THE UNIVERSITY of EDINBURGH

Energy-Efficient LTE Transmission Techniques — Introducing Green Radio from Resource Allocation Perspective

Rui Wang

A thesis submitted for the degree of Doctor of Philosophy.
The University of Edinburgh.
March 2011

Abstract

Energy consumption has recently become a key issue from both environmental and economic considerations. A typical mobile phone network in the UK may consume approximately 40-50 MW, contributing a significant proportion of the total energy consumed by the information technology industry. With the worldwide growth in the number of mobile subscribers, the associated carbon emissions and growing energy costs are becoming a significant operational expense, leading to the need for energy reduction. The Mobile VCE Green Radio Project has been launched, which targets to achieve 100x energy reduction of the current wireless networks by 2020. In this thesis, energy-efficient resource allocation strategies have been investigated taking the LTE system as an example.

Firstly, theoretical analysis of energy-efficient design in cellular environments is provided according to the Shannon Theory. Based on a two-link scenario the performance of simultaneous transmission and orthogonal transmission for network power minimization under the specified rate constraints is investigated. It is found that simultaneous transmission consumes less power than orthogonal transmission close to the base station, but much more power in the cell-edge area. Also, simulation results suggest that the energy-efficient switching margins between these two schemes are dominated by the sum total of their required data rates. New definitions of power-utility and fairness metrics are further proposed, following by the design of weighted resource allocation approaches based on efficiency-fairness tradeoffs.

Apart from energy-efficient multiple access between different links, the energy used by individual base stations can also be reduced. For example, deploying sleep modes is an effective approach to reduce radio base station operational energy consumption. By periodically switching off the base station transmission, or using fewer transmit antennas, the energy consumption of base station hardware may decrease. By delivering less control signalling overhead, the radio frequency energy consumption can also be reduced. Simulation results suggest that up to 90% energy reduction can be obtained in low traffic conditions by employing time-domain optimization in each radio frame. The optimum on/off duty cycle is derived, enabling the energy consumption of the base station to scale with traffic loads. In the spatial-domain, an antenna selection criterion is proposed, indicating the most energy-efficient antenna configuration with the knowledge of users' locations and quality of service requirements. Without time-domain sleep modes, using fewer transmit antennas could outperform full antenna transmission. However, with time-domain sleep modes, using all available antennas is generally the most energy-efficient choice.

Declaration of Originality

I hereby declare that the research recorded in this thesis and the thesis itself was composed and originated entirely by myself in the School of Engineering at the University of Edinburgh.

Rui Wang

Acknowledgements

First of all, I would like to express my appreciation to my supervisors Dr. John S. Thompson, Prof. Harald Haas and Prof. Norbert Görtz for offering me the opportunity to pursue this project and for their excellent guidance and support during the course of my Ph.D. research. They are perfect gentlemen who are always nice, polite, responsible, and considerate. I have learned so much from them and working together with them is the most precious experience in my life.

I would like to thank the technical and financial support of the Mobile VCE UK. The regular scheduled industrial steering meetings and researchers' workshops keep me up-to-date with the latest research information in Green Communications. The technical feedbacks from my industrial reviewers are so precious for improving my future research.

I am also grateful to my industrial supervisors Danny Kershaw and David Bull for their valuable advice on low power design during my six month period in ARM Ltd. The joint research project between ARM Ltd. and Edinburgh University provides me further insights into the energy-efficient design from circuit perspective.

My families have been the most important support for me. I could not have reached this important milestone without their encouragement, which helped me conquer difficulties and complete my work. Also, I would like to thank Robert De Bold for sharing with me our plans and dreams for the future. His love, support and understanding during the last stages of my research have been of great importance for the completion of this thesis.

Last but definitely not the least, I would like to thank all my friends and colleagues at IDCOM. There is so much fun during the last four years of my stay in Edinburgh University. I will never forget this sweet memory.

Contents

Declaration of Originality . . . iii
Acknowledgements . . . iv
Contents . . . v
List of Figures . . . viii
List of Tables . . . x
Acronyms and Abbreviations . . . xi
Nomenclature . . . xiv

1 Introduction **1**
1.1 Motivation . . . 1
1.2 Thesis Structure . . . 4
1.3 Contributions . . . 5

2 Energy Efficient Communication Systems **8**
2.1 Introduction . . . 8
2.2 Green Radio . . . 9
2.2.1 Potentials and Targets . . . 9
2.2.2 Relevant Techniques . . . 11
2.2.3 Energy Efficiency Metrics . . . 13
2.3 Wireless Communication Environment . . . 14
2.3.1 Cellular Environment . . . 14
2.3.2 Mobile Radio Propagation . . . 16
2.4 Capacity . . . 17
2.4.1 Ergodic Capacity . . . 18
2.4.2 Outage Capacity . . . 20
2.4.3 Capacity Region . . . 20
2.5 Radio Resource Management . . . 22
2.5.1 Power Allocation Policies . . . 22
2.5.2 Adaptive Modulation and Coding . . . 24
2.5.3 Multiple Access Techniques . . . 27
2.5.4 Scheduling Strategies . . . 28
2.5.5 Fairness Concepts . . . 29
2.5.6 Constraints and Tradeoffs . . . 30
2.5.7 DTX and DRX . . . 31
2.6 MIMO Systems . . . 32
2.6.1 Space Time Block Code . . . 33
2.6.2 Spatial Multiplexing . . . 35
2.6.3 MIMO in Interference Environments . . . 36
2.7 Summary . . . 37

3 Two-Link Study of Energy-Efficient Resource Allocation **39**
3.1 Introduction . . . 39

3.2 Two-Link Model Assumptions . . . 40
3.3 Energy-Efficient Multiple Access . . . 42
3.3.1 Simultaneous Transmission . . . 43
3.3.2 Orthogonal Transmission . . . 44
3.3.3 Orthogonal Transmission Scheme Extensions . . . 45
3.3.4 Scheme Selection with Link Data Rate Constraints . . . 46
3.3.5 Scheme Selection with Sum Data Rate Constraints . . . 47
3.4 Performance Analysis and Tradeoffs . . . 49
3.4.1 Energy Consumption Rating . . . 49
3.4.2 Efficiency-Fairness Tradeoffs . . . 51
3.4.3 Outage Probability . . . 52
3.5 Soft Frequency Reuse in Energy-Efficient Design . . . 55
3.5.1 Multiple-Link Scenarios . . . 55
3.5.2 SFR for Maximizing Spectral Efficiency . . . 56
3.5.3 SFR for Improving Energy Efficiency . . . 58
3.6 Summary . . . 60

4 Time-Domain Base Station Sleep Mode Optimization 62
4.1 Introduction . . . 62
4.2 Base Station Energy Efficiency Model . . . 64
4.3 Traffic Profiles in Network . . . 67
4.4 Innovations of Time-Domain Sleep Mode Design . . . 69
4.4.1 Possible Energy Reductions of Sleep Modes . . . 70
4.4.2 Sleep Modes with Resource Management . . . 73
4.5 Basic Sleep Mode Modeling for SISO System . . . 74
4.5.1 Model Assumptions . . . 74
4.5.2 On/Off Time Optimization . . . 76
4.5.3 Performance Analysis of SISO Sleep Modes . . . 78
4.6 Summary . . . 81

5 Sleep Modes with MIMO Systems 83
5.1 Introduction . . . 83
5.2 LTE Downlink Transmission . . . 84
5.2.1 LTE Physical Layer Structure . . . 85
5.2.2 LTE Control Signalling Process . . . 87
5.2.3 LTE Control Overhead Resizing . . . 91
5.3 Sleep Modes in a MIMO Environment . . . 94
5.3.1 Optimization Issues for MIMO Sleep Modes . . . 94
5.3.2 Model Assumptions . . . 97
5.3.3 Antenna Selection for Non-Sleep Modes . . . 99
5.3.4 Performance Analysis of MIMO Sleep Modes . . . 101
5.3.5 Practical Issues of Spatial-Domain Antenna Sleep Modes . . . 103
5.4 Other Discussions . . . 105
5.4.1 Multi-User Sleep Modes with Frequency-Domain Scheduling . . . 106
5.4.2 Sleep Modes in VoIP . . . 108
5.4.3 Legacy Handsets . . . 108
5.5 Summary . . . 109

6 Conclusions **110**
6.1 Contributions 110
6.2 Future Work 113

A Appendix: Publication List **116**

References **117**

List of Figures

2.1 Mobile network traffic and revenue growth [24] . . . 10
2.2 Cellular network power consumption [24] . . . 11
2.3 Base station energy budget model [25] . . . 12
2.4 Evolution of wireless and wireline user data rates [41] . . . 15
2.5 Signal degradation mechanisms in the mobile wireless environment [42] . . . 16
2.6 Capacity as a function of bandwidth B after [45] . . . 19
2.7 Capacity as a function of SNR after [45] . . . 19
2.8 The boundary of rate pairs achievable by CDMA with SIC and TDMA after [45] 22
2.9 Capacity in Rayleigh fading environment for single user case . . . 23
2.10 Constellation diagrams of QPSK and 16-QAM . . . 25
2.11 Adaptive modulation and coding for frequency-domain scheduling [56] . . . 26
2.12 Multiuser diversity gain for Rayleigh fading after [57] . . . 28
2.13 Diagram of a MIMO communication system [65] . . . 33
2.14 Diagram of MIMO spatial multiplexing after [66] . . . 36

3.1 Two-link line model with mutual interference. In the figure BS_1 and UE_1 denote the base station and mobile user equipment in cell 1; whereas BS_2 and UE_2 denote the base station and mobile user equipment in cell 2. . . . 41
3.2 Required power of ST and OT with $R_1 = R_2 = 1$ bps/Hz . . . 46
3.3 Scheme selection for sum power minimization with $R_1 + R_2 = 12$ bps/Hz and R_1 uniformly distributed in the range [0,12] with 20 samples . . . 47
3.4 Scheme selection under different sum rate constraints ($R_1 = R_2$). ST is preferred below the curve, whereas OT is preferred above the curve. . . . 49
3.5 ECR under different bandwidth allocation for OT scheme . . . 51
3.6 Power, outage and delay tradeoffs with $R_1 = 1.5$ bps/Hz, $R_2 = 0.5$ bps/Hz . . . 54
3.7 Soft frequency reuse for interference mitigation [90] . . . 56
3.8 Flexible frequency allocation for a simple three link scenario . . . 58
3.9 Power-efficient ST/OT boundary locations for three-link scenarios . . . 59

4.1 GSM base station efficiency model [111] . . . 65
4.2 Suggested LTE base station efficiency model . . . 66
4.3 Normalized traffic in 3G network in London [118] . . . 67
4.4 Daily RNC level hourly data volume deviation [41] . . . 68
4.5 Cells' data volume contribution to total RNC data volume [41] . . . 69
4.6 Demonstration of traffic-dependent power control . . . 70
4.7 LTE general frame structure [119] . . . 71
4.8 LTE frame structure assumption for control signalling and data transmission . . . 71
4.9 Energy gains of time-domain base station sleep modes . . . 72
4.10 DL control signalling assumptions [119] . . . 75
4.11 Energy consumption for different numbers of active subframes . . . 79
4.12 Non-sleep/sleep mode energy consumption under different ALFs . . . 80

4.13 Energy consumption under various user's locations with fixed ALF. 81

5.1 LTE downlink control signalling structure [122] 85
5.2 Antenna port precoding for 4×4 MIMO [124] 88
5.3 CCE aggregation for PDCCH (a) Ergodic capacity of LTE control signalling channels (b) Probabilities of different CCE aggregation levels for PDCCH . . . 90
5.4 Averaged control overhead size under specific user distance 93
5.5 Control overhead size under different system bandwidth for SISO 94
5.6 Spectral efficiency of MIMO spatial multiplexing 95
5.7 Comparisons of energy consumption between SISO and MIMO 96
5.8 Energy efficiency regions under different antenna configurations 100
5.9 Energy consumption of MIMO sleep modes with 40 dB SINR 102
5.10 Energy consumption of MIMO sleep modes with 20 dB SINR 102
5.11 Antenna sleep patterns within 2×2 MIMO system 104
5.12 Antenna sleep patterns within 4×4 MIMO system 105
5.13 SISO and 4×4 MIMO control overhead size for multi-user scenarios 107

List of Tables

3.1 LTE physical layer model parameters 42
3.2 Power allocation algorithm with constraints 54

4.1 Estimated power consumption for base stations [113] 66
4.2 Model parameters for LTE downlink transmission 74

5.1 Estimated size of control signalling overhead for single user case 92
5.2 Multi-user scheduling incorporating sleep modes 107

Acronyms and Abbreviations

ACG:	Amplitude Carving Greedy
ADC:	Analog to Digital Converter
ALF:	Average Load Factor
AMC:	Adaptive Modulation and Coding
ASK:	Amplitude-Shift Keying
AP:	Access Point
AWGN:	Additive White Gaussian Noise
BABS:	Bandwidth Assignment Based on SNR
BG:	Bearer Gateway
BS:	Base Station
CCE:	Control Channel Element
CDMA:	Code Division Multiple Access
CSI:	Channel State Information
DAC:	Digital to Analog Converter
DCI:	Downlink Control Information
DRX:	Discontinuous Reception
DSP:	Digital Signal Processing
DTX:	Discontinuous Transmission
ECG:	Energy Consumption Gain
ECR:	Energy Consumption Rating
ERG:	Energy Reduction Gain
FDD:	Frequency Division Duplex
FDMA:	Frequency Division Multiple Access
FFR:	Fractional Frequency Reuse
FI:	Fairness Index
FRF:	Frequency Reuse Factor
FSK:	Frequency-Shift Keying
GSM:	Global System for Mobile Communications
HARQ:	Hybrid Automatic Repeat Request

HSDPA: High Speed Downlink Packet Access
ICT: Information and Communication Technology
LOS: Line of Sight
IP: Internet Protocol
ISI: Inter Symbol Interference
LDPC: Low-Density Parity-Check
LF: Load Factor
LTE: Long Term Evolution
MAC: Medium Access Control
MIB: Master Information Block
MIMO: Multiple-Input Multiple-Output
MISO: Multiple-Input Single-Output
OFDM: Orthogonal Frequency Division Multiplexing
OFDMA: Orthogonal Frequency Division Multiple Access
OSI: Open Systems Interconnection
OPEX: Operational Expense
OT: Orthogonal Transmission
PA: Power Amplifier
PBCH: Physical Broadcast Channel
PCFICH: Physical Control Format Indicator Channel
PCI: Physical Cell Identity
PDCCH: Physical Downlink Control Channel
PDSCH: Physical Downlink Shared Channel
PFI: Power Fairness Index
PHICH: Physical HARQ Indicator Channel
PI: Position Information
PMCH: Physical Multicast Channel
PRB: Physical Resource Block
PSK: Phase-Shift Keying
PSU: Power Supply Unit
QAM: Quadrature Amplitude Modulation
QoS: Quality of Service
QPSK: Quadrature Phase-Shift Keying

RBS:	Radio Base Station
RCG:	Rate Craving Greedy
RE:	Resource Element
REG:	Resource Element Group
RF:	Radio Frequency
RNC:	Radio Network Controller
RS:	Reference Signal
SC-FDMA:	Single Carrier Frequency Division Multiple Access
SDMA:	Space Division Multiple Access
SFR:	Soft Frequency Reuse
SIB:	System Information Block
SIC:	Successive Interference Cancelation
SIMO:	Single-Input Multiple-Output
SINR:	Signal to Interference Plus Noise Ratio
SISO:	Single-Input Single-Output
SM:	Spatial Multiplexing
SNR:	Signal to Noise Ratio
ST:	Simultaneous Transmission
STBC:	Space-Time Block Code
TDD:	Time Division Duplex
TDMA:	Time Division Multiple Access
TOC:	Top of Cabinet
UE:	User Equipment
UMTS:	Universal Mobile Telecommunications System
UT:	User Terminal
VCE:	Virtual Centre of Excellence
VoIP:	Voice over IP
WLAN:	Wireless Local Area Networks

Nomenclature

$P_{\text{operational}}$	operational power consumption
E_{total}	total system energy consumption
E_{embodied}	embodied energy consumption
t_{act}	active operating time
T	throughput
P	transmit power
E_{pro}	energy consumption of the proposed system
E_{b}	energy consumption of the baseline system
B	bandwidth
P_{r}	received power
y	received signal
x	transmitted signal
σ	channel noise amplitude
N_0	noise density
λ	channel state information amplitude
g	channel gain due to path loss, shadowing and fading *etc*
C	channel capacity
R	data rate
P_{out}	outage probability
$\mathbb{P}[\quad]$	probability operation
I	number of active links
τ	time instance
κ	bits per modulated symbol
ϱ	modulation alphabet
$[\quad]^*$	complex conjugate
N	number of receiver antennas
M	number of transmit antennas
$C_{N\times M}$	MIMO capacity with M transmit antennas and N receiver antennas
h	channel gain amplitude due to fading

O_{rd}	maximum spatial multiplexing order
υ	number of information bits
β	information bit
$\mathbf{I}_N$	N by N identity matrix
$\mathbb{E}[\quad]$	expectation operation
$\mathbf{H}$	channel state matrix due to fading
D	inter-site distance
d	transceiver distance
d'_{BP}	fiducial distance
L	path loss in decibel
f_{c}	carrier frequency
h_{BS}	base station antenna height
h_{MS}	mobile station antenna height
c	speed of light
T_0	temperature
K	Kelvin constant
$\top$	infinity
δf	subcarrier bandwidth
N_{sub}	number of subcarriers
ζ	number of subcarriers allocated to a specific user
f	frequency band ratio allocated to a specific user
n_{p}	optimum subcarrier allocation in terms of sum energy minimization
t	time ratio allocated to a specific user
t_{p}	optimum time allocation in terms of sum energy minimization
$E(t)$	energy consumption per second with TDMA
$E(f)$	energy consumption per second with FDMA
P_s	probability that orthogonal transmission is more energy-efficient
$\mathbb{R}$	network sum data rate
n_{u}	optimum subcarrier allocation in terms of power-utility minimization
n_{f}	optimum subcarrier allocation in terms of maximizing fairness
w	weighting factor of power-fairness tradeoff
G_{s}	channel gain due to shadowing in decibel
g_{s}	channel gain due to shadowing

n_{w}	optimum subcarrier allocation under weighted power-fairness tradeoff
ι	standard derivation of shadowing
D_{el}	delay constraint
$\bar{P}$	average system power constraint
P_{m}	peak power constraint
χ	number of subcarriers allocated to the cell edge users
J	cell edge junction
G	channel gain due to path loss in decibel
x_{s}	control signalling overhead size
R_{d}	required data rate for user information transmission
t_{d}	time for delivering user data in a radio frame
P_{d}	power for delivering user data in a radio frame
t_{f}	time length per radio frame
t_{s}	time for control signalling in a radio frame
E	energy consumption per frame for the baseline system
N_{act}	active number of subframes in a radio frame
P_{ds}	power for delivering user data in a radio frame with time-domain sleep modes
t_{ds}	time for delivering user data in a radio frame with time-domain sleep modes
t_{ss}	time for control signalling in a radio frame with time-domain sleep modes
E_{s}	energy consumption per frame with time-domain sleep modes
N_{opt}	optimum number of active subframes per frame for minimizing energy
$\mathbb{LW}[\quad]$	Lambert-W function operation
$\mathbb{F}[\quad]$	general function operation
N_{min}	minimum number of active subframes in a radio frame
R_{m}	achievable data rate in a radio frame under peak power constraint
R_{ergodic}	ergodic data rate per radio frame under peak power constraint
ρ	receiver signal to interference plus noise ratio
$\bar{\rho}$	average receiver signal to interference plus noise ratio
I_{int}	co-channel interference power
π	number of downlink control information bits
α	control channel element aggregation level
η	number of resource elements in a control channel element
δT	symbol time

N_{PRB}	number of physical resource block in a radio frame
g_0	channel gain threshold for truncated channel inversion
$E_{\mathrm{x,N\times M}}$	energy used for control signalling with $N \times M$ MIMO
$E_{\mathrm{d,N\times M}}$	energy used for data transmission with $N \times M$ MIMO
$R_{\mathrm{eq},NM/N'M'}$	data rate for equal energy consumption with $N \times M$ and $N' \times M'$ MIMO
$R_{\max}$	maximum data rate supported by a specific modulation
$r_{\mathrm{eq},m}$	data rate delivered by the spatial channel m
$h_{N\times M,m}$	$N \times M$ MIMO channel gain amplitude due to fading on spatial channel m

Chapter 1
Introduction

This thesis is devoted to the energy-efficient resource allocation design taking the long term evolution (LTE) system as an example. It focuses on the downlink transmission techniques from radio base stations (RBS) to mobile user terminals (UT). In the introductory chapter, the origin and motivation of this work will be provided in Section 1.1. The overview of the organization of the remaining chapters will be presented in Section 1.2. Finally, the main contributions of this thesis will be summarized in Section 1.3.

