the BS and serve the user in T_2. In this case, the required $\left(\frac{E_b}{N_0}\right)$ to achieve the spectral efficiency of orthogonal scheme, $\mathrm{C}^{\mathrm{orthg}}$, in the low SNR regime can be given by

$$\left(\frac{E_b}{N_0}\right)^{\mathrm{orthg}}_{\mathrm{dB}} \approx 10\log_{10}\left(\frac{2\ln 2}{\alpha G A_d}\right) + 2\mathrm{C}^{\mathrm{orthg}} \times 10\log_{10} 2 \tag{4.10}$$

where α $(0 < \alpha < 1)$ denotes a fraction of the total transmit power available for orthogonal transmission. In (4.10) $A_b = GA_d$ represents the mean power of the backhaul channel.

Proof. See Appendix C. □

- Non-orthogonal scheme:

In the scenario two uses are receiving data in T_2, one with single MFemtcell and another directly with the BS. For a fair comparison, the BS transmit power is also shared equally between the MFemtocell and the direct transmission user. The

required $\left(\frac{E_b}{N_0}\right)$ to achieve the spectral efficiency, $\mathrm{C}^{\text{non-orthg}}$, for non-orthogonal scheme in the low SNR regime is given by

$$\begin{aligned}\left(\frac{E_b}{N_0}\right)_{\text{dB}}^{\text{non-orthg}} &\approx 10\log_{10}\left(\max\left\{\frac{4\ln 2}{(\alpha G+1)A_d}, \frac{2\ln 2}{(1-\alpha)A_a + A_d)}\right\}\right) \\ &+ \frac{3(G^2\alpha^2+1)\kappa(|h_d|)}{(G\alpha+1)^2}\mathrm{C}^{\text{non-orthg}} \end{aligned} \tag{4.11}$$

where A_a denotes the mean power of the access channel.

Proof. See Appendix D. □

4.3.2 High SNR regime

- Direct transmission scheme:

The dependence between $\left(\frac{E_b}{N_0}\right)$ and $\mathrm{C}^{\text{direct}}$ for the direct transmission link with the absence of the MFemtocell can be calculated by

$$\left(\frac{E_b}{N_0}\right)_{\text{dB}}^{\text{direct}} \approx \mathrm{C}^{\text{direct}} 10\log_{10} 2 - 10\log_{10}(\mathrm{C}^{\text{direct}}) + 2.5067 + 10\ln(A_d). \tag{4.12}$$

Proof. See Appendix E. □

- Orthogonal scheme:

Assuming that the user is within close proximity to the respective MFemtocell, then the relationship between $\left(\frac{E_b}{N_0}\right)$ and $\mathrm{C}^{\text{orthg}}$ is characterised by

$$\begin{aligned}\left(\frac{E_b}{N_0}\right)_{\text{dB}}^{\text{orthg}} &\approx \mathrm{C}^{\text{orthg}} 10\log_{10} 2 - 10\log_{10}(\mathrm{C}^{\text{orthg}}) \\ &+ \left(-\log_2(2\alpha G A_d) + \frac{\Upsilon}{\ln 2}\right) 10\log_{10} 2 \end{aligned} \tag{4.13}$$

where Υ is the Euler-Mascheroni constant.

Proof. See Appendix F. □

- Non-orthogonal scheme:

Again, we assume here there are two users served by the BS and an MFemtocell at T_2. The total power transmitted in the both time slots is limited to P. The relationship between $\left(\frac{E_b}{N_0}\right)$ and $C^{\text{non-orthg}}$ in the high SNR regime is then characterised by

$$\begin{aligned}\left(\frac{E_b}{N_0}\right)^{\text{non-orthg}}_{\text{dB}} &\approx C^{\text{non-orthg}} 10\log_{10} 2 - 10\log_{10}(C^{\text{non-orthg}}) \\ &+ \left(-0.5\log_2(\alpha G A_d) + \frac{\Upsilon}{\ln 2}\right) 10\log_{10} 2. \end{aligned} \tag{4.14}$$

Proof. See Appendix G. □

4.3.3 Simulation Results and Discussions

To verify the derived equations with different resource partitioning schemes, we performed simulations with a single BS, an MFemtocell, and two users. The distance between the MFemtocell and its user is much smaller than the distance between the MFemtocell and the BS. The macrocell user or direct transmission user was placed far away from the MFemtocell so that the interference from the MFemtocell to the direct transmission user can be neglected or considered as background noise. In the non-orthogonal scheme, both users can receive data from the MFemtocell and BS. Both A_d and A_a are equal to 1 while backhaul gain G is equal to 8 dB. Furthermore, α is equal to 70% for orthogonal and non-orthogonal schemes.

Fig. 4.3 presents the spectral-energy efficiency trade-off for the direct, orthogonal, and non-orthogonal schemes. In the low SNR regime, the orthogonal partitioning scheme provides better energy efficiency (i.e., needs less energy) than both the non-orthogonal partitioning scheme and direct transmission scheme with the same spectral efficiency. However, the gap between the orthogonal and non-orthogonal partitioning schemes starts to decrease as we move to the high SNR regime. The decrease in the gap is due to the fact that the slope of the orthogonal scheme

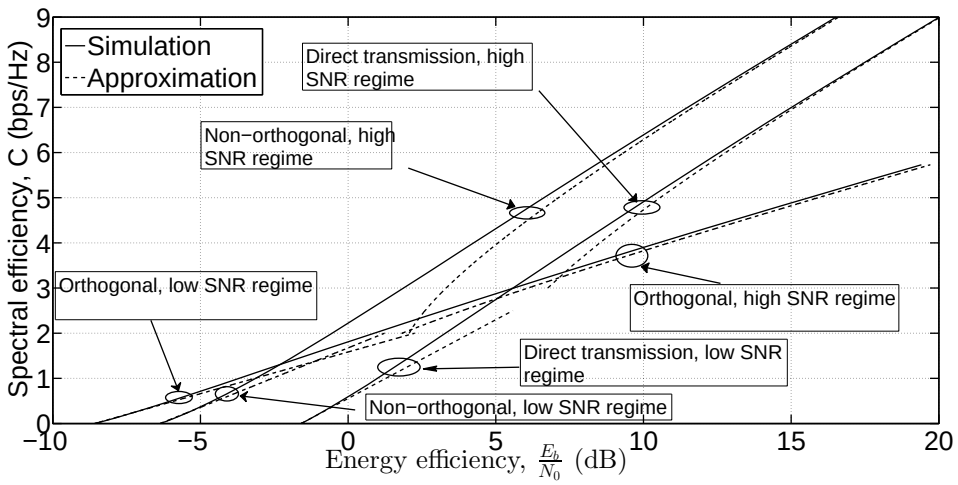

Figure 4.3: Spectral efficiency vs. energy efficiency for an SU MFemtocell with orthogonal and non-orthogonal resource partitioning schemes.

is less than the slope of the non-orthogonal scheme. We can also notice that the direct transmission scheme has better energy efficiency than the orthogonal partitioning scheme when the spectral efficiency is more than 3 bps/Hz. This is because when the underlying channel condition is good (high spectral efficiency), it is more energy efficient for the BS to transmit signals directly to a user with one time slot rather than to transmit to a user with an MFemtocell. Furthermore, in both SNR regimes, the non-orthogonal partitioning scheme provides better spectral efficiency than the direct transmission scheme with the same energy consumption because of the spectrum sharing between the BS and the MFemtocell. As we can see, the simulation results match well the derived closed-form expressions in high and low SNR regimes.

The maximum spectral efficiency can be achieved when backhaul spectral efficiency, C_{b}, and the access spectral efficiency, C_{a}, are equal. In this case, the optimum value for the power fraction α^* can be calculated according to

$$\alpha^* = \frac{|h_{\mathrm{a}}|^2}{|h_{\mathrm{a}}|^2 + |h_{\mathrm{b}}|^2}. \tag{4.15}$$

Hence, the spectral efficiency can be enhanced by using the optimum power fraction to allocate the power between the BS and the MFemtocell. Fig. 4.4 depicts the spectral-energy efficiency trade-off for the two partitioning schemes using the

optimum and fixed power allocation schemes. We can see that in the case of the orthogonal scheme, there is a 3 dB improvement in the energy efficiency by using the optimum power allocation compared with using the fixed power allocation. Furthermore, the slope of the spectral efficiency for the orthogonal scheme is not changed in both power allocation mechanisms. This is due to the fact that the slope does not depend on the power fraction α. We can also notice that when the spectral efficiency is less than 4 bps/Hz, the non-orthogonal scheme provides better energy efficiency by adopting the optimum power allocation scheme rather than adopting the fixed power allocation scheme. However, when the spectral efficiency is larger than 4 bps/Hz, the non-orthogonal scheme needs more energy to achieve a given spectral efficiency by using the optimum power allocation scheme compared with using the fixed power allocation scheme. This is due to the fact that there is strong intra-cell interference from the BS to the users.

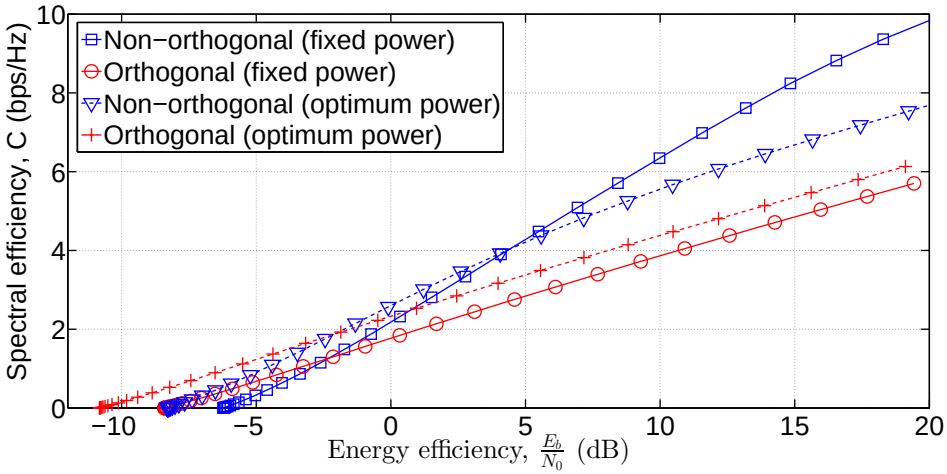

Figure 4.4: Spectral efficiency vs. energy efficiency for an SU MFemtocell with orthogonal and non-orthogonal resource partitioning schemes having different power control schemes.

4.4 Spectral Efficiency Analysis for MU MFemtocells with Multi-User Scheduling

In this section, we consider an MU MFemtocell network with one BS (single cell), multiple MFemtocells and multiple users. The spectrum is split into orthogonal resource blocks (RBs) for OFDMA-based cellular systems. These RBs are shared by different users by using opportunistic resource allocation in both frequency and time domains. Multiuser scheduling is assumed here where the macrocell users and MFemtocells are served over S RBs (or sub-channels), indexed by s=1,...,S, based on the well-known MAX-SINR and proportional fairness (PF) scheduling policies [105]. The BS and all MFemtocells transmit with fixed power per RB. To support the opportunistic scheduling, the BS gathers the channel quality indicator (CQI) from all users and MFemtocells. The users within an MFemtocell will feedback this information to the MFemtocell only. By using the MAX-SINR scheduler, the BS will assign a RB s to a user n having the highest instantaneous rate at a sub-frame t. i.e.,

$$\bar{n}_s = \underset{n \in \mathcal{N}^{\mathrm{d}}}{\operatorname{argmax}} R_n(t,s), \quad s = 1, ..., S \tag{4.16}$$

where $R_n(t,s) \propto \gamma_{n_{(\mathrm{D})}}(t,s)$ in the orthogonal scheme (or $\acute{\gamma}_{n_{(\mathrm{D})}}(t,s)$ in the non-orthogonal scheme) is the instantaneous achievable rate on RB s for a user n and $\operatorname{argmax} f$ is an operator that gives the index at which the vector f has the maximum value. In the PF scheduling case, the scheduler allocates the RB s to a user $n \in \mathcal{N}^{\mathrm{d}}$ according to the following criterion:

$$\bar{n}_s = \underset{n \in \mathcal{N}^{\mathrm{d}}}{\operatorname{argmax}} \frac{R_n(t,s)}{\bar{R}_n(t)} \quad s = 1, ..., S \tag{4.17}$$

where $\bar{R}_n(t)$ is the average delivered rate in the past, measured over a fixed window of observation. It can be calculated using an average filtering [105], which will be updated using the following formula

$$\bar{R}_n(t) = (1 - \frac{1}{T})\bar{R}_n(t-1) + \frac{1}{T} \sum_{s=1}^{S} R_n(t,s)\ \bar{I}_n(t,s) \tag{4.18}$$

where T is the time window constant, $\bar{I}_n(t,s)$ is a binary indicator that is set to 1 if user n is scheduled on RB s at time t and to 0 otherwise.

The communication over the BS-MFemtocell links takes place over a dedicated time-frequency zone, as shown in Fig. 4.2 (a) and Fig. 4.2 (b). Moreover, the same scheduling algorithm is used for the BS to schedule MFemtocells and direct transmission users. Within the j_{th} MFemtocell, it is assumed that the users ($\mathcal{M}_j$) are served according to round-robin policy. In case of the orthogonal scheme, it is assumed that a fraction of the spectrum ϕ, $0 < \phi < 1$, is allocated exclusively for direct transmissions in the second portion of the time, as shown in Fig. 4.2(a). Whereas in the non-orthogonal scheme, the BS and MFemtocells can utilise the whole spectrum to serve their users, as shown in Fig. 4.2(b).

The achievable spectral efficiency on the direct transmission link on time t can be calculated by

$$C_{\text{d}}(t) = \begin{cases} \frac{1}{2B} \sum\limits_{n \in \mathcal{N}^{\text{d}}} \sum\limits_{s=1}^{\phi S} R_n(t,s)\ \bar{I}_n(t,s), & \text{orthogonal} \\ \frac{1}{2B} \sum\limits_{n \in \mathcal{N}^{\text{d}}} \sum\limits_{s=1}^{S} R_n(t,s)\ \bar{I}_n(t,s), & \text{non orthogonal.} \end{cases} \tag{4.19}$$

The achievable spectral efficiency on the access link can be given by

$$C_{\text{a}}^{j}(t) = \begin{cases} \frac{1}{2B} \sum\limits_{m_j \in \mathcal{M}_j} \sum\limits_{s=1}^{(1-\phi)S} R_{m_j}(t,s)\ \bar{I}_{m_j}(t,s), & \text{orthogonal} \\ \frac{1}{2B} \sum\limits_{m_j \in \mathcal{M}_j} \sum\limits_{s=1}^{S} R_{m_j}(t,s)\ \bar{I}_{m_j}(t,s), & \text{non orthogonal} \end{cases} \tag{4.20}$$

where $R_{m_j}(t,s)$ is the instantaneous achievable rate for an access user m_j. However, the rates on access link between MFemtocells and their users are truncated by the achievable spectral efficiency of the backhaul link for MFemtocell j, i.e.,

$$C_{\text{b}}^{j}(t) = \frac{1}{2B} \sum_{s=1}^{S} R_j(t,s)\ \bar{I}_j(t,s) \tag{4.21}$$

where $R_j(t,s)$ is the instantaneous achievable rate over the backhaul link for an MFemtocell j. As a result, the total system spectral efficiency (in bps/Hz/cell)

after allocating all RBs to the selected users, including MFemtocell users, can be calculated according to

$$C_{\text{sys}}(t) = \sum_{j=1}^{J} \min[C_{\text{b}}^{j}(t-\tau), C_{\text{a}}^{j}(t)] + C_{\text{d}}(t) \tag{4.22}$$

where τ is the time required to decode, buffer, and re-encode the incoming data from the backhaul links. The first term in (4.22) stands for the achievable spectral efficiency of data flow from the BS to users through an MFemtocell and the second term represents the spectral efficiency for direct transmission users. To get an efficient resource usage for MFemtocell deployment, the spectral efficiency over the backhaul and access links should be equal, i.e., $C_{\text{b}}^{j}(t-\tau) = C_{\text{a}}^{j}(t)$. It is worth mentioning that to improve the spectral efficiency, another form of spectrum reuse scheme can be adopted. For example, multiple MFemtocells can use a common set of sub-channels simultaneously to serve their users. Both orthogonal and non-orthogonal resource partitioning schemes can benefit from MFemtocell spectrum reuse to improve spectral efficiency. This can only work, however, if multiple MFemtocells are located large distances apart or the coverage of each MFemtocell is limited to a small area by using a directive antenna.

4.4.1 Simulation Results and Discussions

The performance of the MFemtocell in the cellular system is evaluated in this section. A frequency-selective fading channel with 6 taps is used. Without loss of generality, The LTE frame structure is considered, which consists of blocks of 12 contiguous sub-carriers in the frequency domain and 7 orthogonal frequency-division multiplexing (OFDM) symbols in the time domain. One sub-frame (1 ms) is regarded as a scheduling period. Dynamic system level simulator which is compliant with 3GPP LTE specification [106] is used. The simulations are based on the Monte Carlo method, which consists of multiple snapshots. In each snapshot, the direct transmission users are distributed randomly and independently within the coverage of the BS. In addition to that, each snapshot having 20000

sub-frames, which are divided into units of 10. Each 10 time slots are further divided equally so that the first 5 time slots are allocated to the T_1 and the second 5 time slots are allocated to the T_2. The carrier bandwidth is fixed at 10 MHz with 50 RBs. All users are equipped with a single antenna while the MFemtocells have two antennas working in diversity mode. A full BS buffer is considered where there are always buffered data ready for transmission for each node. The users inside an MFemtocell experience 5 dB penetration loss when receiving a signal from the BS. Other relevant simulation parameters are summarised in Table 4.1.

Table 4.1: MFemtocell Simulation Parameters.

Parameter	**BS-MFmeto**	**BS-user**	**MFemto-user**
Antenna height (m)	20	20	2
Shadowing (dB)	8	8	4
Antenna gain (dBi)	18	18	5
Bandwidth (MHz)	10	10	10
Transmit power (dBm)	46	46	35
Spectrum sharing (ϕ)	100%	50%	50%
Time sharing	50%	50%	50%
Distance D (m)	$\geq$40	$\geq$50	$\leq$15
path loss (dB)	$128.1 + 37.6 \log_{10}\left(\frac{D}{1000}\right)$	$128.1 + 37.6 \log_{10}\left(\frac{D}{1000}\right)$	$127 + 30 \log_{10}\left(\frac{D}{1000}\right)$

Fig. 4.5 compares the average spectral efficiency of the orthogonal and non-orthogonal partitioning schemes using MAX-SINR and PF scheduling algorithms in an MFemtocell-enhanced system as a function of a percentage of users that associate with MFemtocells. Here, it has been assumed that the total number of users, K, and MFemtocells, J, are assumed to be 50 and 3, respectively. The MFemtcells are mounted inside buses which carry 10 users each. The MFemtocells are distributed randomly within the coverage of the BS in each snapshot. Regardless of using either orthogonal and non-orthogonal partitioning schemes, adopting the concept of the MFemtocell has a positive impact on the overall spectral efficiency. In addition to that, increasing the percentage of users that communicate through the MFemtocell leads to an increase in the overall spectral efficiency in comparison with the case where all the users are communicating

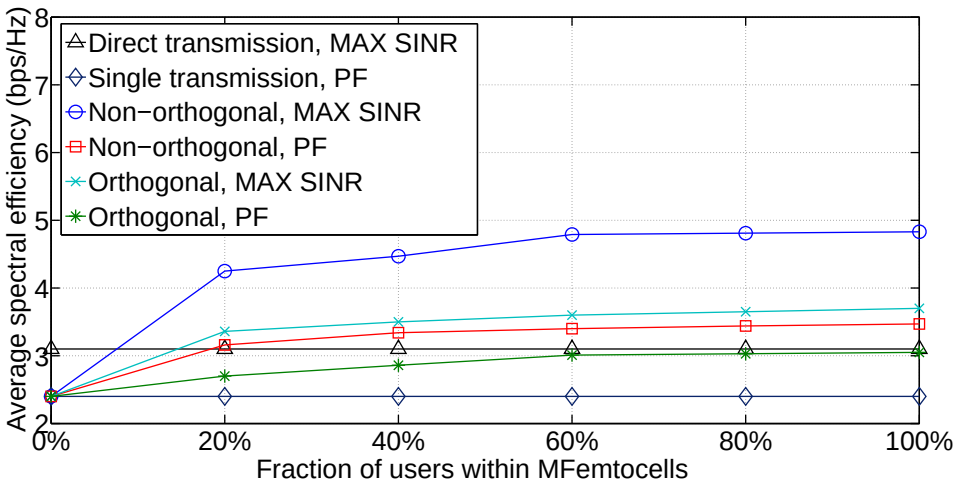

Figure 4.5: Average spectral efficiency of MU MFemtocells with multi-user scheduling and resource partitioning schemes (K=40, S=50, P_{BS}=46 dBm, P_{MF}=35 dBm, G =8 dB).

directly to the BS, i.e., $U^d = K$. This is because of the MU diversity gain, offered by opportunistic scheduling for users and the MFemtocells, as well as the backhaul gain over the direct transmission that the MFemtocell can offer. The simulation results demonstrate that there is a better spectral efficiency gain which can be achieved through sharing the spectrum between the BS and the MFemtocell in the same cell in the case of non-orthogonal transmission scheme. Although, the access users experience additional interference contributed from the BS, however, their performance does not get impacted in a noticeable way since they have a good quality channel to the MFemtocell station with high signal power. We can consider the spectral efficiency with the MAX-SINR scheduler as an upper bound performance for MFemocell systems because the BS always selects an MFemtocell and/or a user which has the best channel condition. However, it is noticeable that the gap between the MAX-SINR and PF scheduling algorithms in the non-orthogonal partitioning scheme is much larger than that in the orthogonal partitioning scheme. Although MAX-SINR scheduling gives a more consistent performance, it is at the cost of degrading the performance of cell-edge users. It is very important that MFemtocell deployment does not com-

promise the performance of the direct users, thus PF scheduling is more suitable in this case.

Fig. 4.6 depicts the system spectral efficiency of the two partitioning schemes as functions of transmit power of the BS. Here, we assume that there is one MFemtocell which is located 200 m distance apart from the BS and carries traffic for 10 users. There are also 40 direct transmission users which are randomly located within the coverage of the BS. We can notice that both schemes provide better performance, as compared to the direct transmission reference system. However, when the BS transmit power is sufficiently high , i.e., when $\frac{P_{\mathrm{BS}}}{BN_0}$ is larger than 14 dB, the direct transmission scheme offers better spectral efficiency than the orthogonal partitioning scheme. This is due to the fact that in the high SNR regime, it is worth transmitting directly to a user using one time slot rather than using two time slots through an MFemtocell. This also means that the orthogonal partitioning scheme may not be suitable when the MFemtocells are moving near the BS.

In wireless communication systems, the BS will require some information from the end users in order to allocate the radio resources efficiently or make handover operation. The end users in LTE cellular systems, for instance, are needed to feed back

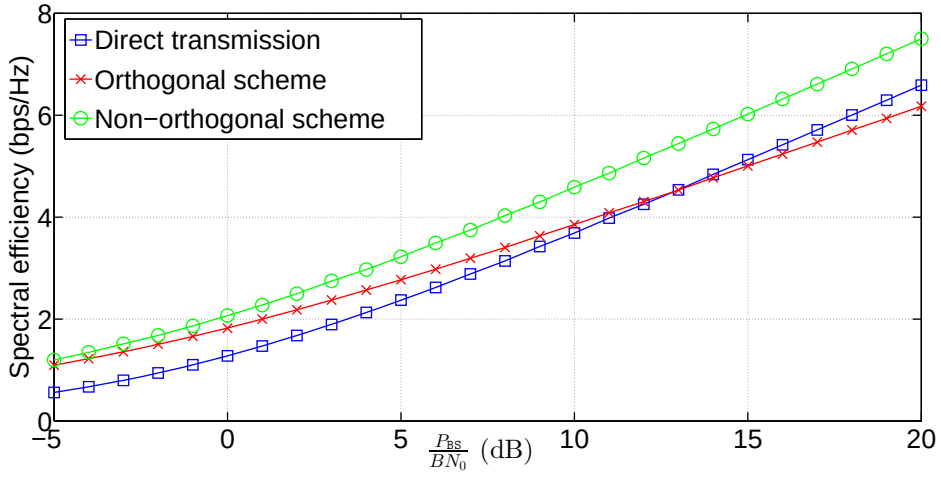

Figure 4.6: Spectral efficiency as a function of $\left(\frac{P_{\mathrm{BS}}}{BN_0}\right)$ (K=40, S=50, P_{BS}=46 dBm, P_{MF}=35 dBm, G=8 dB).

their downlink channel quality to the BS to support the channel-dependent scheduling and link adaptation. This information is forwarded to the BS over dedicated uplink control signalling channels. More control signalling means more sacrifice in system performance. As mentioned earlier, the MFemtocell can decrease significantly the uplink control signalling information of the direct transmission users. To demonstrate this fact, we analyse the uplink control signalling overhead in LTE cellular systems with and without MFemtocell deployment. The signalling overhead is the percentage of required control signalling with respect to the reference direct transmission. Here, we use discrete adaptive modulation schemes including quaternary QPSK, 16-QAM, and 64-QAM. Each modulation scheme with a different coding rate is supported when the uplink SNR is above a predefined SNR threshold. The SNR thresholds have been taken from [107]. The size of the control signal in unit of symbol/sub-frame is determined by [107]

$$Z_{\text{ctrl}} = \left\lceil \frac{N_{\text{sys}} \times CR}{M_{\text{mod}}} \right\rceil \tag{4.23}$$

where N_{sys}, M_{mod}, and CR are the number of control signalling bits, bit per symbol, and the coding rate of the selected modulation and coding scheme (MCS), respectively. Fig. 4.7 presents the reduction in the control signalling in the uplink backhaul link as a function of the number of users, within a vehicle, that communicate through an MFemtocell. As it is clearly shown, the greater a percentage of users that register to an MFemtocell, the greater a reduction in the control signalling overhead, compared with that of the reference direct transmission scheme, i.e., 0% of users communicate through an MFemtocell. Thus, with MFemtocell deployment, all users within an MFemtocell only need to send their control signalling to the MFemtocell rather than to the BS. The MFemtocell will cut off all these messages and send only its feedback to the BS. The control signalling information between users and the MFemtocell does not add to the total system control signalling overhead. The amount of reduction in signalling overhead can then be replaced with useful data and hence improvement in the users and system throughput. We can also notice that for any case of registered users, the control signal overhead tends to be saturated after a certain number of users within the

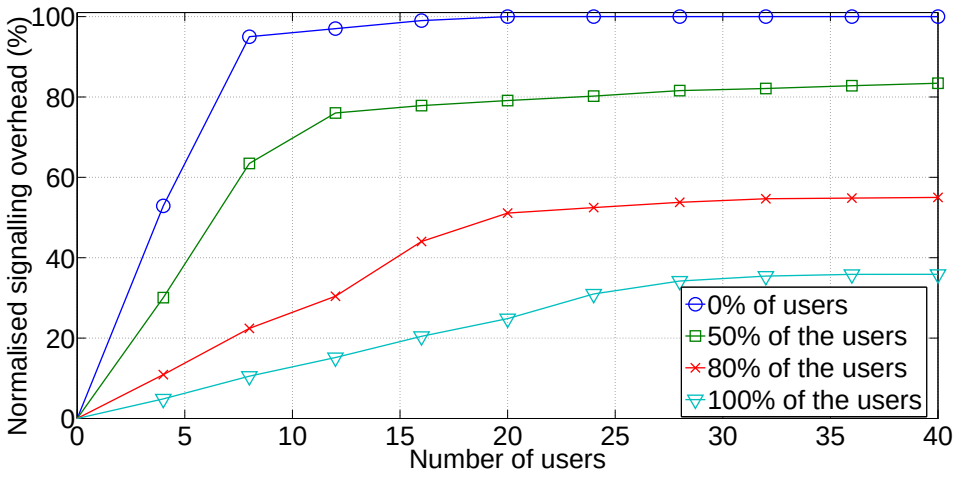

Figure 4.7: Normalised signalling overhead as a function of the number of users within an MFemtocell (J=10, S=50).