1.1 Motivation

Climate change and the need to reduce energy consumption are today widely acknowledged. In 2007 China overtook the US as the top CO_2 emitter, whilst its economic growth continues unabated at 10% *pa*. For a 24-hour UK digital broadcast TV network the power consumption is about 2MW. For comparison, a typical mobile phone network in the UK may consume about 40 MW, excluding the power consumed by the users' handsets. For comparison, mobile networks in China today will have about 10x the number of subscribers, and proportionately higher power consumption reflecting the size of the country; as subscriber numbers grow these figures will increase. The carbon footprint will become an increasingly important constraint in future system design. Put simply, there is a need on environmental grounds to reduce the energy requirements of radio access networks. From an operator's perspective, reduced energy consumption translates directly to the bottom line: lower operating expenditure (OPEX). These are the key drivers of the Green Radio programme [1].

In order to reduce the energy consumption of mobile phone networks, relevant techniques can be either derived from the network architecture aspect or the protocol stack aspect. The network architecture aspect refers to changing the existing network structures to offer potential reduction in energy consumption. For example, moving the access network closer to the user using femtocells enables a reduction in the required transmit power, although this is achieved at the expense of increased backhaul requirements. Suitable architectures and topologies which

trade off these factors to achieve a net gain must be identified. Regarding this, [2] discussed four fundamental tradeoffs in mobile wireless networks: deployment efficiency and energy efficiency tradeoff, spectrum efficiency and energy efficiency tradeoff, bandwidth-power tradeoff, and delay-power tradeoff. With the help of these tradeoffs, key network performance/cost indicators can be analyzed together. In [3], the effects of cell size on energy saving and system capacity are analyzed, which suggests small-cell based future mobile communication systems have the potential to improve energy efficiency. However, system interference needs to be migrated for the sake of energy conservation. An increased density of wireless access devices require the effective utilization of multiple access techniques. In [4], a novel frequency-domain multiple access scheme for uplink transmission is presented, providing better performance than the single carrier frequency division multiple access (SC-FDMA) strategies in terms of power efficiency, robustness, flexibility and scalability. In [5], energy efficient resource allocation is analyzed based on time division multiple access (TDMA) over fading channels. Furthermore, incorporating relay into the network architecture also has the potential to achieve energy savings. Exploitation of location and mobility information with regard to relays can enhance these benefits. In [6], the energy efficiency by deploying different relay schemes such as amplify and forward, decode and forward, and block Markov coding are compared with that of direct transmission. In [7], a new approach of relay selection using a threshold-based transmission protocol is proposed for a system with multiple amplify and forward relays and users.

Techniques across protocol stack will also help with energy conservation. This refers to either reducing the required radiated power to achieve certain quality of service (QoS), or reducing the overall power to achieve the required radiated power. For the former case, effective radio resource management and signal processing may reduce the radiated power requirement to maintain the required QoS. In [8], cognitive radio is used to deliver Green Communications. By cognitively determining the assignment and use of the available spectrum, the modulation level can be adapted to minimize energy consumption. In [9], a game-theoretic model is proposed to study the cross-layer problem of joint power and rate control with quality of service constraints in multiple access networks. For the latter case, power and energy efficient hardware implementation of the necessary functionality will facilitate a reduction in total energy consumed relative to the radiated power. In [18], Nokia has developed a new antenna technology which incorporates elements of the base station within the antenna, thus reducing the footprint of base stations and lowering power consumption. In [19], a high energy efficiency multimode broadband power amplifier is proposed for the LTE base stations.

Although various techniques can contribute to energy-efficient communication system design, here we only focus on the resource allocation problem for LTE downlink transmission. Since the design of the LTE physical layer is heavily influenced by the requirements for high peak transmission rate (100 Mbps downlink/50 Mbps uplink), spectral efficiency, and multiple channel bandwidths (1.25-20 MHz), orthogonal frequency division multiplex (OFDM) was selected as the basis for the LTE physical layer to fulfill these requirements. The unit of resource assignment is a physical resource block (PRB), which is 10 ms (1 radio frame) in the time-domain and 180 KHz (12 subcarriers) in the frequency domain. Within a specific cell, different users are assigned different PRBs, thus interference can be avoided if perfect orthogonality is assumed among different subcarriers. For users in different cells, co-channel interference may still exist if the same frequency bands are used for transmission. For energy-efficient design, an appropriate allocation of subcarriers and frequency-domain scheduling across multiple cells can help reduce the overall network interference. For example, geographically separating the two active links in the neighboring cells will generally improve signal to interference plus noise ratio (SINR). To maximize system reuse efficiency, soft frequency reuse (SFR) techniques are also suggested in LTE [10]. Although interference mitigation has been studied for many years, most of these research targets to improve system throughput, or spectral efficiency. Further analysis is required from an energy efficiency perspective.

Apart from the considerations in the frequency-domain, energy saving can also be obtained from the time-domain. Instead of focusing on the multi-cell cooperation, here we start with analyzing LTE control signalling. In LTE, 1 radio frame is composed of 10 subframes, each of which is 1 ms in time. For each subframe, about 5% to 30% of the time is used for control signaling, which generally includes the signals for channel estimation, user-specific resource assignment, system configuration and synchronization *etc*. In the first release of LTE, these signals are always delivered in each subframe even there is no user data for transmission. Therefore, potential energy saving can also be achieved by gating off the control signalling which is not necessary in the time-domain. During the inactive time period, the energy used for supporting the relevant hardware operation may also decrease. This is similar to the time-domain sleep modes in [11] and [12], where the mobile user terminals can be switched among different modes based on the current traffic loads. The difference is that here we focus on switching off base stations rather than user terminals. This is due to the base station energy contributing to a large proportion of the total network energy consumption. Designing base station time-domain sleep modes may result in large energy reduction gain. Also, it is expected that the energy used

for operating the base station circuits and cooling etc will reduce correspondingly.

Further, the sleep mode idea can be applied in the spatial-domain as well. As multiple-input multiple-output (MIMO) system is now widely deployed. It could be argued that switching off antennas could conserve energy. From the circuit energy consumption perspective, power amplifiers which are often equipped on each individual antenna generally consume lots of energy. Thus gating off radio antennas can reduce the energy for operating antenna circuits. However, as MIMO can improve spectral efficiency and enable high data rate transmission based on spatial diversity and multiplexing, whether switching off antennas could reduce radio frequency (RF) energy consumption is still an open problem.

1.2 Thesis Structure

As discussed above, co-channel interference, excessive control signalling and relevant hardware energy consumption are the main sources that waste energy in LTE. Taking these into consideration, in this thesis we will study the resource allocation problem from the frequency, time and spatial-domain perspectives. The thesis structure is summarized as follows:

Chapter 2

This chapter starts with introducing the concept of Green Radio, specifying the potentials and targets of energy efficient communication system design. It provides a literature review of the current energy efficiency metrics, which are essential for defining the optimization problems in the following technical chapters. Then, some background information about cellular communication, mobile channel propagation models and the Shannon capacity is introduced. After that, some basic theoretical methods of power allocation are described for spectral efficiency maximization, following by discussions of scheduling, efficiency-fairness tradeoffs, system constraints and discontinuous transmission/reception modes from a practical perspective. Finally, we present some basic introduction to MIMO systems.

Chapter 3

This chapter studies the effects of frequency allocation on network energy minimization. Assuming a two link scenario, several multiple access techniques for achieving the same data rate requirements are compared in terms of energy consumption. Analysis and proofs are provided to illustrate whether non-orthogonal transmission or orthogonal transmission is more energy-efficient. After that, new metrics of energy efficiency and fairness are proposed. The tradeoffs

among energy efficiency, fairness and outage probability are further discussed. Finally, we extend our study to a multiple link scenario and demonstrate the idea of deploying energy-efficient soft frequency reuse.

Chapter 4

This chapter describes the effects of switching off radio base stations during certain time periods on energy conservation. Starting with the analysis of base station energy consumption models and network traffic profiles, the potential of deploying time-domain base station sleep modes is summarized. Then, analysis regarding how to choose the sleep mode duty cycle is presented based on the assumption of a slow fading scenario with a single user single-input single-output (SISO) system. Finally, simulations demonstrate the significant energy reduction gain of the proposed scheme.

Chapter 5

This chapter extends the sleep mode idea described in Chapter 4 into the spatial-domain. Instead of switching off a number of subframes in the time-domain, individual transmit antennas can be gated off depending on the network traffic profile. An detailed estimate of control signalling overhead size is provided based on the LTE physical layer structure. Then, the energy consumption of SISO, 2×2 MIMO and 4×4 MIMO is compared based on the energy tradeoff of LTE control signalling and user data transmission. Assuming known channel conditions and user data rate requirements, an antenna selection criterion is further proposed. In addition, the combined spatial-time sleep modes are simulated, followed by a brief discussion of multi-user scenario extensions and scheduling strategies. Finally, some practical application issues such as how to deploy sleep modes in the delay-sensitive Voice over IP (VoIP) application, and how to incorporate them into legacy handsets are suggested.

Chapter 6

Finally, some conclusion remarks are presented and possible future work is described.

1.3 Contributions

Contributions from this research for energy-efficient resource management in LTE systems can be summarized as follows:

- Mathematical analysis and modeling of two possible frequency allocation scenarios: non-

orthogonal transmission and orthogonal transmission based on a two-link scenario. Assuming the same rate targets and channel conditions, it is proved that the time division multiple access (TDMA) based orthogonal transmission has the same energy efficiency but higher peak power requirement compared to the frequency division multiple access (FDMA) based orthogonal transmission. Also, non-orthogonal full bandwidth transmission outperforms FDMA-based orthogonal transmission at near-cell-center area in terms of energy efficiency. This distance generally depends on the sum network data rate requirements [13].

- Proposed a new energy efficiency metric 'power consumption per bit' and its corresponding metric 'power fairness index' [13]. Compared with the old metrics, data rates are considered to guarantee that reducing energy consumption would not sacrifice users' quality of service. Also, the new metric can be easily combined with other embodied and operational energy consumption terms.

- Time-domain base station sleep mode optimization for the single user SISO scenario. By switching off a number of subframes per frame according to the current traffic profile, energy reduction can be obtained by delivering less control signalling overhead. Based on the base station energy consumption model, the energy used by operating antenna power amplifiers (PA), cooling the base station, processing the baseband signals *etc* are also expected to be reduced. The optimum time-domain sleep mode duty cycle is derived in a closed-form. Simulation results suggest up to 90% energy reduction gains [14].

- Estimated the LTE control signalling overhead size for different antenna configurations and control channel element (CCE) aggregation levels. Assuming a fixed transmission bandwidth, calculated the proportion each control channel contributes, including reference signals (RS), synchronization signals, the physical downlink control channel (PDCCH), the physical broadcast channel (PBCH), the physical control format indicator channel (PCFICH), the physical HARQ indicator channel (PHICH) and the physical multicast channel (PMCH) *etc*. The amount of reference signals depends on the number of transmit antennas, which contributes a high proportion of the total overhead. The amount of PDCCHs depends on the CCE aggregation levels, or channel state information, which also occupies a large proportion [15].

- Spatial-domain sleep mode design by gating off a number of transmit antennas. With known channel state information (CSI) and data rate requirements, the optimum antenna

configurations are identified to minimize energy consumption. By gating off individual transmit antennas, the corresponding energy consumed by the antenna power amplifiers can also be reduced [15] [16].

- Provided analysis of the combined spatial-time sleep modes. Simulation results suggest that without the time-domain sleep mode, switching off radio antennas can achieve significant energy reduction. With the time-domain sleep mode being enabled, the obtained energy gains of further deploying spatial-domain optimization are trivial. Due to the complexity issues, using all available antennas is generally a good choice [15] [17].

Chapter 2
Energy Efficient Communication Systems

2.1 Introduction

As wireless is becoming all-pervasive, the vision of the future information society is growing more complex, moving from being an add-on to wireline communications to being an essential enabler for multi-service ubiquitous communications. In this circumstance, intelligent design for system optimization under different service requirements is imminently needed. Regarding network and transport layers, an auto-configurable internet protocol (IP) based network is suggested, adding great flexibility to future communication system design. For lower layers, efficient adaptation and optimization strategies for wireless resource management are also essential under various QoS requirements [20].

Apart from these changes in the communication domain, its impacts on the environment must be minimized. The potential market size and social acceptance always affect the emergence of new inventions and technologies. With the expansion of population, the number of mobile users are also increasing, implying a growth in demand for communications services and increased capacity. Besides, it is noticed that the population density is geographically-dependent. Thus, to achieve higher technical and economic performance, different strategies may be applicable to different scenarios. For high-density areas, increased number of links will cause severe interference, which significantly degrades quality of service. High energy consumption also challenges the power limits of communication devices. For low-density areas, inefficient use of wireless devices makes communication expensive, due to the high cost in enabling and maintaining effective long distance communications.

At the same time, wireless technologies are continually evolving, adapting to the new requirements in modern communications. For example, the LTE standard is designed to support a wide range of terminals and high data rates. It also supports flexible carrier bandwidths based on a simplified and IP-based structure [21] [22]. These new network features provide us more

space for designing intelligent software solutions such as energy-efficient resource allocation and packet scheduling strategies.

Based on the above points, the concept of Green Radio is proposed, aiming at economically and environmentally efficient system design in modern communications [16]. Two aspects have been focused on for Green Radio to date. One is to identify a green network architecture, which means designing a low power wireless network and backhaul that still provides good QoS. The other one is to identify the best radio techniques across all layers of the protocol stack that collectively address the aspiration of achieving better power and energy efficiency. Here, our interests are particularly drawn to the resource management aspect.

Starting with introducing the potential for Green Radio and targets for energy conservation in Section 2.2, this chapter will provide some relevant background information for our energy-efficient resource allocation schemes later in the thesis. In Section 2.3, we will summarize the key points of wireless communication environments. In Section 2.4, the concepts of channel capacities will be introduced, following by a brief description of resource management. Finally, in Section 2.5 we will describe the basics of MIMO system.

2.2 Green Radio

Capturing the rapid increase of information and communication technology (ICT) energy consumption, this section describes the objectives of Green Radio project and the relevant solutions. A brief introduction to energy efficiency metrics is also provided.

2.2.1 Potentials and Targets

Currently there are 4 billion mobile phone users in the world. With the global population expansion and the prevalence of mobile phone devices, energy consumption and carbon dioxide emissions will become a major concern of wireless network industries. The Mobile Virtual Centre of Excellence (VCE) 2020 Vision documents state that a typical mobile phone network in the UK consumes approximately 40 MW, excluding the power consumed by the users' handsets [1]. The worldwide telecommunications industry is currently responsible for 183 million tones or 0.7% of the total carbon dioxide emissions, a figure which is increasing at a rapid rate [23].

Meanwhile, the high energy prices will increase the cost of operating cellular systems. As shown in Figure 2.1, the network traffic including voice and data is expected to increase exponentially over time in the next few years. This will lead to an exponential increase of the network energy cost. However, the revenue is expected to grow only linearly. The diverging expectations for traffic and revenue growth suggest that we must reduce the energy use per bit carried very significantly.

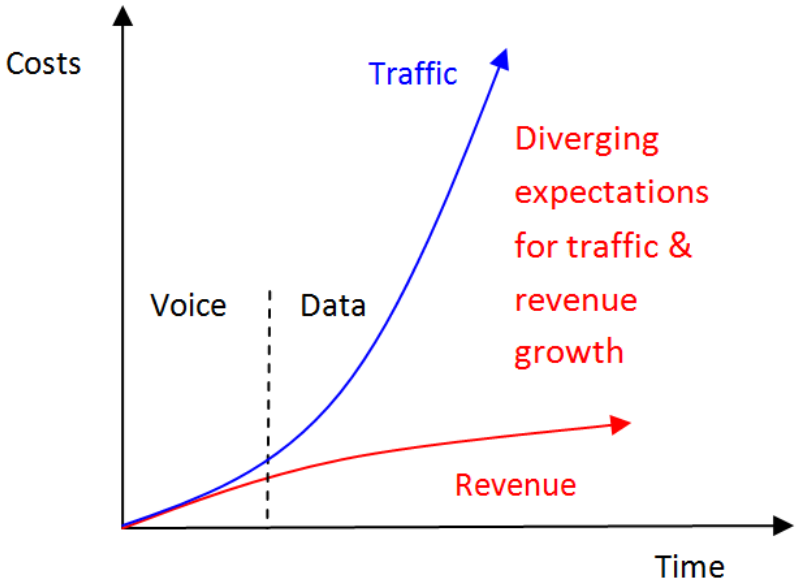

Figure 2.1: *Mobile network traffic and revenue growth [24]*

In order to develop more energy efficient and environmentally friendly wireless networks, the Core 5 research programme of Mobile VCE decided to focus on the Green Radio concept. Using high speed downlink packet access (HSDPA) as the baseline system, it aims to deliver high data rate services with a 100x reduction in power consumption over current wireless communication networks. This includes life time and embodied energy impacts, which refer to the energy used to make the product, bring it to market, and dispose of it over the whole life cycle. Since the proposed wireless access systems need to support multimedia service data rates at much lower transmission energy levels, they will not come simply from more traditional research on single aspects of the physical layer but will require holistic thinking. By 2020, Vodafone aspirates the CO_2 emissions can be reduced by 50% [1].

In order to hit this target, we start with analyzing cellular network energy consumption. Based on the results shown in Figure 2.2, radio base stations are identified to be the most energy-

consuming components, which use about 57% of the total network power. The energy spent on mobile switching, core transmission, data center and retail is about 20%, 15%, 6% and 2% respectively. Therefore, currently most energy-efficient designs are targeted towards radio base stations.

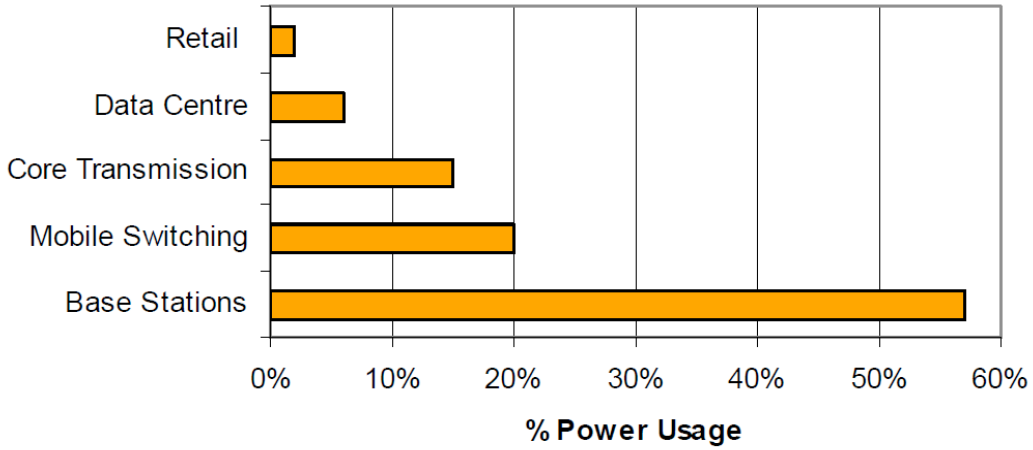

Figure 2.2: *Cellular network power consumption [24]*

2.2.2 Relevant Techniques

In order to design energy-efficient radio base stations, an energy budget model which calculates the overall energy consumption and losses in a cellular network during data transmission over a period of time is proposed [25]. The model takes into account the total energy consumption in various parts of the communications system and the total energy consumption over the equipment's lifetime, terms as the embodied energy. As shown in Figure 2.3, the RBS energy consumption generally includes the embodied energy and the operational energy. The operational energy refers to the energy used for operating the RBS, which scales proportionally with the operating time. Using $P_{\text{operational}}$ to denote the operational power consumption per unit time, Figure 2.3 suggests some essential factors which may contribute to this part of power consumption, including traffic-dependent resource management, radio antenna patterns, cell size, user density and interference mitigation *etc*. In general, this part of energy consumption is heavily dependent on the communication system design, which is the energy term Mobile VCE Green Radio project expects to reduce.

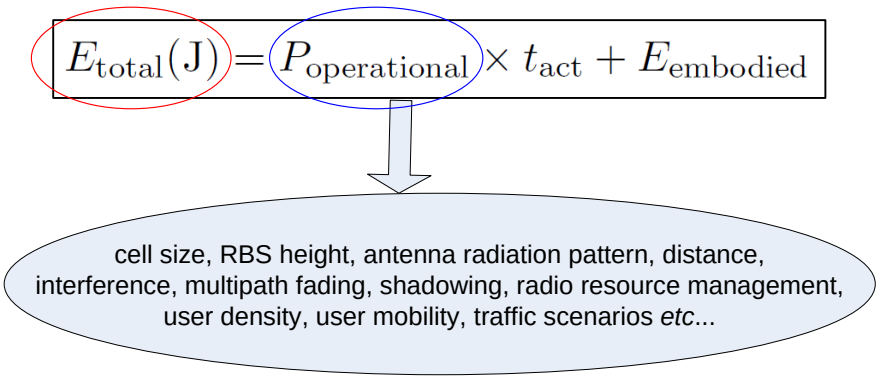

Figure 2.3: *Base station energy budget model [25]*

Regarding this, some efforts have been targeted towards the RBS architectural aspect. For example, in the paper [26], the problem of whether deploy large-cell or small-cell is discussed, indicating energy-efficient cell configuration. In [27], a new store carry and forward relay scheme is proposed to reduce network energy consumption. At the same time, there are some efforts based on physical layer and media access control (MAC) layer design. In [28], interference cancellation techniques are studied for MIMO systems. In [29], the bandwidth expansion mode is proposed for LTE advanced, enabling us to further lower energy consumption.

In this thesis, we focus on physical and MAC layer design for energy-efficient radio resource management. From a theoretical point of view, the difference between energy minimization and spectral maximization needs to be distinguished. This is based on the analysis of Shannon capacity formula and optimization functions. Apart from this, some new techniques which bring innovative ideas into resource allocation solutions need to be considered. For example, [30] analyzed different power consumption models for different base station types in heterogeneous cellular networks. These models may have potential influence on energy-efficient resource management, since the operational energy consumption of hardware devices and embodied energy term may also be included for energy consumption optimization. In [31], new solutions for energy-efficient power amplifier design are presented, which will further impact on the energy consumption models.

2.2.3 Energy Efficiency Metrics

To enable the assessment and comparison of the energy or power efficiency in telecommunication networks, suitable metrics need to be identified. In general, the energy and power metrics have to take into account the following factors and parameters [32]:

- The network load (as affected by the fraction of active network elements) and the transmission resources used (such as the bandwidth, time and power);
- Network architecture and network topology;
- The quality of service requirements (including reliability, latency and throughput as specified at a given open system interconnection (OSI) layer);
- User behaviors, profiles and mobility models;
- The physical layer interface, signaling, transmission format and the protocol stack used;
- The traffic models and patterns;
- Communication and application scenarios and contexts.

Currently, the most commonly used energy efficiency metric is to minimize the total power consumption in the network [33]. [34] proposed novel resource allocation schemes which minimize the power consumption of mobile stations in orthogonal frequency division multiple access (OFDMA) systems. In [35], a resource allocation strategy combining sub-carriers, bits and power allocation for a multi-user network is documented to minimize power consumption of base stations. However, with different service requirements, it is difficult to find a corresponding fairness measure. If we aim to improve power fairness by assigning equal power to all the links, it is obviously meaningless as different link may have different data rate requirements, peak power and delay constraints *etc*.

Taking this issue into consideration, another power efficiency metric is proposed by summing up the 'data rate per unit power achieved' in that network [36]. Combining energy minimization with spectral maximization, a novel resource allocation scheme based on game theory is discussed. Similarly, [37] has employed this metric for performance evaluation in an *ad-hoc* network operating in time division duplex (TDD) mode. Regarding fairness definitions, it is noticed that aiming to achieve the equal 'rate per unit power achieved' on all links is also a

useful criterion, since it makes sense to allocate more power to the user with higher data rate requirements.

Moreover, we can extend this concept to minimize the sum 'power consumed per bit' of all the links. The energy consumption rating (ECR) is defined as $\mathrm{ECR} = P/T$ W/bps, where T is the maximum throughput in bits per second (bps) reached during the measurement, and P is the measured transmit power in Watts while performing the measurement test. The lower the ECR, the less energy is consumed to transport the same amount of data. The ECR metric can be used to measure energy efficiency of the telecommunication equipments as well as telecommunication networks. The ECR can be also considered to be a payload-normalized power efficiency of the equipment or the network. Compared with the metric of maximizing 'data rate per unit power achieved', this metric also takes the throughput T into consideration, thus having maintained users' quality of service requirements while minimizing energy. Further, this metric can be easily extended to include the hardware energy consumption as well.

For all the energy efficiency metrics mentioned above, we can use them to set optimization objectives. In order to verify the system design, another two metrics are proposed. The first one is energy consumption gain (ECG), which is the quotient of the proposed system energy consumption divided by the baseline system energy consumption $\mathrm{ECG} = E_{\mathrm{pro}}/E_{\mathrm{b}}$. The other one is energy reduction gain (ERG), which is defined as the energy saving of the proposed system divided by the energy consumption of the baseline system $\mathrm{ERG} = (E_{\mathrm{b}} - E_{\mathrm{pro}})/E_{\mathrm{b}}$. Here, E_{pro} and E_{b} are the energy consumption of the proposed system and baseline system respectively.

2.3 Wireless Communication Environment

This section introduces the basic concepts of cellular environments and mobile radio propagation. It also provides some historical background of the communication system evolution, indicating the trends for designing new wireless technologies.

2.3.1 Cellular Environment

The first land mobile communication systems were based on wide area transmission [38]. Each base station had to provide coverage for large geographical zones. Calls of customers leav-

ing a zone had to be dropped and re-established in a new zone. Such systems suffered from low-capacity and high-transmit power requirements for mobile transceivers [39]. Due to these drawbacks, the cellular mobile communication system was introduced by enabling frequencies used in one cell to be reused under certain conditions in other cells to increase capacity.

The first generation (1G) cellular mobile communication system was introduced in late 1970s and early 1980s. It is characterized by analogue (frequency modulation) voice transmission and limited flexibility [40]. Capacity increase was one of the main motivations for introducing the second generation (2G) systems in early 1990s. It also enhanced security and integrated voice and data owing to the use of digital technology. An example of the 2G system in its later stage is the global system for mobile communications (GSM). With the demand for high and variable bit-rate multimedia services, the third generation (3G) systems such as wideband code division multiple access (CDMA) and universal mobile telecommunications system (UMTS) were designed. Although 3G systems can in theory sustain high bit-rates over the air interface due to its wideband carriers, it can not meet the rapid increase of traffic during these years. Higher data rate requirements call for new standards to be established. As shown in Figure 2.4, the LTE system is now being deployed to provide high throughput, low latency and high spectral efficiency.

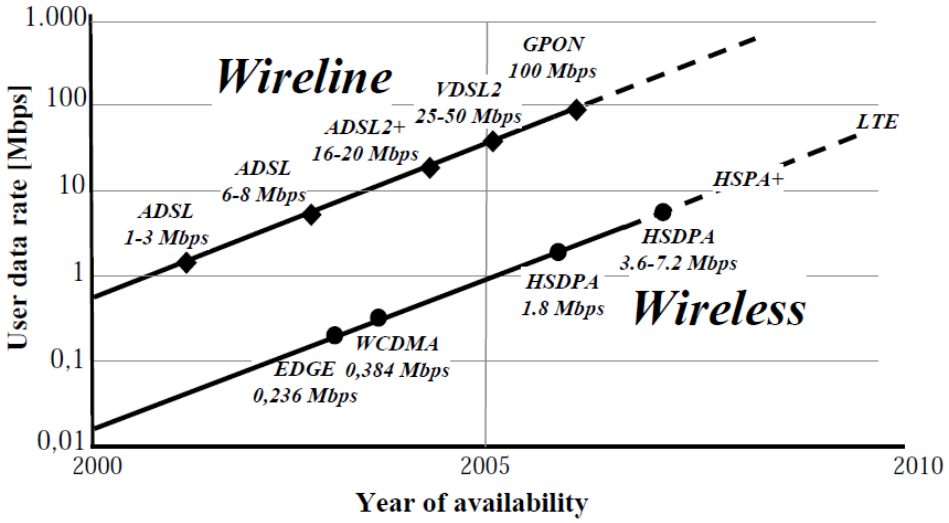

Figure 2.4: *Evolution of wireless and wireline user data rates [41]*

Since the cellular environment enables spectrum reuse, co-channel interference exists when multiple links are operating on the same spectrum bands. At the receiver, the signal to noise

plus interference ratio is defined as the receiver side signal power divided by the sum of noise and interference power. With increased interference or decreased SINR, the likelihood of successfully decoding the delivered information will degrade. Thus under the framework of cellular communications, spectrum allocation is designed to minimize this effect. The concept of soft frequency reuse (SFR) is introduced for interference mitigation, which allows the flexible sharing of subcarriers orthogonally or non-orthogonally among the neighboring cells. Frequency reuse factor (FRF) is further defined as the rate at which the same frequency can be used in the network. It equals to the number of cells which cannot use the same frequencies for transmission. Common values for the frequency reuse factor are 1/3, 1/4, 1/7, 1/9 and 1/12.

2.3.2 Mobile Radio Propagation

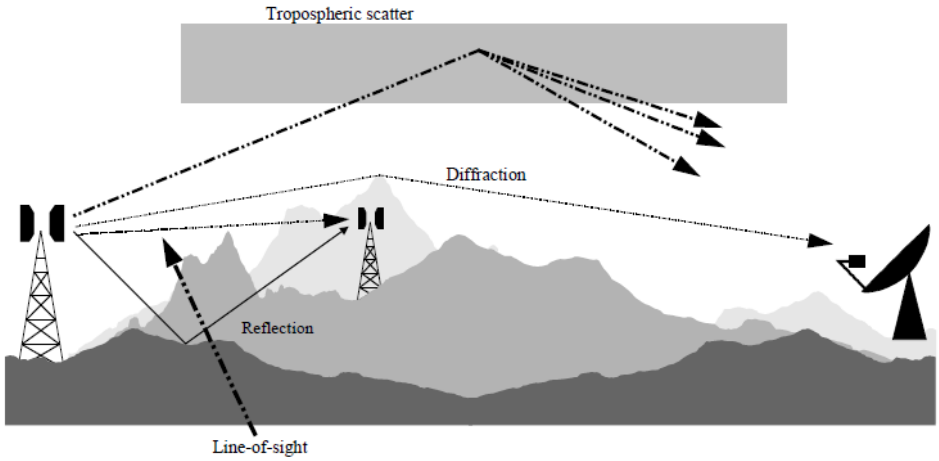

Figure 2.5: *Signal degradation mechanisms in the mobile wireless environment [42]*

The mobile radio environment is particularly hostile to reliable communications. There are many reasons for this, several of which are illustrated in Figure 2.5 which shows a conceptualized cell in a wireless network. The difficulties that can be expected in wireless communications include

- Diffraction: occurs typically when the radio wave passes over the edge of a building or through a small gap; diffraction can occur multiple times for a single path.
- Reflection: the mechanism of the signal being reflected off surfaces such as buildings, vehicles or trees.

- Long time delay: a situation where signals can arrive at the receiver with significantly different time delays causing inter symbol interference (ISI) as information from the previous data transmission may arrive at the receiver when it is expecting information about the next data transmission.

Due to the wireless propagation conditions experienced on the radio channel, the design of cellular communication systems is quite challenging. The pathloss, shadowing and multi-path fading are the three main propagation effects of the mobile radio channel. The path loss describes the average signal attenuation as a function of distance between the transmitter and receiver, which includes the free-space attenuation, transmit power and antenna gain. Shadowing describes slow signal fluctuations, which are typically caused by large structures, such as big buildings, obstructing the propagation paths. Fast fading is caused by the fact that signals propagate from transmitter to receiver through multiple paths, which can add at the receiver constructively or destructively depending on the relative signal phases [40]. Further, fading can be divided into flat fading and frequency-selective fading. While the former refers that the fading gain is constant across the bandwidth, the later refers to the coherence bandwidth of the channel is smaller than the bandwidth of the signal.

When designing cellular communication systems and particularly when planning the deployment of such systems, one will have to account for these propagation phenomena appropriately. Different statistical models are required for different scenarios. For example, in [43], WINNER II channel models generally include indoor office, indoor to outdoor, urban micro-cell, urban macro-cell, rural macro-cell and moving network scenarios. In the fourth generation LTE systems, a LTE channel model is also proposed in [44]. Using these channel models, we can simulate the performance of the new designs towards any specific scenario.

2.4 Capacity

This section studies the performance limits in communication systems. Based on the point-to-point communication scenario, the concepts of ergodic capacity, outage capacity and capacity region are described.

2.4.1 Ergodic Capacity

The framework for studying performance limits in communication is information theory, which was invented by Claude Shannon in 1948. He showed that with intelligent coding, one can communicate at a strictly positive rate but at the same time with as small an error probability as desired. However, there is a maximum rate, called the capacity of the channel. If one attempts to communicate at rates above the channel capacity, then it is impossible to drive the error probability to zero. The capacity of a channel is the basic measure of a communication system performance, it is the maximum rate of communication for which arbitrarily small error probability can be achieved [45].

Firstly, assume a wireless channel with bandwidth B Hz, received power P_{r} Watts and additive white Gaussian noise with power spectral density $N_0/2$. Following the passband-baseband conversion and sampling at rate $1/B$, this can be represented by a discrete time complex baseband channel:

$$y[\tau] = x[\tau] + \sigma[\tau], \tag{2.1}$$

where the noise σ follows an independent and identical distribution (i.i.d.) over time instances τ. The capacity of an AWGN channel can be represented as

$$C_{\mathrm{awgn}}(P_{\mathrm{r}}, B)(\mathrm{bps}) = B \log_2 \left(1 + \frac{P_{\mathrm{r}}}{N_0 B}\right). \tag{2.2}$$

Noting the signal to noise ratio $SNR = P_{\mathrm{r}}/(N_0 B)$, the AWGN channel capacity can be rewritten as

$$C_{\mathrm{awgn}}(P_{\mathrm{r}}, B)(\mathrm{bps}) = B \log_2 (1 + \rho). \tag{2.3}$$

As the formulas suggest, P_{r} and B are the two basic components of the communication channel. We would like to check how the capacity depends on the received power P_{r} and bandwidth B. Based on the logarithmic function

$$\begin{aligned} \log_2(1 + \varpi) &= \varpi \log_2 e \quad \varpi \approx 0 \\ \log_2(1 + \varpi) &= \log_2 \varpi \quad \varpi \gg 1 \end{aligned} \tag{2.4}$$

where ϖ is a general variable, we can plot Figure 2.6 and Figure 2.7. As B increases, the capacity increases monotonically and approaches the asymptotic limit corresponding to the $\varpi \approx 0$ case; as P_{r} or ρ increase, the capacity increases logarithmically corresponding to the $\varpi \gg 1$

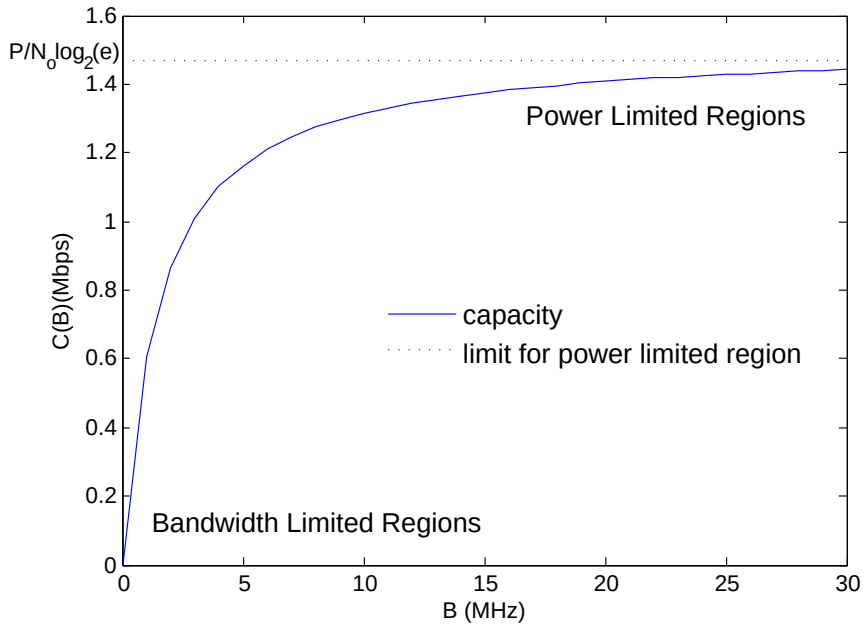

Figure 2.6: *Capacity as a function of bandwidth B after [45]*

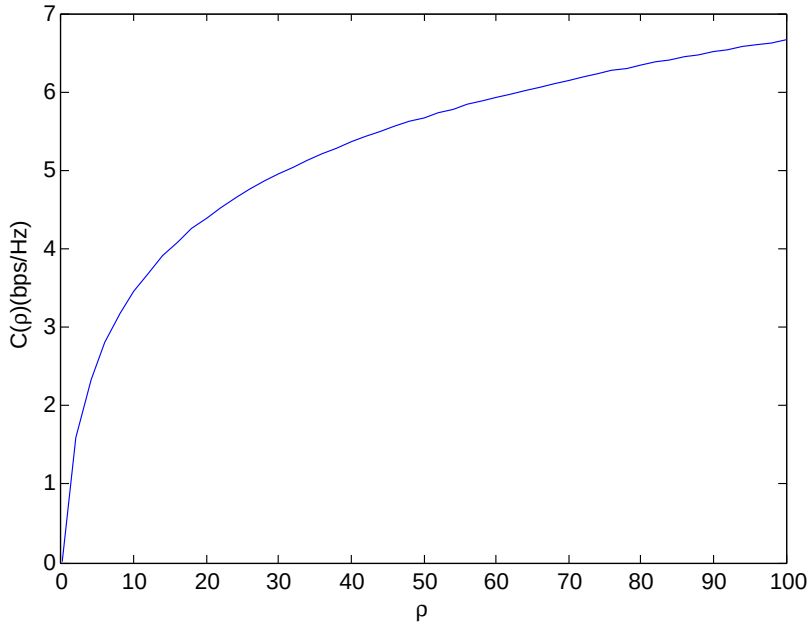

Figure 2.7: *Capacity as a function of SNR after [45]*

case. Because the wireless channel varies over time due to the factors mentioned in Section 2.3.2, for a fixed transmit power P, the received power and the achieved capacity are time-dependent. Therefore, the concept of ergodic capacity is proposed in Equation (2.5) by taking the expectation of the instantaneous channel capacities, where $g[\tau]$ and $P_{\mathrm{r}}[\tau]$ represent the in-

stantaneous channel gain and received power depending on the time instance τ respectively. Ergodic capacity describes the average performance of a communication link.

$$C_{\text{ergodic}}(P_{\text{r}}[\tau], B)(\text{bps}) = \mathbb{E}\left[B\log_2\left(1+\frac{P_{\text{r}}[\tau]}{N_0 B}\right)\right] = \mathbb{E}\left[B\log_2\left(1+\frac{Pg[\tau]}{N_0 B}\right)\right]. \quad (2.5)$$

2.4.2 Outage Capacity

In the remainder of the thesis, the Shannon bandwidth is normalized to 1. Supposing the transmitter encodes data at a rate R bps/Hz, if the channel attenuation $g[\tau]$ is such that $\log_2\left(1+\frac{Pg[\tau]}{N_0 B}\right) < R$, then whatever the code used by the transmitter, the decoding error probability cannot be made arbitrarily small [45]. The system is said to be in outage, and the outage probability is

$$P_{\text{out}}(R) = \mathbb{P}\left[\log_2\left(1+\frac{Pg[\tau]}{N_0 B}\right) < R\right]. \quad (2.6)$$

By constraining the maximum value of the acceptable outage probability P_{out}, we can define the outage capacity: the maximum data rate which can be delivered while guaranteeing the decoding error probability is less or equal to P_{out}. While the ergodic capacity can be used to describe the average channel quality, the outage capacity is more application-dependent, since different applications deploy different decoding mechanisms, allowing different bit error rates. When there are multiple users being scheduled, different users can accept different outage probabilities, which impacts on the resource allocation decisions. Typical communication systems can accept 1% to 10% P_{out}. Also, while the ergodic capacity is usually used for designing the systems with elastic traffic, the outage capacity is a common metric for delay-limited applications.