MFemtocells. This is due to the fact that there is a limitation to the number of users that can use the uplink control channel. In the case of 50% (or 80%) of registered users, the control signalling overhead consists of the control signal that comes from both the direct and access users.

4.5 Chapter Summary

This chapter has introduced the MFemtocell concept, which can be a potential candidate for the 5G cellular system, and highlighted the advantages of this idea. It has studied the performance of two resource partitioning schemes which can be used for the MFemtocell deployed scenarios. In this context, the spectral efficiency and energy efficiency for an SU MFemtocell system have been analysed first. The analysis has demonstrated the beneficial impact of using the non-orthogonal scheme in both low and high SNR regimes. Following that, the spectral efficiency of two resource partitioning schemes has been studied in the presence of opportunistic scheduling in a system level with a multi user and multi MFemtocell scenario. Finally, the performance of the MFemtocell deployment is investigated in term of signalling overhead reduction.

Chapter 5

Spectral and Energy Efficiency Analysis for Cognitive Radio Networks

5.1 Introduction

Radio spectrum refers to the electromagnetic spectrum that equivalents to radio frequencies lower than around 300 GHz. It is employed to transmit information using wireless communication media. Typically, radio spectrum is a precious and rare resource. The spectrum allocation is regulated by government agencies, for example, the federal communications commission (FCC) in the USA and the office of communications (Ofcom) in the UK. Today, the radio spectrums are congested and there are limited new spectrum bands available for commercial use. The pattern in Fig. 5.1 indicates how overcrowded the spectrum allocation is in the USA, and it is evident to even the casual observer that the radio spectrum has been fully reserved. Despite this fact, FCC has revealed that a significant amount of the radio spectrum in the United States is underutilised during the day [109]. Likewise, Ofcom has reported that the spectrum utilisation shows huge temporal and spatial variations ranging from 15 to 85%, as it is indicated in Fig. 5.2. It is generally agreed that the current policy of the spectrum usage is inefficient. The fixed spectrum allocation and the exclusive use of the spectrum does not allow us to exploit the spectrum to its full advantage.

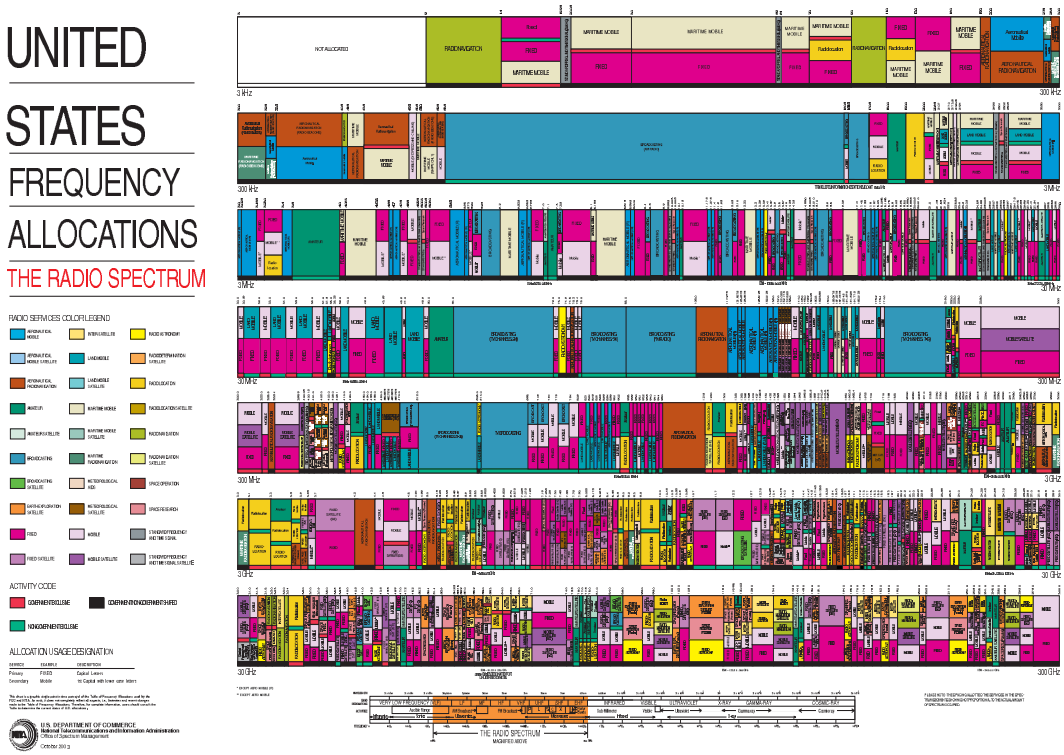

Figure 5.1: U.S. frequency allocations chart [108].

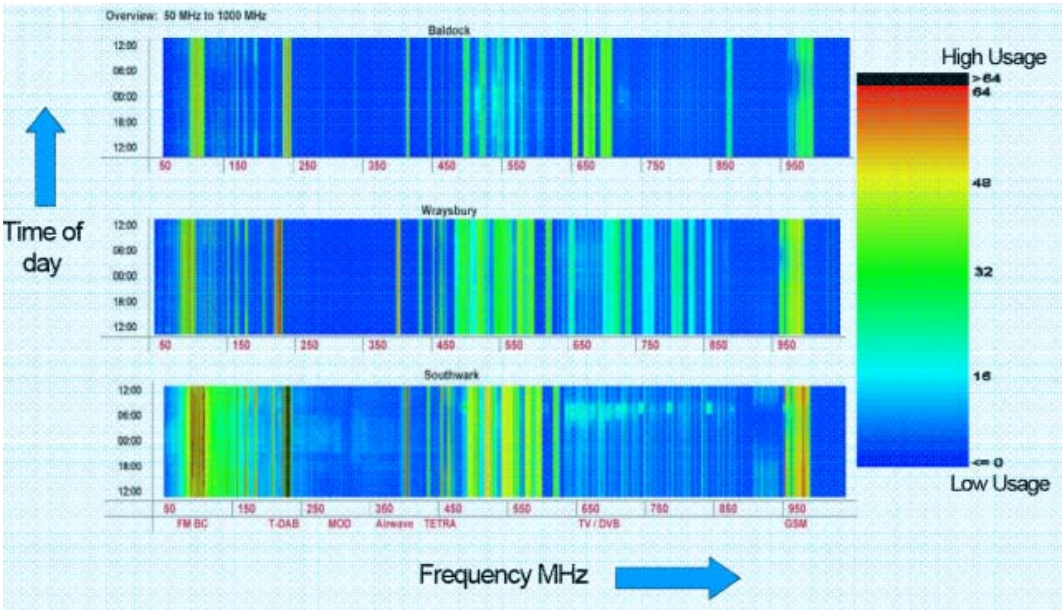

Figure 5.2: Spectrum occupancy measurements in a rural area (top), near Heathrow airport (middle) and in central London (bottom) [109].

By adopting an approach that is both dynamic and adaptable, it is possible then to use the spectrum to a fuller potential. This ignited the research activities to improve the usage of the highly sought-after radio spectrum and as a result, the CR concept has been introduced [110, 111].

5.1.1 Cognitive Radio Technology

CR is an innovative radio technique that aims to utilise the radio spectrum more efficiently by intelligently exploiting the licensed spectrum. Hence, a CR network, i.e., secondary network, shares radio spectrum owned by a licensed network, i.e., primary network. The secondary network is authorised of dynamically and autonomously adapting its radio operating parameters to coexist with the primary network, providing that the performance primary network is protected or above a certain level of quality. CR can be classified under two categories, namely interference-free and interference-tolerant CR networks [112]. In the interference-free CR networks, secondary transmitters (STs) can only use those spectrums which are not occupied by PRs [113, 114]. The IEEE 802.22 working group has chosen the interference-free CR mode as the enabling technique for the first standardised CR network, known as wireless regional area network (WRAN) [115]. Whereas in the interference-tolerant CR networks, the STs can share the whole spectrum as long as the interference to PRs is kept below a threshold and without causing any outage on primary network operation. If the aim is to apply an interference-tolerant CR then information is required from the PR on how much interference it can tolerate across the spectrum. This is known as the interference temperature limit [116]. Therefore, it is essential that CR systems acquire the interference temperature limit, in real time feedback, from the PRs. To this end, some modification to the primary system is unavoidable. In this chapter, we focus on the spectral and energy efficiency for the interference-tolerant CR networks. There were various studies that analysed the spectral efficiency of CR networks at the link level [117–120] as well as at the system level [24, 121, 122]. For SU CR networks, the spectral efficiency for AWGN channels was derived in [117] under

an assumption of average power constraint. In [118], the spectral efficiency of a CR channel was analysed against various fading channel distributions. The authors of [119] studied the spectral efficiency of a CR channel under different power allocation policies. In [120], ergodic and outage capacities of CR channels were evaluated under both peak and average interference power constraints. The spectral efficiency for both the link level and system level cooperative CR networks was studied in [121]. However, the power control of STs did not consider the interference threshold that the PRs can tolerate. In [122], the authors derived the average throughput of a system level CR network and studied its asymptotic behavior. However, the analysis was limited to a single PR. The spectral efficiency for hybrid CR networks was studied in [24] under average interference power constraint. Hybrid CR allows a network to be simultaneously both primary and secondary networks, thus gaining the advantages of both networks. Hybrid CR networks can be adopted in cellular networks to explore additional bands and enhance the spectral efficiency. It is noticeable that all the aforementioned studies focused on analysing the spectral efficiency but neglected to study the spectral-energy efficiency trade-off which is an increasingly important area nowadays. Therefore, our novel contributions in this chapter are summarised as follows:

1) We compare the spectral-energy efficiency trade-off in the low and high SNR regimes when transmitting a signal under average power constraint with transmitting a signal under peak power constraint while keeping the interference on PR below an acceptance level for both.

2) We propose a CR-based cellular network where a secondary network shares a spectrum that belongs to an indoor system. The spectral efficiency for the proposed network with multiple primary and secondary users is analysed using extreme value theory. The analysis will highlight the impact of the MU diversity gain of both the primary and secondary users on the achievable spectral efficiency.

3) We evaluate the spectral-energy efficiency trade-off of CR-based cellular

network using a general analytical framework all SNR values using peak-power interference constraint. The framework takes into account the numbers of primary and secondary receivers, transmit power, and interference threshold.

The remainder of this chapter is organised as follows. Section 5.2 describes the SU model of the proposed study and study the resulting spectral-energy efficiency trade-off, along with numerical results and discussions. The MU CR spectral efficiency and energy efficiency of multiple cognitive links are subsequently derived in Section 5.3. Finally, summary of the chapter is drawn in Section 5.4.

5.2 SU CR Spectral-Energy Efficiency Trade-off

In this section, we consider an SU CR channel. It consists of an interference-tolerant secondary transmitter-receiver pair that shares a spectrum with a primary transmitter-receiver pair [118–120], as it is shown in Fig. 5.3. We assume a point-to-point flat fading channel that is corrupted by AWGN. All nodes in this model are assumed to be equipped with a single antenna. The channel between the ST and secondary receiver (SR) is defined as the cognitive channel, while the channel between the ST and the PR is defined as the interference channel. The cognitive and interference channel gains are denoted by g_c and g_i, respectively. They are random variables drawn from an arbitrary continuous distribution with an expected value of unity and they are mutually independent. The ST is assumed to have perfect knowledge of the instantaneous CSI for the cognitive and interference channels. It is further assumed that the interference from the primary transmitter (PT) to the SR is considered as background noise. There are two constraints that the ST has to take into the account before transmitting a signal to the SR. The first constraint is the allowable received peak interference power at the PR. This constraint is essential in CR networks in order to avoid harmful interference at the PR. The second constraint is the available power that

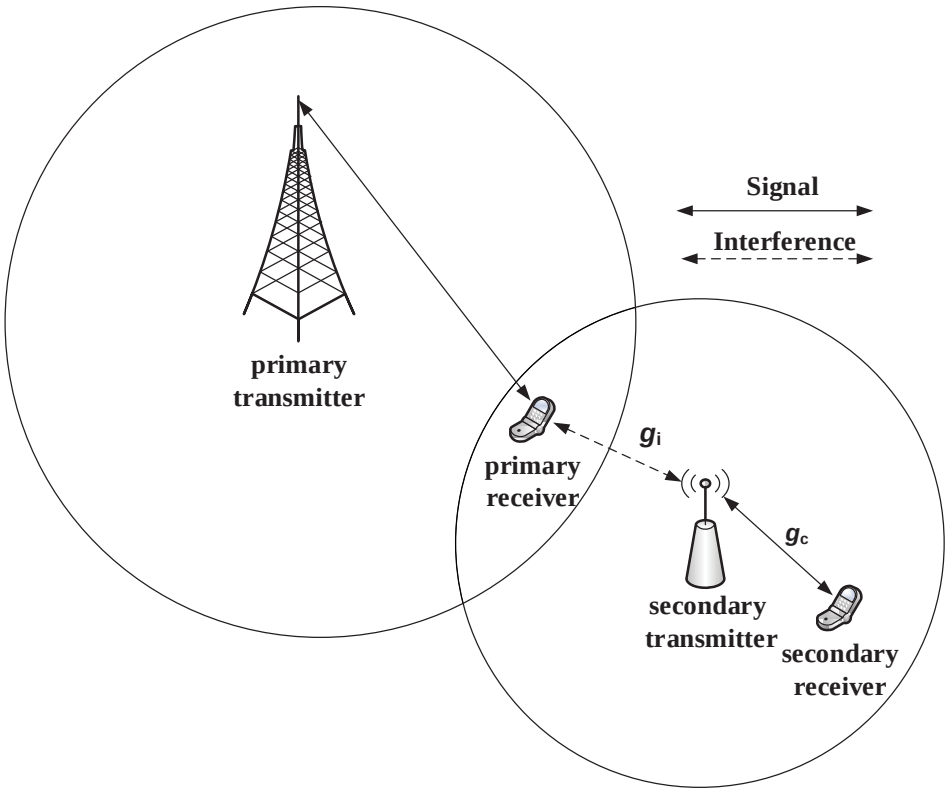

Figure 5.3: SU CR System model.

the ST has. In this paper, we consider two types of power constraint which are the average and peak transmit power constraints.

5.2.1 Fading Channels with Average Transmit Power Constraint

If we consider a CR channel under the average transmit power and peak interference power constraints, the spectral efficiency in this case can be calculated by

$$C = \max_{\gamma_s(g_c,g_i)\geq 0} \mathbb{E}[\log_2\left(1+\frac{g_c\gamma_s(g_c,g_i)}{N_0}\right)] \tag{5.1}$$

$$\text{s.t.} \quad \mathbb{E}[\gamma_s(g_c,g_i)] \leq \gamma_{\text{avg}} \tag{5.2}$$

$$g_i\gamma_s(g_c,g_i) \leq Q \tag{5.3}$$

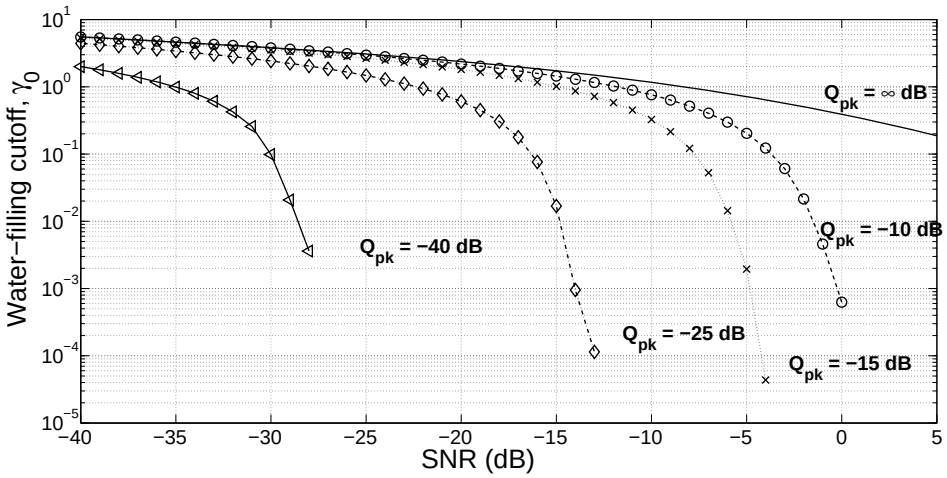

Figure 5.4: Cut off values γ_0 versus SNR with optimum power allocation.

where $\gamma_s(g_c, g_i)$, γ_{avg}, and Q are the instantaneous transmit power, allowed average transmit power, and peak received interference power that the PR can tolerate, respectively, and $\mathbb{E}[\cdot]$ denotes the statistical expectation. The optimum power allocation can then be expressed by

$$\gamma_s^*(g_c, g_i) = \min\left\{\left(\frac{1}{\gamma_0} - \frac{N_0}{g_c}\right)^+, \frac{Q}{g_i}\right\} \tag{5.4}$$

where γ_0 is the water-filling cutoff value that can be calculated from the constraint in (5.2) and $(x)^+ = \max\{0, x\}$. Numerical optimisation is required to get the optimum value of γ_0. Fig. 5.4 shows different value of the cutoff γ_0 value versus SNR under different peak interference constraints, Q. Unlike in the primary network, where only the CSI of the PR is required at the PT, the CSIs of both the PR and SR are needed at the ST as inputs for the power allocation algorithm. Depending upon the gains of the two types of channel, the CR transmission resides in different modes. No communication is allowed as long as the CR channel gain is below the cutoff value, i.e., $g_c \leq \gamma_0$. The classical water-filling algorithm can be adopted if $\left(\frac{1}{\gamma_0} - \frac{N_0}{g_c}\right) \leq \frac{Q}{g_i}$. If $\left(\frac{1}{\gamma_0} - \frac{N_0}{g_c}\right) > \frac{Q}{g_i}$, the transmit power is equal to $\frac{Q}{g_i}$.

Theorem 2. *Under average transmit and peak interference power constraints, the minimum energy efficiency required for reliable information over the cognitive*

channel is given by

$$\left(\frac{E_b}{N_0}\right)_{min} = \frac{\ln 2}{g_{c(max)}} \tag{5.5}$$

Proof. See Appendix H. □

where $g_{c(\max)}$ is the supremum of a random variable g_c. Unsurprisingly, the minimum energy is only affected by the cognitive channel while the interference channel has no influence on it. If the cognitive channel, for instance, is an AWGN channel, $\left(\frac{E_b}{N_0}\right)_{\min}$ will be equal to -1.59 dB. For the Rayleigh fading channel, $\left(\frac{E_b}{N_0}\right)_{\min} = 0$ ($-\infty$ dB) as the fading channel gain is unbounded, i.e., $g_{c(\max)} = \infty$. Fig. 5.5 presents the spectral-energy efficiency trade-off when the cognitive channel is under Rayleigh and AWGN channels against different interference channel fading distributions. Here, Q is assumed to be -5 dB. We can see that $\left(\frac{E_b}{N_0}\right)_{\min}$ depends only on the fading statistics of the cognitive channel regardless of the distribution of the interference channel, which verifies *Theorem* 1. Better energy efficiency can be obtained when the cognitive channel follows a Rayleigh distribution due to additional gain in the fading distribution. In the high SNR regime,

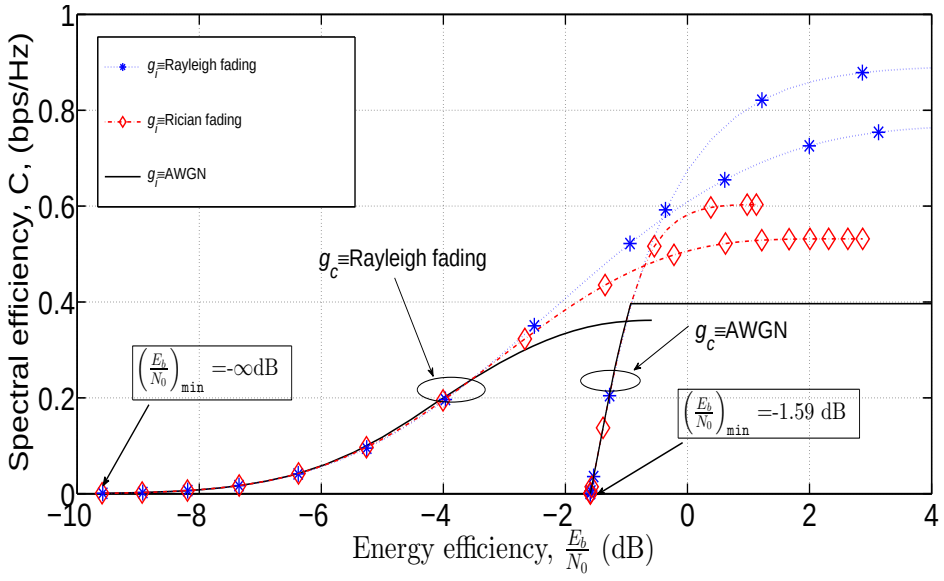

Figure 5.5: Per-link spectral-energy efficiency trade-off with different fading distributions of the interference channel in the low SNR regime ($Q = -5$, Rician factor $\mathbb{K}$=5).

Fig. 5.6 shows the spectral-energy efficiency trade-off when g_c and g_i are both changing according to Rayleigh fading distributions. It is worth noting that the SNR regime in which the cognitive channel resides is not only decided by the transmit power but also by the interference threshold Q. Regardless Q and the fading distribution of the cognitive and interference channels, the slope of the spectral efficiency goes to 0 as $\left(\frac{E_b}{N_0}\right)$ grows. The reason is that for CR channel the spectral efficiency is limited by interference threshold of the primary channel, i.e., even without Gaussian noise it achieves a bounded spectral efficiency $C_{\max}$. If the cognitive and interference channels follow Rayleigh distributions, then $C_{\max}$ is equal to

$$C_{\max} = \frac{\log_2(\frac{Q}{N_0})}{(1 - \frac{N_0}{Q})}. \tag{5.6}$$

The detailed derivation is given in Appendix I. Hence, $C_{\max}$ is characterised by Q and it is independent of the fading distribution of the cognitive and interference channels. However, we can notice that If Q is high enough, the spectral-energy efficiency trade-off can be approximated by

$$\left(\frac{E_b}{N_0}\right)_{\text{dB}} \approx C \times 10\log_{10} 2 - 10\log(C) + 2.5067 \qquad \forall C < C_{\max}. \tag{5.7}$$

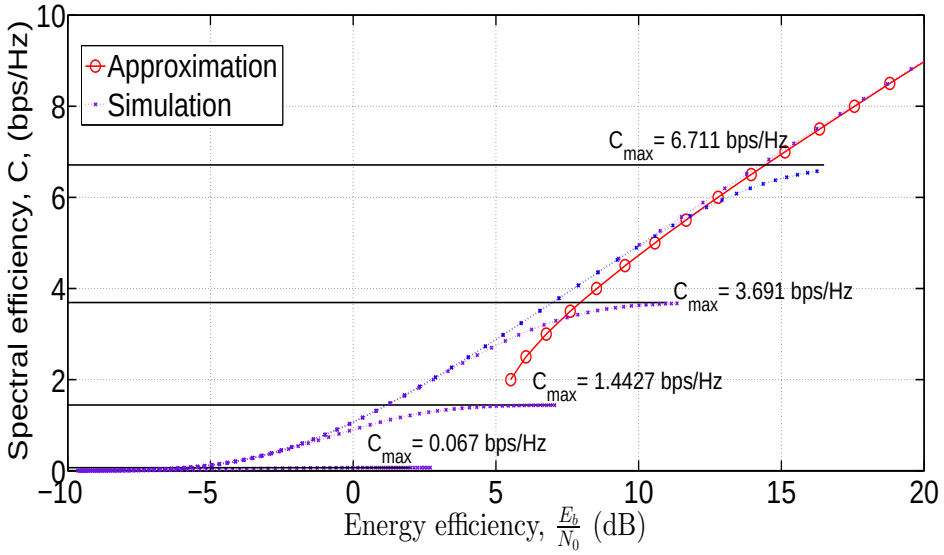

Figure 5.6: Per-link spectral-energy efficiency trade-off for different interference threshold values in the high SNR regime (g_c and $g_i \equiv$ exponential distribution).

Expression (5.6) is similar to spectral-energy efficiency trade-off approximation that can be applied to the single primary channel [39]. The only difference is that $C_{\max}$ has no limited value for the primary channel while it is limited for the cognitive channel.

5.2.2 Fading Channels with Peak Transmit Power Constraint

The optimum power allocation in this case is equal to

$$\gamma_s^*(g_c, g_i) = \min\left\{\gamma_{\text{pk}}, \frac{Q}{g_i}\right\} \tag{5.8}$$

where γ_{pk} is the peak transmit power of the ST. Unlike (5.4), only g_i is required as input for power policy. That makes it more straight forward as it only requires the CSI of the interference channel as an input. The minimum energy efficiency can be calculated by

$$\left(\frac{E_b}{N_0}\right)_{\min} = \lim_{\gamma_{\text{pk}}\to 0} \frac{\mathbb{E}[\gamma_s^*(g_c, g_i)]}{N_0 \mathbb{E}[\log_2(1 + \frac{\gamma_s^*(g_c,g_i)g_c}{N_0})]}. \tag{5.9}$$

In the low SNR regime, $\mathbb{E}[\gamma_s^*(g_c, g_i)] = \gamma_{\text{pk}}$. When we take this into consideration, $\left(\frac{E_b}{N_0}\right)_{\min}$ is always equal -1.59 dB for all types of cognitive channel fading distribution.

Fig. 5.7 compares spectral-energy efficiency trade-off when the ST transmits a signal under the average and peak power constraints. In the low SNR regime, transmitting a signal with average power constraint provides better energy efficiency than transmitting a signal with peak power constraint. Moreover, reliable communication is no possible for $\left(\frac{E_b}{N_0}\right)$ <-1.59 dB when transmitting a signal under peak power constraint. This is due to that fact that the power policy with peak power constraint does not benefit from the available energy at those moment in which the cognitive channel fading is exceedingly high. In the high SNR regime, both power policies behave similarly and approach the same maximum spectral efficiency $C_{\max}$ since the transmit power for both policies is controlled by $\frac{Q}{g_i}$. Therefore, (5.7) can also be applied in the high SNR regime if Q is high enough.

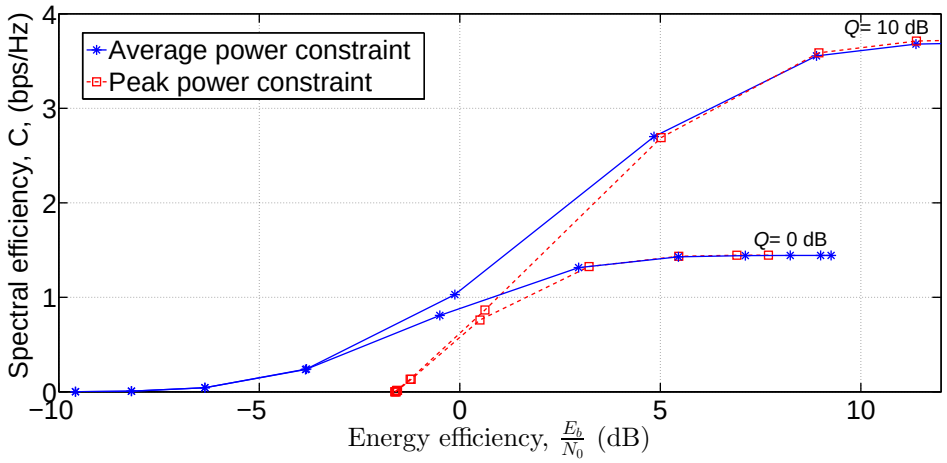

Figure 5.7: Per-link spectral-energy efficiency trade-off for CR channel with average and peak power constraints (Rician factor $\mathbb{K}$=5).

5.3 MU CR Spectral-Energy Efficiency Trade-off

In this section, we will study the spectral and energy efficiency for a CR-based cellular network. The intention here is not to build a complete cellular network using the concept of CR, but rather to enhance the spectral efficiency of the cellular networks for a short period of time by sharing a spectrum that belongs to another licensed network. We assume that a CR network consists of a single ST, i.e., macro BS, that transmits signals to multiple SRs. The CR network shares a spectrum owned by indoor primary network. The primary network also consists of multiple PRs, i.e., primary indoor access points (APs). The SRs and PRs are indexed by $n \in \mathcal{N} = \{1, ..., N\}$ and $k \in \mathcal{K} = \{1, ..., K\}$, respectively. The SRs and PRs are uniformly distributed in a cell of radius d and a cell of radius D ($d \leq D$), respectively, as it is shown in Fig. 5.8. The downlink transmissions of the CR network are considered and assumed to occur in the uplink transmission of the primary network. There are many advantages for sharing the spectrum of the uplink transmissions of an indoor network. First, since the primary network is assumed to be an indoor one, the mutual interference between the primary and secondary networks will be scaled down because of the penetration losses.

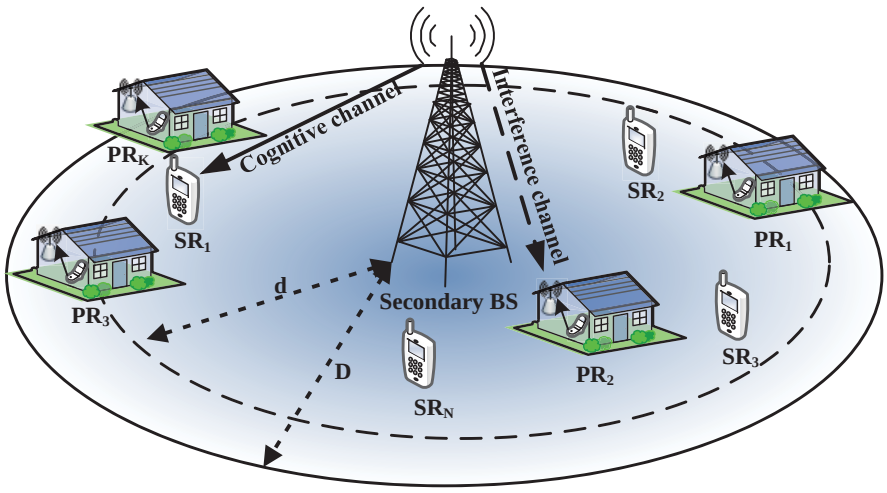

Figure 5.8: System model of CR-based cellular secondary BS, multiple primary receivers, and multiple secondary receivers.

Secondly, as the PRs, are all fixed in position, this offers an opportunity to easily detect them by the ST. Hence, the STs can detect the pilot channel broadcasted from indoor PRs and decide how many PRs with which they are surrounded [123]. The ST can then rely on channel reciprocity and estimate the channel coefficient of the interference channel using injected pilots in the uplink channel of the PRs. Finally, it is also possible that the interference CSI (ICSI) is sent from all PRs along with their identities and collected by a certain central unit. In fact, using a separate wireline control channel that broadcasts the interference measured over a broadband connection is a very practical solution. Before the secondary network can utilise the spectrum, it must register itself with the central unit first in order to be updated regarding the ICSI. However, the PRs do not necessarily need to identify each registered ST. The ICSI can inform the STs regarding the status of the worst aggregate interference that a PR suffers. The STs can also use ICSI as an alternative way to estimate the channel status to that PR and regulate their transmit power accordingly. In this work, we assume that there is only one registered CR network with a single secondary cell.

5.3.1 The Distribution of the Channel Gain

The cognitive channel power gain between the ST and the n_{th} SR is denoted by $g_c(n)$, while the interference channel gain between the ST and the k_{th} PR is denoted by $g_i(k)$. The cognitive and interference channels experience pathloss, shadowing, and multi-path fading. The focus in this section will be on the cognitive channel. However, the same analysis can be applied to the interference channel. The pathloss is a function of the distance r between the ST and the n_{th} SR, and can be expressed by $g_{\mathtt{p}}(n) = Ar^{\beta}(n)$, with β representing the pathloss exponent. The propagation coefficient A includes parameters related to antenna height, antenna gain, path-loss frequency dependence, and, in the case of the interference channel, the indoor loss. The combined channel gain $g_c(n)$ is given by $g_c(n) = \frac{g_m(n)g_s(n)}{g_{\mathtt{p}}(n)}$, where $g_s(n)$ and $g_m(n)$ represent the power gain of the shadowing and multi-path fading of the n_{th} SR, respectively. We use the composite channel model for both shadowing and fast fading [124]. This model is the result of the multiplication of the log-normal distribution with the Nakagami distribution. Thus, the composite channel gain can be approximated by log-normal distribution [124, pp. 132], i.e.,

$$f_h(x) = \frac{\xi}{\sqrt{2\pi}\sigma x} \exp\left\{-\frac{(10\log_{10}(x)-\mu)^2}{2\sigma^2}\right\} \tag{5.10}$$

where $\xi = \frac{10}{\log(10)}$. The mean μ and variance σ^2 can be given by

$$\mu = \left(\sum_{k=1}^{m-1}\frac{1}{k} - \ln(m)\right) + \mu_{\Omega} \tag{5.11}$$

$$\sigma^2 = \sum_{k=0}^{\infty}\frac{1}{(m+k)^2} + \sigma_{\Omega}^2 \tag{5.12}$$

respectively, where μ_{Ω} and σ_{Ω}^2 are the mean and variance of the shadowing, respectively, while m represents the Nakagami fading factor. If we recall $g_c(n) = \frac{g_m(n)g_s(n)}{g_{\mathtt{p}}(n)}$ and (5.10), the distribution of $g_c(n)$ can be expressed by

$$\begin{aligned} f_{g_c}(g) = &\frac{\mathtt{B}e^{\left(\frac{2}{a}(\mu-\xi\log g)\right)}}{g}\text{erfc}\left(\frac{a(L_{\max}+\xi\log g-\mu)-2\sigma^2}{a\sqrt{2\sigma^2}}\right) \\ &-\frac{\mathtt{B}e^{\left(\frac{2}{a}(\mu-\xi\log g)\right)}}{g}\text{erfc}\left(\frac{a(L_{\min}+\xi\log g-\mu)-2\sigma^2}{a\sqrt{2\sigma^2}}\right) \end{aligned} \tag{5.13}$$

where $L_{\max}$ and $L_{\min}$ are the maximum and minimum path-loss values in dB, respectively, and they depend on the cell radius d and minimum distance to the ST $d_{\min}$. In (5.13), $a = \xi\beta$ and $\mathtt{B}$ is a constant given $\mathtt{B}$ is a constant given by

$$\mathtt{B} = \frac{\xi e^{\left(\frac{2\sigma^2}{a^2}\right)}}{aA^2(d^2 - d_{\min}^2)}. \tag{5.14}$$

The detailed derivation of (5.13) is given in Appendix J.

5.3.2 Interference Constraint and CR Power Control

In order to keep a certain level of performance for the primary network, the aggregate interference I_k at any PR must always be below a predefined threshold, i.e., Q. The interference I_k that a PR experiences consists of the aggregate interference I_p from all transmit nodes in the primary network as well as the interference $I_c = \gamma_s g_i(k)$ from the CR network, where γ_s is the ST transmit power. In other words, the aggregate interference at a PR can be calculated by

$$I_k = I_p + I_c \leq Q \tag{5.15}$$

It is assumed that the adjacent indoor APs are using orthogonal radio recourses to avoid strong interference among them. Moreover, the interference between non-adjacent indoor APs, i.e., I_p, can be negligible or considered as background noise. This is because signals that come from a primary user should travel through at least two walls to reach the other primary APs [125]. However, $I_c = \gamma_s g_i(k)$ is dominate compared with I_p because typically macro BS transmits a signal with high power and then its signal could be high enough to propagate through the walls of the building where the PR is deployed and generate interference. Given there are many PRs, it is important to ensure that ST transmit power γ_s should always be tightly controlled to avoid harmful interference on a PR which has the maximum channel gain toward ST. This will inevitably protect the other PRs and (5.15) remains true for all PRs. Therefore, the transmit power of ST is controlled

according to

$$\gamma_s = \begin{cases} \frac{Q}{\max\limits_{k\in\mathcal{K}} g_i(k)}, & \max\limits_{k\in\mathcal{K}} g_i(k) > \frac{Q}{\gamma_{\rm pk}} \\ \gamma_{\rm pk}, & \max\limits_{k\in\mathcal{K}} g_i(k) \leq \frac{Q}{\gamma_{\rm pk}} \end{cases} \tag{5.16}$$

where $\gamma_{\rm pk}$ is the peak transmit power. The above power control is similar to (5.8) which demonstrates the power control under peak power constraint. It is more suitable for the proposed system to use the power policy with peak power constraint rather than power policy with average power constraint for many reasons. Firstly, in typical cellular networks, the BS has a limited maximum power that it can transmit with. The power control with average power constraint does not take into account this limitation. Secondly, to get as much benefit as possible of CR network, the SRs would usually be close to the ST, and so they can be within high SNR regime. This means that any gain of power control under average power constraint is minor. Finally, the power control with peak power constraint is more straight forward as it requires g_i only as an input rather than g_i and g_c. To this end, the ST can request the worst ICSI, which belong to the surrounding PRs, from the central unit. The ST then uses this ICSI to estimate the channel status and regulate the transmit power. Alternatively, or concurrently, the ST can use injected uplink pilots to estimate the channel status assuming that the channel is reciprocal.

5.3.3 Spectral Efficiency Analysis

The ST schedules SRs in orthogonal mode to avoid the intra-cell interference. In this work, time division multiple access (TDMA) is assumed by which the ST chooses an SR whose CSI implies the largest channel gain among all other SRs. In this case, the received SINR, $\acute{\gamma}$, for the scheduled SR is equal to

$$\acute{\gamma} = \frac{\gamma_s g_c(n^*)}{I} = \begin{cases} \frac{Q}{I}\frac{X}{Y}, & Y > \frac{Q}{\gamma_{\rm pk}} \\ \frac{\gamma_{\rm pk}}{I}X, & Y \leq \frac{Q}{\gamma_{\rm pk}} \end{cases} \tag{5.17}$$

where X and Y are random variables that represent $\max\limits_{n\in\mathcal{N}}\{g_c(n)\}$ and $\max\limits_{k\in\mathcal{K}}\{g_i(k)\}$, respectively, and n^* is the index for an SR who has the maximum value of the

channel gain. In (5.17), I is interference plus noise power. The spectral efficiency can then be evaluated by

$$C_{\text{sys}} = \int \log_2(1+\acute{\gamma}) f(\acute{\gamma}) d\acute{\gamma} \tag{5.18}$$

where $f(\acute{\gamma})$ is the probability density function (PDF) of $\acute{\gamma}$. By adopting the extreme value theory [126], $f(\acute{\gamma})$ converges to

$$\begin{aligned} f(\acute{\gamma}) \rightarrow & \frac{\mathrm{K}_1 \acute{\gamma}^{-1.5}}{2} \exp\left(-\mathrm{K}_1 \acute{\gamma}^{-0.5} - \mathrm{K}_2\right) \\ - & \frac{\mathrm{K}_1 \mathrm{K} \acute{\gamma}^{-1}}{2\left(\mathrm{K}\acute{\gamma}^{0.5}+1\right)} \exp\left(-\mathrm{K}_1 \acute{\gamma}^{-0.5} - \mathrm{K}\mathrm{K}_1\right) \\ + & \frac{\mathrm{K}}{2\left(\mathrm{K}+\acute{\gamma}^{-0.5}\right)^2 \acute{\gamma}^{\frac{3}{2}}} \left(1 - \exp\left(-\mathrm{K}_1 \acute{\gamma}^{-0.5} - \mathrm{K}\mathrm{K}_1\right)\right). \end{aligned} \tag{5.19}$$

See Appendix K for the detailed derivation. In (5.19),

$$\mathrm{K} = \left(\frac{I\delta_p}{Q\delta_c}\right)^{0.5} \tag{5.20}$$

$$\mathrm{K}_1 = \left(\frac{\gamma_{\text{pk}}\delta_c}{I}\right)^{0.5} \tag{5.21}$$

$$\mathrm{K}_2 = \left(\frac{\gamma_{\text{pk}}\delta_p}{Q}\right)^{0.5} \tag{5.22}$$

where δ_c and δ_p are the scale parameters for the cognitive and interference channels, respectively. A numerical integration of (5.18) provides convenient spectral efficiency evaluation.

5.3.4 Spectral-Energy Efficiency Trade-off

The average energy efficiency is given by $\left(\frac{E_b}{N_0}\right)_{\text{sys}} = \frac{\bar{\gamma}_{\text{avg}}}{N_0 C_{\text{sys}}}$, where $\bar{\gamma}_{\text{avg}}$ is the average transmit power. From (5.16), we have

$$\begin{aligned} \bar{\gamma}_{\text{avg}} = \mathbb{E}[\gamma_s] &= \int \gamma_s(y) f_Y(y) dy \\ &= \int_{\frac{Q}{\gamma_{\text{pk}}}}^{\infty} \frac{Q}{2} \delta_p^{0.5} y^{-2.5} \exp - \left(\frac{\delta_p}{y}\right)^{0.5} dy \\ &+ \gamma_{\text{pk}} \exp\left[-\left(\frac{\delta_p \gamma_{\text{pk}}}{Q}\right)^{0.5}\right] = P_1 + P_2 \end{aligned} \tag{5.23}$$

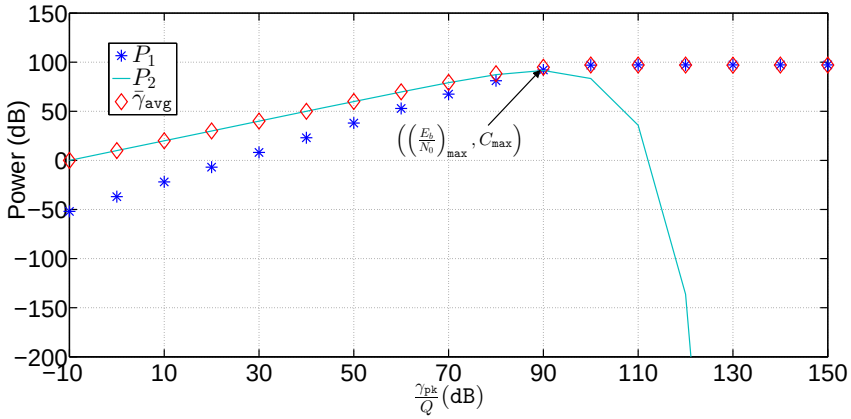

Figure 5.9: The average transmit power as a function of $\frac{\gamma_{pk}}{Q}$ (N =20, Q=0 dB, D=1000 m, $\frac{I}{N_0} = 1$, A_{dB}=128 dB).

where $f_Y(y)$ is the PDF of Fréchet distribution (refer to (K.4) in Appendix K). The transmit power can be modeled as a summation of two power components, i.e., P_1 and P_2. Fig. 5.9 shows how P_1, P_2, and $\bar{\gamma}_{avg}$ change as a function of $\frac{\gamma_{pk}}{Q}$. We can notice that P_1 is a monotonously increasing function while P_2 is a waterfall curve with respect to $\frac{\gamma_{pk}}{Q}$. Hence, $\bar{\gamma}_{avg}$ is dominated by P_2 until a point, after which $\bar{\gamma}_{avg}$ is relatively constant and dominated by P_1. This gives the conclusion that $\bar{\gamma}_{avg}$ can be approximated by

$$\bar{\gamma}_{avg} \approx \max(P_1, P_2). \tag{5.24}$$

If we allow the ST to transmit to the best user who has the maximum channel gain, then the average energy efficiency is given by

$$\left(\frac{E_b}{N_0}\right)_{sys} = \begin{cases} \frac{\gamma_{pk} \exp\left[-\left(\frac{\delta_p \gamma_{pk}}{Q}\right)^{0.5}\right]}{N_0 C_{sys}}, & C_{sys} < C_{max} \\ \left(\frac{E_b}{N_0}\right)_{max}, & \text{otherwise} \end{cases} \tag{5.25}$$

where C_{max} and $\left(\frac{E_b}{N_0}\right)_{max}$ are the maximum spectral efficiency and energy efficiency that the CR network can reach, respectively. They are given by

$$C_{max} = \int \log(1+\acute{\gamma}) \frac{\mathrm{K}}{2\left(\mathrm{K} + \acute{\gamma}^{-0.5}\right)^2 \acute{\gamma}^{\frac{3}{2}}} d\acute{\gamma} \tag{5.26}$$

$$\left(\frac{E_b}{N_0}\right)_{max} = \frac{P_1}{N_0 C_{max}} = \frac{\int_{\frac{Q}{\gamma_{pk}}}^{\infty} \frac{Q}{2} \delta_p^{0.5} x^{-2.5} \exp - \left(\frac{\delta_p}{x}\right)^{0.5} dx}{N_0 C_{max}}. \tag{5.27}$$

See Appendix L for the detailed derivation of (5.26). We can notice that $C_{\max}$ and $\left(\frac{E_b}{N_0}\right)_{\max}$ can be considered as new parameters that characterise the CR network. In (5.25), γ_{pk} can be numerically calculated as a function of C_{sys} from (5.18) and (5.19).

So far, the theoretical framework, which numerically evaluates the spectral-energy efficiency trade-off, is established for MU CR networks. In the following two subsections, we will focus on low and high SNR regimes for the proposed network.

Low SNR regime

If we recall (2.2), $\left(\frac{E_b}{N_0}\right)_{\min}$ is giving by

$$\left(\frac{E_b}{N_0}\right)_{\min} = \lim_{\bar{\gamma}_{\text{avg}}\to 0} \frac{\bar{\gamma}_{\text{avg}}}{N_0\mathbb{E}[\log_2(1+\frac{\bar{\gamma}_{\text{avg}}X}{I})]} \tag{5.28}$$

where the expectation is with respect to the random variable X. Applying L'Hopital's Rule into (5.28), $\left(\frac{E_b}{N_0}\right)_{\min}$ can then be calculated by

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{I\ln 2}{N_0\mathbb{E}[\max\limits_{n\in\mathcal{N}}\{g_c(n)\}]}. \tag{5.29}$$

We can conclude that in a very noisy region, the minimum energy efficiency in the CR network is characterised by the cognitive channels of the best SR. Hence, the interference channels have no impact on $\left(\frac{E_b}{N_0}\right)_{\min}$. The slope of the spectral efficiency versus $\frac{E_b}{N_0}$ is given by

$$S_0 = \frac{2\mathbb{E}^2[\max\limits_{n\in\mathcal{N}}\{g_c(n)\}]}{\mathbb{E}[\max\limits_{n\in\mathcal{N}}\{g_c(n)\}^2]} = \frac{2}{k(X)} \tag{5.30}$$

with $k(X)$ is the kurtosis of X.

High SNR regime

In the high SNR regime, the available transmit power and Q are important. If Q is set to be high, the spectral and energy efficiency converges to $C_{\max}$ and $\left(\frac{E_b}{N_0}\right)_{\max}$, respectively. Using (2.5), The spectral efficiency slope of the cognitive

network in this case is equal to 1. The horizontal penalty with respect to the AWGN channel, $\left(\frac{E_b}{N_0}\right)_{\text{penalty}}$, is given by

$$\left(\frac{E_b}{N_0}\right)_{\text{penalty}} = \mathbb{E}\left[\log_2\left(\frac{1}{X}\right)\right]. \tag{5.31}$$

See Appendix M for the detailed derivation.

5.3.5 Numerical Results and Discussions

This section presents the simulation results of the spectral and energy efficiency for an MU-CR network. The simulation is based on the Monte Carlo method, which consists of 10^6 channel realisations. The analysis is carried out with the following parameters: $\frac{I}{N_0} = 1$, $\beta = 3.7$, the indoor loss is 8 dB, and A=120 dB (128 dB in the case of the interference channel).

Fig. 5.10 shows the spectral efficiency of the CR network as a function of the number of SRs. We assume a reasonable transmit power that can be used in a typical cellular network. The number of PRs is assumed to be 20.

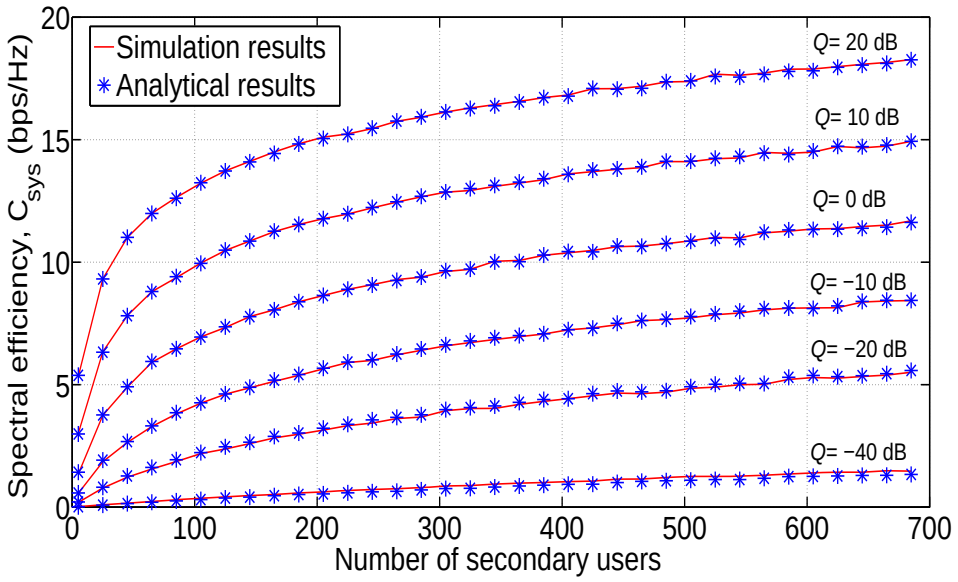

Figure 5.10: The average network spectral efficiency of the CR network as a function of the number of the secondary users (γ_{pk}=43 dB, K=20, D=1000 m, $\frac{I}{N_0} = 1$, A_{dB}=128 dB).

We can notice from this figure that the maximum spectral efficiency is still achievable even with a reasonable transmit power and does not need to be unlimited. The spectral efficiency of the CR network improves with the increase of the number of the SRs due to the additional gain that comes from the MU diversity. The theoretical results obtained from the numerical integration of (5.26) agree with the simulation results. It is worth mentioning that these results represent an ideal situation assuming that there are no constraints on the users' ability to feedback the channel to the BS. Of course, in reality, this is unlikely to be the case, as the number of users able to feedback at any one time is limited by the feedback resources available for this purpose. Nevertheless, this figure us an insight into the potential upper bound if the best spectral efficiency was achieved for the CR system.

The increase in the spectral efficiency is, however, sensitive to the interference threshold and the number of the PRs. Fig. 5.11 shows the impact of the number of the PRs on the spectral efficiency of the CR network. The number of the SRs is assumed to be 20 in this example. As shown, the spectral efficiency decreases quickly with the increase of the number of PRs. It indicates that adopting the spectrum sharing with another licensed network is unsuitable if the density of the PRs is relatively very high. However, the spectral efficiency can be improved by relaxing the interference threshold of the primary network, as it is shown in Fig. 5.11. The spectral efficiency can also be improved if the CR network considers only the SRs that are within a short distance to the ST. Hence, Fig. 5.12 shows the spectral efficiency of the CR network as a function of the number of the PRs with different values of d for a given D, assuming $Q = 0\ dB$. Clearly, for the given number of the PRs, the spectral efficiency increases dramatically with the decreasing of d.