2.4.3 Capacity Region

When there are multiple active links, the capacity region is used to describe the feasible sets of data rates which the channel can support. Here we use a two-user downlink scenario, featuring a single transmitter (the base station) sending separate information. The baseband downlink AWGN channel with two users is

$$y_i[\tau] = \lambda_i[\tau]x_i[\tau] + \sigma_i[\tau] \quad i = \{1, 2\} \quad (2.7)$$

where the complex Gaussian noise σ_i is independent and identically distributed and drawn from a zero-mean normal distribution with variance N_0B. λ_i denotes the instantaneous channel state information satisfying $g_i[\tau] = \lambda_i[\tau]^2$. The scalar $y_i[\tau]$ is the received signal of user i at time instance τ for both users $i = \{1, 2\}$. With two links sharing the same frequency band, interference exists which will degrade the receiver side signal to noise ratio. Successive interference cancellation (SIC) can be introduced, which allows the good link to decode the data intended for the bad link. Then the good link can subtract out the transmit power of the bad link and decode its data. For example, if $g_1 < g_2$, with successive interference cancellation and the transmit power $P = P_1 + P_2$, the following rate pairs can be achieved [45]

$$\begin{cases} C_1(\text{bps/Hz}) = \log\left(1 + \dfrac{P_1 g_1}{P_2 g_2 + N_0 B}\right) \\ C_2(\text{bps/Hz}) = \log\left(1 + \dfrac{P_2 g_2}{N_0 B}\right) \\ \text{s.b.} \quad C_1 + C_2 \leq \log\left(1 + \dfrac{P_1 g_1 + P_2 g_2}{N_0 B}\right). \end{cases} \tag{2.8}$$

Successive interference cancellation is usually used for code division multiple access. With SIC, a larger capacity region can be achieved compared with the orthogonal scheme such as time division multiple access. In Equation (2.9), $t \in (0, 1)$ is used to denote the time proportion assigned to user 1, the achieved data rate pairs can be represented as

$$\begin{cases} C_1(\text{bps/Hz}) = t \log\left(1 + \dfrac{P_1 g_1}{t N_0 B}\right) \\ C_2(\text{bps/Hz}) = (1-t) \log\left(1 + \dfrac{P_2 g_2}{(1-t) N_0 B}\right). \end{cases} \tag{2.9}$$

As shown in Figure 2.8, a two-user asymmetric downlink AWGN channel with the user SNR equal to 0 and 20 dB is assumed. In the orthogonal scheme TDMA, both the power and time splits are jointly optimized to compute the capacity boundary. We observe that the performance of CDMA with SIC is better than that of the orthogonal scheme. Reference [45] further points out that the gap in performance of these two schemes is more pronounced when the asymmetry between the two users deepens.

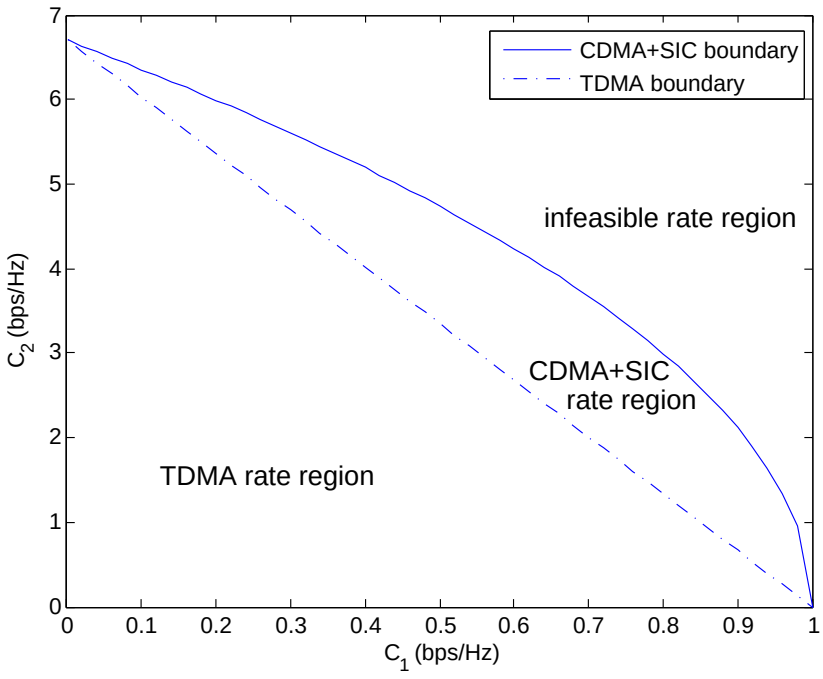

Figure 2.8: *The boundary of rate pairs achievable by CDMA with SIC and TDMA after [45]*

2.5 Radio Resource Management

Based on information theory, this section will further discuss how to allocate the limited wireless resources such as power, time and bandwidth to achieve high system performance.

2.5.1 Power Allocation Policies

For simplicity, we assume a single user block fading channel with a Rayleigh distribution, where the fading coefficient g changes from one block to another but can be considered static within a block [46]. We can then calculate the ergodic capacity according to Shannon theory in Equation (2.5). To achieve capacity-approaching performance, several power allocation schemes are briefly introduced depending on the available channel gain g. If the channel state information is only known at the receiver, a customary used approach is to allocate constant power through all the fading states. Then, the instantaneous signal to noise ratio could be represented by $\rho = Pg/(N_0B)$. If the channel state information is known at both the transmitter and the receiver, power allocation could be adapted according to the channel statistics to achieve higher efficiency. Subject to the average power constraint $\bar{P}$, we introduce two power allocation policies which are called waterfilling and channel inversion respectively.

The optimal power allocation policy is waterfilling, which is derived from a Lagrange representation by maximizing the spectral efficiency $C(P[\tau], B)(\text{bps/Hz}) = \mathbb{E}\left[\log_2\left(1 + \frac{Pg[\tau]}{N_0 B}\right)\right]$ subject to an average power constraint $\bar{P}$. In waterfilling, the transmitter allocates more power when the channel is good, taking advantage of the better channel conditions, and less or even no power when the channel is poor. Channel inversion is aimed at keeping a constant received SNR irrespective of the channel gain, thus the expected rate R can be achieved no matter what the fading state is. However, the price to pay is that higher power has to be consumed to invert the channel when it is very bad. Especially for Rayleigh fading, channel inversion results in infinite power or else a zero capacity. Thus, truncated channel inversion is further introduced by only inverting the channel when the channel gain g is large than a threshold channel gain g_0.

The capacity of the fading channel with transceiver CSI should be interpreted as a long-term average rate of information, averaged over the fluctuations of the channel. While the waterfilling strategy increases the long-term throughput of the system by transmitting when the channel is good, an important issue is the potential additional delay. Compared to waterfilling, channel inversion is much less power-efficient due to the inversion of bad channels. On the other hand, since the rate of flow of information is constant, the associated delay is independent of the time-scale of channel variation. Thus, one can view the channel inversion strategy as a delay-limited power allocation strategy.

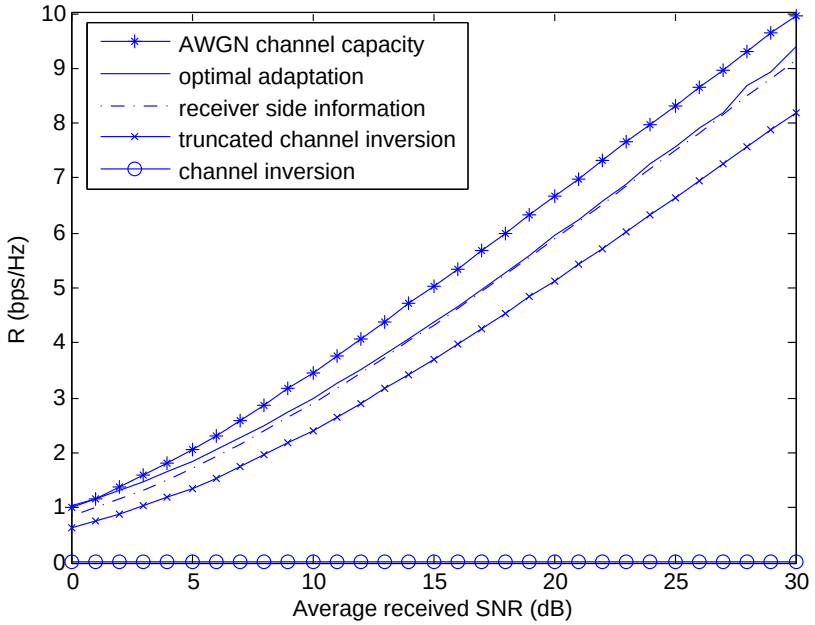

Figure 2.9: *Capacity in Rayleigh fading environment for single user case*

Figure 2.9 compares the performance of different power allocation policies in Rayleigh fading as a function of average received SNR $\bar{\rho}$ [47]. The capacity of the AWGN channel for a given average power is also shown. From the simulation results, we could summarize as follows. First, for this range of SNR values, the capacity of the AWGN channel is largest, so fading usually reduces channel capacity. Secondly, by allocating power according to the fading statistics, the optimal adaptation (waterfilling policy) always performs better than the other policies with only receiver side information (constant power allocation). However, the improvement is quite small, which implies that when the transmission rate is adapted relative to the channel, adapting the power as well yields a negligible capacity gain. Thirdly, truncated channel inversion exhibits a 1 to 5dB penalty while channel inversion without truncation yields zero capacity. This is because infinite average power is required in order to maintain a non-zero rate.

2.5.2 Adaptive Modulation and Coding

While the Shannon capacity formula has provided the theoretical upper bound of reliable transmission over wireless channels, in practice the data delivery is achieved by deploying specific coding and modulation schemes. Modulation is the process of conveying a message signal, for example a digital bit stream or an analog audio signal, inside another signal that can be physically transmitted. In most applications of modulation the carrier signal is a sine wave, which is completely characterized by its amplitude, its frequency, and its phase relative to some point in time. Basically, there are two forms of modulation: analog modulation refers to the modulation is applied continuously in response to the analog information signal, while in digital modulation an analog carrier signal is modulated by a digital bit stream. The most fundamental digital modulation techniques are based on keying:

- Phase-shift keying (PSK): a finite number of phases are used.
- Frequency-shift keying (FSK): a finite number of frequencies are used.
- Amplitude-shift keying (ASK): a finite number of amplitudes are used.
- Quadrature amplitude modulation (QAM): a finite number of at least two phases, and at least two amplitudes are used.

In QAM, an inphase signal and a quadrature phase signal are amplitude modulated with a finite number of amplitudes and then summed together. It can be seen as a two-channel system, each

channel using ASK. The resulting signal is equivalent to a combination of PSK and ASK. In all of the above methods, each of these phases, frequencies or amplitudes are assigned a unique pattern of binary bits, thus can effectively deliver digital information. For example, if $\kappa = 2$ bits are assigned to represent each phase, frequency or amplitude, there will be an alphabet consists of $\varrho = 2^{\kappa} = 4$ composed binary symbols $\{00, 01, 10, 11\}$. For $\varrho = 4$ phase shift keying, it is termed as QPSK. For $\varrho = 16$ quadrature amplitude modulation, it is named as 16-QAM. Figure 2.10 shows the constellation diagrams of QPSK and 16-QAM.

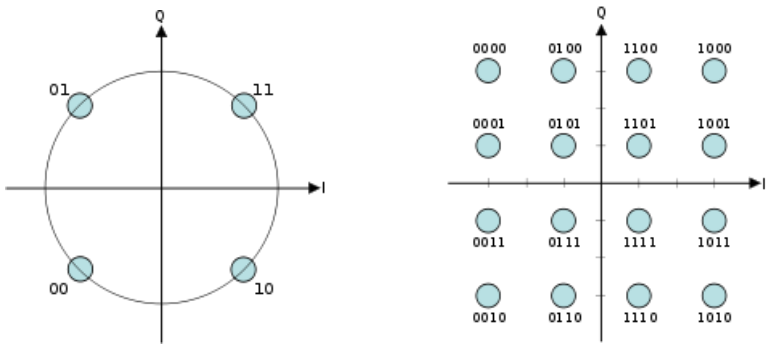

Figure 2.10: *Constellation diagrams of QPSK and 16-QAM*

Channel coding is a system of error control for data transmission, whereby the sender adds systematically generated redundant data to its messages. The carefully designed redundancy allows the receiver to detect and correct a limited number of errors occurring anywhere in the message without the need to ask the sender for additional data. Since Shannon's original work, lots of effort has been spent on designing efficient coding schemes to enable data transmission techniques approaching the Shannon bound for bit rate. However, there is a fundamental difference between the power-limited regime with low SNR and the high SNR bandwidth-limited regime. Early work focused on the power-limited regime. In this domain, binary codes suffice, and inter symbol interference due to non-flat channel responses is rarely a serious problem. As early as the 1960's, sequential decoding [48] of binary convolutional codes [49] was shown to be an implementable method for achieving the cutoff-rate , which at low SNR is 3 dB away from the Shannon limit. In the bandwidth-limited regime there was essentially no practical progress beyond uncoded multilevel modulation until the invention of trellis-coded modulation in the 1970's and its widespread implementation in the 1980's [50]. An alternative route to trellis-coded modulation in this regime is via multilevel coding, which was also introduced

during the 1970's [51]. Both trellis-coded modulation and multilevel coding are based on the concept of set partitioning, which are closely related to the bit-interleaved coded modulation firstly introduced by Zehavi [52]. Currently, the invention of turbo codes [53] and the rediscovery of low-density parity-check (LDPC) codes [54] have created tremendous excitement. These schemes operate successfully at rates well beyond the cutoff rate, within tenths of a decibel of the Shannon limit, in both the low SNR and high SNR regimes.

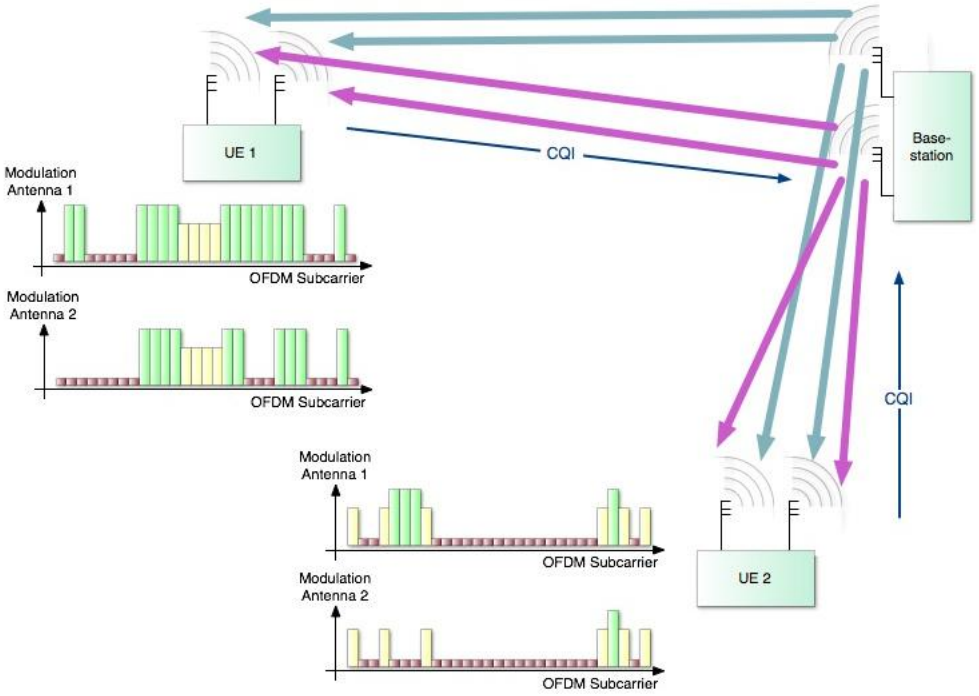

Figure 2.11: *Adaptive modulation and coding for frequency-domain scheduling [56]*

In cellular communication systems, the quality of a signal received by a UE depends on the path loss, shadowing, fading and noise *etc*. In order to improve system capacity, peak data rate and coverage reliability, the signal transmitted to and by a particular user is modified to account for the signal quality variation through a process commonly referred to as link adaptation. Traditionally, CDMA systems have used fast power control as the preferred method for link adaptation. Recently, adaptive modulation and coding (AMC) have offered alternative link adaptation methods that promise to raise the overall system capacity [55]. AMC provides the flexibility to match the modulation coding scheme to the average channel conditions for each user. With AMC, the power of the transmitted signal can be held constant over a frame interval, and the modulation and coding format is changed to match the current received signal quality

or channel conditions. Adaptive modulation is often used together with scheduling, either in time, frequency or spatial-domain to provide an efficient use of wireless channels. Figure 2.11 has demonstrated the deployment of AMC for the frequency-selective fading channels in LTE according to the CSI feedback. Here, the pink bars represent QPSK modulation, the yellow bars represent 16-QAM and the green bars represent 64-QAM.

2.5.3 Multiple Access Techniques

Multiple access allows several terminals connected to the same base station to transmit or receive over the shared wireless resources. It is based on multiplexing methods, which are in this context provided by the physical layer. These are the four fundamental types of channel access schemes [45]:

- Frequency division multiple access: provides different frequency bands to different data streams.
- Time division multiple access: provides different time slots to different data streams.
- Code division multiple access: based on spread spectrum technology where each transmitter is assigned a code to allow multiple users to be multiplexed over the wireless channel.
- Space division multiple access (SDMA): based on using smart antenna technology and differing spatial locations of mobile units within the cell to offer attractive performance enhancements. Thus the radiation pattern of the base station, both in transmission and reception, is adapted to each user to obtain highest gain in the direction of that user.

With the above multiple access techniques, multiuser diversity can be obtained by scheduling at either the transmitter or the receiver. A common opportunistic user scheduling is as follows: the transmitter selects the best user among candidate receivers according to the qualities of each channel between the transmitter and each receiver. Traditionally, this selection process is based on time division multiple access by only transmitting to the best user in each time slot [57].

Figure 2.12 illustrates that an increased number of users can benefit the sum network capacity. When there are many users that fade independently, at any one time there is a high probability that one of the users will have a strong channel. By allowing only that user to transmit,

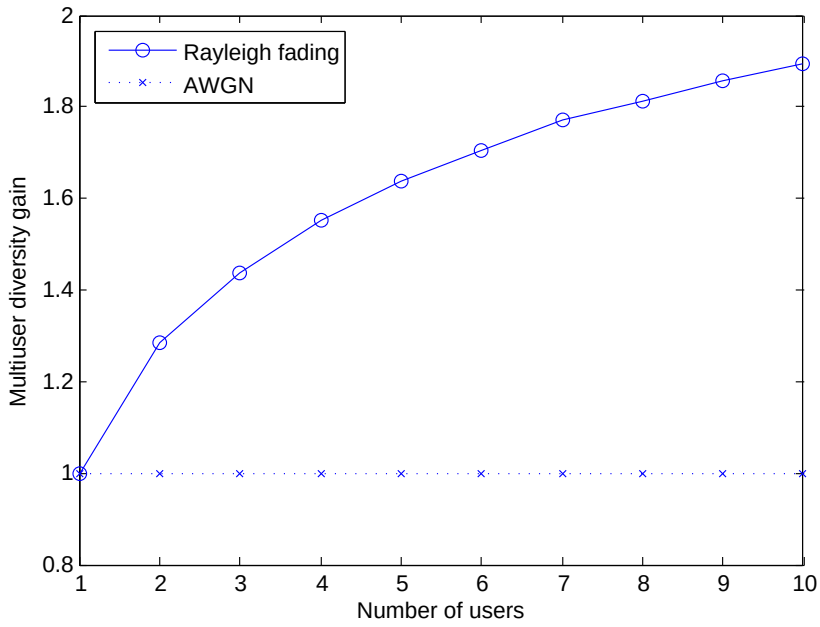

Figure 2.12: *Multiuser diversity gain for Rayleigh fading after [57]*

the shared channel resource is used in an efficient manner and the total system throughput is increased.

To characterize this phenomenon, the so-called multiuser diversity gain is defined to be the instantaneous channel signal to noise ratio of the time instances over which the user is scheduled for transmission, divided by the average channel SNR of that user [58]. Figure 2.12 shows the multiuser diversity gains for Rayleigh fading channels and AWGN channels. As opposed to the diversity techniques for single user case, it is noted that multiuser diversity improves system performance by exploiting channel fading. Here the channel fluctuations due to fading ensure that with high probability there is a user with its channel strength much larger than the mean level, thus large and fast variations become the preferred characteristic of wireless channels.

2.5.4 Scheduling Strategies

The key elements of the base network related to scheduling are bearer gateways (BG), access points (AP) and user terminals (UT). Considering a single cell downlink direction, BG is where the packet traffic enters the radio access network. Traffic conditioning and packet marking are performed by the BG. The scheduler is located at the AP in this architecture. Downlink

and uplink traffic can have separate schedulers. In the scheduler, the requirements of different services as well as the instantaneous radio channel conditions are considered. By properly allocating the finite resources over different channel states and different users, diversity gain could be achieved to improve efficiency. UT refers to the mobile users with different QoS requirements, which further induce some constraints for optimization. In this system, scheduling algorithms play a central role in determining the overall performance, which is normally characterized by high throughput and fair resource allocation among all the mobile users. Recently, various design approaches have been proposed to address these issues. In [59], some desired characteristics for an ideal scheduler were discussed, which include efficient link utilization, delay bound, fairness, throughput, implementation complexity, energy consumption and scalability.

2.5.5 Fairness Concepts

In mobile communications, where a set of resources is to be shared by a number of users, fair allocation is very important. However, the concept of fairness is ambiguous, depending on whether 'fairness' refers to fair allocation of air link resources or fair service among the users. If fair resources are allocated among the users, it is difficult to guarantee users' QoS since they have different average channel qualities and application requirements. On the other hand, if fair service is targeted, lots of power and bandwidth resources are consumed by bad channel conditions which leads to capacity inefficiency.

There are many scheduling algorithms which take fairness into consideration. A simple one is called the proportional fair scheduler [57], designed to meet the challenges of delay and fairness constraints while harnessing multiuser diversity. It schedules a user when its instantaneous channel quality is high relative to its own average channel conditions, which means data packets are transmitted to a user when the channel is near its peak gain. Multiuser diversity benefits can still be extracted because channels of different users fluctuate independently. If there is a sufficient number of users in the system, most likely there will be a user near its peak at any one time. Other strategies such as max-min fair scheduler and weighted fair scheduler could also be used depending on the application requirements.

Fairness could also be explained as fair decision making, since the resource allocation problem is similar to the bargaining problem in cooperative game theory, where a set of decision makers bargain among themselves for some shared resources. For each user, a utility function and an

minimal initial utility which is required by the user without any cooperation in order to enter the game are defined. For all the users, these utility functions and initial points further form a feasible set of utilities to represent the service qualities of these users. In this region, we can obtain the Nash bargain solution based on some unanimous agreement on the utility point [60]. Furthermore, several axiomatic properties of Nash bargaining theory are summarized in [61], indicating the framework for defining fair resource allocation strategies. For instance, proportional fair resource allocation can be interpreted as a symmetric Nash bargaining game when the object of fair share is the throughput and the minimum required rate is zero [62].

In [63], a quantitative measure of fairness which is called fairness index is defined. A good fairness measure should be bounded, continuous, independent of scale and able to be applied to any number of users. Jain's fairness index (FI) is defined as follows

$$\mathrm{FI} = \frac{(\sum_{i=1}^{I} R_i)^2}{I \sum_{i=1}^{I} R_i^2} \tag{2.10}$$

This index measures the 'equality' of the allocated R for user $i = \{i = 1, 2, \cdots I\}$, where R often represents the data rate. It is noticed that if all users get the same value of R_i, then the fairness index is 1, and the system is 100% fair. As the disparity increases, fairness decreases and a scheme which favors only a selected few users has a fairness index near to 0.

2.5.6 Constraints and Tradeoffs

Since practical constraints should also be considered in resource management, we need a holistic view for designing energy-efficient resource allocation schemes. The constraints usually refers to the physical constraints placed by hardware devices, the quantitative energy consumption limits during each process *etc*. Identifying the key points which can improve overall energy efficiency is essential for future communication system design. Here, we will take a constant-rate end-to-end scenario as an example.

Firstly, from the users' perspective, high data rates are always desired. However, with limited power and bandwidth in practice, the maximum supported rate in each frame is clearly finite. If the desired rate is too high to be supported in every frame, error-free transmission can not be guaranteed. The tradeoffs among power, bandwidth, delay and error probability are obvious.

Secondly, from the network operators' perspective, a fair share of the limited wireless resources

among different users is essential. However, the resource allocation decisions made based on certain fairness metrics may not be the optimum decisions for maximizing users' spectral efficiency. For example, assuming a multi-user downlink scenario with base station peak power constraint, allocating more power to the users near the base station will generally benefit network spectral efficiency, however, the resource share in the network is unfair. This is the so-called near-far problem. Obviously, there are tradeoffs between fairness and spectral efficiency.

We need to notice that the above tradeoffs also exist in variable rate systems. They are the basic tradeoffs according to information theory, however, based on the various standpoints other tradeoffs may also be considered. For example, high traffic loads usually require a large amount of data to be delivered within short time. However, in practical networks high traffic loads also leads to higher waiting times per packet due to the limited buffering capability of the service. The more packets there are, the larger the average delay exists. For delay-sensitive applications such as a voice call, call blocking may happen which gives a bad user experience. Apart from this example, scheduling strategies could also be designed from a revenue point of view. Different criteria result in different optimization problems, and maximizing profit may not be directly related to maximizing throughput. Further, we need to notice there are tradeoffs between spectral efficiency and energy efficiency. While the spectral efficiency maximization problem provides a solution for the efficient use of fixed RF power, it is usually different from the solutions for minimizing energy consumption, as other operational energy consumption may also be included for optimization in the latter case. A good example is the discontinuous transmission (DTX) and reception (DRX) mode, which is already widely deployed for mobile terminals. More details will be provided in the following context to illustrate resource management for energy-efficient design.

2.5.7 DTX and DRX

Low user equipment energy consumption has been of prime interest ever since the emergence of handheld mobile terminals more than 15 years ago. As a consequence, energy-efficient UE discontinuous transmission and reception have been important components of cellular systems for many years. Generally speaking, DTX and DRX refer to the mobile device and the network negotiating phases in which data transfer occurs. During other times the device turns its transmitter/receiver off and enters a low power state. Usually, a function designed into the communication protocol allows this to happen by informing the wireless devices how the trans-

mission is structured. For example, the headers in the radio frame structure usually contain address details, thus the devices can listen to these headers in each frame to decide whether the transmission is relevant to them or not. In this case, the receiver only has to be active at the beginning of each slot to receive the header, conserving battery life. Other techniques include polling, whereby the device is placed into standby for a given amount of time and then a beacon is sent by the access point or base station periodically which indicates if there is any waiting data for it. This is used in 802.11 wireless networks when compatible access cards and access points negotiate a power saving mode arrangement. A hybrid of the above techniques could be used in reality as well.

In parallel to the requirements for low UE energy consumption, the possibility for low network energy consumption is also becoming more important. One reason is the simple fact that, for many operators, the cost of the energy needed to operate the system constitutes a non-negligible part of the overall operating expense. Today, mobile broadband is growing rapidly in developing countries where electricity grid connections are often expensive, unreliable, and have long installation lead times. As a result the number of off-grid base station sites, mostly powered by diesel generators, is currently growing which leads to the increased costs in emerging markets. To reduce the operational cost of base stations, [64] discussed the possibility to introduce extended cell DTX for the purpose of an energy-efficient base station operation mode. It proposed the idea to reduce the number of reference signals for empty cells which have no active user terminals. This can be made possible by introducing UE mobility measurements used for cell search on the secondary synchronization signals rather than on reference signals. Also, the reduced transmission time will lead to the less energy consumption on power amplifiers. Regarding the base station discontinuous transmission mode, although the idea and its potential issues have been raised, more efforts are required towards standardization. The open questions such as when to switch off the base station and how to design the corresponding control signalling need to be further addressed.

2.6 MIMO Systems

In this section, the basic concept of multiple-input multiple-output system is introduced. Taking the space-time block code and the spatial multiplexing as examples, the capacities of MIMO system are derived. Further, for the multi-user MIMO scenarios, this allows the base station to transmit or receive the information from multiple users. Interference can also be mitigated by

enabling the multiple antennas transmitting in a coordinated manner.

2.6.1 Space Time Block Code

As discussed in Section 2.4 and 2.5, diversity is traditionally achieved by repeating the transmitted symbols in time, in frequency or using multiple antennas at the receiver. Diversity gain is thus defined as the increase in SNR due to some diversity scheme, or how much the transmit power can be reduced when a diversity scheme is introduced without a performance loss. For the spatial-domain case, the diversity gain is compounded to the array gain, which also consists of an increase in average received SNR due to the coherent combination of received signals. In addition to the multiple receiver antennas, if multiple transmit antennas are also equipped, a MIMO system can be obtained as shown in Figure 2.13.

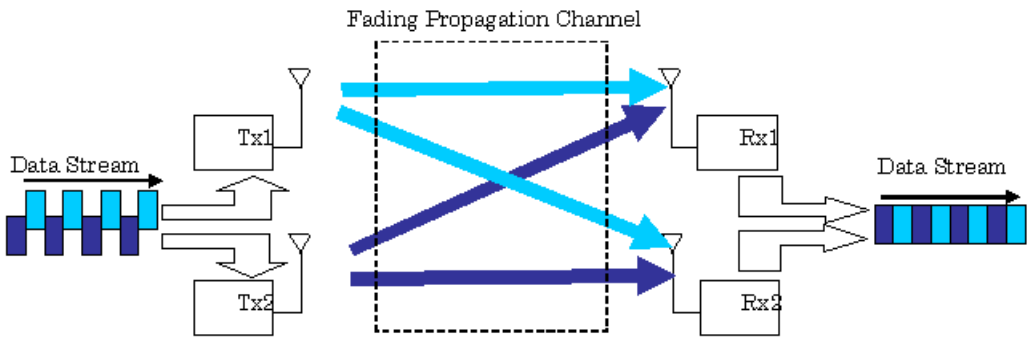

Figure 2.13: *Diagram of a MIMO communication system [65]*

With multiple transmit and receiver antennas, transmission reliability can be enhanced. Space-time block code (STBC) is a technique used in wireless communications to transmit multiple copies of a data stream across a number of antennas and to exploit the various received versions of the data to improve the reliability of data transmission. The fact that the transmitted signal must traverse a potentially difficult environment with scattering, reflection and refraction effects and may then be further corrupted by thermal noise in the receiver means that some of the received copies of the data will be better than others. This redundancy results in a higher chance of being able to use one or more of the received copies to correctly decode the received signal. In fact, space time coding combines all the copies of the received signal in an optimal way to extract as much information from each of them as possible.

The Alamouti code is the simplest space-time block code. It was designed for a two-transmitter

antenna system and has the coding matrix as

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix}, \tag{2.11}$$

where the horizontal direction denotes the spatial-domain and the vertical direction denotes the time-domain. This means that at a given symbol period, two signals are simultaneously transmitted from the two antennas. The signal transmitted from antenna 1 is denoted by x_1 and from antenna 2 by x_2. During the next symbol period signal $-x_2^*$ is transmitted from antenna 1, and x_1^* signal is transmitted from antenna 2, where $[\quad]^*$ is the complex conjugate operation. At the receiver side, assuming that fading is constant across two consecutive symbols, the received signal y at time instance τ and $\tau + 1$ will be

$$\begin{cases} y[\tau] = \lambda_1[\tau]x_1[\tau] + \lambda_2[\tau]x_2[\tau] + \sigma[\tau] \\ y[\tau+1] = -\lambda_1[\tau+1]x_2[\tau]^* + \lambda_2[\tau+1]x_1[\tau]^* + \sigma[\tau+1] \end{cases} \tag{2.12}$$

This can be rewritten in a matrix form as

$$\mathbf{Y} = \begin{bmatrix} y[\tau] \\ y^*[\tau+1] \end{bmatrix} = \begin{bmatrix} \lambda_1[\tau] & \lambda_2[\tau] \\ \lambda_2[\tau+1]^* & -\lambda_1[\tau+1]^* \end{bmatrix} \begin{bmatrix} x_1[\tau] \\ x_2[\tau] \end{bmatrix} + \begin{bmatrix} \sigma[\tau] \\ \sigma[\tau+1]^* \end{bmatrix}. \tag{2.13}$$

By inverting Equation (2.13), the delivered data can be estimated by removing the channel effects of the received signals, which is called channel equalization. Among the various equalization techniques, linear maximal ratio combining is often used. The Alamouti code is the only space-time block code with a code rate of unity. It is linear in the transmitted symbols as well. Apart from these, the Alamouti code is the only optimal space-time block code in terms of capacity

$$C_{N\times M}(\text{bps/Hz}) = \log_2\left(1 + \frac{\bar{\rho}}{M}\sum_{m=1}^{M}\sum_{n=1}^{N}|h_{nm}|^2\right), \tag{2.14}$$

where M denotes the number of transmit antennas and N denotes the number of receiver antennas. $\bar{\rho}$ is the averaged receiver side signal to noise ratio. The scalar h_{nm} is the channel gain due to Rayleigh fading, which is a random variable with unit mean power and variance. Compared with single-input single-output systems, diversity gains are obtained in the spatial-domain.

2.6.2 Spatial Multiplexing

Similar to the diversity gain, the spatial multiplexing (SM) gain is also defined as the capacity increase with no additional power or bandwidth by deploying multiple-input multiple-output systems. The spatial multiplexing gain can be realized by transmitting independent information from the individual antennas. The enormous values of the spatial multiplexing gain potentially achieved by MIMO techniques have had a major impact on the introduction of MIMO technology in wireless systems.

Spatial multiplexing is a transmission technique in MIMO wireless communication to transmit independent and separately encoded data signals, often called streams, from each of the multiple transmit antennas. Therefore, the space dimension is reused or multiplexed more than one time. If the transmitter is equipped with M antennas and the receiver has N antennas, the maximum spatial multiplexing order (the number of streams) is $O_{\mathrm{rd}} = \min(M, N)$ if a linear receiver is used [45]. This means that O_{rd} streams can be transmitted in parallel, ideally leading to an O_{rd} increase of the spectral efficiency. The practical multiplexing gain can be limited by spatial correlation, which means that some of the parallel streams may have very weak channel gains.

To illustrate SM operation, we assume the number of transmit antennas $M = 3$ and there are $\upsilon = 6$ bits of information $(\beta 1, \beta 2, \beta 3, \beta 4, \beta 5, \beta 6)$ for transmission. Figure 2.14 shows how the spatial multiplexing works. Firstly, the information bits are divided into $M = 3$ substreams of data $(\beta 1, \beta 4)$, $(\beta 2, \beta 5)$, $(\beta 3, \beta 6)$. Then each substream of data is multiplied with the same carrier frequency in order to transmit them via three separate antennas. Meanwhile, at the receiver side, each substream will derive from 3 spatial channels from different transmit antennas. A total of 9 spatial channels will be at the receiving antennas due to the multipath environment. By combining the information from these spatial channels, data substreams will be demultiplexed and decoded in order to get back to the original data stream.

With spatial multiplexing, the ergodic capacity with receiver side CSI can be written as

$$C_{N\times M}(\mathrm{bps/Hz}) = \mathbb{E}\left[\log_2 \det\left(\mathbf{I}_N + \frac{\bar{\rho}}{M}\mathbf{H}\mathbf{H}^*\right)\right]. \tag{2.15}$$

The function $\mathbb{E}[\ \ ]$ denotes the expectation operation, and $\mathbf{I}_N$ is the $N \times N$ identity matrix. The notation $\mathbf{H}$ is an $N \times M$ matrix with each of its elements h_{nm} representing the channel

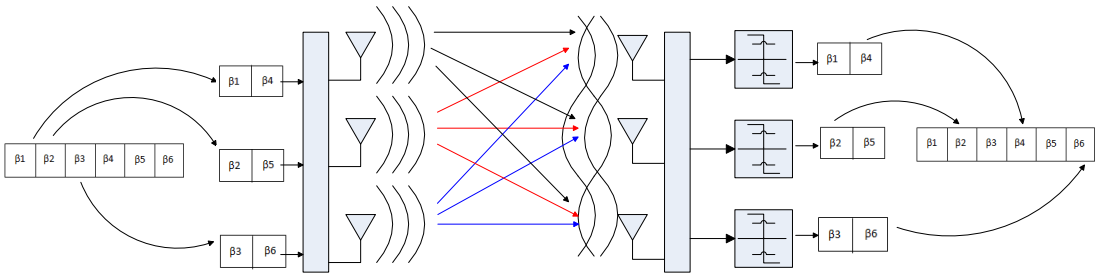

Figure 2.14: *Diagram of MIMO spatial multiplexing after [66]*

gain due to Rayleigh fading from transmit antenna m to receiver antenna n as

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1M} \\ h_{21} & h_{22} & \dots & h_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ h_{N1} & h_{N2} & \dots & h_{NM} \end{bmatrix}. \tag{2.16}$$

Based on Equation (2.15), we notice that the capacity is approximately $O_{\mathrm{rd}} \log_2 \bar{\rho}$ at high SNR regime, where a gain in spatial degrees of freedom can be obtained [45]. While at low SNR regime, the capacity can be simplified as $N\bar{\rho} \log_2 e$, where a receive beamforming gain is achieved. For the high SNR scenario, the achieved capacity is almost O_{rd} times of that under a SISO scenario.