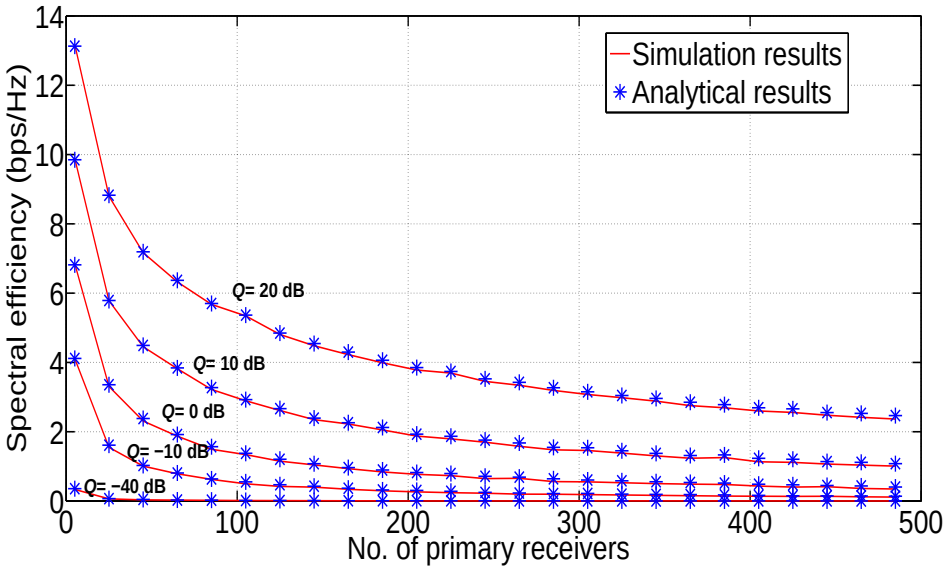

Figure 5.11: The average network spectral efficiency of the secondary network as a function of the primary receivers with different values of Q (γ_{pk}=43 dB, N=20, D=1000 m, $\frac{I}{N_0} = 1$, A_{dB}=128 dB).

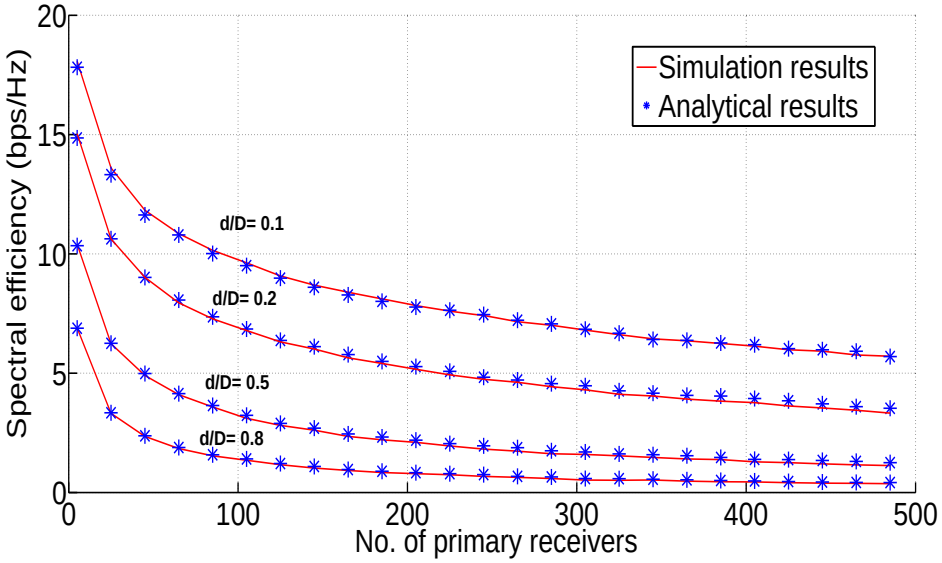

Figure 5.12: The average network spectral efficiency of the secondary network as a function of the primary receivers with different values of $\frac{d}{D}$ (γ_{pk}=43 dB, N=20, D=1000 m, $\frac{I}{N_0} = 1$, A_{dB}=128 dB).

Fig. 5.13 shows the spectral-energy efficiency trade-off for the CR network. The numbers of SRs and the PRs are assumed to be 30 and 20, respectively. As we can see, for a given Q, the spectral-energy efficiency trade-off curve tends to a point that corresponds to the maximum spectral efficiency, i.e., $C_{\max}$. It also indicates that the high SNR asymptotic tool is valid only if the CR network is working below its maximum spectral efficiency. The minimum energy efficiency, defined in (5.29), is the same for all curves.

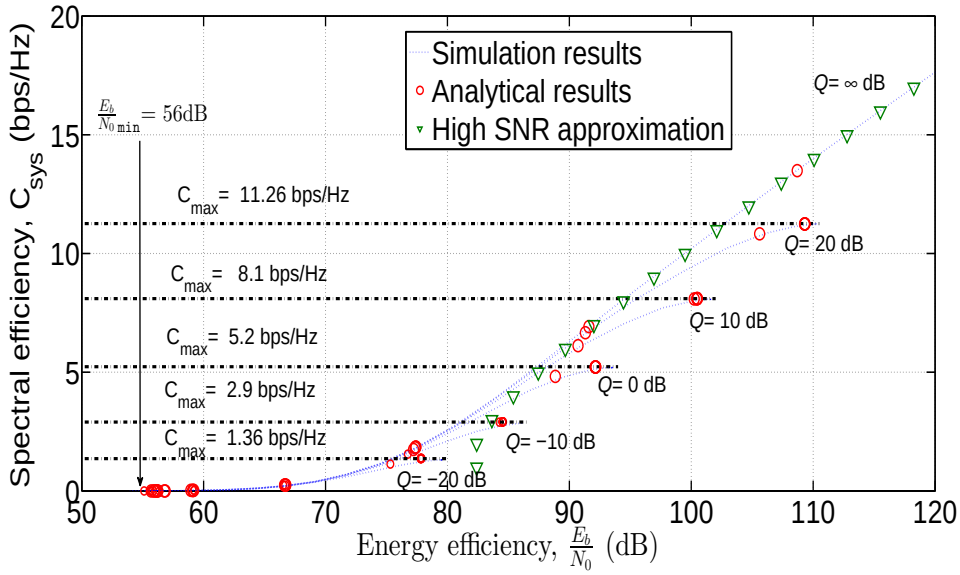

Figure 5.13: Per-network spectral-energy efficiency trade-off for CR network with different values of Q (N=30, K=20, D=1000 m, $\frac{I}{N_0} = 1$, A_{dB}=128 dB).

The impact of the MU diversity gain on the spectral-energy trade-off is pointed out in Figs. 5.14 and 5.15. Thus, Fig. 5.14 illustrates the spectral-energy efficiency trade-off as a function of the number of the SRs. We can notice that, for a given spectral efficiency, increasing the SRs improves the energy efficiency. This improvement is because the horizontal penalty spectral-energy efficiency trade-off and $\left(\frac{E_b}{N_0}\right)_{\min}$ are characterised by the cognitive channels only. Here, it has been assumed that the CPU power consumption has a minor impact on the total power consumption, giving that the CPU' computational performance per watt is high. The impact of the number of the PRs on the spectral-energy efficiency trade-off,

however, has a different aspect. Thus, from Fig. 5.15, the number of PRs has impact only on the maximum spectral efficiency but not on the horizontal penalty nor $\left(\frac{E_b}{N_0}\right)_{\min}$.

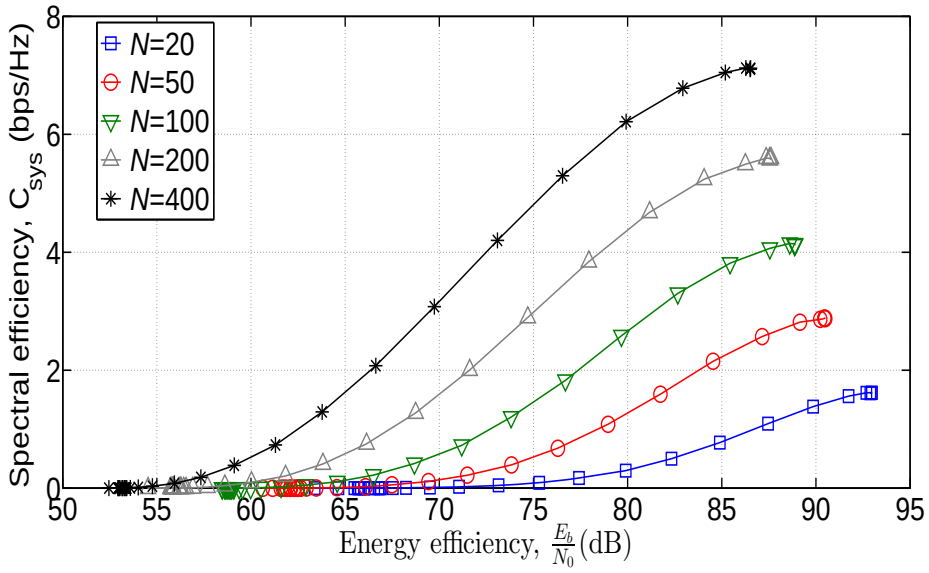

Figure 5.14: Per-network spectral-energy efficiency trade-off as a function of the number of secondary receivers (K=50, D=1000 m, $\frac{I}{N_0} = 1$, A_{dB}=128 dB).

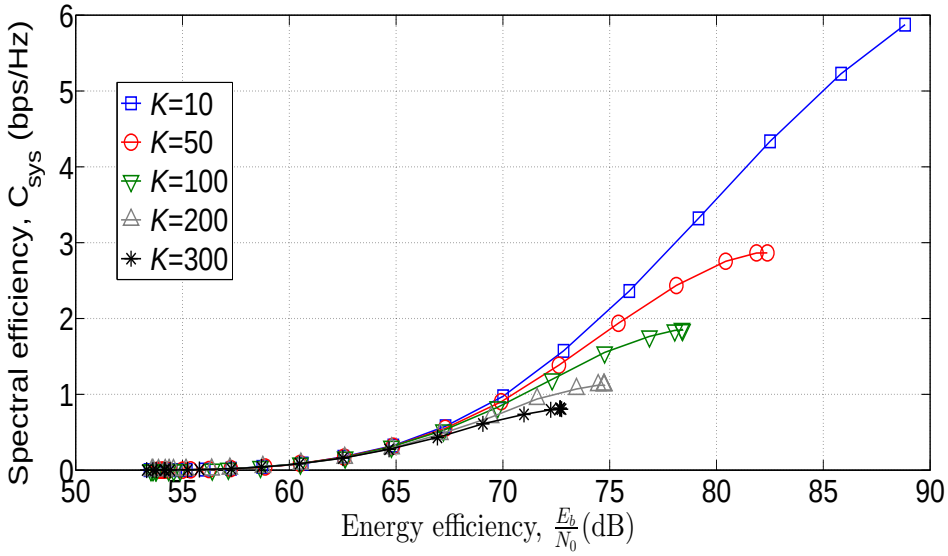

Figure 5.15: Per-network spectral-energy efficiency trade-off as a function of the number of primary receivers (N=50, D=1000 m, $\frac{I}{N_0} = 1$, A_{dB}=128 dB).

5.4 Chapter Summary

This chapter has investigated the spectral and energy efficiency in interference-tolerant CR networks. It has first studied the spectral-energy efficiency trade-off for an SU CR network under transmit power and interference constraints. Here, it has compared the case when the CR transmitter is transmitting a signal with average power and when transmitting a signal with peak power constraint in the low and high SNR regimes. The chapter has also analysed the spectral and energy efficiency for a system level CR network with multiple SRs and PRs. In this scenario, by the use of the extreme value theory, the analysis has highlighted the impact of the number of SRs, the number of PRs, the interference threshold, and the distance where SRs are located, on the performance of the CR network.

Chapter 6

Spectral-Energy Efficiency Trad-eoff for Multi-User Spatial Modulation in Massive MIMO Networks

6.1 Introduction

SM is a novel MIMO technique that has been proposed for low complexity implementation of MIMO systems without degrading the system performance [127]. Instead of simultaneously transmitting multiple data streams from the available antennas, SM encodes part of the information data using the spatial position of each transmit antenna in the antenna array. Therefore, only one transmit antenna is active at any time instance while other antennas are idle. In addition to the usual signal constellation diagram, the antenna array plays the role of a second constellation diagram, the so-called spatial constellation diagram. This can be used to increase the transmission data rate compared to single-antenna wireless systems. A block of information bits is split into two sub-blocks, each of base-2-logarithm number of bits. The first sub-block identifies the active antenna from a set of transmit antennas, while the second sub-block selects the symbol from the signal constellation diagram that will be sent from that active antenna. Therefore, SM is a combination of space shift keying (SSK) and amplitude/phase

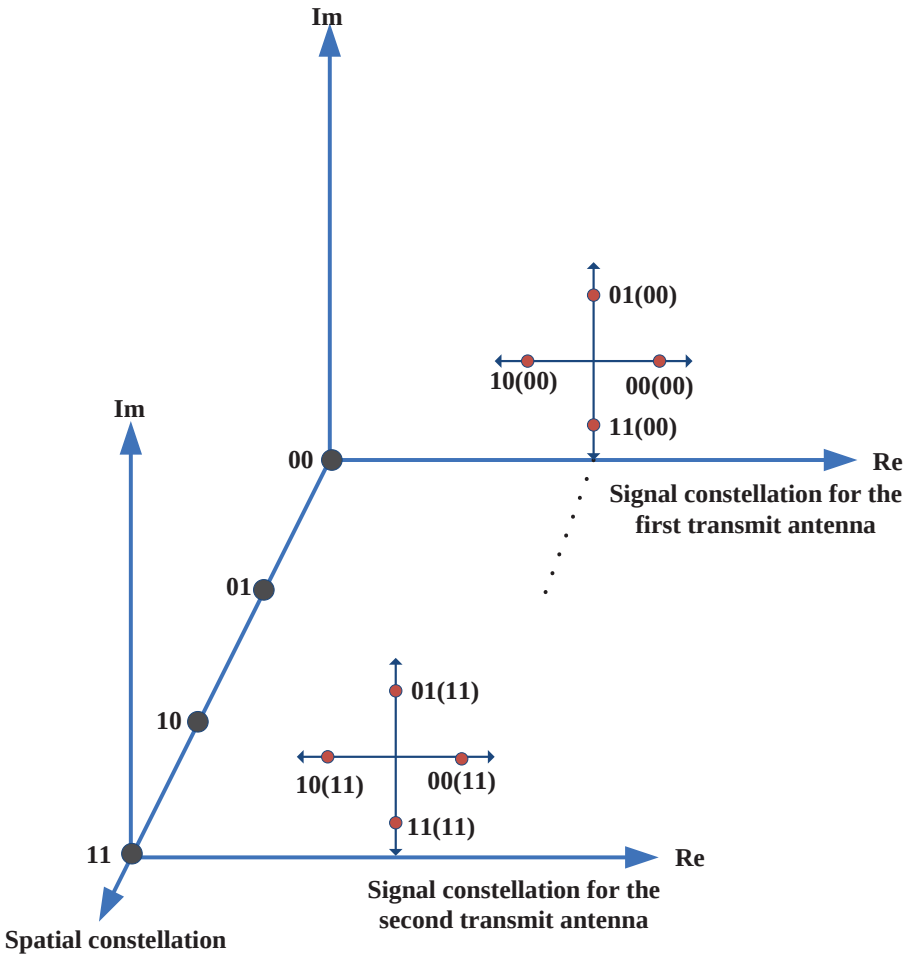

Figure 6.1: SM constellation diagram using four transmit antennas and QPSK modulation scheme.

modulation. Fig. 6.1 shows the SM constellation diagram with 4 transmit antennas and QPSK modulation scheme as an example. The receiver must determine which of the antennas was selected for transmission in order to decode that part of the information and uses it as a base for decoding the rest of the information. The receiver can then employ optimal ML detection to decode the received signal [128]. Using the SM approach, the inter-antenna interference and the need for synchronising the transmit antennas are avoided since the received signal comes from a single antenna.

The performance of SM has been studied extensively in the last few years to understand the impact of different channel fading propagation on the end-to-end

error probability. The SM scheme excels is in the feasibility to design a very efficient and simpler decoder compared with the conventional point-to-point SU-MIMO schemes such Vertical-Bell laboratories layered space-time (V-BLAST) and maximum ratio combining (MRC) [128–131]. A comprehensive survey of SM is available in [33] where many advantages of the SM have been highlighted. Despite these advantages, there are two important limitations to the SM scheme. The first limitation is that the SM fails in offering additional degrees of freedom or multiplexing gain. The multiplexing gain is the asymptotic gain that can be achieved gain over single-antenna channels [132]. This is because in SM scheme only one antenna is active for sending the information. The second limitation in the SM is that the encoding/decoding processes of SM require the number of transmit antennas to be in a power of two.

One way to overcome these two limitations whilst keeping similar SM encoding/decoding schemes is to use the so-called generalised SM (G-SM) schemes [133, 135]. The primary intention of G-SM is to modulate the information bits onto both the signal constellation diagram and combinations of multiple active antennas, thus increasing the degrees of freedom with unrestricted number of transmit antennas. The G-SM in [133] offered more diversity gain because of transmitting the same complex symbol via the activated antennas. The diversity gain refers to the improvement in the reliability or total number of channel fading gains that one can average this to [132]. The authors of [135] proposed another version of G-SM, named multiple active antennas SM (MA-SM). Unlike [133], the MA-SM allows to transmit multiple symbols on the activated antennas, thus offering more multiplexing gain. In the MA-SM, a different constellation plane is required for each antenna group in order to maximise the minimum distance between two codewords. This is done by rotating the constellation diagram on an angle for each plane. The optimum angles are calculated by an extensive computer search. Hence, the complexity in the transmission scheme is high, especially with a larger number of transmit antennas. In this chapter, we intend to take this further by proposing an enhanced G-SM based scheme that offers less complexity than MA-SM.

Another way to overcome the limitations of SM is by adopting the SM decoder into a MU-MIMO environment. This potentially can be an potential opportunity for the 5G cellular system especially with the deployment of the massive MIMO system, in which the number of transmit antennas is large. Recently, there were some works that studied SM-based MU-MIMO system [136–139]. The work of [136,137] focused at adopting SM on the uplink multiple access schemes. In [138], the authors applied SM for an MU-OFDMA system. Only [139] considered a system that is capable of serving multiple users on the same time/frequency resources and then each user decodes the received signal using SM-like decoder in donwlink. In that work transmit symbols are encoded by channel impulse responses (CIRs) of selected active antennas and then using zero-forcing (ZF) beamforming to eliminate the MU interference. However, an important drawback in this work is that the proposed scheme was devised under the assumption that the transmitting signals function under an average power constraint. This potentially means that the instantaneous power can be high. Moreover, the deployment scenario was necessarily limited to comply with the dimensional restriction, i.e., the number of active transmit antennas should be equal to the number of total received antennas. In this chapter, we will propose an enhanced version of the MU-SM transmission scheme in which (BD) linear precoding technique is first employed to construct parallel channels and then the SM scheme is employed individually on each channel. From the transmitter perspective, unlike SM, the precoding column index is now the additional source to convey information. Employing such a system, a higher degrees of freedom with low complexity decoder can be achieved. We will also investigate the trade-off between the spectral and energy efficiency for MU-SM system in this chapter. In this chapter, the spectral efficiency is defined as the number of bits per second over a given bandwidth (in bps/Hz) while the energy efficiency is defined as the number of bits per joule over a given bandwidth (in bit/joules/Hz). Therefore, we can summarised contribution in this chapter as follows:

1) We propose a novel G-SM based transmission scheme that offers more de-

grees of freedom than the original SM with reasonable complexity at the transmitter and receiver sides. The proposed scheme does not necessarily require more antennas at the transmitter side compared to the original SM. The complexity of the transmission scheme in the transmitter and receiver sides is much simpler than that of the MA-SM.

2) We propose a transmission scheme that will enable the SM decoder to be adopted in MU-MIMO systems. BD precoding is adopted to prevent the inter-antenna interference. The performance of the proposed system is compared with that of other MU-MIMO systems.

3) We study the spectral-energy efficiency trade-off analysis of SM-based MU-MIMO systems in the presence of a large number of antennas under a more practical PCM. Under such a set up, the transmit power is optimized in order to achieve optimum energy efficiency for a specific spectral efficiency.

The rest of this chapter is organised as follows. Section 6.2 describes the system model of the pre-coded SM and evaluates its performance. Section 6.3 introduces the MU-SM system model. Section 6.4 analyses the resulting performance in terms of bit error rate (BER), achievable rate, and energy efficiency along with numerical results and discussions. Finally, Section 6.5 summarises the chapter.

6.2 Pre-coded SM

In this section, we focus on a point-to-point SU-MIMO system that consists of a transmitter, i.e., BS, with N_t antennas, and a receiver, equipped with N_r antennas, as shown in Fig. 6.2. The transmitter divides the number of transmit antennas into M groups, each group has $\tilde{\mathrm{M}}$ antennas. The total number of the groups should be in a power of 2, however, there is no condition on the total number of antennas. Thus, the same antenna collection as in the G-SM and MA-SM schemes can be adopted. The total number of bits that can be transmitted

per given channel is given by $b = \log_2(M) + \tilde{\mathrm{M}}\bar{m}$ with $\bar{m}$ is the number of bits per modulated symbol. A random sequence of independent bits of size of b enters the pre-coded SM mapper which maps them onto a transmit vector $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$, i.e.,

$$\mathbf{x} = [\underbrace{0 \;\; .. \;\; 0}_{1} \;\; \underbrace{0 \;\; .. \;\; 0}_{m-1} \;\; \underset{\substack{\uparrow \\ m_{\text{th}} \text{ group}}}{\bar{\mathbf{x}}} \;\; \underbrace{0 \;\; .. \;\; 0}_{m+1} \;\; .. \;\; \underbrace{0 \;\; .. \;\; 0}_{M}]^T \tag{6.1}$$

where $\bar{\mathbf{x}} = \mathbf{Qa} \in \mathbb{C}^{\tilde{\mathrm{M}} \times 1}$ is a vector that contains the pre-coded symbols, and $\mathbf{a}$ is a vector that contains modulated symbols from M-array constellation. In (6.1), m_{th} represents the selected antenna group that the pre-coded symbols are transmitted. The transmitter exploits the antenna group index as additional means for

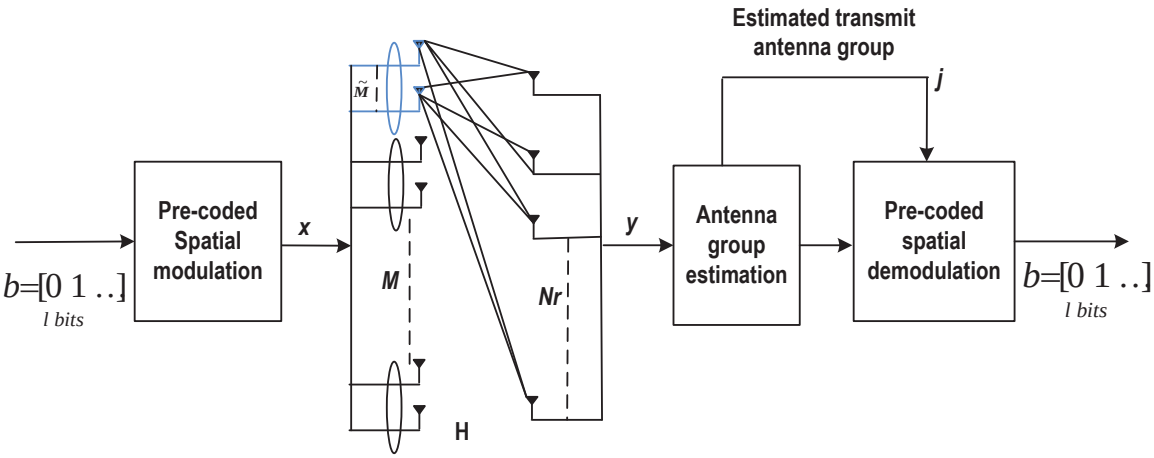

Figure 6.2: Pre-coded SM system model.

information. The transmit signal $\mathbf{x}$ is normalised in a such that each active antenna has unit transmit power. The received signal of the channel when the signal $\mathbf{x}$ is transmitted through the MIMO channel $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is

$$\mathbf{y} = \sqrt{\frac{\rho}{\tilde{\mathrm{M}}}}\mathbf{H}\mathbf{x} + \mathbf{n}. \tag{6.2}$$

where ρ is the average SNR, $\mathbf{n}$ is AWGN vector with zero-mean and unit variance per dimension at the receiver output. $\mathbf{H}$ can be written as a set of vectors which corresponds to channel gains between a transmit antenna group and receive

antennas, i.e,

$$\mathbf{H} = [\bar{\mathbf{H}}_1 \quad \bar{\mathbf{H}}_2 \quad ... \quad \bar{\mathbf{H}}_M]. \tag{6.3}$$

The entries of $\bar{\mathbf{H}}_m = [\bar{\mathbf{h}}_{m,1} \quad \bar{\mathbf{h}}_{m,2} \quad ... \quad \bar{\mathbf{h}}_{m,\tilde{\mathrm{M}}}] \in \mathbb{C}^{N_r \times \tilde{\mathrm{M}}}$ are modelled as complex independent and identically distributed (i.i.d) Gaussian random variables with zero-mean and unit-variance.

6.2.1 Pre-coded SM detector

The most significant role of the detector is how to estimate the active antenna group. One possible idea to do so is to project the received signal $\mathbf{y}$ into subspaces $\mathbf{P}_m^{\perp}$ that is orthogonal to subspace $\mathbf{v}_m$, where $\mathbf{v}_m$ is the subspaces that is spanned by a set of vectors $\bar{\mathbf{H}}_m$ [134]. The detection for the antenna group is performed by finding the minimum norm among the resulting projected vectors, i.e.,

$$j = \arg\min\{||\mathbf{P}_m^{\perp}\mathbf{y}||_F^2\} \tag{6.4}$$

where $\mathbf{P}_m^{\perp}$ is the orthogonal projection whose range is $\bar{\mathbf{H}}_m^{\perp}$. See Appendix N for more details about generating matrix $\mathbf{P}_m^{\perp}$. Once the index of antenna group j is estimated, the transmit modulated symbols can be estimated as follow

$$\hat{\mathbf{a}} = \acute{Q}(\mathbf{Q}^H\mathbf{W}_{m=j}\mathbf{y}) \tag{6.5}$$

where $\acute{Q}(\cdot)$ is constellation slicing function and $\mathbf{W}_{m=j}$ is the weight matrix. The performance of the decoder can be further improved by employing the MMSE pre-equalisation to reduce the impact of the noise. In this case, the weight matrix is given by

$$\mathbf{W}_m \quad = (\bar{\mathbf{H}}_m^H\bar{\mathbf{H}}_m + \sigma^2\mathbf{I})^{-1}\bar{\mathbf{H}}_m^H. \tag{6.6}$$

6.2.2 Pre-coded Matrix

Unlike in [135], rotating the constellation diagram is not required in the proposed scheme. Instead, the modulated symbols are pre-coded by a random matrix $\mathbf{Q}$

before transmitting. The use of pre-coded operation is to improve the separation of the different layers, i.e., reducing the inter-stream interference. The pre-coding matrix is generated in a way that the columns are orthogonal with each other, i.e., $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$. One way to generate an orthogonal matrix is to apply the QR decomposition on a random matrix $\mathbf{A}$ [59]. The entries of $\mathbf{A}$ are Gaussian random variables with zero-mean and unit-variance. The transmitter and the receiver both have a pre-coded matrix $\mathbf{Q}$ that is decided once the transmission begins and can then be fixed for the rest of the transmission. Without using the unitary matrix $\mathbf{Q}$, the receiver is only able to decode part of transmitted symbol that is based on active antenna group.