2.6.3 MIMO in Interference Environments

As discussed above, multiple antennas at the transmitter and the receiver can provide diversity gain as well as increased data rates through space-time signal processing. Alternatively, sectorization or smart antenna array techniques can be used to provide directional antenna gain at the transmitter or at the receiver. This directionality can increase the cell range, reduce channel delay spread and flat fading, and suppress interference among users [45]. Although interference typically arrives at the receiver from different directions, directional antennas can exploit these differences to null or attenuate interference arriving from the given directions, thereby increasing system capacity. Exploiting the reflected multipath components of the signal arriving at the receiver requires an analysis of multiplexing-diversity-directionality tradeoff. Whether it is best to use the multiple antennas to increase data rates through multiplexing, increase robustness to

fading through diversity, or reduce channel delay spread and interference through directionality is a complex tradeoff decision that depends on the overall system design as well as on the environment [67].

2.7 Summary

In this chapter, the relevant background materials on Green Radio, channel capacity, resource allocation and MIMO systems are documented, providing the essential fundamentals for the energy-efficient design in the following chapters.

Firstly, the background and targets of the Mobile VCE Green Radio project are introduced. Generally, it is because the revenue increase cannot meet the energy increase in the communication industry. Apart from these economic reasons, environmental considerations such as the need for reducing CO_2 emission are also important. By 2020, the Mobile VCE targets to achieve 100x energy reduction by designing more energy-efficient network architectures and radio resource allocation techniques.

As Green Radio is a new concept, we start with interpreting it from a theoretical point of view. Based on the characteristics of wireless time-varying channels in cellular systems, the Shannon capacity formula is introduced. This theory provides us the upper bounds for reliable transmission: providing a power constraint, the maximum data rate we can deliver over a wireless channel can be calculated. The concepts of ergodic capacity, outage capacity and capacity region are summarized, all of which are widely used metrics for communication system design.

In order to improve the system performance and approach the Shannon capacity in a real communication network, various resource allocation techniques are summarized for different scenarios. With full channel state information, waterfilling is an optimum approach which achieves capacity. However, obtaining full CSI at both the transmitter and the receiver sides is usually not practical. Some sub-optimal power allocation schemes such as channel inversion may be used depending on different channel statistics. Also, the Shannon capacity is derived for ideal coded system assumptions, which is a continuous function in terms of SNR. In practice, transmission is conveyed by using certain coding and modulation schemes in discrete formats. There is always a performance gap between realistic communication scheme and Shannon capacity.

Although the efficient use of wireless resources can generally improve system performance,

fairness needs to be maintained. In practice when we design a communication system, there are also some system constraints that need to be considered depending on the deployed hardware and specific user application requirements. The optimization objectives for both the spectral maximization and energy minimization are formed as constrained problems with various trade-offs. Theoretically, for RF energy consumption, some spectral-efficient design solutions can also be applied for energy-efficient design. However, they can be easily differentiated by taking other sources of operational energy into consideration, such as the energy consumed by antenna power amplifiers. The DTX/DRX modes are introduced as one example for practical energy-efficient design. To maximize the spectral efficiency, the preferred transmission mode should activate all available time resources. However, switching off the network devices in the time-domain may help with energy conservation if hardware operational energy is considered.

Finally, multiple-input multiple-output systems are described. The capacities for STBC and SM are summarized, which can be treated as the basic forms of MIMO schemes. Under these schemes, spatial diversity and multiplexing gain can be obtained respectively. While diversity generally contribute to improve the error probability, multiplexing is used for increasing the data rate. In general there is a tradeoff between spatial diversity and multiplexing. Further, we briefly introduced the multi-user MIMO concept, where with well designed MIMO beamforming patterns, interference among users can be mitigated.

The above mentioned concepts and materials are closely related to the design of energy-efficient resource allocation schemes, which thus provide essential background knowledge for the following technical chapters. In Chapter 3, we will introduce how to assign bandwidth resources to multiple links to improve network energy efficiency. The discussions are based on the Shannon equations to provide theoretical analysis for energy-efficient multiple access techniques.

Chapter 3
Two-Link Study of Energy-Efficient Resource Allocation

3.1 Introduction

With the expansion of the global population and the prevalence of wireless devices, interference among multiple active links will become a dominant problem, causing serious QoS degradation for future wireless communications [68]. Although various interference mitigation techniques are designed to cope with that effectively, these methods generally require high complexity transceiver devices [69] [70]. Meanwhile, in order to maintain users' quality of service during the mutually-interfered transmission process, more transmit power might be consumed. Taking these issues into consideration, modern communication industries have shifted their interests to looking for economical system design strategies rather than simply improving QoS with more advanced techniques [71].

Here, 'economical design' takes into account both the network operation requirements and the requirement to minimize energy consumption for environmental purposes. This could be interpreted together with geographic demographics. In high-density city areas, interference could be quite severe. For the reliable reception of the transmitted information, high transmit power might be required which will cause huge CO_2 emission [72]. The energy loss due to enabling and maintaining communications takes a great portion of the total energy consumption. How to apply green technologies to reduce interference and improve system energy efficiency is an important research topic.

Regarding implementations, intelligent resource allocation and scheduling algorithms are required under the definition of energy efficiency metrics. These are often related with certain coding, modulation or diversity techniques in the physical and MAC layers [73]. While network operators are generally interested in achieving an overall energy efficiency, upper layers are also involved in the design processes, referring to network architectures and routing protocols *etc* [74]. Meanwhile, modern communications are becoming more complex, with various

QoS requirements placed by different user terminals at different times [75]. We also need to take these application layer requirements into consideration. Therefore, application-dependent energy-efficient solutions via flexible cross-layer design are required [76] [77].

In order to design intelligent resource allocation strategies, several approaches are considered. Traditional methods are based on forming objective functions under specified constraints, resulting in various optimal and suboptimal solutions in different scenarios [78] [79]. Modern methods such as game-theoretic and utility-based approaches are also proposed recently for intelligently allocating limited wireless resources to different services [80] [81]. Compared with the throughput maximization problem which employs capacity as an effective measure, designing new power-efficient strategies under Green Radio framework is still an open problem. Here we only focus on lower layer resource allocation problems under the assumed network architecture and QoS requirements. Our goal is to provide some basic insights into energy-efficient design in interference environments.

In this chapter, Section 3.2 will describe a simple mutually-interfered two-link model with the users' locations constrained on the line between the two neighboring base stations. In Section 3.3, we will compare the performance of simultaneous transmission (ST) and orthogonal transmission (OT) schemes in an FDMA framework with the same QoS requirements. Further, we will discuss how to combine them for sum power minimization [13]. In Section 3.4, by setting different optimization objectives, some QoS tradeoffs among power, fairness, delay and outage probabilities will be investigated under the assumed rate constraints. Finally, in Section 3.5 the results found in this work will be compared with the existing soft frequency reuse techniques for multi-link scenario implementations.

3.2 Two-Link Model Assumptions

In this section, we focus on the impacts of different resource allocation schemes on power reduction in an interference environment. By ignoring the energy consumed in transceiver circuits and only considering the RF component, we investigate the problem of minimizing total RF energy consumption in the network under specific rate constraints.

To simplify the problem, we consider a two-cell system in which two neighboring base stations communicate with mobile terminals over a coverage area as shown in Figure 3.1. For the sake of simple presentation, we focus on downlink transmission. Here, BS_i and UE_i are used to

denote the base station and mobile user in cell i $(1 \leq i \leq 2)$ respectively. The data sent by base station BS_i is transmitted with power P_i. Mobile user UE_i can only move on the line between the two base stations within cell i. D is the inter-site distance, and d_{ij} represents the distance between the base station in cell i and the mobile user in cell j $(1 \leq i, j \leq 2)$.

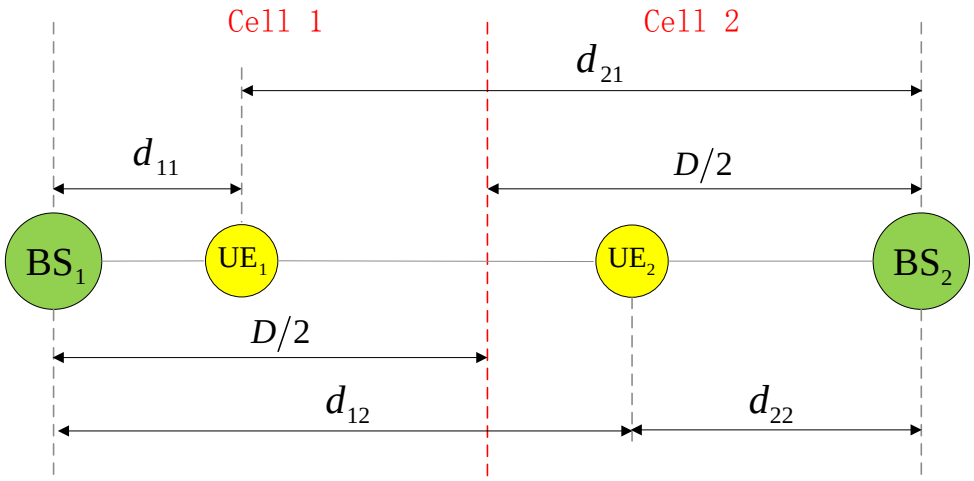

Figure 3.1: *Two-link line model with mutual interference. In the figure BS_1 and UE_1 denote the base station and mobile user equipment in cell 1; whereas BS_2 and UE_2 denote the base station and mobile user equipment in cell 2.*

To simulate a real wireless communication environment with mutual interference, we choose the system parameters according to the WINNER II urban macro-cell channel model [43]. By ignoring large-scale shadowing and small-scale fading of wireless channels, the channel gains g_{ij} are only determined by path loss, which are computed in relation to d_{ij} $(1 \leq i, j \leq 2)$. For a line of sight (LOS) scenario, this transceiver distance also results in different representations of WINNER II channel models based on a comparison with the fiducial distance d'_{BP}. When the transceiver distance is smaller than d'_{BP}, the path loss can be modeled as [43]

$$L_{ij}(\mathrm{dB}) = A \log_{10}(d_{ij}) + B + C \log_{10}(f_{\mathrm{c}}/5), \tag{3.1}$$

where f_{c} represents the carrier frequency, and A, B, C are the model coefficients. By assigning different values to A, B and C, this path loss model can be fitted into most scenarios including indoor LOS, urban micro-cell LOS, large indoor hall, suburban LOS, typical urban LOS, rural

Parameters	Values
Carrier frequency $f_{\rm c}$	5GHz
Speed of light c	3×10^8m/s
Cell diameter D	800m
Total bandwidth B	20MHz
Subcarrier spacing δf	15KHz
Number of subcarrier $N_{\rm sub}$	1200
BS antenna height $h_{\rm BS}$	25m
User antenna height $h_{\rm MS}$	1.5m
Operation temperature T_0	290 K

Table 3.1: *LTE physical layer model parameters*

macro-cell LOS *etc*. However, if the transceiver distance is larger than $d'_{\rm BP}$, more complicated path loss model is required. Using the parameters shown in Table 1, this fiducial distance could be calculated as

$$d'_{\rm BP}({\rm m}) = \frac{4(h_{\rm BS}-1)(h_{\rm MS}-1)f_{\rm c}}{c} = \frac{4\times(25-1)\times(1.5-1)\times(5\times10^9)}{3\times10^8} = 800. \quad (3.2)$$

Since the cell diameter D is within this distance, the path loss under the corresponding channel model for urban macro-cell LOS scenario could be represented as

$$L_{ij}({\rm dB}) = 26\log_{10}(d_{ij}) + 39 + 20\log_{10}(f_{\rm c}/5) = 26\log_{10}(d_{ij}) + 39 + 20\log_{10}(5/5). \quad (3.3)$$

If we ignore the antenna gains, the channel gains are only determined by the transceiver distances as

$$g_{ij} = 10^{\frac{-L_{ij}}{10}} = 10^{\frac{-26\log_{10}(d_{ij})-39-20\log_{10}(5/5)}{10}} = \frac{d_{ij}^{-2.6}}{10^{3.9}}. \quad (3.4)$$

Thus, we can calculate the channel gains g_{11}, g_{12}, g_{21} and g_{22} according to the above two-link line model. Under the operation temperature T_0, the noise density in the two cells could also be obtained according to

$$N_0({\rm W/Hz}) = KT_0 = 1.38\times10^{-23}\times290 = 4.0\times10^{-21}. \quad (3.5)$$

3.3 Energy-Efficient Multiple Access

In this section, we formulate ST and OT schemes to reduce network power consumption under specified rate constraints. To reduce the required power on each link, ST is designed based on

maximizing bandwidth reuse efficiency without link cooperation. Since this selfish behavior may cause severe co-channel interference and degrade the performance on other links, OT is also studied to indicate the relative superiorities of the two schemes in terms of sum power minimization.

3.3.1 Simultaneous Transmission

Simultaneous transmission means that the both base stations deliver information to their mobile users simultaneously using the full time slot and the total bandwidth [82]. It could be treated as a non-orthogonal TDMA/FDMA scheme with maximum available time/bandwidth assigned to both links. When the data rate on both links are specified, the power needed at each base station could be uniquely determined. According to Shannon theory, this can be represented as

$$\begin{cases} R_1 = \log_2(1 + \dfrac{P_1 g_{11}}{N_0 B + P_2 g_{21}}) \\ R_2 = \log_2(1 + \dfrac{P_2 g_{22}}{N_0 B + P_1 g_{12}}). \end{cases} \tag{3.6}$$

So the power required on each link could be written as

$$\begin{cases} P_1 = \dfrac{\frac{g_{21} N_0 B (2^{R_1}-1)(2^{R_2}-1)}{g_{11} g_{22}} + \frac{N_0 B (2^{R_1}-1)}{g_{11}}}{1 - \frac{g_{21} g_{12} (2^{R_1}-1)(2^{R_2}-1)}{g_{11} g_{22}}} \\ P_2 = \dfrac{\frac{g_{12} N_0 B (2^{R_1}-1)(2^{R_2}-1)}{g_{11} g_{22}} + \frac{N_0 B (2^{R_2}-1)}{g_{22}}}{1 - \frac{g_{21} g_{12} (2^{R_1}-1)(2^{R_2}-1)}{g_{11} g_{22}}}. \end{cases} \tag{3.7}$$

As the allocated power should be finite and nonnegative, the rate requirements could not be satisfied when

$$(2^{R_1} - 1)(2^{R_2} - 1) \geq \frac{g_{11} g_{22}}{g_{12} g_{21}}. \tag{3.8}$$

The terms on the left hand side of Equation (3.8) are a function of rate, while the terms on the right hand side is a function of channel gains. Thus, we could treat $\frac{g_{11} g_{22}}{g_{12} g_{21}}$ as a measure of the achievable rate in the network under simultaneous transmission. Meanwhile, because this metric is the quotient of direct-link and cross-link channel gains, it reveals that the achievable rate is location-dependent, while increasing the transmission power is not an effective method due to mutual interference between the two links. A similar representation is also found in [83], which is derived by multiplying the SINR of the two links for power term elimination. In addition, a game-theoretic approach for spectrum maximization under power constraints is

discussed in [84], which uses $(g_{11}g_{22})/(g_{12}g_{21})$ as an optimum policy indicator. Thus, for both power minimization and spectral maximization problems, $(g_{11}g_{22})/(g_{12}g_{21})$ is a useful metric for system analysis. Furthermore, satisfying Equation (3.8) is a necessary but not sufficient condition in practical systems, because base stations usually have finite maximum power limits.

3.3.2 Orthogonal Transmission

The FDMA-based orthogonal transmission scheme, which is often implemented as orthogonal frequency division multiple access, is designed to avoid mutual interference by allocating non-overlapping frequency bands to the two links [85]. While the total bandwidth for data transmission is divided into N_{sub} subcarriers with subcarrier-spacing δf where $N_{\text{sub}}\delta f < B$ due to the guard band. A weighting factor $f_1 = \zeta\delta f/B, 0 \leq \zeta \leq N_{\text{sub}}$ is defined as the fraction of bandwidth assigned to link 1, so link 2 takes the remaining fraction $f_2 = (N_{\text{sub}} - \zeta)\delta f/B$. The transmission rates could be represented as

$$\begin{cases} R_1 = f_1 \log_2(1 + \dfrac{P_1 g_{11}}{N_0 B f_1}) \\ R_2 = f_2 \log_2(1 + \dfrac{P_2 g_{22}}{N_0 B f_2}). \end{cases} \tag{3.9}$$

Thus the power assignment under specified rate requirements is calculated as

$$\begin{cases} P_1 = \dfrac{f_1 N_0 B(2^{\frac{R_1}{f_1}} - 1)}{g_{11}} \\ P_2 = \dfrac{f_2 N_0 B(2^{\frac{R_2}{f_2}} - 1)}{g_{22}}. \end{cases} \tag{3.10}$$

We try to minimize the sum power requirement by properly selecting the value of ζ for subcarrier assignment as

$$\min\left(P(\zeta)\right) = \min\left(\frac{\zeta\delta f N_0(2^{\frac{R_1 B}{\zeta\delta f}} - 1)}{g_{11}} + \frac{(N_{\text{sub}} - \zeta)\delta f N_0(2^{\frac{R_2 B}{(N_{\text{sub}}-\zeta)\delta f}} - 1)}{g_{22}}\right). \tag{3.11}$$

By taking the derivative in terms of ζ and setting $\partial P(\zeta)/\partial\zeta = 0$, the stationary point n_{p} can be obtained as

$$\frac{2^{\frac{R_1 B}{n_{\text{p}}\delta f}}\left(1 - \frac{R_1 B \ln 2}{n_{\text{p}}\delta f}\right) - 1}{g_{11}} = \frac{2^{\frac{R_2 B}{(N_{\text{sub}}-n_{\text{p}})\delta f}}\left(1 - \frac{R_2 B \ln 2}{(N_{\text{sub}}-n_{\text{p}})\delta f}\right) - 1}{g_{22}}. \tag{3.12}$$

Under the constraint $0 \leq \zeta \leq N_{\text{sub}}$ ($N_{\text{sub}}\delta f < B$), simulation experience suggests that only a single solution for n_{p} exists satisfying Equation (3.12). As $\partial^2 P(\zeta)/\partial\zeta^2 > 0$, this stationary point is a global minimum, illustrating that the minimum energy consumption could be uniquely achieved when $\zeta = n_{\text{p}}$.

3.3.3 Orthogonal Transmission Scheme Extensions

The OT scheme could also be applied in a TDMA framework by allowing only one link to transmit at any time with full bandwidth [86]. In each time unit, a weighting factor $t_1 = t$ $(0 \leq t \leq 1)$ is defined to represent the fraction of time assigned to link 1, while link 2 occupies the other fraction $t_2 = 1 - t$. This could be written as

$$\begin{cases} R_1 = t_1 \log_2(1 + \dfrac{P_1 g_{11}}{N_0 B}) \\ R_2 = t_2 \log_2(1 + \dfrac{P_2 g_{22}}{N_0 B}). \end{cases} \tag{3.13}$$

Under the rate constraints, the corresponding power needed during their active time is

$$\begin{cases} P_1 = \dfrac{N_0 B(2^{R_1/t_1} - 1)}{g_{11}} \\ P_2 = \dfrac{N_0 B(2^{R_2/t_2} - 1)}{g_{22}}. \end{cases} \tag{3.14}$$

In order to calculate sum energy consumption during that time unit, we multiply the link power with its corresponding transmission time. Thus the energy consumption of TDMA-based OT scheme can be formulated as

$$\min\left(E(t)\right) = \min\left(\frac{tN_0 B(2^{R_1/t} - 1)}{g_{11}} + \frac{(1-t)N_0 B(2^{R_2/(1-t)} - 1)}{g_{22}}\right). \tag{3.15}$$

Also, based on the discussions in Section 3.3.2, with full time transmission the energy consumption per second of FDMA-based OT is written as

$$\min\left(E(f)\right) = \min\left(\frac{fN_0 B(2^{R_1/f} - 1)}{g_{11}} + \frac{(1-f)N_0 B(2^{R_2/(1-f)} - 1)}{g_{22}}\right). \tag{3.16}$$

Thus, we can conclude that by limiting the transmission time on each link in TDMA, the sum energy consumption to achieve the specified data rates for the two orthogonal schemes are equivalent. As lots of hardware such as antenna power amplifier usually have peak power

constraints, lowering the instantaneous peak power is also important for the practical design. Regarding this, the TDMA-based OT tends to have the higher peak power than FDMA-based OT. In practice, the implementation of TDMA or FDMA depends on the specified system peak power constraints and delay constraints.