6.2.3 Simulation Results and Discussions

Let us evaluate the performance of pre-coded SM by comparing its performance with other MIMO schemes, such the SM, MA-SM, and V-BLAST. Here, a flat Rayleigh fading is assumed AWGN is assumed. The receiver, but not the transmitter, is assumed to have a full channel knowledge. The BER performance of these systems was evaluated as a function of the average SNR per receive antenna against various MIMO schemes, using Monte Carlo simulations. The V-BLAST system uses MMSE detection with ordered successive interference cancellation decoding, assuming that the receiver has a knowledge of SNR. The SM system uses the maximum likelihood (ML) detector. The BER curves of 10 bits transmitted with 5 receive antennas are plotted in Fig. 6.3. In this example, the pre-coded SM and MA-SM use $N_t = 4$, $\tilde{\tilde{\mathrm{M}}} = 2$ and 16-QAM modulation, the V-BLAST uses $N_t = 2$ and 32-QAM modulation, and the SM system with two different antenna configurations. As it is shown, the pre-coded SM provides SNR gains of 2 dB, 1 dB, 4 dB, and 6 dB over MA-SM, V-BLAST, SM with $N_t = 8$, and SM with $N_t = 4$, respectively, at BER value of 10^2.

The V-BLAST system outperforms the MA-SM when SNR>18 dB and the

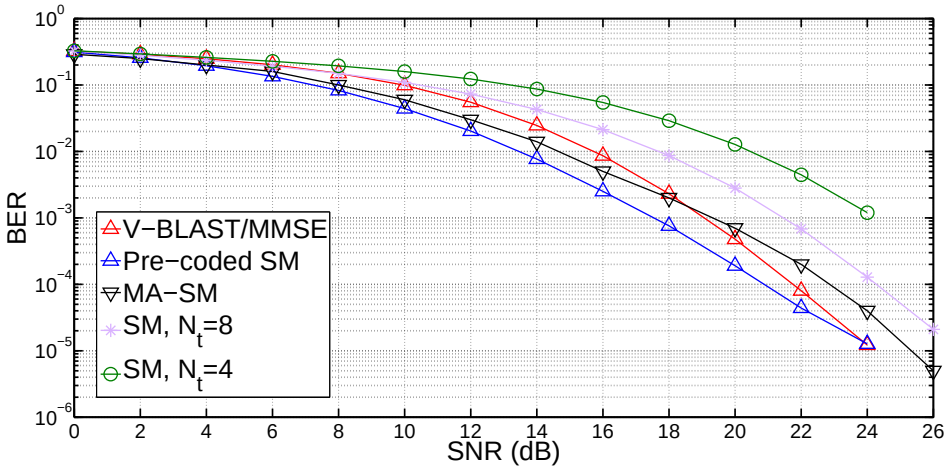

Figure 6.3: BER performance comparison of transmitting b= 10 bits among different SU-MIMO schemes with N_r=5 (Pre-coded SM and MA-SM use N_t=4 and $\tilde{\mathrm{M}} = 2$, V-BLAST uses N_t=2, and SM uses N_t= 8 and 4).

performance of pre-coded SM at higher SNR values (SNR>24 dB). This is because at the high SNR regime, the error is dominated by the estimation of the active antenna group.

The proposed pre-coded SM scheme can also be used to reduce the complexity of the SM at high spectral efficiency requirements, assuming that there is a large number of transmit antenna. The pre-coded SM can be near-optimal detector that is based upon the projection-based antenna group detection. In this case, $n = 1$ and $Q = 1$. Fig. 6.4 presents the comparison of projection-based detection SM over optimal SM [128] and sub-optimal SM [127]. In [127], maximum receive ratio combining (MRRC) is used to detect the index of the active transmit antenna. However, the CSI has to be known at the transmitter (CSIT) in order to normalise the channel gain and to get similar results to [127]. Our comparison also includes the case of sub-optimal SM but without CSIT. To have a fair comparison, we assume that $b = 10$ with $N_t = 16$ and $N_r = 4$ antennas for the all cases. We can see that, with projected-based SM, there is minor performance degradation at the high SNR regimes (SNR $>$ 18 dB). However, the pre-coded SM gains 6 dB over sub-optimal detector (with CSIT) of [127].

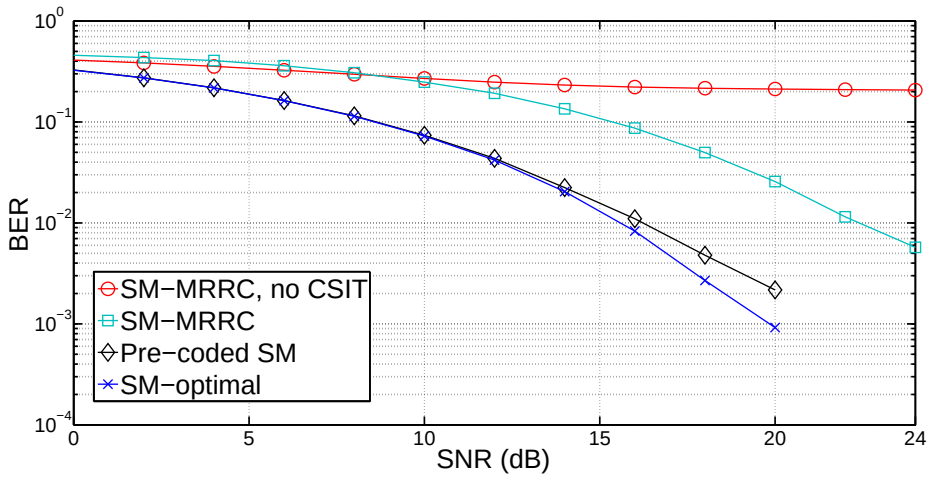

Figure 6.4: BER performance comparison between pre-coded SM and different SM decoding schemes (b= 10 bits, N_t=8, and N_r=4).

6.3 Multi-user SM System Model

In this section, we consider a broadcast channel consisting of a single BS, with Nt antennas, and K receivers, each equipped with Nr_k antennas, as it is shown in Fig. 6.5. For simplicity and without loss of generality, we assume that all users can have the same number of receive antennas and therefore the index k can be omitted (i.e., $Nr_1 = Nr_2 = Nr_k = Nr$). The transmit symbol of user k is denoted as a $\tilde{\mathrm{M}}_k$-dimensional vector $\bar{\mathbf{x}}_k$, which has the following form

$$\bar{\mathbf{x}}_k = [0 \ldots 0 \underbrace{a}_{j_{\text{th}} \uparrow \text{position}} 0 \ldots 0]^T \tag{6.7}$$

where a is the complex modulated symbol. Similar to SM scheme, the number of transmit bits for each user b_k can be adjusted in two different ways, either by changing the signal modulation order of a or by changing the location j where modulated symbol a is positioned, providing that the length of vector $\bar{\mathbf{x}}_k$ is in a power of 2.

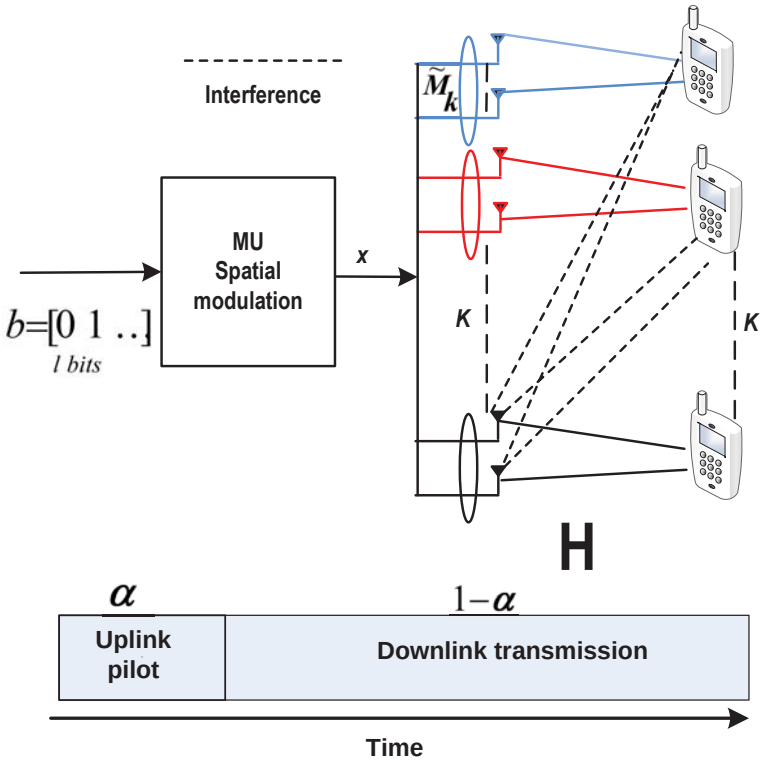

Figure 6.5: Multi-user SM system model.

The vector $\bar{\mathbf{x}}_k$ has unity power constraint, i.e., $\mathbb{E}[\bar{\mathbf{x}}_k^H \bar{\mathbf{x}}_k] = 1$. The BS embeds all vectors from all users and constructs a vector $\mathbf{x} \in \mathbb{C}^{\left(\sum_{k=1}^{K} \tilde{\mathrm{M}}_k\right) \times 1}$, i.e.,

$$\mathbf{x} = [\bar{\mathbf{x}}_1 \quad ... \quad \bar{\mathbf{x}}_k \quad ... \quad \bar{\mathbf{x}}_K]^T. \tag{6.8}$$

In MU-MIMO system, the BS simultaneously transmits to multiple users causing co-channel interference to occur among the users. In a such case, the channel knowledge of each user at the BS is essential for mitigating the interference. Hence, CSI of any user is required to be entirely known at the BS in order to produce pre-coding vector that aids in interference mitigation. One way to accomplish this is by using common downlink pilot symbols that are broadcast from the BS in order to permit the users to estimate their CSIs. The CSIs from all users are then feedback via an uplink control channel to the BS. This scheme is well-known as frequency-division duplexing (FDD) in which the downlink phase occurs in a different spectrum band from that of the uplink phase. In a such

arrangement, however, overheads of the channel estimation at the receiver and of the training required to collect the CSI at the transmitter both grow linearly with the number of transmit antennas and the number of users. Consequently, the expected improvement from the large-scale MIMO can be limited significantly by such overheads. An alternative solution is to exploiting channel reciprocity where the CSI can be obtained from uplink pilot signals by using a time division duplexing (TDD) protocol. Hence, in TDD systems, both uplink and downlink phases occur in the same spectrum band but at different time components, each carrying information and pilot signals [18–20]. One major issue with TDD is that it is time consuming to re-farm the spectrum bands that have been standardised to FDD systems to the TDD protocol [140]. Therefore, this study has accounted for the fact that the BS and users are perfectly synchronised and operate according to a particular case of TDD. As it is indicated in Fig. 6.5, there are set up ratios of uplink and downlink phases, denoted by α and $1-\alpha$, respectively. Unlike the traditional TDD system, only pilot signals are transmitted in the uplink phase period. The proposed scheme will likewise benefit the FDD systems in reaching high-data rate in the downlink despite the fact of dedicating the α portion of time of the downlink band for the pilot signals. It is assumed that the channels $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ that experience by different users are statistically independent. Let us define the network channel matrix and pre-coding vectors as $\mathbf{H} \in \mathbb{C}^{KNr \times Nt}$ and $\mathbf{V} \in \mathbb{C}^{Nt \times K\left(\sum_{k=1}^{K} \tilde{\mathrm{M}}_k\right)}$ as

$$\begin{aligned} \mathbf{H} &= [\mathbf{H}_1^T \ \ ... \ \ \mathbf{H}_2^T \ \ \mathbf{H}_K^T]^T \\ \mathbf{V} &= [\mathbf{v}_1 \ \ ... \ \ \mathbf{v}_2 \ \ \mathbf{v}_K]^T \mathbf{\Lambda}^{\frac{1}{2}} \end{aligned} \tag{6.9}$$

where $\mathbf{\Lambda}^{\frac{1}{2}}$ is a diagonal matrix whose elements are to scale the transmit power on each column and satisfy the total power $P_{\max}$ constraint. Next step is to design a precoding vector $\mathbf{V}$ that eliminates the co-channel interference while generation a similar channel situation to let the user adopt the same SM decoder for recovering the information. An effective means to construct complete diagonalisation is by using BD-based linear precoding, discussed in the following sub-section.

6.3.1 BD-based SM Algorithm

The key idea of BD is to eliminate all MU interference by designing a pre-coding vector $\mathbf{v}_k \in \mathbb{C}^{Nt \times \tilde{\mathrm{M}}_k}$ so that $\mathbf{H}_i \mathbf{v}_k = 0$ for all $i \neq k$ [141]. Now, if we define a matrix $\widetilde{\mathbf{H}}_k$ as

$$\widetilde{\mathbf{H}}_k = [\mathbf{H}_1^T \ \ ... \ \ \mathbf{H}_{k-1}^T \ \ \mathbf{H}_{k+1}^T \ \ ... \ \ \mathbf{H}_K]^T \tag{6.10}$$

$\mathbf{v}_k$ should lie in the null space of $\widetilde{\mathbf{H}}_k$ and then the zero-interference constraint is fulfilled. BD-based precoding has, however, a dimension restriction which is $\sum_{k=1}^{K} Nr_k \leq Nt$. This condition can be easily to meet in the massive MIMO system. The pre-coding vector $\mathbf{v}_k$ can be constructed using the singular value decomposition (SVD) of matrix of $\widetilde{\mathbf{H}}_k$, i.e.,

$$\widetilde{\mathbf{H}}_k = \widetilde{\mathbf{U}}_k \widetilde{\mathbf{\Sigma}}_k [\bar{\mathbf{v}}_k \ \ \mathbf{v}_k]^H \tag{6.11}$$

where $\mathbf{v}_k$ contains the last $\tilde{\mathrm{M}}_k$ columns of singular vectors and $\bar{\mathbf{v}}_k$ contains the first columns $(Nt - \tilde{\mathrm{M}}_k)$ singular vectors. The required magnitude of computations in (6.11) makes the SVD operation very onerous, especially when there is a large number of transmitting antennas. The complexity of the BD can be reduced significantly by adopting QR decomposition method instead [59]. This can be done by building a projection whose range is $\widetilde{\mathbf{H}}_k$, i.e., (N.2), followed by QR decomposition operation. Table 6.1 illustrates the computational cost in term of number of flops of QR-BD precoding which has been assumed in this work [59]. Thus, the QR decomposition on a matrix $\mathbf{A}$ of $N_t \times N_r$ size using Gram-Schmidt orthogonalisation or fast given method takes about $12N_t^2(N_r - \frac{N_t}{3})$ flops, while matrix inverse on $(\mathbf{A}^H\mathbf{A})^{-1}$ takes $6N_t^3$ flops. Furthermore, although the water-filling operation is performed in real-valued domain, it does not have a fixed complexity. However, in the worst case, water-filling over M eigenvalues needs up to at most $2N_t^2 + 6N_t$ flops [59].

Assuming that the channels $\mathbf{H}_k$ experienced by different users are statistically

Table 6.1: QR-BD Precoding Computational Cost

QR-BD	Step 1	$12K(K-1)^2N_r^2\left(\frac{(N_t-(K-1)}{3}\right)$
	Step 2	$8KN_rN_t(N_t-(K-1)N_r)$
	Step 3	$6K(9N_r^3+8N_r^2(N_T-(K-1)N_r)+4N_r(N_t-(K-1)N_r)^2)$
	Step 4	$2N_t^2+6N_t$
	Step 5	$KN_r+2KN^2+KN_rN_t(N_t-(K-1)N_r)$

independent, then

$$\mathbf{HV}=\begin{bmatrix}\mathbf{H}_1\mathbf{v}_1 & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & \mathbf{H}_K\mathbf{v}_K\end{bmatrix}$$

where $\bar{\mathbf{H}}_k=\mathbf{H}_k\mathbf{v}_k=[\bar{\mathbf{h}}_{k,1}\quad\bar{\mathbf{h}}_{k,2}\quad\dots\quad\bar{\mathbf{h}}_{k,\tilde{\mathrm{M}}_k}]\in\mathbb{C}^{Nr_k\times\tilde{\mathrm{M}}_k}$ represents the effective channel after precoding process that the receiver k only needs to estimate to decode the received signals based on any SM-based receiver. The received signal $\mathbf{y}_k\in\mathbb{C}^{Nr_k\times1}$ at the k'th user is given by

$$\mathbf{y}_k = \mathbf{H}_k\sum_{k=1}^{K}\sqrt{\rho_k}\mathbf{v}_k\mathbf{x}_k+\mathbf{n}_k \tag{6.12}$$

$$= \sqrt{\rho_k}\bar{\mathbf{H}}_k\bar{\mathbf{x}}_k+\mathbf{n}_k \tag{6.13}$$

where ρ_k is the transmit power for the k_{th} user. From (6.13), we can notice that the receive signal is similar to that signal received using original SM scheme. Although BS can rely on channel reciprocity and estimate the complete channel matrix, each user is however required to calculate its effective channel matrix $\bar{\mathbf{H}}_k$.

6.4 Spectral-Energy Efficiency Trade-off

6.4.1 Achievable Rate for SU-SM

Let us first discuss the achievable information rate (in bps/Hz) of point-to-point SU-SM scheme which is nevertheless subject to interpretation. It is interesting to know that the achievable rate of SM system has been analysed even before the

invention of SM. Thus, the authors of [142] discussed the achievable rate that can be obtained when there are a number of transmit antennas and cycling through all transmit antennas in sequence, using only one antenna at a time to transmit the signal. The achievable ergodic rate R_{SM} for a such system is given by

$$R_{\text{SM}} = \mathbb{E}\left[\frac{1}{N_t}\sum_{i=1}^{N_t}\log_2(1+\rho||\mathbf{h}_i||_F^2)\right] \tag{6.14}$$

where $\mathbf{h}_i \in \mathbb{C}^{Nr\times 1}$ represents a column from the channel matrix. The exception in (6.14) is taken with respect to the Frobenius norm of the channel column vector. In [143], the authors placed that the achievable rate is the sum of two terms, one account for signal constellation R_1 and another for spatial constellation diagrams R_2. However, assuming Gaussian input signals, then the impact of R_2 is very minor, thus the achievable rate is lower bonded by the R_1, i.e., (6.14).

Proposition 1. *In SM, if the number of transmit antenna is very large $N_t \to \infty$, the instantaneous rate approaches the ergodic rate, i.e.,*

$$\frac{1}{N_t}\sum_{i=1}^{N_t}\log_2(1+\rho||\boldsymbol{h}_i||_F^2) \xrightarrow{yields} R_{SM}. \tag{6.15}$$

Proving the *Proposition* 1 can be simply achieved by applying the law of a large number theorem to (6.14). Now, for a given N_t, by taking the expectation over all channels and assuming independent and identically distributed fading, then achievable rate of SM and SIMO are almost the same. This implies that, assuming one antenna at the receiver side, the rate of single-input-single-output (SISO) is a lower bound of the SM rate, i.e., $R_{\text{SM}} = R_{\text{SISO}}$. Therefore, even SM is capable of improving the transmission rate compared with SISO/MISO, The degrees of freedom is still limited to SISO/MISO.

6.4.2 Achievable Rate for Pre-coded SM

The achievable rate of proposed pre-coded SM, which is also applied to G-SM based schemes, is given by

$$R_{\text{pre-coded SM}} = \mathbb{E}\left[\frac{1}{M}\sum_{m=1}^{M}\left[\sum_{i=1}^{r}\log\left(1+\frac{\rho}{\tilde{\tilde{\text{M}}}}\lambda_i\right)\right]\right] \tag{6.16}$$

where $\lambda_1 \dots \lambda_m$ are the eigenvalues of $\bar{\mathbf{H}}_m \bar{\mathbf{H}}_m^H \in \mathbb{C}^{r \times r}$, while $r = \min(\tilde{\mathrm{M}}, N_r)$ is the additional degrees of freedom that proposed pre-coded SM can offer. In (6.16) the expectation is taken with respect to probability distribution of the ordered eigenvalue. Therefore, the rate is now bounded by the rate of $r \times r$ MIMO system.

6.4.3 Achievable Rate for MU-SM

Diagonalising the channel matrix using BD precoding provides multiple parallel independent SU-MIMO channel. Hence, by deploying SM-based decoding scheme, the achievable rate for each user is given by

$$R_k = \frac{1}{\tilde{\mathrm{M}}_k} \sum_{i=1}^{\tilde{\mathrm{M}}_k} \log_2(1 + \rho ||\bar{\mathbf{h}}_{k,i}||_F^2). \tag{6.17}$$

Assuming that there is K active users each with N_r antennas, with $N_t \geq KN_r$ and taking the summation of achievable rate to the parallel users yields the total achievable rate of the MU-SM system, i.e.,

$$\begin{aligned}
R_{\text{MU-SM}} &= (1-\alpha)\mathbb{E}\left[\sum_{k=1}^{K} R_k\right] && (6.18)\\
&\leq (1-\alpha)\mathbb{E}\left[\sum_{i=1}^{KN_r} \log\left(1 + \frac{\rho}{KN_r}\lambda_i\right)\right] && (6.19)\\
&\leq (1-\alpha)\int KN_r \log\left(1 + \frac{\rho}{KN_r}\lambda\right) f_{N_t,N_r}(\lambda)d\lambda && (6.20)\\
&= (1-\alpha)\int K \log\left(1 + \frac{\rho}{K}\lambda\right) f_{N_t=1,N_r}(\lambda)d\lambda && (6.21)
\end{aligned}$$

where $f_{N_t,N_r}(\lambda)$ is the distribution of unordered eigenvalues and it is given by [144]

$$f_{N_t,N_r}(\lambda) = \frac{1}{j}\sum_{i=1}^{j} \frac{i!}{(i+n-j)!}\left[L_i^{n-l}(\lambda)\right]^2 \lambda^{n-j} e^{-\lambda} \tag{6.22}$$

with $j = \min(N_T, KN_r)$ and $n = \max(N_T, KN_r)$, and $L_i^{n-m}(\lambda)$ is associated Laguerre polynomials of order i. It is obvious that rate of MU-SM is inferior to the optimum BD-based MIMO capacity. Inequalities in (6.19) and (6.20) hold because not all eigenvalues are utilised within each user's channel dimension while equality (6.21) holds by applying $j = 1$ and $n = N_r$. We can notice that

the degrees of freedom is increased by factor K. Moreover, selecting MU-SM or other MU-MIMO schemes means the trade-off between diversity and multiplexing gains. This because, in MU-SM, the transmit power is concentrated on a single eigenvalue and that strengthens the SNR received at the users. Whereas in others MU-MIMO schemes, the power is allocated among all available eigenvalues, thus providing a more multiplexing gain to the users.

6.4.4 Energy efficiency

In this sub-section, the spectral and energy efficiency analysis for MU-SM system is studied. The spectral efficiency is defined as the number of bits per second transmitted over a given bandwidth (in bps/Hz) while the energy efficiency is defined as the number of bits per joule unit transmitted over a given bandwidth (in bit/joules/Hz).

6.4.5 Power Consumption Model

Recall Section 2.5 in Chapter 2, the total power consumption P_{total} is given by

$$P_{\text{total}} = \frac{P_r}{\eta_{\text{PA}}} + BR\left(\frac{1}{\eta_{\text{load}}} + \frac{1}{\eta_{\text{BH}}}\right) + P_{\text{BBU-TXB}} + P_{\text{RRU-TXB}} + P_{\text{cpu}} + P_{\text{fix}} + P_{\text{cooling}} \tag{6.23}$$

where P_r, $P_{\text{BBU-TXB}}, P_{\text{RRU-TXB}}, P_{\text{cpu}}, P_{\text{fix}}$ and P_{cooling} refer to the total radiated power, BBU TXB power consumption, RRU TXB power consumption, CPU power consumption, fixed backhaul power consumption, and cooling power consumption, respectively. Also, η_{PA}, η_{load}, and η_{BH} are power amplifier efficiency, load consumption factor (in bps/Watt), and backhaul consumption factor (in bps/Watt), respectively.

6.4.6 Energy Efficiency Optimisation

The objective of optimisation problem is to optimally distributed the power $P_t(\lambda)^*$ on active eigenvalues over the time so then the energy efficiency EE (in bit/HZ

per Joule) can be maximised for a given spectral efficiency SE (in bps/Hz). The optimisation problem $(\mathcal{P})$ can be denoted as

$$(\mathcal{P}) \max_{P_t(\lambda)} EE = \frac{\int K\bar{N} \log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right) f_{\bar{N},N_r}(\lambda)d\lambda}{\frac{K\bar{N}\int P_t f_{\bar{N},N_r}(\lambda)d\lambda}{\eta_{\text{PA}}} + B\int K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right) f_{N_t,N_r}(\lambda)d\lambda\left(\frac{1}{\eta_{\text{load}}}+\frac{1}{\eta_{\text{BH}}}\right)+P_{\text{other}}} \tag{6.24}$$

$$\text{s.t.} \quad \int K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right) f_{\bar{N},N_r}(\lambda)d\lambda \geq SE \tag{6.25}$$

$$\sum_{k=1}^{K\bar{N}} P_t(\lambda_k) \leq P_{\max} \tag{6.26}$$

where N_0 is the noise power and $P_{\max}$ is the maximum sum transmit power that the BS able to use, P_{other} accounts for all other power consumptions i.e., $P_{\text{BBU-TXB}}, P_{\text{RRU-TXB}}, P_{\text{cpu}}, P_{\text{cooling}}$, and P_{fix}, that scaled only with the number of transmit antennas. $\bar{N}(1 \leq N \leq N_r)$ is the degrees of freedom that one user is able to achieve. Here we assumed that $M >> K$. The constraints of (6.28) and (6.29) represents the lower bound of spectral efficiency and the upper bound of sum transmit power. The solution of the above problem is nontrivial since the objective function is nonconcave. The objective function (6.24) is the ratio of two functions to be maximised. This type of optimisation problem is commonly known as fractional programming [80–84]. In Chapter 2, we have leaned an approach that deals with such kind of problems. The approach begins by separating the numerator and denominator with the help of parameter q. Therefore, assuming q is not negative, problem (6.24)–(6.29), can be transferred to another problem $(\bar{P})$ with parametric and non-fractional objective function, i.e.,

$$(\bar{P}) \max_{P_t(\lambda)} \bar{EE} = \int K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right) f_{\bar{N},N_r}(\lambda)d\lambda - q \times \left(\frac{\int P_t f_{\bar{N},N_r}(\lambda)d\lambda}{\eta_{\text{PA}}} + B\int K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right) f_{N_t,N_r}(\lambda)d\lambda\left(\frac{1}{\eta_{\text{load}}}+\frac{1}{\eta_{\text{BH}}}\right)+P_{\text{cons}}\right) \tag{6.27}$$

$$\text{s.t.} \quad \int K\bar{N} \log\left(1 + \frac{P_t(\lambda)}{N_0}\lambda\right) f_{\bar{N},N_r}(\lambda)d\lambda \geq SE \tag{6.28}$$

$$\sum_{k=1}^{K\bar{N}} P_t(\lambda_k) \leq P_{\max} \tag{6.29}$$

We know that, the method to solve problem $\bar{P}$ must ensure that the optimum value of $P_t(\lambda)$ resides between $P_{\min}$ and $P_{\max}$ to satisfy constraints (6.28) and (6.29). Here $P_{\min}$ refers to the minimum power required to achieve the target spectral efficiency SE. Now if we relax the problem $\bar{P}$ by removing constraint (6.29) and applying the Karush-Kuhn-Tucker (KKT) optimality conditions, the optimum solution can be then given by

$$P_t(\lambda)^* = \left[\frac{K\bar{N}\eta_{\text{PA}}(1 - Bq(\frac{1}{\eta_{\text{BH}}} + \frac{1}{\eta_{\text{load}}}) + \nu)}{q} - \frac{N_0}{\lambda}\right]^+ \tag{6.30}$$

Proof. See Appendix O. □

where ν is the nonnegative variables which can be determined by substituting (6.30) into the constraints (6.28). Once achieving this, we must now identify if the optimum solution $P_t(\lambda)^*$ falls within the interval of $[P_{\min}, P_{\max}]$. If it does not, then the solution is infeasible. We refer to the $[P_{\min}, P_{\max}]$ as the feasible region of the sum transmit power. Nevertheless, expression (6.30) has a water-filling fashion in which q is equivalent to water level. The solution can be seen as an energy-efficient version of well-know water-filling power allocation algorithm. Meanwhile, **Algorithm** 3 is employed to determine both q and ν. The algorithm is based on the application Dinkelbach's method [84]. Once both q and ν are calculated, the BS regulates the transmit power on each eigenvalue, providing that the solution is within the feasible region of the su transmit power.

6.4.7 Simulation Results and Discussions

In this section, the performance of MU-SM in terms of average BER, achievable rate, and the spectral-energy efficiency trade-off is evaluated using Monte Carlo simulations. In the simulation, a flat Rayleigh fading that is corrupted by AWGN

Algorithm 3

Require: $P_t \Leftarrow P_t^{(0)}$, $M, K, \bar{N}, f_{\bar{N},N_r}(\lambda)$.

1: **while** $|Z| > \delta$ **do**

2: $\quad q = \frac{\int K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right)f_{\bar{N},N_r}(\lambda)d\lambda}{\frac{\int P_t f_{\bar{N},N_r}(\lambda)d\lambda}{\eta_{\text{PA}}}+B\int K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right)f_{N_t,N_r}(\lambda)d\lambda\left(\frac{1}{\eta_{\text{load}}}+\frac{1}{\eta_{\text{BH}}}\right)+P_{\text{other}}}$

3: $\quad$ **Find** ν by solving

4: $\quad \int K\bar{N}\log\left(\frac{K\bar{N}\eta_{\text{PA}}(1-Bq(\frac{1}{\eta_{\text{BH}}}+\frac{1}{\eta_{\text{load}}})+\nu)\lambda}{q\bar{N}_0}\right)f_{\bar{N},N_r}(\lambda)d\lambda = SE$

5: $\quad P_t^* = \left[\frac{K\bar{N}\eta_{\text{PA}}(1-Bq(\frac{1}{\eta_{\text{BH}}}+\frac{1}{\eta_{\text{load}}})+\nu)}{q} - \frac{N_0}{\lambda}\right]^+$

6: $\quad Z = \mathbb{E}\left[K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right)\right] - q\left(\frac{\mathbb{E}[P_t]}{\eta_{\text{PA}}} + \mathbb{E}\left[K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right)\right]\left(\frac{B}{\eta_{\text{load}}}+\frac{B}{\eta_{\text{BH}}}\right)+P_{\text{cons}}\right)$

7: **end while**

8: $q \Leftarrow q^*, \nu \Leftarrow \nu^*$.

is assumed. The receive antennas are assumed to be separated wide enough to avoid correlation.

In Fig. 6.6, the robustness of the proposed MU-SM based on BD precoding scheme is studied against others MU-MIMO schemes. We assume that there are 5 active users each with 4 receive antennas. The number of transmit bits for each user, i.e., b_k, is assumed to be 10 bits. We will compare the performance of MU-SM with the performance of SU-SM, ZF-based MU-MIMO, and BD-based MU-MIMO with two different detection algorithms, i.e., V-BLAST and ZF. For the V-BLAST scheme, the SNR is assumed to be known at each user when using MMSE for the signal detection. In the case of SU-SM, the BS picks up the best user which has the maximum norm channel gain using a TDMA scheme. For a fair comparison, all users are assumed to have the same average SNR. Furthermore, although a better performance can be achieved with water-filling, considering equal power allocation makes the comparison, in this example, between SU-SM and MU-SM fair. All schemes show approximately similar performance at the low SNR (SNR<5 dB), assuming a fixed transmission rate per user. MU-SM scheme outperforms both BD-ZF and ZF MU-MIMO when SNR> 5dB. Additionally, MU-SM starts to show noticeably better performance compared with BD-based with V-BLAST detection scheme specifically when SNR> 10dB. We can also

notice that, at high SNR values, the users in an MU-SM system lags by 1 dB behind that of the TDMA-based SU-SM scheme. However, when we consider that in MU-SM scheme there is actually 5 active users,

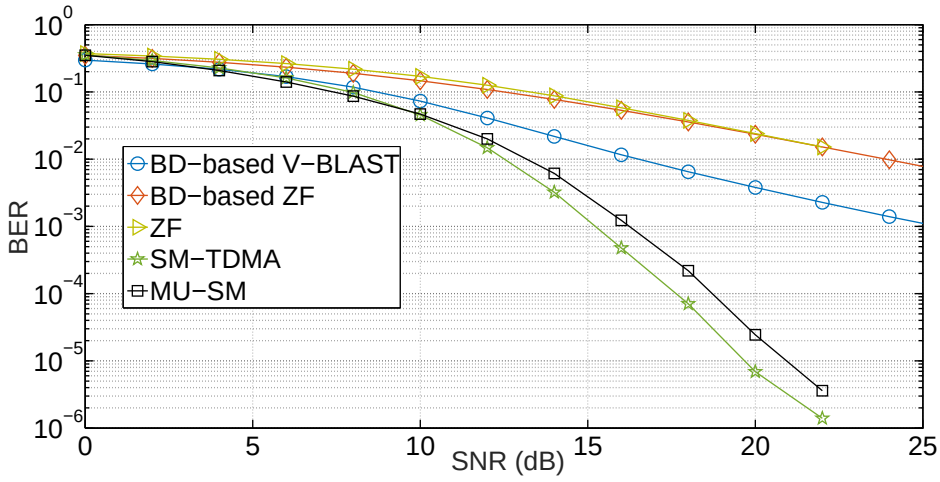

Figure 6.6: BER performance comparison among different MU-MIMO schemes ($K = 5$, b_k= 10 bits, and N_r=4).

each receiving 10 bits, we can then realize that this loss of 1 dB is inconsequential compared to the gain achieved to the total transmission rate.

Fig. 6.7 compares the ergodic achievable rate achieved by MU-SM compared with the SU-SM and BD-based MU-MIMO schemes. The results are obtained with $K = 4$, $N_t = 8$, and $N_r = 2$. The ratios of uplink and downlink phases are assumed to be 10% and 90%, respectively. The simulation results shows that MU-SM scheme achieves a significant part of the BD-based MU-MIMO scheme achievable rate and bound in (6.20) is tight for low and medium SNRs. As SNRs goes to high SNR, the ergodic achievable rates of MU-SM and BD-based MU-MIMO schemes both increase with fixed slope. As SNR increases, the achievable rate is noticeable compared with SU-SM because of the additional degrees of freedom that the proposed scheme offers.

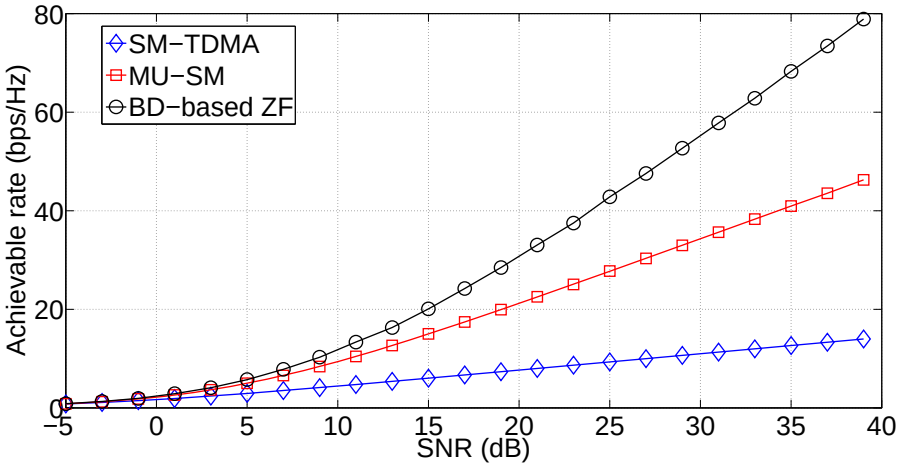

Figure 6.7: Ergodic achievable rate of MU-SMS, TDMA-based SM, and BD-based ZF MU-MIMO in Rayleigh fading channel (N_t=8, $K = 4$, $N_R = 2$, $\tilde{\mathrm{M}}_k = 2$).

Fig. 6.8 shows the achievable spectral efficiency (in bps/Hz per user) and energy efficiency for different number of transmit antennas N_t under MU-SM scheme. For each point, the transmit power is optimized according to energy efficiency iterative water-filling method. In this example, only pathloss is considered for the large scale fading model and the pathloss model set to $128.1+37.6\log_{10}(d)$, with d is the distance in kilometer. The carrier frequency, transmission bandwidth, and noise density N_0 are set to 2GHz, 100MHz, and -174dBm/Hz, respectively. The simulation parameters that are relevant to the PCM are given in Table 6.2. These parameters has been taken from prior works of [47, 50–57]. The optimum energy efficiency can be achieved when the spectral efficiency per given user is 5.67 bps/HZ and N_t is 44. We can also notice that achieving spectral efficiency of more than 8 bps/Hz per user is not possible due to the constraint on maximum transmit power.

Table 6.2: Multi-user SM Simulation Parameters.

	Parameter	Value		Parameter	Value
BBU	Performance per Watt: η_{cpu}	14 GFLOPS /Watt	**RRU**	Power efficiency: η	0.34
	Load consumption factor: η_{load}	1 Gbps/Watt		Figure-of-merit: FOM	10^{-8}
	RRM: P_{RRM}	1E-3 Watt		LMA power gain: η_{LAN}	10
	Circuit power: $P_{\text{BBU-crt}}$	11 Watt		Noise power: BN_0	-90dBm
Backhaul	Fixed power: P_{fix}	5 Watt		Circuit power: $P_{\text{RRU-crt}}$	13 Watt
	Load factor: η_{BH}	0.4 Gbps/Watt	**Cooling**	Power: P_{cooling}	8%

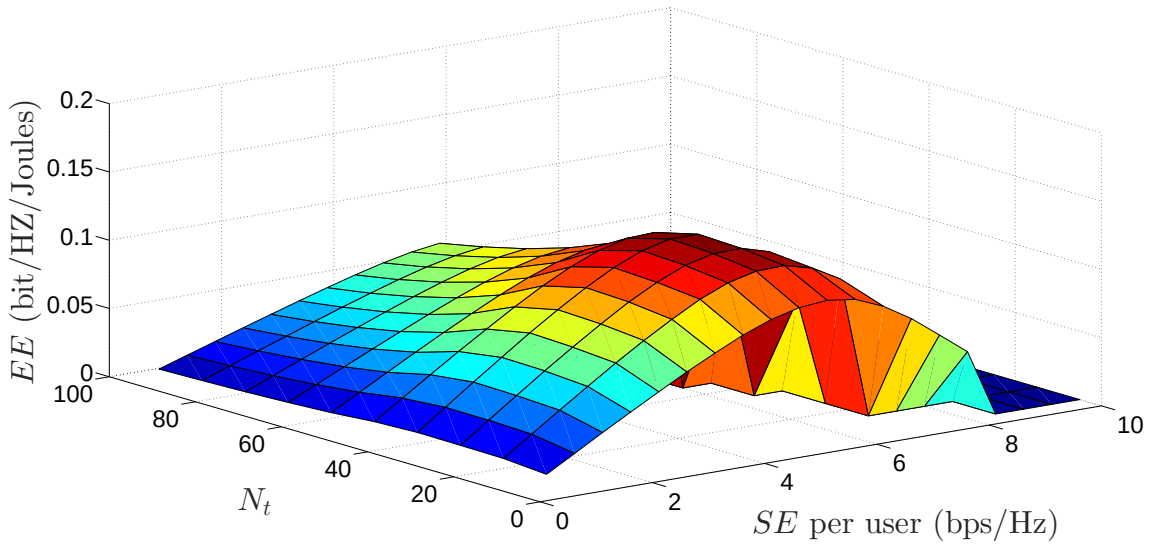

Figure 6.8: Spectral-energy efficiency trade-off against different numbers of transmit antennas, $N_t > K + 1$, $N_r = 2$, $P_{\max} = 100$ Watt, and $\tilde{\text{M}}_k = 2$.

6.5 Chapter Summary

This chapter has proposed two SM-like transmission schemes which can be considered as promising transmission schemes and likely candidates for the next-generation MIMO-based cellular systems because of their ability to offer more degrees of freedom. It has first proposed the pre-coded SM to be an additional option that suits SU-MIMO systems. The chapter has also proposed a transmission structure that enables the SM scheme to be adopted in an MU-MIMO system, thus improving the spectral efficiency using the same low-complex SM detector. Finally, the chapter has evaluated the spectral-energy efficiency trade-off for the proposed multi user SM system, using an iterative energy-efficient water-filling algorithm to optimise the transmit power.

Chapter 7

Conclusions and Future Work

The emergence of easy-to-use smartphones has lead to an exponential increase in annual data traffic. To address this challenge, an innovative cellular system (i.e., 5G) is required. The 5G mobile cellular system should be designed with new technologies to be able to creatively manage the traffic demands from different perspectives. To achieve this intended objective, however, a certain prerequisite is important to consider. Each year there is an expected traffic growth level based upon a data traffic forecast. Therefore, it is important to specify the appropriate spectral efficiency predicted to manage this level of traffic, whilst as the same time being mindful of the energy consumption used. In other words, the spectral-energy efficiency tradeoff tool is required to analyse the energy consumption of a wireless technology for a given required spectral efficiency. Using this as a guide, new technologies should be developed with the objective of enhancing the energy efficiency, with the proviso that the spectral efficiency delivered is able to manage the level of traffic predicted in a forecast graph.

In this thesis, we analyse the spectral-energy efficiency trade-off for four advanced technologies, i.e., OFDMA, MFemtocell, CR, and enhanced SM-based schemes. In this concluding chapter, we summarise the key findings of the chapters and suggest some interesting ideas for future research directions.

7.1 Summary of Results

Chapter 3 has studied the sub-channel and bit allocation scheme which minimises the energy per transmitted bit for MU-OFDMA systems. An optimisation problem with an energy efficiency objective function can be reformulated to an equivalent problem with a non-fractional function. The equivalent problem can then be solved by using a parametric program. This reduces the complexity of the original problem. By exploiting the integer linear fractional programming, a novel iterative RRA algorithm has been proposed. The objective of the proposed algorithm is to minimise the system energy efficiency subject to a given spectral efficiency and QoS constraints. It is worth mentioning that the proposed algorithm can also be tweaked to be applied to the energy-efficient optimisation problem with an objective of maximising the number of bits per joule. The simulation results have demonstrated the potential benefits of having multiple users in enhancing the energy efficiency. The simulation results have also highlighted the impact of the users' location on the achieved energy efficiency. More interestingly, the simulation results have shown that the users' QoS requirement has a minor impact on the energy efficiency if a large spectral efficiency is needed.

Chapter 4 has introduced the concept of the MFemtocell. It has studied the spectral and energy efficiency for the MFemtocell system with two resource partitioning schemes. It has been shown that the energy efficiency can be enhanced by using the non-orthogonal scheme for the moderate and high SNR values. The chapter has also investigated the spectral efficiency of two resource partitioning schemes in a system level MFemtocell in the presence of opportunistic scheduling. It has been shown that the spectral efficiency can be improved by employing the non-orthogonal partitioning scheme to share the spectrum between the BS and MFemtocells. The purpose of using the non-orthogonal transmission scheme is that the access users are considered indoor and that they are within close proximity to the MFemtocell station. Although the non-orthogonal scheme avoids the interference, it comes at the price of reduced spectral efficiency. This makes the non-orthogonal transmission scheme more attractive than the orthogonal trans-

mission scheme, even in the absence of any kind of coordination between the BS and MFemtocells. This chapter has also demonstrated that the implementation of MFemtocells can reduce the signalling overhead resulting in an enhanced system performance. In summary, the MFemtocell can be considered as a potential candidate technology to be deployed in 5G mobile cellular system, to increase the performance for users within vehicles as well as the overall system.

Chapter 5 has investigated the spectral and energy efficiency for interference-tolerant CR networks. The initial analysis has studied the spectral-energy efficiency trade-off for an SUCR network under transmit power and interference constraints. In the low SNR regime, transmitting a signal with average power constraint provides better energy efficiency than transmitting a signal with peak power constraint. In addition to that, the interference channel has no impact on $\left(\frac{E_b}{N_0}\right)_{\min}$ which is required for reliable communications. In the high SNR regime, however, transmitting signals with either power constraint gives the same energy efficiency.

We have also proposed a CR-based cellular network in which a secondary network shares a spectrum belonging to an indoor system. This chapter has also demonstrated that with CR technology, cellular operators can share their spectrum opportunistically with each other to increase the performance of their network. One way to do so is to share a spectrum in the uplink phase of an indoor system. This is indeed an opportunity to make the implementation of the CR-based cellular network more feasible in the near future without the necessity of modification at the end user's handset. An example of a practical application of CR and its integration with existing technology, is the use of carrier aggregation applied with CR to allow greater spectrum accessibility. It manages to remain efficient by using the underutilised spectrum. The challenge is how to practically estimate the interference channels by the STs. Relying on channel reciprocity or broadcasting ICSIs can give some insight to solving this issue.

The spectral and energy efficiency of the proposed network have been analysed. By adopting the extreme value theory, we have derived the spectral efficiency of the MU-CR network under optimal power allocation. We have studied the

impact of MU diversity gain in both the primary and secondary receivers on the spectral and energy efficiency. The spectral efficiency of the CR network is relatively large while the number of PRs is small. The spectral efficiency, however, diminishes rapidly with the increase in the number of PRs. This degradation can be compensated for by relaxing the interference threshold or by increasing the number of SRs that are within a short distance from the ST.

Chapter 6 has worked on the evolution of the SM scheme. The SM standing alone may be incapable of providing higher spectral efficiency and may not be a potential solution in managing the increasing data traffic. To overcome this limitation and keep the same level of low decoding complexity, we have proposed two SM-like transmission schemes. Both can be considered as promising transmission proposals and likely candidates for the 5G mobile cellular system. The first scheme, called pre-coded SM, suits SU-MIMO systems. Here, the simulation results have demonstrated that, for a given high transmission rate, the pre-coded SM outperforms the optimal SM, MA-SM and V-BLAST schemes. The second scheme allows the same SM-like decoder to be adapted in the MU-MIMO environment with the aid of BD precoding. Using BD precoding is an attractive technique that transforms the MU-MIMO channel into parallel SU-MIMO channels. In this case, the transmitter transmits signals simultaneously to multiple users. The users will then able to decode the received signal using the original SM scheme. This provides more degrees of freedom whilst keeping the decoding complexity at its low point. Hence, it is able to achieve a higher spectral efficiency by using the same low-complex SM decoder. This chapter has also examined the spectral-energy efficiency for the proposed MU-SM system by adopting the energy-efficient water-filling algorithm. The analysis has been conducted using a much more realistic PCM. The simulation results have shown that an optimum energy efficiency with high spectral efficiency can be achieved by deploying massive MIMO with a relatively large number of active users.

Overall, the studies in Chapter 3 through Chapter 6 have demonstrated how new advanced radio technologies, such as MFemtocell, CR and SM-based massive MIMO, are required to offer greater spectral efficiency. The studies have shown how the spectral-energy efficiency trade-off is an important tool that should be taken into account when designing the 5G mobile cellular system. This tool helps to understand and minimise the energy consumption whilst providing a certain spectral efficiency which is able to manage a certain level of traffic. The research conducted in this thesis is of great practical significance in assisting the deployment of the next-generation network in future green wireless communication systems.

7.2 Future Research Topics

For the investigation on the energy-efficient resource allocation algorithm, this thesis has only focused on the Dinkelbach paramedic approach to solve the optimisation problem. However, it is worth expanding the investigation to include more fractional programming approaches, e.g., parameter-free convex optimisation, in the future work. In investigation of such techniques, the focus would be on finding the most sub-optimal algorithm that can find a near-optimal solution for the resource allocation whilst keeping the complexity as low as possible.

For the MFemtocell concept, the proposed scheme is based on the relaying method to carry the traffic over the backhaul link. It is worth applying other backhauling methods and compare their performance with each other. Another interesting future work is how to allow MFemtocell architecture to be integrated into the intelligent transportation system (ITS). This is especially beneficial for assisting with vehicle traffic light control in relaying the important traffic messages through the cellular system.

For the work on the CR network, this thesis considers only one registered CR network with a single secondary cell. One possible future research direction would be to extend the work by considering multiple registered CR networks that are

all trying to share a spectrum which belongs to a primary network. These CR networks must make sure that the aggregated interference is below the allowable level. It would be worthwhile to investigate the impact of using a coordination scheme for the CR network usage of the primary network spectrum on the system performance, and compare it to a random allocation scheme, i.e., the absence of any kind of coordination.

Furthermore, the work in this thesis assumes that the massive MIMO system uses the TDD scheme to enable the acquisition of CSIs to be achieved via the uplink pilot signals. However, the assumption here is that the uplink pilot signals are fully orthogonal among the users. Lack of fully orthogonal pilots and the growth of pilot resources along with the increase in the number of users create a bottleneck that limits achieving the full benefit of the massive MIMO system. An interesting area for future work is to combine the massive MIMO system with device-to-device (D2D) communication. In such a case, the creation of a single virtual massive MIMO is possible where the BS is communicating with a group of users, virtually seen as one user that is equipped with large-scale antennas. D2D communication is an innovative technology which enables direct communication between nearby mobiles without the need of a cellular infrastructure. The D2D has already been considered to be a part of the 3GPP LTE-Advanced cellular system. Using this system can be deployed in both the TDD and FDD systems. It is worth investigating the spectral-energy efficiency trade-off for this system and compare the outcome with one of the MU massive MIMO systems that uses a precoding scheme to mitigate the inter-user interference.

Appendix A

Proof of Theorem 1

Let the vector **S** is the feasible solution for problem P, then we have

$$q^* = \frac{f(x^*)}{g(x^*)} \leq \frac{f(x)}{g(x)}, \quad \forall x \in \mathbf{X} \tag{A.1}$$

this leads to

$$f(x) - q^* g(x) \geq 0 \tag{A.2}$$

and concludes a fact that

$$\min_x \{f(x) - q^* g(x)\} = 0. \tag{A.3}$$

Now if the x^* is the optimum solution, then the following is right:

$$f(x) - q^* g(x) \geq f(x^*) - q^* g(x^*) = 0 \tag{A.4}$$

which proves **Theorem** 1.

Appendix B

Derivation of (4.9)

The achievable spectral efficiency C, of the direct transmission scheme is given by

$$C = \mathbb{E}\left[\log_2\left(1 + \frac{|h_d|^2 P_{\mathrm{BS}}}{BN_0}\right)\right] \tag{B.1}$$

where $\mathbb{E}[\cdot]$ is the expectation operator. If we assume that $|h_{\mathrm{d}}|$ follows a Rayleigh fading distribution, $|h_d|^2$ has an exponential distribution (i.e., $e^{-A_d x}$, $\forall x > 0$) with $E[|h_d|^2] = A_d$ denoting the mean power of the channel. By letting SNR $\equiv \frac{P_{\mathrm{BS}}}{BN_0}$, (2.2) and (2.3) can be written as

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{\ln 2}{\mathbb{E}[|h_d|^2]} = -1.59 + 10\log_{10}(A_d) \quad \mathrm{dB} \tag{B.2}$$

$$S_0 = \frac{2}{\kappa(|h_d|)} \tag{B.3}$$

respectively. In (B.3), $\kappa(|x|)$ is kurtosis of random variable x defined as [39]

$$\kappa(|x|) = \frac{\mathbb{E}[|x|^4]}{\mathbb{E}[|x|^2]^2}. \tag{B.4}$$

For Rayleigh fading, $\kappa(|x|)$=2 and consequently $S_0 = 1$. We can notice that $\left(\frac{E_b}{N_0}\right)_{\min}$ is the same in AWGN and Rayleigh fading channels, but the slope, S_0, has a 2.5 dB difference. Applying (B.2) and (B.3) into (2.1), (4.9) is then obtained.