3.3.4 Scheme Selection with Link Data Rate Constraints

Simulations are performed to compare ST and FDMA-based OT schemes for sum power minimization. With full time slot transmission, the energy per packet transmission is proportional to power in the following discussions. Minimizing power is equivalent of reducing energy. Due to the geometric duality of the two links, here we fix link 2 and investigate how the performance changes with d_{11}.

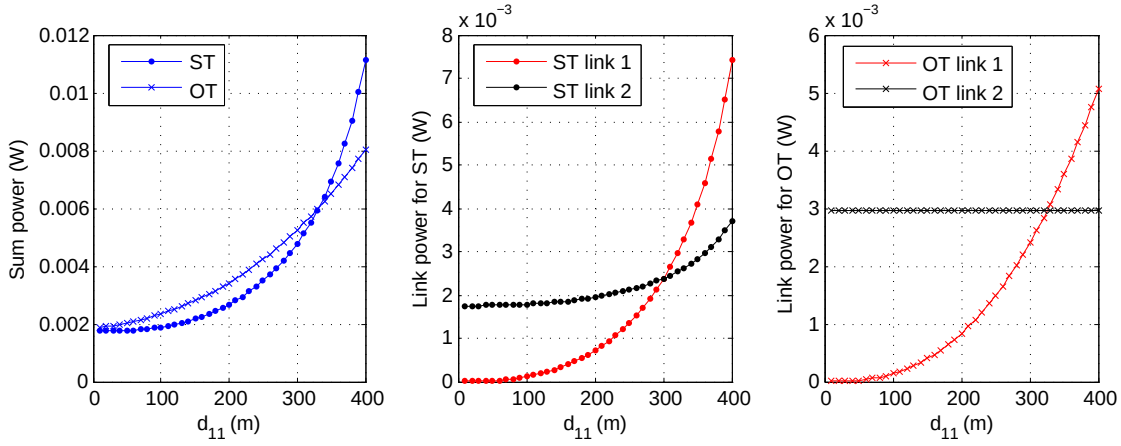

Figure 3.2: *Required power of ST and OT with* $R_1 = R_2 = 1$ *bps/Hz*

As a numerical example, we assume that the rate constraints on both links are 1 bps/Hz, and user 2 is located with a distance $d_{22} = 300$ m to base station 2. Firstly, under the parameters in Table 1, when user 1 moves from cell center to cell edge, Figure 3.2 suggests that the required sum power increases monotonically. For the ST scheme, the assigned power of both links increases with d_{11}. The required power of link 1 starts from 0 when $d = 0$ m and dramatically increases with its transceiver distance. This illustrates that in order to maintain the required rate constraints, more power is needed due to their increased mutual interference. For the OT scheme, only the power allocated for link 1 increases with d_{11}. Besides that, an interesting observation is the intersection of the two sum energy consumption curves, which indicates that different schemes may be adopted in different scenarios for the purpose of network energy

reduction. Intuitively, because the two links are interfering with each other in the ST scheme, simultaneous transmission is generally preferred when the users are near their cell centers, while the OT scheme is preferred when the users approach to their cell edges.

3.3.5 Scheme Selection with Sum Data Rate Constraints

We further investigate a selection criterion for the ST/OT allocation schemes for sum power minimization under sum rate constraints. Our numerical results suggest that the intersections of ST and OT curves are closely related with the sum rate constraints in the network.

Based on the location-dependent superiority of the ST and OT schemes, their intersections could also be treated as the boundary between selecting one scheme over another. In practice, obtaining all the division points under specific rate constraints is very intensive. Instead, we could relax our conditions by defining sum rate constraints in the network, and then investigating the relation between users' locations and scheme selection.

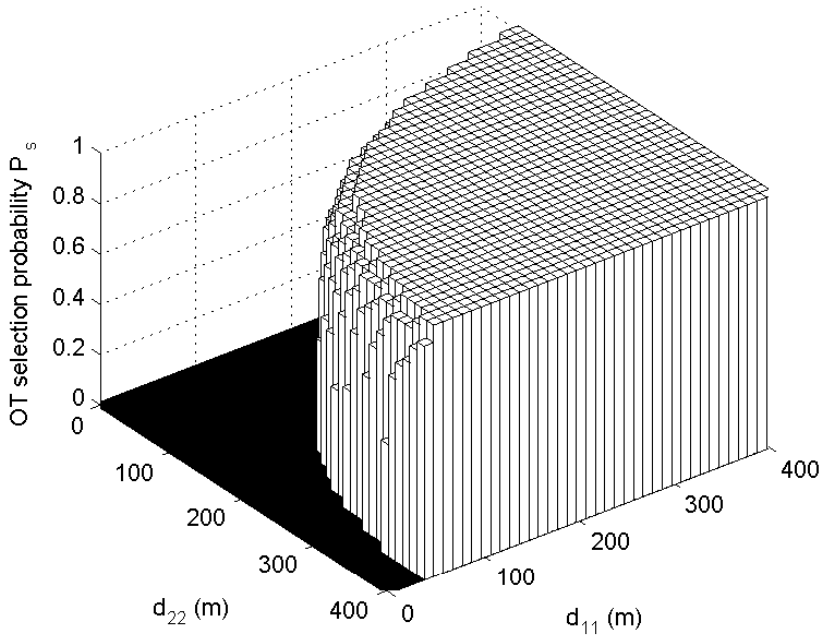

Figure 3.3: *Scheme selection for sum power minimization with $R_1 + R_2 = 12$ bps/Hz and R_1 uniformly distributed in the range [0,12] with 20 samples*

Under a sum rate constraint $R_1 + R_2 = \mathbb{R}$ ($\mathbb{R}$ is a real constant), the sum energy consumption

under ST and OT is compared. For ST, the energy consumption is set to infinity if the rate requirements could not be satisfied. For OT, the optimum frequency division factor ζ is numerically evaluated for sum power minimization. For instance, by constraining $R_1 + R_2 = 12$ bps/Hz and changing R_1 uniformly in the range $[0, 12]$ with 20 samples, we could determine the area where OT is more power-efficient than ST for achieving the same desired sum data rates. The associated probability P_s is plotted, indicating the proportions that OT is more power-efficient with all the assumed date rate pairs. This results in a location-based scheme selection criterion as shown in Figure 3.3.

As the figure suggests, the OT scheme is selected when users are far away from cell centers, while the ST scheme is preferred when users are near cell centers. This is due to the increased mutual interference in the ST case when the users approach cell edges. In addition, it is noticed that the 'boundary area' with $0 < P_s < 1$ in Figure 3.3. is quite narrow. This area represents that the preferred scheme depends on the specified rates on both links rather than the sum rate constraints alone. Further, Figure 3.3. suggests that changing R_1 under the constraint $R_1 + R_2 = 12$ bps/Hz only has very minor effects on scheme selection. Although the proportions of the sum rate occupied by link 1 and link 2 vary linearly, the selected OT region is almost static. Under other sum rate constraints, this property also holds as summarized in Remark 1.

Remark 1: Based on a two-link line model, the boundary between the ST and FDMA-based OT schemes being selected is dominated by the users' locations under sum rate constraints.

To get an impression of the boundary shape, we could simply assume $R_1 = R_2$ and investigate the ST/OT scheme selection criterion under different sum rate constraints $\mathbb{R}$.

As shown in Figure 3.4, the selected OT region expands with an increased value of $\mathbb{C}$ from cell edges to cell centers. At low rates, the ST scheme outperforms the OT scheme in most of the cell-center area. While at high rates, the OT scheme is usually the best in terms of sum power minimization. In this circumstance, the selection criterion not only illustrates that OT is more power-efficient than ST, but also indicates the higher probabilities of failing to satisfy the assumed rate constraints by using the ST scheme. In a fading environment, the channel gains become stochastic, and the optimum selection criterion for sum power minimization will be time-dependent. However, depending on the variance of the fading components, the criterion depicted in Remark 1 could also be modified to operate on time averaged channel conditions.

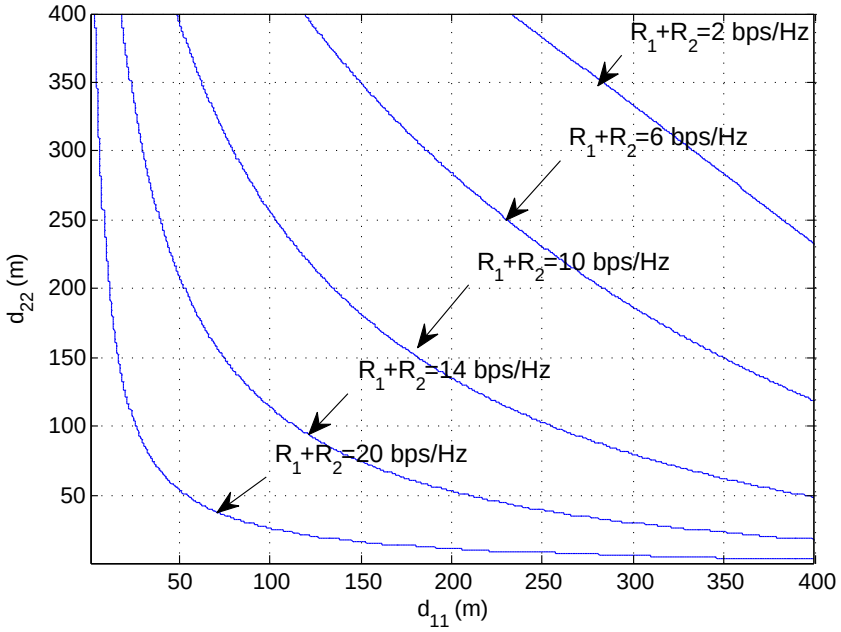

Figure 3.4: *Scheme selection under different sum rate constraints* ($R_1 = R_2$). *ST is preferred below the curve, whereas OT is preferred above the curve.*

3.4 Performance Analysis and Tradeoffs

While reducing energy consumption for Green Radio is becoming a hot topic for future communications, lots of practical issues need to be studied. In this section, various tradeoffs and constraints are taken into account to define power-efficient resource allocation schemes. Besides data rate, other application-dependent QoS requirements such as delay and error probability are involved. Meanwhile, the fairness problem among multiple links, and the physical constraints set by hardware devices are also imposed on system design.

3.4.1 Energy Consumption Rating

Although network operators aim to use wireless resources economically and efficiently by reducing network energy consumption, users' QoS needs to be maintained on all the links. Unlike the sum power reduction problem in Section 3.3, here the energy consumption on each link is scaled by its corresponding rate based on ECR definition in Section 2.2.3. Thus minimizing the sum 'power assigned per bit' could be interpreted as reducing the energy consumption on each link under given rate constraints, or maximizing the throughput on each link under given power constraints. Assuming there are I links in the network, the system energy efficiency can

be represented as

$$\mathrm{ECR} = \sum_{i=1}^{I} \frac{P_i}{R_i}. \tag{3.17}$$

Minimizing ECR could reflect our objective of reducing sum energy consumption under QoS constraints. Associated with this, a fairness metric called the power fairness index (PFI) is also defined as

$$\mathrm{PFI} = \frac{(\sum_{i=1}^{I} \frac{P_i}{R_i})^2}{I \sum_{i=1}^{I} (\frac{P_i}{R_i})^2}. \tag{3.18}$$

This fairness index is defined under the framework of Jain's fairness by choosing P_i/R_i as our objective for optimization [63]. As we are trying to minimize the sum 'power assigned per bit' in the network, we also desire that the ratio P_i/R_i $(1 \leq i \leq I)$ be fairly achieved among all the links. This is because according to the approximation of $\log$ functions in Equation (2.4), we could obtain

$$R_i = \log_2(1+\rho_i) \approx \begin{cases} \log_2 \rho_i & \rho_i >> 1 \\ \rho_i \log_2 e & \rho_i \approx 0. \end{cases} \tag{3.19}$$

While $\rho_i = P_i g/(N_0 B + I)$, it is found that for the ST scheme under fixed channel conditions, the achieved rate is proportional to the consumed power in the low-power regime, resulting in a fixed P_i/R_i value; while in the high-power regime, the achieved rate increases logarithmically with the consumed power, and P_i/R_i will dramatically increase with the power used on that link. For the OT scheme, although there is no co-channel interference, the approximation of $\log$ functions is also valid according to Equation (3.22). Therefore, achieving fair P_i/R_i could put limits on the likelihood of high-power links, reducing the probabilities of reaching maximum power constraints. Here, the fair share of 'power assigned per bit' is denoted by a high PFI value, which should be within the range $[1/I, 1]$. Using these two metrics, resource allocation strategies can be designed according to network operators' requirements. For example, we assume there are two users in the current cell, one is close to the base station with good channel condition and the other is near the cell edge with poor channel quality. The required sum power of all the active links in the cell is constrained by a system peak power limit. For operating expense, energy cost is proportional to the power which has been allocated. Assuming users are charged in proportion to their data rates being serviced, we would like to allocate power proportional to the data rate for each link. In this case, maximizing PFI is applied. In contrast, if we just want to achieve the highest energy efficiency, it is better to allocate more power to the user with good channel quality and less power to the user on the cell edge.

3.4.2 Efficiency-Fairness Tradeoffs

For the OT scheme described in Section 3.2.2, the optimum number of subcarriers n_{p} is selected to minimize sum energy consumption. Here, we emphasize the scenario that the two links have quite different rate requirements. Correspondingly, we switch our objective to minimize the sum 'power assigned per bit' ECR.

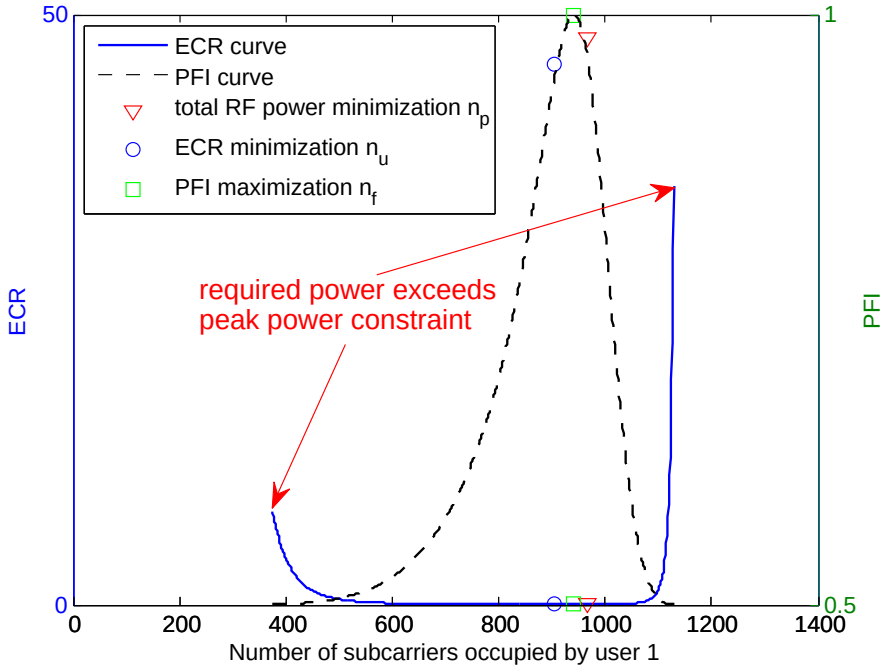

Figure 3.5: *ECR under different bandwidth allocation for OT scheme*

For instance, we set $R_1 = 5$ bps/Hz and $R_2 = 1$ bps/Hz based on the two-link line model of Section 3.1. The location of the two users are fixed at $d_{11} = 200$ m and $d_{22} = 300$ m. The peak power constraint of each link is assumed to be $P_1 = P_2 = 40$ W. Following the parameters in Table 1, Figure 3.5 is plotted for examining the achieved ECR under different subcarrier assignments n.

As shown in Figure 3.5, the ECR curve in terms of ζ is convex with its unique minimum point achieved at n_{p}. Because the power allocated on each link is scaled by its corresponding rate, the optimum operating points n_{u} under ECR minimization, and n_{p} under sum power minimization are different. Also, because $R_1 > R_2$ and $d_{11} < d_{22}$, the number of subcarriers assigned to link 1 is greater than that of link 2, indicating that the link with higher data rate requirements or better channel conditions should be allocated more resources under the assumption of ECR

minimization. Furthermore, we notice that the curve around n_{u} is relatively flat, with a scale of 10^{-2} on the achieved ECR. While for the points far away from n_{u}, the value of ECR grows dramatically. It illustrates that slightly moving the operating points would not cause a great increase in ECR. By adjusting the value of d_{11} and d_{22}, it is found that the above features still hold for different users' locations.

Apart from ECR minimization, we also concerned about the fair sharing of P_i/R_i among different links. According to the PFI definition for a two-link scenario, the fairest point $\mathrm{PFI} = 1$ is obtained when $P_1/R_1 = P_2/R_2$. Under this constraint, we could also find the unique solution n_{f} for OT to achieve its highest fairness. By taking both ECR and PFI into consideration, a weighted resource allocation approach is proposed based on power efficiency and fairness tradeoffs. With a weighting factor w $(0 \leq w \leq 1)$, we can calculate the bandwidth occupied by link 1 for the weighted approach as

$$n_{\mathrm{w}} = wn_{\mathrm{u}} + (1-w)n_{\mathrm{f}}. \tag{3.20}$$

Thus, changing n_{w} from n_{f} to n_{u} means an improvement in power efficiency but a loss in fairness. Based on this efficiency-fairness tradeoff, different strategies could be designed according to the network operation policy.

3.4.3 Outage Probability

This part illustrates the tradeoffs among energy consumption, delay constraints and outage probabilities. Simulation results suggest that the probabilities of failing to satisfy the rate requirements decrease with either relaxed delay constraints or short-term power constraints.

Because of the time-varying characteristics of wireless fading channels, a block fading channel model is now employed for analysis. By dividing the time axis into small discrete time steps τ compared to the channel coherence time, we could assume that the fading factors are fixed in each time slot but are randomly reselected for each different time slots [45]. Here, we choose large-scale shadow fading using the log-normal distribution to describe environmental variations between different cells. The channel gains affected by shadowing could be represented as a function of time τ as $g_{\mathrm{s}}(\tau) = 10^{G_{\mathrm{s}}(\tau)/10}$, where $G_{\mathrm{s}}(\tau)$ is in normal distribution [43]. Depending on the practical environment, the channel gains due to path loss, large-scale shadowing and small-scale fading could be multiplied together to model real communication channels.

Apart from the parameters shown in Table 3.1, we further assume that the locations of the two users are fixed. Here, both path loss and log-normal shadowing with zero mean and standard derivation $\iota = 4$ dB are considered. In addition, we set a short-term power constraint on each link per 1 or 2 time blocks, which represents the maximum power allowed for transmission during that time period [87]. Under the specified data rates, if the required power on that link exceeds its power limit, outage occurs due to the failure to achieve the specified QoS. For a long-term process, the outage probability P_{out} is often treated as a measure of system performance, related to the packet error rate under certain coding and modulation schemes. Here, the selected time span for applying short-term power constraint equals to the delay constraint, which is application-dependent and measured by the number of time slots that the application being used could tolerate. The system delay constraint is usually specified by the number of time slots for constructing an appropriate coding scheme [88].

Based on the model discussed above, we assume that both of the links are delay-limited with the fixed amount of information delivered flexibly over every D_{el} time slots. For example, under the delay constraint $D_{\text{el}} = 1$ time slot, R bps/Hz information needs to be delivered within each time slot. If the delay constraint is relaxed to $D_{\text{el}} = 2$ time slots, $2R$ bps/Hz can be flexibly transmitted within every two time slots, making the traffic relatively elastic.