Appendix C

Derivation of (4.10)

Refereeing to Fig. 4.2 (a), the MFemtocell will receive the data in the T_1 and then transmit the data again to user in the T_2. The achievable spectral efficiency C over two time slots is limited to minimum capacity, i.e.,

$$C = \min\{C_b, C_a\} \tag{C.1}$$

where C_b and C_a are the spectral efficiency of the backhaul and access links, respectively. However, by considering that the distance between the access user and MFemtocell is much shorter that the distance between the BS and MFemtocell, we can then expect that the spectral efficiency is always limited by the backhaul spectral efficiency. The spectral efficiency of the backhaul link can be calculated by

$$C_b = \frac{1}{2}\mathbb{E}\left[\log_2\left(1 + \frac{\alpha|h_b|^2 P}{BN_0}\right)\right]. \tag{C.2}$$

The factor $\frac{1}{2}$ in (C.2) is due to the fact that the transmission to the end user occurs in two successive time slots. Moreover, both the BS and MFemtocell transmit only half of the time. Using (2.2), $\left(\frac{E_b}{N_0}\right)_{\min}$ can be calculated by

$$\left(\frac{E_b}{N_0}\right)_{\min} = \left(\frac{E_b}{N_0}\right)^b_{\min} = \frac{2\ln 2}{\alpha G A_d}. \tag{C.3}$$

The minimum normalised energy is characterised by the backhaul gain and the fraction of the transmit power allocated to the BS. By substituting (2.3), the slope of the spectral efficiency for the orthogonal scheme is expressed by

$$S_0 \quad = \quad (S_0)_b = \frac{1}{\kappa(|h_{\mathrm{d}}|)}. \tag{C.4}$$

Substituting (C.3) and (C.4) into (2.1), (4.10) can be obtained.

Appendix D

Derivation of (4.11)

In the non-orthogonal partitioning scheme, both BS and MFemtocell can share the spectrum of T_2 to serve direct transmission and access users but at the price of introducing additional interference to both users. The access user, however, would experience the burden of the majority of the interference. Hence, the impact of the MFemtocell interference on the direct transmission users can be ignored and considered as background noise by assuming that the MFemtocell and the user are a long distance apart. Nevertheless, the achievable spectral efficiency is equal to

$$C_{\text{non-orthog}} = \min\{C_b, C_a\} + C_d \tag{D.1}$$

where $C_d = \frac{1}{2}\mathbb{E}\left[\log_2\left(1 + \frac{|h_d|^2 P_{\text{BS}}}{2BN_0}\right)\right]$ is the spectral efficiency of the direct transmission link. Here, it has been assumed that the transmit power of the BS is shared equally between the MFemtocell and direct transmission user. Again, the factor $\frac{1}{2}$ is because the direct transmission happens only in the second time slot. The spectral efficiency over the access link, C_a, and the backhaul link are given by

$$\begin{aligned} C_a &= \frac{1}{2}\mathbb{E}\left[\log_2\left(1 + \frac{(1-\alpha)|h_{\text{a}}|^2 P}{|h_{\text{d}}|^2 P_{\text{BS}} + BN_0}\right)\right] \\ &= \frac{1}{2}\mathbb{E}\left[\log_2\left(1 + \frac{(1-\alpha)|h_{\text{a}}|^2 \text{SNR}}{|h_{\text{d}}|^2 \text{SNR} + 1}\right)\right] \end{aligned} \tag{D.2}$$

and

$$C_b = \frac{1}{2}\mathbb{E}\left[\log_2\left(1 + \frac{\alpha|h_{\text{b}}|^2 P}{2BN_0}\right)\right] \tag{D.3}$$

respectively. Substituting (D.1)–(D.3) into (2.2), $\left(\frac{E_b}{N_0}\right)_{\min}$ can be obtained as

$$\begin{aligned}\left(\frac{E_b}{N_0}\right)_{\min} &= \frac{\ln 2}{\frac{\partial\left(\frac{1}{2}\mathbb{E}\left[\min\left\{\log_2\left(1+\frac{(1-\alpha)|h_{\rm a}|^2{\rm SNR}}{|h_{\rm d}|^2{\rm SNR}+1}\right),\log_2(1+0.5\alpha|h_{\rm b}|^2{\rm SNR})\right\}+\log_2(1+|h_d|^2{\rm SNR})\right]\right)}{\partial{\rm SNR}}} \\ &= \max\left\{\frac{4\ln 2}{(\alpha G+1)A_d}, \frac{2\ln 2}{(1-\alpha)A_a+A_d)}\right\}, \quad {\rm SNR}=0 \qquad \text{(D.4)}\end{aligned}$$

where $A_a = \mathbb{E}[|h_a|^2]$ denotes the mean power of the access channel. The wide-band slope S_0 can be obtained by substituting (D.1)–(D.3) into (2.3) and can be expressed by

$$\begin{aligned}S_0 &= \frac{\frac{\partial\left(\mathbb{E}\left[\min\left\{\log_2\left(1+\frac{(1-\alpha)|h_{\rm a}|^2{\rm SNR}}{|h_{\rm d}|^2{\rm SNR}+1}\right),\log_2\left(1+0.5\alpha|h_{\rm b}|^2{\rm SNR}\right)\right\}+\log_2\left(1+|h_d|^2{\rm SNR}\right)\right]\right)}{\partial{\rm SNR}}}{-\frac{\partial^2\left(\mathbb{E}\left[\min\left\{\log_2\left(1+\frac{(1-\alpha)|h_{\rm a}|^2{\rm SNR}}{|h_{\rm d}|^2{\rm SNR}+1}\right),\log_2(1+0.5\alpha|h_{\rm b}|^2{\rm SNR})\right\}+\log_2(1+|h_d|^2{\rm SNR})\right]\right)}{\partial{\rm SNR}^2}} \\ &= \frac{(G\alpha+1)^2}{(G^2\alpha^2+1)\kappa(|h_{\rm d}|)}, \quad {\rm SNR}=0. \qquad \text{(D.5)}\end{aligned}$$

Putting (D.4) and (D.5) into (2.1) leads to (4.11).

Appendix E

Derivation of (4.12)

For direct transmission user, in the high SNR regime, the slope of spectral efficiency is equivalent to

$$S_{\infty} = \lim_{\mathrm{SNR}\to\infty} \mathrm{SNR}\frac{\partial\left(\log_2(1+|h_{\mathrm{d}}|^2\mathrm{SNR})\right)}{\partial \mathrm{SNR}}. \tag{E.1}$$

The asymptotic penalty in the high SNR regime in (2.6) can be calculated by

$$\begin{aligned}\left(\frac{E_b}{N_0}\right)_{\text{penalty}} &= \lim_{\mathrm{SNR}\to\infty} \log_2\left(\frac{\mathrm{SNR}}{1+|h_{\mathrm{d}}|^2\mathrm{SNR}}\right) \\ &= -\mathbb{E}[\log_2(|h_{\mathrm{d}}|^2)]. \end{aligned}\tag{E.2}$$

Since $|h_{\mathrm{d}}|^2$ has an exponential probability density function, i.e., $e^{-A_d x}$, (E.2) can be re-written as

$$\left(\frac{E_b}{N_0}\right)_{\text{penalty}} = -\int_0^{\infty} e^{-A_d x}\log_2(x)dx = \frac{\gamma}{\ln 2} + \ln(A_d) \tag{E.3}$$

where Υ is the Euler-Mascheroni constant. Putting (E.1) and (E.3) into (2.4), (4.12) can be obtained.

Appendix F

Derivation of (4.13)

For the orthogonal scheme, the slope in the high SNR regime is equal to

$$S_\infty = \lim_{\mathrm{SNR}\to\infty} \mathrm{SNR}\frac{\partial\left(\min\left\{\frac{1}{2}\log_2(1+2\alpha G|h_d|^2\mathrm{SNR}), \frac{1}{2}\log_2(1+2(1-\alpha)|h_\mathrm{a}|^2\mathrm{SNR})\right\}\right)}{\partial \mathrm{SNR}} \tag{F.1}$$

while $\left(\frac{E_b}{N_0}\right)_{\text{penalty}}$ is given by

$$\begin{aligned}\left(\frac{E_b}{N_0}\right)_{\text{penalty}} &= \max\left\{\lim_{\mathrm{SNR}\to\infty}\log_2\left(\frac{\mathrm{SNR}}{1+2\alpha G|h_d|^2\mathrm{SNR}}\right), \lim_{\mathrm{SNR}\to\infty}\log_2\left(\frac{\mathrm{SNR}}{1+2(1-\alpha)|h_\mathrm{a}|^2\mathrm{SNR}}\right)\right\} \\ &= \max\left\{-(\log_2(2\alpha G A_d)+\frac{\gamma}{\ln 2}), -(\log_2(2(1-\alpha)A_a)+\frac{\gamma}{\ln 2})\right\}.\end{aligned} \tag{F.2}$$

Substituting (F.1) and (F.2) into (2.4), (4.13) can be obtained.

Appendix G

Derivation of (4.14)

For the non-orthogonal scheme, the slope of the spectral efficiency is give by

$$S_\infty = \lim_{\mathrm{SNR}\to\infty} \mathrm{SNR} \frac{\left\{ \begin{array}{c} \partial\left(\mathbb{E}\left[\min\left\{\log_2\left(1+\frac{(1-\alpha)|h_\mathrm{a}|^2\mathrm{SNR}}{|h_\mathrm{d}|^2\mathrm{SNR}+1}\right), \log_2\left(1+\alpha|h_\mathrm{b}|^2\mathrm{SNR}\right)\right\}\right]\right) \\ +\log_2\left(1+|h_d|^2\mathrm{SNR}\right) \end{array} \right\}}{\partial \mathrm{SNR}}. \tag{G.1}$$

Also, (2.6) can be calculated as

$$\left(\frac{E_b}{N_0}\right)_{\text{penalty}} = -0.5\log_2(\alpha G A_d) + \frac{\gamma}{\ln 2}. \tag{G.2}$$

Substituting (G.1) and (G.2) into (2.4), (4.14) can be obtained.

Appendix H

Proof of Theorem 5.5

The minimum energy efficiency $\left(\frac{E_b}{N_o}\right)_{\min}$ occurs when γ_{avg} approaches zero, i.e.,

$$\left(\frac{E_b}{N_0}\right)_{\min} = \lim_{\gamma_{\text{avg}}\to 0} \frac{\gamma_{\text{avg}}}{N_0 C} = \lim_{\gamma_{\text{avg}}\to 0} \frac{\mathbb{E}[\gamma_s^*(g_c, g_i)]}{N_0 \mathbb{E}[\log_2(1 + \frac{\gamma_s^*(g_c,g_i)g_c}{N_0})]} \tag{H.1}$$

where the expectation in the denominator and nominator is with respect to two random variables, i.e. g_i and g_c. We can notice from (5.4) that γ_{avg} vanishes when γ_0 approaches $g_{c(\max)}$. Then, (H.1) can be re-written as

$$\lim_{\gamma_0\to g_{c(\max)}} \frac{\int_{\gamma_0}^{g_{c(\max)}} \left(\int_0^{\frac{Q}{\frac{1}{\gamma_0}-\frac{N_0}{g_c}}} \left(\frac{1}{\gamma_0} - \frac{N_0}{g_c}\right) + \int_{\frac{Q}{\frac{1}{\gamma_0}-\frac{N_0}{g_c}}}^{\infty} \frac{Q}{g_i} \right) f(g_c)f(g_i)dg_c dg_i}{N_0 \int_0^{g_{c(\max)}} \left(\int_0^{\frac{Q}{\frac{1}{\gamma_0}-\frac{N_0}{g_c}}} \log_2\left(\frac{g_c}{\gamma_0 N_0}\right) + \int_{\frac{Q}{\frac{1}{\gamma_0}-\frac{N_0}{g_c}}}^{\infty} \log_2(1 + \frac{g_c Q}{g_i N_0}) \right) f(g_c)f(g_i)dg_c dg_i}. \tag{H.2}$$

Applying L'Hpital's Rule into (H.2) and following Leibniz integral rule, $\left(\frac{E_b}{N_o}\right)_{\min}$ is calculated by

$$\left(\frac{E_b}{N_o}\right)_{\min} = \lim_{\gamma_0\to g_{c(\max)}} \frac{\left(\frac{1}{\gamma_0^2}\right) \int_{\gamma_0}^{g_{c(\max)}} \int_0^{\frac{Q}{\frac{1}{\gamma_0}-\frac{N_0}{g_c}}} f(g_c)f(g_i)dg_c dg_i}{\frac{1}{\gamma_0 \ln 2} \int_{\gamma_0}^{g_{c(\max)}} \int_0^{\frac{Q}{\frac{1}{\gamma_0}-\frac{N_0}{g_c}}} f(g_c)f(g_i)dg_c dg_i}. \tag{H.3}$$

Expression (5.5) can then be obtained after applying $\gamma_0 \to g_{c(\max)}$.

Appendix I

Derivation of (5.6)

Let us first rewrite (5.1) as

$$C = \int_0^\infty \int_0^\infty \log_2\left(1+\frac{g_c\gamma_s^*(g_c,g_i)}{N_0}\right) f(g_c)f(g_i)dg_c dg_i. \quad \text{(I.1)}$$

In the high SNR regime, for sufficiently large transmit power, $\gamma_s(g_c,g_i))^*$ will be equal to $\frac{Q}{g_i}$. Let $x=\frac{g_c}{g_i}$ and $y=g_i$. In this case, $f(x)=\int_0^\infty y f_{g_c}(xy)f_{g_i}(y)dy$. If both the cognitive and interference channels follow Rayleigh distribution (i.e., both g_c and g_i would be exponentially distributed with unit-mean), $f(x) = \int_0^\infty y\exp(-xy)\exp(-y)dy = \frac{1}{(1+x)^2}$. Therefore,

$$\begin{aligned} C &= \int_0^\infty \log_2\left(1+\frac{xQ}{N_o}\right) f(x)dx \\ &= \int_0^\infty \log_2\left(1+\frac{xQ}{N_o}\right)\frac{1}{(1+x)^2}dx. \end{aligned} \quad \text{(I.2)}$$

By applying integration by parts method to (I.2), (5.6) can be obtained.

Appendix J

Derivation of (5.13)

The distance from the ST to a given SR is an independent and identical random variable following a uniform distribution. The probability that $R \leq r$ holds is given by [145]

$$F_d(r) = \begin{cases} \frac{2r}{d^2 - d_{\min}^2} & d_{\min} \leq r \leq d \\ 0 & \text{otherwise} \end{cases} \tag{J.1}$$

where $d_{\min}$ $(0 \leq d_{\min} \leq d)$ is the minimum distance between the ST and a SR. The cumulative distribution function (CDF) of pathloss in the dB scale can be calculated by

$$F_L(y) = \Pr\{a \log(Ar) \leq y\} = \int_{d_{\min}}^{\frac{e^{\left(\frac{y}{a}\right)}}{A}} \frac{2r}{d^2 - d_{\min}^2} dr. \tag{J.2}$$

The PDF, i.e., $\frac{dF_L(r)}{dr}$, is then given by

$$f_L(y) = \frac{2e^{\left(\frac{2y}{a}\right)}}{aA^2(d^2 - d_{\min}^2)} \tag{J.3}$$

where $a = \xi\beta$. The channel gain $g_c(n) = \frac{S(n)}{g_{\text{P}}(n)}$, where $S(n) = g_m(n)g_s(n)$ is a random variable derived from (5.10) to represent the combined shadowing and fading channel status. The CDF of $g_c(n)$ can be expressed by

$$F_{g_c}(g) = \int_{L_{\min}}^{L_{\max}} \frac{2e^{\left(\frac{2y}{a}\right)}}{aA^2(d^2 - d_{\min}^2)} \left\{ \frac{1}{2} + \frac{1}{2}\text{erfc}\left(\frac{\xi \log(ge^{\xi^{-1}y}) - \mu}{\sqrt{2\sigma^2}} \right) \right\} dy \tag{J.4}$$

By differentiating (J.4) with respect to the random variable g and by applying Leibniz's rule, the PDF of $g_c(n)$ can be calculated by

$$f_{g_c}(g) = \frac{2\xi}{g\sqrt{\pi}aA^2(d^2 - d_{\min}^2)\sqrt{2\sigma^2}} e^{\left(\frac{2}{a}(\mu - \xi\log g)\right)} \int_{L_{\min}+\xi\log g-\mu}^{L_{\max}+\xi\log g-\mu} e^{\left(\frac{2w}{a}\right)} e^{-\left(\frac{w^2}{2\sigma^2}\right)} dw \tag{J.5}$$

$$= \mathrm{B}\frac{e^{\left(\frac{2}{a}(\mu-\xi\log g)\right)}}{g}\left[\mathrm{erfc}\left(\frac{aw-2\sigma^2}{a\sqrt{2\sigma^2}}\right)\right]_{L_{\min}+\xi\log g-\mu}^{L_{\max}+\xi\log g-\mu}$$

where we have used $\frac{d}{dx}\mathrm{erfc(x)} = \frac{2}{\sqrt{\pi}}e^{-x^2}$ and $w = (\xi\log(g) + y - \mu)$. By some manipulations, (5.13) can be obtained.

Appendix K

Derivation of (5.19)

To derive an explicit expression for the PDF of the $\acute{\gamma}$, we have to first find the distribution of $X = \max_{n=1,...,N}\{g_c(N)\}$ and $Y = \max_{k=1,...,K}\{g_i(K)\}$. Since $[g_c(1), g_c(2), ..., g_c(n)]$ and $[g_i(1), g_i(2), ..., g_i(k)]$ are independent random variables that are drawn from a common distribution, i.e., (5.13), then the CDF of the maximum for any size of the samples is equal to $[F_{g_c}(g)]^N$ and $[F_{g_i}(g)]^K$, respectively [146]. Using these two new distributions may not provide understandable results. However, the distribution of the maximum function can be tracked using the extreme-value theory, as we will see in the following Lemma.

Lemma 1. Let $z_1, z_2, ..., z_n$ be independent and identically distributed (i.i.d.) random variables drawn from a common CDF $F_Z(z)$. By setting $Z = \max_{n=1,...,N}\{z_n\}$, there exist a sequence of constants $\bar{\lambda}, \bar{\lambda}, \bar{\delta}$ and some non-degenerate distribution function $\bar{H}_Z(\bar{\lambda}, \bar{\delta}, \bar{\zeta})$ such that f(z) converges to the distribution $\bar{H}_Z(\bar{\lambda}, \bar{\delta}, \bar{\zeta})$. The distribution $\bar{H}_Z(\bar{\lambda}, \bar{\delta}, \bar{\zeta})$ is called generalised extreme value distribution (GEVD) [146] and it is equal to

$$\bar{H}_Z(\bar{\lambda}, \bar{\delta}, \bar{\zeta}) \quad = \quad \exp - \left[1 + \bar{\zeta}\left(\frac{z - \bar{\lambda}}{\bar{\delta}}\right)\right]^{\frac{-1}{\bar{\zeta}}} \tag{K.1}$$

where $\bar{\lambda}$, $\bar{\delta}$, and $\bar{\zeta}$ are the location, scale, and shape parameters, respectively.

The GEVD inherently contains the three well-known extreme value distributions, i.e., Gumbel, Weilbull, and Fréchet distributions. If $\bar{\zeta} > 0$, then the distribution

converges to Weibull. If $\bar{\zeta} < 0$, it converges to Fréchet distribution. If $\bar{\zeta} = 0$, then it converges to Gumbel distribution. It has been found that the following condition is sufficient to determine if $F_X(x)$ and $F_Y(y)$ belong to type II or Fréchet distribution domain of attraction [126, Theorem 1.6.1]:

$$\lim_{t\to\infty} \frac{tf(t)}{1-F(t)} = \alpha > 0. \tag{K.2}$$

Therefore,

$$F_X(x) \;\rightarrow\; \exp - \left(\frac{x-\lambda_c}{\delta_c}\right)^{\bar{\beta}} \tag{K.3}$$

$$F_Y(y) \;\rightarrow\; \exp - \left(\frac{y-\lambda_p}{\delta_p}\right)^{\bar{\beta}} \tag{K.4}$$

where $\lambda_c = \bar{\lambda}_c - \frac{\bar{\delta}_c}{\bar{\zeta}_c}$, $\delta_c = \frac{\bar{\delta}_c}{\bar{\zeta}_c}$, $\lambda_p = \bar{\lambda}_p - \frac{\bar{\delta}_p}{\bar{\zeta}_p}$, $\delta_p = \frac{\bar{\delta}_p}{\bar{\zeta}_p}$, and $\bar{\beta} = -\frac{1}{\bar{\zeta}}$. By applying Theorem 9.5 of [146], δ_c and δ_p can be calculated by $\delta_c = F_{g_c}^{-1}\left(1-\frac{1}{N}\right)$ and $\delta_p = F_{g_i}^{-1}\left(1-\frac{1}{K}\right)$, respectively. Here, λ_c and λ_p both approach 0, and $\bar{\beta}$ always equals -0.5 for all cases. Due to the complicated distribution of channel gain, finding closed from expressions for δ_c and δ_p is difficult. In this work, however, we adopt the ML method to estimate the parameters of (K.1), and consequently δ_c and δ_p [146]. Other methods to estimate these parameters can be found in [147]. Now, the conditional distribution of $\acute{\gamma}$, i.e., $\Pr\{\acute{\gamma} \le \acute{\gamma}_0\}$, can be expressed by

$$\begin{aligned} F(\acute{\gamma}) &= \exp\left\{-\left(\frac{\gamma_{\rm pk}\delta_c}{I\acute{\gamma}}\right)^{\bar{\beta}} - \left(\frac{\delta_p\gamma_{\rm pk}}{Q}\right)^{\bar{\beta}}\right\} \\ &+ \bar{\beta}\delta_p^{\bar{\beta}} \int_{\frac{Q}{\gamma_{\rm pk}}}^{\infty} y^{-(\bar{\beta}+1)} \exp\left\{-\left(\frac{Q\delta_c}{I\acute{\gamma}y}\right)^{\bar{\beta}} - \left(\frac{\delta_p}{y}\right)^{\bar{\beta}}\right\} dy \\ &= \exp\left\{-\left(\mathrm{K}_1\acute{\gamma}^{-\bar{\beta}}\right) - \mathrm{K}_2\right\} + \frac{\mathrm{K}\acute{\gamma}^{\bar{\beta}}}{\left(\mathrm{K}\acute{\gamma}^{\bar{\beta}}+1\right)}\left(1 - \exp\left[-\left(\mathrm{K}_1\acute{\gamma}^{-\bar{\beta}} + \mathrm{K}\mathrm{K}_1\right)\right]\right) \end{aligned} \tag{K.5}$$

where $W = \left(\frac{Q\delta_c}{I\acute{\gamma}y}\right)^{\bar{\beta}}$. Expression (5.19) is obtained by differentiating $F_{\acute{\gamma}}(\acute{\gamma}_0)$ with respect to $\acute{\gamma}_0$.

Appendix L

Derivation of (5.26)

In (K.5), if we assume $\bar{\gamma} \to \infty$, then the term $\exp\left\{-\left(\mathrm{K}_1 \acute{\gamma}^{-\beta}\right)\right\}$ vanishes. Therefore, the PDF of the random variable $\acute{\gamma}$ is given by

$$f(\acute{\gamma}) \approx \frac{\mathrm{K}}{2\left(\mathrm{K}+\acute{\gamma}^{-0.5}\right)^2 \acute{\gamma}^{\frac{3}{2}}} \tag{L.1}$$

which is then used to calculate $C_{\max}$, as expressed by (5.26).

Appendix M

Derivation of (5.31)

From (2.6), we have

$$\left(\frac{E_b}{N_0}\right)_{\text{penalty}} = \lim_{\text{SNR}\to\infty}\left(\log_2(\text{SNR}) - \frac{\mathbb{E}[\log_2(1+\text{SNR}X)]}{S_\infty}\right) \tag{M.1}$$

where the expectation is with respect to the random variable X. Knowing that $\int_0^\infty f(X)dX = 1$, (M.1) can be re-written as

$$\left(\frac{E_b}{N_0}\right)_{\text{penalty}} = \int_0^\infty \log_2\left[\lim_{\text{SNR}\to\infty}\frac{\text{SNR}}{(1+\text{SNR}X)}\right] f(X)dX. \tag{M.2}$$

Applying L'Hpital's rule into (M.2) leads to

$$\left(\frac{E_b}{N_0}\right)_{\text{penalty}} = \int_0^\infty \log_2\left(\frac{1}{X}\right) f(X)dX \tag{M.3}$$

which is the expected value of $\log_2\left(\frac{1}{X}\right)$ and (5.31) is then obtained.

Appendix N

Projection Matrix

The matrix can be considered as a projection matrix when it satisfies the properties of idempotence and Hermitian symmetry [148], i.e.,

$$\mathbf{P}^2 = \mathbf{P}$$
$$\mathbf{P}^H = \mathbf{P}. \tag{N.1}$$

The eigenvalue of a projection are either 0 or 1. However, that does not mean every matrix whose eigenvalues are 0 or 1 is a projection. If we assumed that $\bar{\mathbf{H}}_m$ is disjoint, then building a projection whose range is $\bar{\mathbf{H}}_m$ is given by [148]

$$\mathbf{P}_m \;=\; \bar{\mathbf{H}}_m(\bar{\mathbf{H}}_m^H\bar{\mathbf{H}}_m)^{-1}\bar{\mathbf{H}}_m^H = \bar{\mathbf{H}}_m\mathbf{W}_m \tag{N.2}$$

where $\mathbf{W}_m \in \mathbb{C}^{N_r\times n}$ is the weight matrix. This requires $n+t \leq N_r$, with t ($t \geq 1$) is an integer number. The orthogonal projection whose range is $\bar{\mathbf{H}}_m^{\perp}$ (orthogonal to a subspace $\mathbf{v}_m$, which is spanned by $\bar{\mathbf{H}}_m$) can then be calculated by

$$\mathbf{P}_m^{\perp} \;=\; \mathbf{I} - \bar{\mathbf{H}}_m\mathbf{W}_m. \tag{N.3}$$

The received signal $\mathbf{y}$ and the corresponding $\bar{\mathbf{H}}_m$ span in the same subspace. Therefore, by ignoring the impact of the noise, $\mathbf{y}$ also lies in the null space of the corresponding orthogonal projection $\mathbf{P}_m^{\perp}$, as shown in Fig. N.1 In this case,

$$\mathbf{P}_{m=j}^{\perp}\mathbf{y} = 0 \tag{N.4}$$

with j is the correct antenna group index.

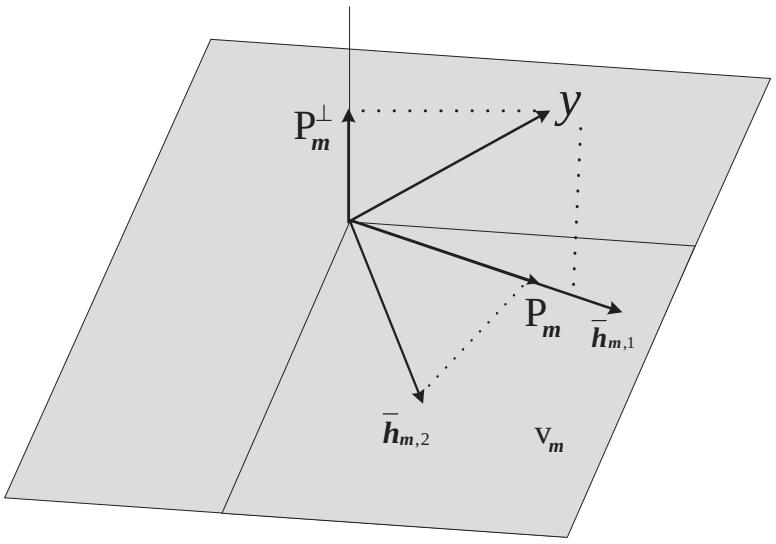

Figure N.1: A schematic representation of a projection operation: $\mathbf{P}_m$ is the projection whose range is $\bar{\mathbf{H}}_m$, $\mathbf{P}_m^{\perp}$ is the orthogonal projection, $\mathbf{y}$ is projected onto a subspace $\mathbf{v}_m$ that contains vector $\bar{\mathbf{H}}_m$.

Appendix O

Derivation of (6.30)

By introducing the dual variable associated with the spectral efficiency constraint, the partial Lagrangian of this problem is expressed as

$$\begin{aligned} L(P_t,\nu) &= \mathrm{E}\left[K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right)\right] \\ &- q\left(\frac{\mathrm{E}[P_t]}{\eta_{\mathrm{PA}}}+\mathrm{E}\left[K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right)\right]\left(\frac{1}{\eta_{\mathrm{load}}}+\frac{1}{\eta_{\mathrm{Bh}}}\right)+P_{\mathrm{other}}\right) \\ &- \nu\left(\mathrm{E}\left[K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right)\right]-SE\right) \end{aligned} \tag{O.1}$$

where ν is the nonnegative dual variable associated with the constraint

$$\mathrm{E}\left[K\bar{N}\log\left(1+\frac{P_t(\lambda)}{N_0}\lambda\right)\right] \geq SE. \tag{O.2}$$

The dual function is then expressed as

$$\max_{P_t} L(P_t,\nu). \tag{O.3}$$

It is certain that the duality gap is zero for the convex optimisation problem addressed here, and thus solving its dual problem is tantamount to solving the problem of (6.27)–(6.29). According to the KKT conditions [149], however, the optimal solutions needs to satisfy (6.28). Next if we differentiate the Lagrangian and set the derivative equal to zero, (6.30) can be obtained.

Bibliography

[1] European Mobile Industry Observatory, GSMA, Nov. 2011.

[2] WWRF, L. Sorensen and K. E. Skouby, *User scenarios 2020*, Report, July 2009, [online]. Available: http://www.wireless-world-research.org.

[3] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. S. Thompson, I. Ku, C.-X. Wang, T. A. Le, M. R. Nakhai, J. Zhang, and L. Hanzo, "Green radio: radio techniques to enable energy efficient wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 46–54, June 2011.

[4] D. Lister, Vodafone Group Research & Development, "An operators view on green radio," in *Proc IEEE GreenComm'9*, June 2009.

[5] Commission of the European Communities, Staff Working Document, "Exploiting the employment potential of ICTs", Apr. 2012.

[6] GSMA, "What is the impact of mobile telephony on economic growth?, "A report from GSM association, Nov. 2012. [Online]. Available: http://www.gsma.com.

[7] 3GPP, TS 36.201 V8.3.0, "Evolved universal terrestrial radio access (E-UTRA); LTE physical layer - general description," Rel. 8, Mar. 2009.

[8] GSA, "LTE-A Network Evolution?, "A report from GSA, July 2015. [Online]. Available: http://www.gsacom.com/.

[9] A. Hashimoto, H. Yorshino, and H. Atarashi, "Roadmap of IMT-Advanced development," *IEEE Microw. Mag.*, vol. 9, no. 4, pp. 80–88, Aug. 2008.

[10] 3GPP, TS 36.104 V10.0.0, "Evolved universal terrestrial radio access (E-UTRA); base station (BS) radio transmission and reception," Rel. 10, Oct. 2010.

[11] Cisco, "Cisco visual networking index: global mobile data traffic forecast update, 2013–2018," White Paper, Feb. 2014. [Online]. Available: http://www.cisco.com.

[12] Ericsson, "Ericsson mobility report," Report, June 2011. [online]. Available: http://www.ericsson.com.

[13] Nokia Siemens Networks, "2020: Beyond 4G: Radio Evolution for the Gigabit Experience," White Paper, 2011. [Online]. Available: http://networks.nokia.com.

[14] A. Bleicher, "Millimeter waves may be the future of 5G phones," *IEEE Spectrum*, Aug. 2013.

[15] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sept. 2008.

[16] H. Haas, "Wireless data from every light bulb," TED Website, Aug. 2011. [Online]. http://www.ted.com.

[17] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122-130, Feb. 2014.

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

[19] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, 2013.

[20] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2012.

[21] M.D. Renzo, H. Haas, and P.M. Grant, "Spatial modulation for multiple-antenna wireless systems: a survey" *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, Dec. 2011.

[22] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.

[23] C.-X. Wang and S. Wu, "Massive MIMO channel measurements and modeling: advances and challenges" *IEEE Wireless Commun. Mag.*, submitted for publication.

[24] X. Hong, C.-X. Wang, M. Uysal, X. Ge, and S. Ouyang, "Capacity analysis of hybrid cognitive radio networks with distributed VAAs," *IEEE Trans. Veh. Tech.*, vol. 59, no. 7, pp. 3510–3523, Sept. 2010.

[25] PBL Netherlands Environmental, "Trends in global CO_2 emissions," Background Studies Rep., Jul. 2012. [Online]. Available: http://www.pbl.nl/en/publications/2012

[26] Gartner, "Gartner estimates ICT industry accounts for 2 percent of global CO2 emissions," Press Release, Apr. 2007. [Online]. Available: http://www.gartner.com.

[27] IDATE, "Green telecom calling for a better future," News 453, Jan. 2009. [Online]. Available: http://www.idate.org/en/News/Green-Telecom_553.html.

[28] Z. Hasan, H. Boostanimehr, and V. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 4, pp. 524–540, Oct. 2011.

[29] A. D. Domenico, E. C. Strinati, and A. Capone, "Enabling green cellular networks: A survey and outlook, *Elseiver Computer Commu.*, vol. 37, pp. 5–24, 2014.

[30] F. Cao and Z. Fan, "The tradeoff between energy efficiency and system performance of femtocell deployment," in *Proc. IEEE ISWCS'10*, Sept. 2010, pp. 315–319.

[31] Y. Hou and D. Laurenson, "Energy efficiency of high QoS heterogeneous wireless communication network," in *Proc. IEE VTC'10-Fall*, Sept. 2010.

[32] M. Jada, M. Hossain, J. Hamalainen, and R. Jantti, "Impact of femtocells to the wcdma network energy efficiency," in *Proc. IEEE IC-BNMT'10*, Oct. 2010, pp. 305–310.

[33] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities and implementation", *Proc. of the IEEE*, vol. 102, no. 1, pp. 56–103, Jan. 2014.

[34] T. Chen, H. Kim, and Y. Yang, "Energy efficiency metrics for green wireless communications," in *Proc IEEE WCSP'10*, Oct. 2010.

[35] H. M. Kwon and T. G. Birdsall, "Channel capacity in bits per joule," *IEEE J. Ocean. Eng.*, vol. 11, no. 1, pp. 97–99, Jan. 1986.

[36] V. Rodoplu and T. H. Meng, "Bits-per-joule capacity of energy-limited wireless networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 857–865, Mar. 2007.

[37] G. Auer, V. Giannini, I. Godor, P. Skillermark, M. Olsson, M.A. Imran, D. Sabella, M.J. Gonzalez, C. Desset, and O. Blume, "Cellular Energy Efficiency

Evaluation Framework," in *Proc. IEEE VTC'11-Spring*, Budapest, Hungary, May 2011.

[38] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[39] S. Shamai and S. Verdu, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[40] A. Lozano, A. Tulino, and S. Verdu, "Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2527–2544, Oct. 2003.

[41] Y. Yao, X. Cai, and G. Giannakis, "On energy efficiency and optimum resource allocation of relay transmissions in the low-power regime," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2917–2927, Nov. 2005.

[42] J. Gomez-Vilardebo, A. Perez-neira, and M. Najar, "Energy efficient communications over the AWGN relay channel," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 32–37, Jan. 2010.

[43] X. Cai, Y. Yao, and G. Giannakis, "Achievable rates in low-power relay links over fading channels," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 184--194, Jan. 2005.

[44] O. Oyman and M. Win, "Power-bandwidth trade-off in multiuser relay channels with opportunistic scheduling," in *Proc. IEEE ALLERTON'08*, Illinois, USA, Sept. 2008, pp. 72–78.

[45] N. Jindal, "High SNR analysis of MIMO broadcast channels," in *Proc. ICSIT'05*, Adelaide, SA, Sept. 2005, pp. 2310–2314.

[46] A. Lozano, A. Tulino, and S. Verdu, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.

[47] G. Auer, O. Blume, V. Giannini, I. Godor, M. Imran, Y. Jading, E. Katranaras, M. Olsson, D. Sabella, P. Skillermark, and W. Wajda, "Energy efficiency analysis of the reference systems, areas of improvements and target breakdown", INFSO-ICT-247733 EARTH, ver. 2, 2012. [Online]. Available: http://www.ict-earth.eu/.

[48] Y. Chen, S. Zhang, S. Xu, and G. Li, "Fundamental trade-offs on green wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 30–37, June 2011.

[49] S. Tombaz, A. Vastberg, and J. Zander, "Energy-and cost-efficient ultra-high-capacity wireless access," *IEEE Wireless Commun. Mag.*, vol. 18, no. 5, pp. 1824, Oct. 2011.

[50] E. Bjrnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: is massive MIMO the answer?," *IEEE Trans. Wireless Commun.*, submitted, 2014. [Online]. Available: http://arxiv.org/abs/1403.6150.

[51] A. Mezghani and J. A. Nossek, "Power efficiency in communication systems from a circuit perspective," in *Proc. IEEE ISCAS11*, May 2011, pp. 1896–1899.

[52] A. Awadelkarim Widaa Ahmed, J. Markendahl, A. Ghanbari, "Toward capacity efficient, cost efficient and power efficient deployment strategy for indoor mobile broadband," in *Proc. ITS European Regional Conference*, Italy, Oct. 2013.

[53] J. Laudon, "Performance/Watt: The new server focus," *ACM SIGARCH Computer Architecture News*, vol. 33, pp. 5–13, Nov. 2005.

[54] S.C. Cripps, "*RF power amplifiers for wireless communications*," Artech House Microwave Libreary, 2nd Edition, June 2006.

[55] B. G. Bathula and J. M. H. Elmirghani, "Green networks: energy efficient design for optical networks," in *Proc. IEEE WOCN'9*, Cairo, Egypt, Apr, 2009.

[56] K. Sato, "Optical Technologies that Enable Green Networks," in *Proc. IEEE ICTON'10*, Munich, Germany, June, 2010.

[57] V. Hanumaiah and S. Vrudhula, "Energy-efficient operation of multicore processors by DVFS, task migration, and active cooling," *IEEE Trans. Comput.*, vol. 63, no. 2, Feb. 2014.

[58] L. Fu, H. Kim, J. Huang, S. C. Liew, and M. Chiang, "*Green communications: Theoretical Fundamentals, Algorithms, and Applications"*, CRC Press, Sept. 2012.

[59] W. Li and M. Latva-aho, "An efficient channel block diagonalization method for generalized zero forcing assisted MIMO broadcasting systems," *IEEE Trans. Wireless Comm.*, vol. 10, no. 3, pp. 739–744, Dec. 2011.

[60] alcatel-lucent, "LTE to 5G," White Paper, 2015. [Online]. Available: http://www.alcatel-lucent.com/solutions/lte-to-5G.

[61] H. Schulze and C. Lueders, *"Theory and applications of OFDM and CDMA wideband wireless communications*." John Wiley, July 2005.

[62] S. Chieochan, and E. Hossain, "Adaptive radio resource allocation in OFDMA systems: A survey of the state-of-the-art approaches," *Wireless Communications and Mobile Computing*, vol. 9, no. 4, pp. 513–527, April 2009.

[63] S. Sadr, A. Anpalagan, and K. Raahemifar "Radio resource allocation algorithms for the downlink of multiuser OFDM communication systems," *Commun. Surveys Tuts.*, vol. 11, no. 3, pp.92–106, July 2009.

[64] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 171–178, Feb. 2003.

[65] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *Proc. IEEE VTC'00-Spring*, vol. 2, May 2000, pp. 1085–1089.

[66] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2726–2737, Nov. 2005.

[67] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit and power allocation," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1747–1758, Oct. 1999.

[68] G. Zhang, "Subcarrier and bit allocation for real-time services in multiuser OFDM systems," in *Proc. IEEE ICC'04*, vol. 5, June 2004, pp. 2985–2989.

[69] L. Xiaowen and Z. Jinkang, "An adaptive subcarrier allocation algorithm for multiuser OFDM system," in *Proc. IEEE VTC'03-Fall*, vol. 3, Oct. 2003, pp. 1502–1506.

[70] I. Kim, I. Park, and Y. Lee, "Use of linear programming for dynamic subcarrier and bit allocation in multiuser OFDM," *IEEE Trans. Veh. Techn.*, vol. 55, no. 4, pp. 1195–1207, July 2006.

[71] Z. Mao and X. Wang, "Efficient optimal and suboptimal radio resource allocation in OFDMA system," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 440–445, Feb. 2008.

[72] G. W. Miao, N. Himayat, G. Y. Li, and D. Bormann, "Energy-efficient design in wireless OFDMA" in *Proc. IEEE ICC'08*, May 2008, pp. 3307–3312.

[73] G. W. Miao, N. Himayat, and G. Y. Li, "Energy-efficient transmission in frequency-selective channels," in *Proc. IEEE GLOBECOM'08*, Nov. 2008.

[74] C. Isheden and G. P. Fettweis, "Energy-efficient multi-carrier link adaptation with sum rate-dependent circuit power," in *Proc. IEEE GLOBECOM'10*, Dec. 2010.

[75] R. S. Prabhu and B. Daneshrad, "An energy-efficient water-filling algorithm for OFDM systems," in *Proc. IEEE ICC'10*, May 2010.

[76] G. Miao, N. Himayat, Y. Li, A. T. Koc, and S. Talwar, "Interference-aware energy-efficient power optimization," in *Proc. IEEE ICC'09*, Dresden, Germany, June 2009, pp. 14–18.

[77] G. W. Miao, N. Himayat, and Y. Li, "Low-complexity energy-efficient OFDMA," in *Proc. IEEE ICC '09*, Dresden, Germany, June 2009.

[78] A. Goldsmith," *Wireless communications*," Cambridge University Press, June 2005.

[79] E. Bajalinov, "*Linear-fractional programming: theory, methods, applications and software*," Springer, 2003.

[80] W. Dinkelbach, "On nonlinear fractional programming", *Management Science*, vol. 13, pp. 492–498, Mar. 1967.

[81] I.M. Stancu-Minasian. *"Fractional Programming: Theory, Methods and Applications"*, Kluwer Academic Publishers, Dordrecht, 1997.

[82] S. Schaible, "Fractional programming," *Zeitschrift für Operations Research*, vol. 27, no. 1, pp. 39–54, Dec. 1983.

[83] S. Schaible and T. Ibaraki, "Fractional programming," *European J. Operational Research*, vol. 12, no. 4, pp. 325338, Apr. 1983.

[84] S. Schaible, "Fractional programmingII: on Dinkelbachs Algorithm," *Management Science*, vol. 22, no. 8, pp. 868–873, Apr. 1976.

[85] 3GPP, TS 25.814 V7.1.0, "Physical layer aspects for evolved UTRA" Rel. 7, Sep. 2006.

[86] J. Laiho, A. Wacker, T. Novosad,"*Radio network planning and optimisation for UMTS,"* John Wiley & Sons, Chichester, 2nd Edition, Dec. 2005.

[87] S.-P. Yeh et al., "Capacity and coverage enhancement in heterogeneous networks," *IEEE Wireless Comm.*, vol. 18, pp. 32–38, June 2011.

[88] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, issue 3, pp. 10–21, June 2011.

[89] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sept. 2008.

[90] D. Feng, C. Jiang, G. Lim, L. Cimini, G. Feng, and G. Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Mag.*, vol. 15, no. 1, pp. 167–178, 2013.

[91] F. Haider, C.-X. Wang, H. Haas, D. Yuan, H. Wang, X. Gao, X.-H. You, and E. Hepsaydir, "Spectral efficiency analysis of mobile femtocell based cellular systems," *Proc. IEEE ICCT'11*, Jinan, China, Sept. 2011, pp. 347–351.

[92] F. Haider, M. Dianati, and R. Tafazolli,"A Simulation based study of mobile femtocell assisted LTE networks," in *Proc. IEEE IWCMC'11*, Istanbul, Turkey, Jul. 2011.

[93] T. Elkourdi and O. Simeone, "Femtocell as a relay: An outage analysis," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4204–4213, Dec. 2011.

[94] Y. Sui, J. Vihriala, A. Papadogiannis, M. Sternad, W. Yang, and T. Svensson, "Moving cells: a promising solution to boost performance for vehicular users," *EEE Commun. Mag.*, vol. 51, no. 6, pp. 62–68, June 2013.

[95] M. Chowdhury, S. Lee, B. Ru, N. Park, and Y. Jang, "Service quality improvement of mobile users in vehicular environment by mobile femtocell network deployment," in *Proc. IEEE ICTC'11*, Sept. 2011, pp. 194–198.

[96] M. Qutqut, M. Feteiha, and H. S. Hassanein, "Outage probability analysis of mobile small cells over LTE-A networks," in *Proc. IEEE IWCMC'14*, Aug. 2014, pp. 1045–1050.

[97] M. Qutqut, F. Al-Turjman, and H. S. Hassanein, "MFW: mobile femtocells

utilizing WiFi: A data offloading framework for cellular networks using mobile femtocells," in *Proc. IEEE ICC'13*, June 2013, pp. 5020–5024.

[98] O. Karimi, J. Liu, and C. Wang, "Seamless wireless connectivity for multimedia services in high speed trains," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 729–739, May 2012.

[99] M. H. Qutqut, "*Mobile small cells in cellular heterogeneous networks*," Ph.D. Thesis. Queens University: Canada, Sep. 2014. [Online]. Available: http://qspace.library.queensu.ca/handle/1974/12538.

[100] C.-X. Wang, X. Hong, X. Ge, X. Cheng, G. Zhang, and J. Thompson, "Cooperative mimo channel models: a survey," *IEEE Commun. Mag.*, vol. 48, no. 2, pp. 80–87, Feb. 2010.

[101] H. Claussen, "Performance of macro-and co-channel femtocells in a hierarchical cell structure," in *Proc. IEEE PIMRC'07*, Sept. 2007.

[102] I. Guvenc, M.-R. Jeong, F. Watanabe, and H. Inamura, "A hybrid frequency assignment for femtocells and coverage area analysis for co-channel operation," *IEEE Commun. Lett.*, vol. 12, no. 12, pp. 880–882, Dec. 2008.

[103] Y. Bai, J. Zhou, and L. Chen, "Hybrid spectrum usage for overlaying LTE macrocell and femtocell," in *Proc. IEEE GLOBECOM'09*, Nov. 2009.

[104] Y. Bai, J. Zhou, and L. Chen, "Hybrid spectrum sharing for coexistence of macrocell and femtocell," in *Proc. IEEE ICCTA'09*, Oct. 2009, pp. 162–166.

[105] A. Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proc. IEEE* VTC'00-Spring, Tokyo, May 2000, PP. 1854–1858.

[106] 3GPP, TS 25.814 V8.1.0, "Physical Layer Aspects for Evolved UTRA (Release 8)," Sept. 2007.

[107] 3GPP, WG1 R1-081853, "Simulation results: linkage between PUSCH and amount of resources for control on PUSCH," May. 2008.

[108] U.S. Department of Commerce, National Telecommunications and Information Administration, Office of Spectrum Management," United States frequency allocations," Oct. 2003. [Online]. Available: http://www.ntia.doc.gov/osmhome/allochrt.pdf.

[109] Federal Communications Commission, "Spectrum policy task force," Tech. Rep., ET Docket, pp. 02–155, Nov. 2002.

[110] J. Mitola and G. Maguire, "Cognitive radio: making software radios more personal," *IEEE Pers. Commun. Mag.*, vol. 6, no. 6, pp. 13–18, Aug. 1999.

[111] C.-X. Wang, H.-H. Chen, X. Hong, and M. Guizani, "Cognitive radio network management," *IEEE Veh. Technol. Mag.*, vol. 3, no. 1, pp. 28–35, Mar. 2008.

[112] X. Hong, C.-X. Wang, H.-H. Chen, and Y. Zhang, "Secondary spectrum access networks: recent developments on the spatial models," *IEEE Veh. Technol. Mag.*, vol. 4, no. 2, pp. 36–43, June 2009.

[113] S. A. Jafar and S. Srinivasa, "Capacity limits of cognitive radio with distributed and dynamic spectral activity," in *Proc. IEEE ICC'06*, Istanbul, Turkey, June 2006, pp. 5742–5747.

[114] D. Ugarte and A. B. McDonald, "On the capacity of dynamic spectrum access enable networks," in *Proc. IEEE DySPAN'05*, Baltimore, USA, Nov. 2005, pp. 630–633.

[115] C. Cordeiro, K. Challapali, D. Birru, Sai Shankar, "IEEE 802.22: the first worldwide wireless standard based on cognitive radios", *IEEE DySPAN'05*, 2005, pp. 328–337.

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

[117] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 471–487, Feb. 2007.

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

[119] X. Kang, Y.-C. Liang, A. Nallanathan, H. Garg, and R. Zhang, "Optimal power allocation for fading channels in cognitive radio networks: ergodic capacity and outage capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 940–950, Feb. 2009.

[120] L. Musavian and S. Aissa, "Capacity and power allocation for spectrum sharing communications in fading channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 148–156, Jan. 2009.

[121] C.-X. Wang, X. Hong, H.-H. Chen, and J. Thompson, "On capacity of cognitive radio networks with average interference power constraints," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1620–1625, Apr. 2009.

[122] J. Hong and W. Choi, "Throughput characteristics by multiuser diversity in a cognitive radio network," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3749–3763, Aug. 2011

[123] B. Wild and K. Ramchandran, "Detecting primary receivers for cognitive radio applications," in *Proc. IEEE DySPAN'05*, Baltimore, MD, Nov. 2005, pp. 124–130.

[124] G. L. Stuber, *"Principles of mobile communication,"* Springer, New York, 3rd Edition, 2011.

[125] Y. Sun, R. P. Jover, and X. Wang, "Uplink interference mitigation for OFDMA femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 614–625, Feb. 2012.

[126] M. R. Leadbetter, G. Lindgren, and H. Rootzen, *Extremes and related properties of random sequences and processes.* Springer-Verlag, New York, 1983.

[127] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation", *IEEE Trans. Veh. Tech.*, no. 4, pp. 2228–2241, July 2008.

[128] J. Jeganathan,A. Ghrayeb,and L.Szczecinski, "Spatial modulation: optimal detection and performance analysis," *IEEE Commun. Lett.*, vol. 12, no.8, pp. 545–547, Aug. 2008.

[129] M. Di Renzo and H. Haas, "Bit error probability of SM-MIMO over generalized fading channels," *IEEE Trans. Veh. Tech.*, vol. 61, no. 3, pp. 1124–1144, Mar. 2012.

[130] R. Y. Mesleh, S. Engelken, S. Sinanovic, and H. Haas, "Analytical SER calculation of spatial modulation," in *Proc. IEEE Symp'08*, Aug. 2008, pp. 272–276.

[131] M. Di Renzo and H. Haas, "Bit error probability of space modulation over Nakagami-m fading: Asymptotic analysis," *IEEE Commun. Lett.*, vol. 15, no. 10, pp. 1026–1028, Oct. 2011.

[132] L. Zheng and D. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple- Antenna Channels," *IEEE Trans. Info. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[133] A. Younis, N. Seramovski, R. Mesleh, and H. Haas, "Generalised spatial modulation," in *Proc. IEEE ASILOMAR'10*, USA, 2010 pp. 1498-1502.

[134] J. Zheng1,, "Spatial modulation detection based on projection decomposition," in *Proc. IEEE ICCC'13*, China, 2013 pp. 309-313.

[135] J. Wang, S. Jia, and J. Song, "Generalised spatial modulation system with multiple active transmit antennas and low complexity detection scheme," *IEEE Trans. Wireless Commun.*, pp. 1605–1615, Mar. 2012.

[136] N. Serafimovski1, S. Sinanovic, M. Di Renzo, and H. Haas, "Multiple access spatial modulation," *EURASIP J. Wireless Commun. and Networking*, Mar. 2012.

[137] N. Serafimovski, S. Sinanovic, A. Younis, M. Di Renzo, and H. Haas, "2-user multiple access spatial modulation," in *Proc. IEEE HeterWMN'11*, Dec. 2011, pp. 343–347.

[138] M. I. Kadir, S. Sugiura, J. Zhang, S. Chen, and L. Hanzo, "OFDMA/SCFDMA aided space-time shift keying for dispersive multi-user scenarios", *IEEE Trans. Veh. Tech.*, vol. 62, no. 1, pp. 408–414, Jan. 2013.

[139] S. Narayanan, M. J. Chaudhry, A. Stavridis, M. Renzo, F. Graziosi, and H. Haas, "Multi-User Spatial Modulation MIMO," in *Proc. IEEE WCNC'14*, pp. 671–676, Apr. 2014.

[140] 3GPP, TS 36.104 V12.5.0, "Evolved universal terrestrial radio access (E-UTRA); base station (BS) radio transmission and reception," Rel. 12, Oct. 2012.

[141] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[142] G. Foschini and M. Gans, "On limits of wireless communication in a fading environment when using multiple antenna," *IEEE Trans. Personal Wireless Commun.*, pp. 311–335, Mar. 1998.

[143] Y. Yang and B. Jiao, "Information-guided channel-hopping for high data rate wireless communication," *IEEE Commun. Lett.*, vol. 12, no. 4, pp. 225–227, Apr. 2008.

[144] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Dec. 1999.

[145] E. Salbaroli, and A. Zanella, "Interference Analysis in a Poisson Field of Nodes of Finite Area," *IEEE Trans. Veh. Tech.*, pp. 1776–1783, Aug. 2008.

[146] E. Castillo, A. S. Hadi, N. Balakrishnan and J. Sarabia, *Extreme value*

and related models with applications in engineering and science, New York, Wiley, 2001.

[147] J. Maritz, and A. Munro, "On the use of the generalized extreme-value distribution in estimating extreme percentiles," Biometrics, vol. 23, no. 1, pp. 79–103, Mar. 1969.

[148] R. Behrens, and L. Scharf, "Signal processing applications of oblique projection operators," *IEEE Trans. Signal Processing*, pp. 1413-1424, June 1994.

[149] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge, UK: Cambridge University Press, 2004.