For the simulations, we choose the ST scheme under average data rate requirements $R_1 = 1.5$ bps/Hz and $R_2 = 0.5$ bps/Hz while fixing the users' locations at $d_{11} = d_{22} = 300$ m. We set the delay constraint on both links as $D_{\text{el1}} = D_{\text{el2}} = 1$ time block for the first case, and gradually increase the assumed short-term power constraint to calculate the corresponding outage probabilities. For the second case we relax the delay constraint on link 2 to $D_{\text{el2}} = 2$ time blocks and compare its performance with the first case according to the algorithm summarized in Table 3.2.

As shown in Figure 3.6, with specified data rates, delay constraints and perfect CSI at the transceivers, the power needed for the ST scheme to achieve these QoS requirements could be determined. Obviously, the outage probabilities P_{out} will decrease with an increased maximum short-term power constraint P_{m}. By setting the same channel gain g and maximum power limit P_{m} on the two links, link 1 which has a higher rate requirement will require more power than link 2, thus exhibiting higher outage probabilities. Also, when the delay constraint of link 2 is increased, the outage probabilities of both link 1 and link 2 decrease. This is because link 2 could adjust its transmission according to time-varying channel conditions. When the channel

Algorithm:

1. Define threshold P_{m} as system short-term power constraint. Using P, R, D_{el} and P_{out} to denote transmit power, data rate, allowed delay and outage probability.
2. Assume the delay constraint $D_{\mathrm{el1}} = D_{\mathrm{el2}} = 1$ applies. For the ST scheme, calculate the required power P_1 and P_2 under R_1 and R_2 respectively.
3. Compare P_1 and P_2 with P_{m} for both links. If $P_1 \leq P_{\mathrm{m}}$ and $P_2 \leq P_{\mathrm{m}}$, transmit simultaneously to both links; if $P_1 > P_{\mathrm{m}}$ or $P_2 > P_{\mathrm{m}}$, switch off one link and calculate the corresponding $\bar{P}_1$ and $\bar{P}_2$ required for the other link.
4. Compare $\bar{P}_1$ and $\bar{P}_2$ with P_{m}. If $\bar{P}_1 > P_{\mathrm{m}}$ and $\bar{P}_2 > P_{\mathrm{m}}$, note that outage happens on both links; if $\bar{P}_1 \leq P_{\mathrm{m}}$ and $\bar{P}_2 > P_{\mathrm{m}}$, or $\bar{P}_1 > P_{\mathrm{m}}$ and $\bar{P}_2 \leq P_{\mathrm{m}}$, just use link 1 / link 2 to transmit and note down outage occurs on link 2 / link 1; if $\bar{P}_1 \leq P_{\mathrm{m}}$ and $\bar{P}_2 \leq P_{\mathrm{m}}$, randomly activate one link and deactivate the other link in the current time instance according to Gaussian distribution.
5. Repeat from 2. for the next time instance to calculate P_{out} for the two links.
6. Assume the delay constraints $D_{\mathrm{el1}} = 1$ and $D_{\mathrm{el2}} = 2$, optimally allocate rate for link 2 to minimize $P_1 + P_2$ every two blocks, and compare the required power on each link with its corresponding short-term power constraint to calculate P_{out}.

Table 3.2: *Power allocation algorithm with constraints*

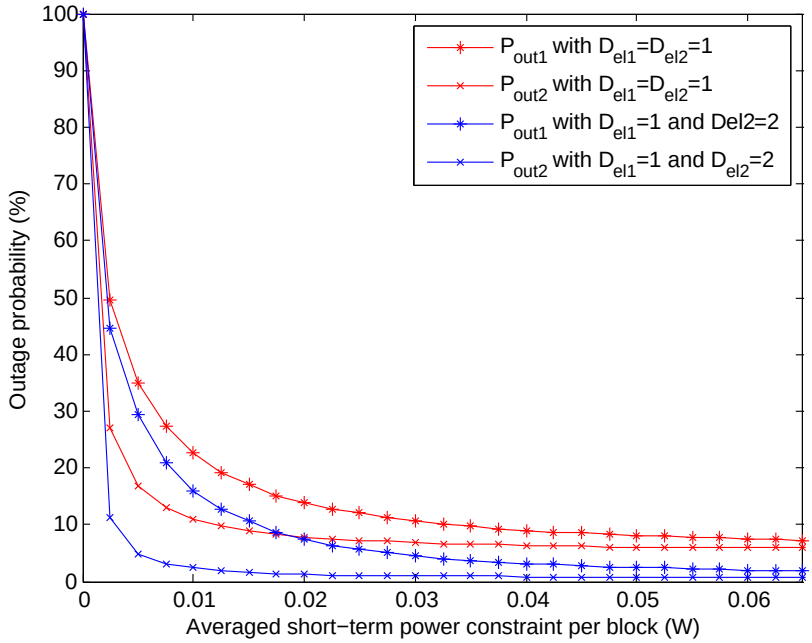

Figure 3.6: *Power, outage and delay tradeoffs with* $R_1 = 1.5$ *bps/Hz,* $R_2 = 0.5$ *bps/Hz*

gains are relatively large, higher rate could be delivered in that time slot. When the channel gains are relatively small, we stop transmission or just use a lower data rate. This flexibility also contributes to link 1 because the two links are mutually interfered. Therefore, Figure 3.6 shows the impacts of relaxing the delay constraints from 1 to 2 time blocks on outage probabilities.

Similarly, this trend holds when we further extend the allowed time delay. For example, in LTE there are ten subframes per frame, outage probability may happen or increase when we use less active subframes to transmit. The above discussion reveals the tradeoffs among rate, power, delay and outage probability. With a relaxed delay constraint, or increased maximum power limit, or decreased rate requirement, the outage probability will decrease, resulting in a lower error rate during the transmission process.

3.5 Soft Frequency Reuse in Energy-Efficient Design

In this section, we extend the above analysis to the multi-link scenarios by introducing the concept of soft frequency reuse. As an effective technique to mitigate interference, SFR is not only able to maximize the cell overall spectral efficiency, but also can significantly improve cell energy efficiency with associated frequency allocation algorithms.

3.5.1 Multiple-Link Scenarios

Based on the above mentioned two link study, it is more energy-efficient to use OT scheme in the cell edge area to avoid inter-cell interference. Since the total bandwidth is divided into two groups, we can also say that the frequency reuse factor (FRF) is 1/2. Similarly, the FRF for the ST scheme is 1. Whether ST or OT is preferred is decided by the users' locations and required data rates.

When there are multiple links, analytical results are not available for subcarrier allocation decisions. Depending on the provided cell traffic profile, it is possible that a hybrid ST/OT scheme may achieve the best energy efficiency. While some links are sharing the same frequency resources, other links may be orthogonal to each other. Since the optimum solution is difficult to obtain, simplified approaches have been proposed by assuming that all users are transmitting simultaneously or orthogonally.

In the LTE system, the soft frequency reuse strategy has been included as a flexible spectrum allocation technique in the standard. Correspondingly, fractional frequency reuse (FFR) has also been introduced for the IEEE 802.20 system. Knowing the user information such as position information (PI) and channel state information, both of these techniques can effectively mitigate inter-cell interference and improve network overall performance. In general, under a limited

power constraint, with users being classified as cell-center users and cell-edge users, FRF=1 is always assigned to the cell-center users and FRF=1/3 or 2/3 is assigned to the cell-edge users for SFR or FFR respectively. Figure 3.7 demonstrates the idea of SFR. While FRF=1 is allocated to the blue cell-center area to maximize spectral reuse efficiency, different frequency bands are orthogonally allocated at cell-edge area with red/green/yellow colours. In LTE, frequency bands could also be subcarriers or resource blocks.

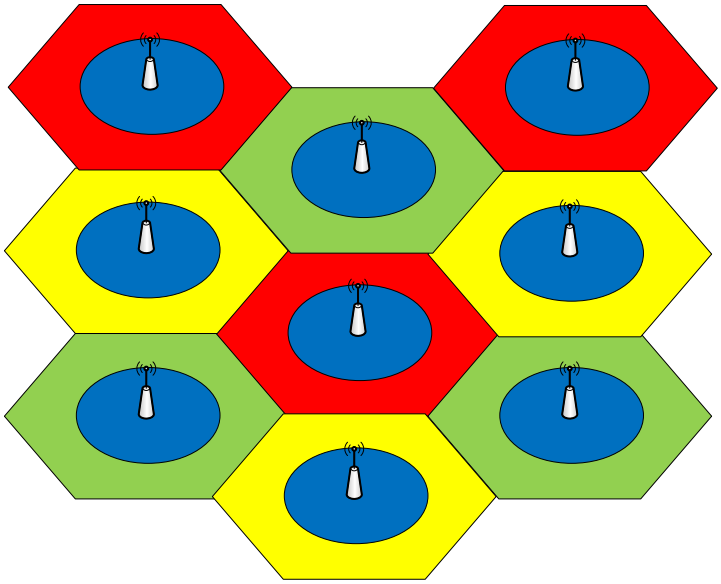

Figure 3.7: *Soft frequency reuse for interference mitigation [90]*

Regarding SFR and FFR, it is important to decide the boundary of cell-center and cell-edge users. This is associated with power control. Generally, for a given network traffic profile, different optimization objectives such as maximizing spectral efficiency and minimizing energy consumption will lead to different boundaries. The allocation process can be either static or dynamic as well.

3.5.2 SFR for Maximizing Spectral Efficiency

Traditionally, SFR is applied for maximizing spectral efficiency. Different algorithms have been developed to improve the performance. In [89], the total frequency band has been divided into a major group and a minor group. While the major group is allocated to the users inside the whole cell, the minor group is only allocated to the cell-center users. Meanwhile, the power assigned to the minor group is defined to be smaller than the power assigned to the major group.

Their quotient is termed as the power ratio. According to the traffic in each cell, the reuse factor can be adapted between 1 and 1/3 by allocating minor group subcarriers and power.

In [90], FRF=1 is initially assigned to all the users within each cell. Then we measure SINR and separate the cell-edge area by identifying those users whose signal to interference plus noise ratio is smaller than a threshold. To improve the performance and allocate appropriate channels, different FRFs are tried for the cell-edge area according to the CSI feedback. Since the process of subcarrier allocation is iterative, this method is not very efficient. In [91], SFR is reviewed by formularizing a function of sum spectral efficiency maximization subject to a sum power constraint in the cell. A sub-optimal scheme is introduced based on the assumption that all the users are allocated the same number of subcarriers. Starting by assigning the current subcarrier to the user with the best channel coefficient, we always allocate the subcarrier to the user with less number of subcarriers than the one with maximum number of subcarriers in its set. Correspondingly, the water-filling power allocation is adopted together with the orthogonal frequency allocation. Meanwhile, FRF=1 is used for the cell-center users. A user is assumed to be located in the cell-center if the path loss difference between its selected cell and its closest neighbor cell is greater than a threshold value.

Besides, industrial companies have proposed different SFR schemes for implementation. According to Ericsson's proposal [92], partial frequency bands are assigned to the cell-edge users with full power, while full frequency bands are allocated to the cell-center users. In Alcatel's proposal [93], a fixed FRF=3/7 is adopted for the cell-edge users, and cell-center users use FRF=1 as well. However, with Siemens [93], only a part of spectrum is used in the cell-center area. In each cell, using N_{sub} and χ to denote the total number of subcarriers and the number of subcarriers allocated to its cell-edge users, only $N_{\mathrm{sub}} - 3\chi$ subcarriers are allocated to the cell-center users, although cell-edge FRF can be adjusted according to the traffic conditions.

Because of the time-varying properties of wireless fading channels, dynamic major group allocation is proposed in [94]. Heavily loaded cells can borrow frequency resources from lightly loaded neighbors. Meanwhile, fairness matrices are defined together with resource allocation based on the proportional fairness concept. Maximizing sum spectral efficiency is replaced by maximizing the quotient of transmitted user data rate divided by required user data rate. In [95], a so-called fairness measure is defined as the standard deviation of the above quotients across all the users within the cell.

3.5.3 SFR for Improving Energy Efficiency

SFR can be used for minimizing network sum energy consumption as well. In [96], τ denotes time instance and I to denote the total number of users in the cell, the optimization equation can be written as

$$\begin{cases} \min \sum_{\tau=1}^{D_{\mathrm{el}i}} \sum_{i=1}^{I} P_i(\tau) \\ \sum_{\tau=1}^{D_{\mathrm{el}i}} R_i(\tau) \geq R_{\min}^{i}. \end{cases} \tag{3.21}$$

Where P_i and R_i represent the power and rate obtained by user i, $R_{\min}^{i}$ is the minimum rate requirement of user i. For frequency allocation, it generally includes bandwidth assignment and subcarrier assignment. Regarding bandwidth assignment, an algorithm called bandwidth assignment based on SNR (BABS) is used to minimize energy consumption. This algorithm is derived from Shannon theory, where increasing bandwidth can improve spectral efficiency, thus wider bandwidth can reduce the energy consumption for achieving the same rate target. The BABS algorithm is based on the principle that allocating each subcarrier to the user which can reduce the most energy consumption. For full bandwidth transmission, this process repeats until all the subcarriers are allocated orthogonally. With flat fading channels, the obtained result converges to the optimal distribution of subcarriers. However, with frequency selective channels, a further process 'subcarrier allocation' is required. In [96], both amplitude carving greedy (ACG) algorithm and rate craving greedy (RCG) algorithm can be deployed for subcarrier assignment.

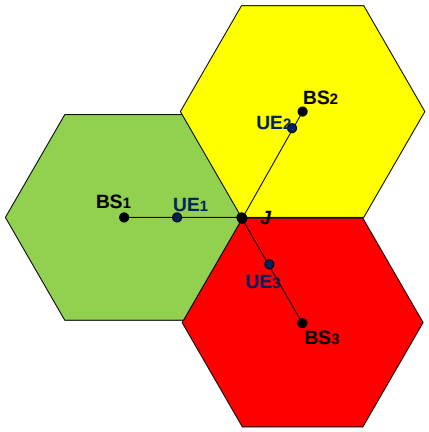

Figure 3.8: *Flexible frequency allocation for a simple three link scenario*

Based on the above algorithms, we extend our analysis in Section 3.4 to a simple three-link scenario. Using BS and UE to denote the base station and user equipment respectively, we assume that all the three users can only move on the link between their base stations and the cell edge junction J. This is shown in Figure 3.8.

If we further assume the transceiver distance follows a random distribution within [0,400], and the three users' sum rate constraint is fixed. By varying the data rates of the three links, the energy-efficient ST/OT boundary is almost static. This also proves our results in Section 3.4 for the two-link scenario. In Figure 3.9, by constraining the users' sum rate constraint as 2.1 bps/Hz, we numerically search for the optimum energy consumption by changing the ST/OT boundary from 0 to 400 m.

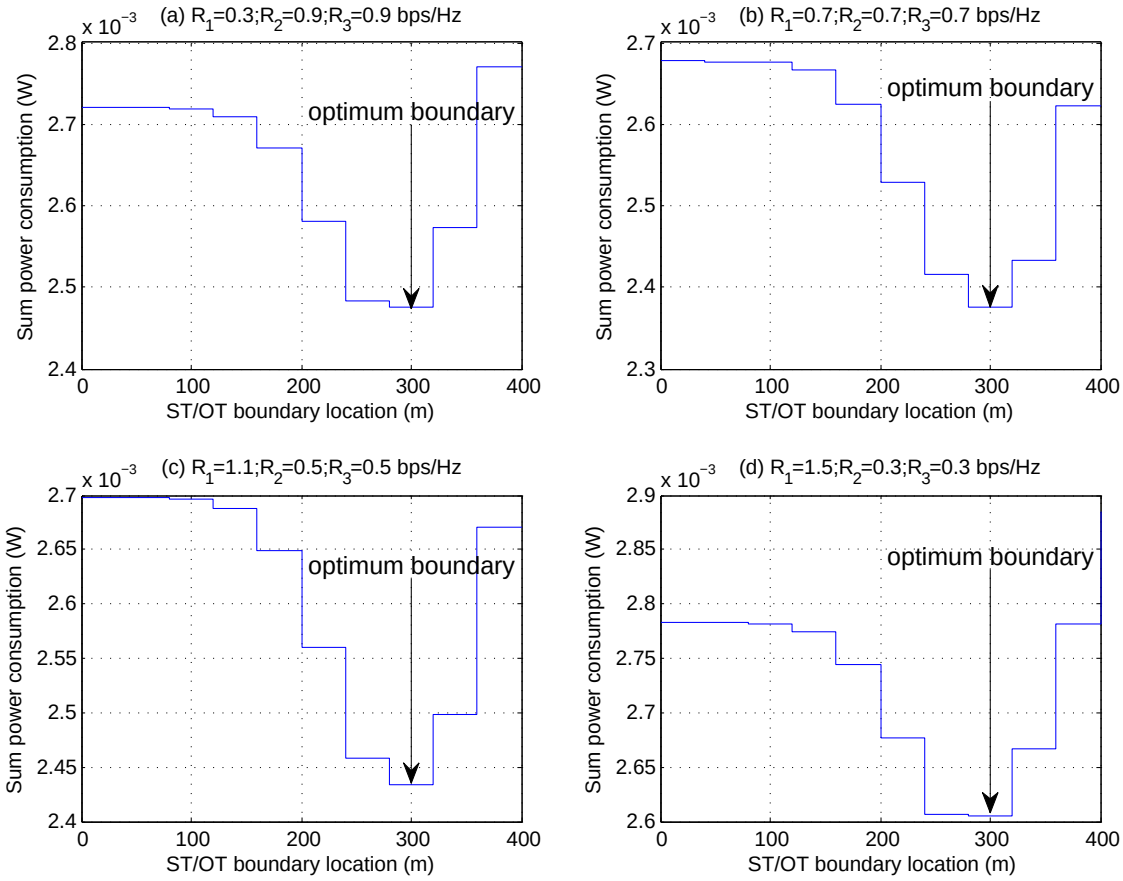

Figure 3.9: *Power-efficient ST/OT boundary locations for three-link scenarios*

As the figure suggests, the minimum energy consumption is obtained when the ST/OT boundary is set around 300 m to the base station. The ST scheme is always deployed when the boundary

is set at 400 m, and the OT scheme is always used when the boundary is set at the base station. With different link data rates, the optimum power allocation is different. In generally, more power is required for transmission when the difference among different links deepens. Using SFR with a proper boundary setting can significantly reduce network energy consumption compared with the traditional ST or OT scheme.

With different sum data rate constraints, similar results can be obtained. Also, we can further relax the constraints on the number of users and their locations. Future research can be carried out by assuming all the users randomly located within the whole cell area.

3.6 Summary

In this chapter, we have discussed power-efficient resource allocation in modern communications under the Green Radio framework. Our emphasis is on considering interference environments. Under the urban macro-cell scenario, a simple two-link line model with the WINNER II channel parameters is assumed for analyzing the properties of different resource assignment schemes. By removing the time-varying characteristics of wireless fading channels and only considering their path loss components, the energy consumption of ST and FDMA-based OT schemes is calculated under the specified rate constraints. Several conclusions are summarized based on their comparisons, which provide us further insights into designing new resource allocation strategies for sum power reduction.

Firstly, our analysis reveals that ST may fail to satisfy the data rate requirements, which could not be improved by power control. Secondly, it is found that the ST/OT scheme selection criterion for sum power reduction is location-dependent. In general, when the two users are close to their cell centers and far away from each other, the ST scheme is more energy-efficient than the OT scheme, and vice versa. This is because in the cell-center area, low power levels are sufficient to satisfy the transmissions with good channel gains, making the interference between the two links relatively weak. Meanwhile, the cross-link channel gains are also poor in this circumstance based on the two-link line model. Thirdly, with the above assumptions, numerical results suggest that the ST/OT selection criterion is dominated by its network rate constraints rather than the specified rate requirements on individual links. Besides these discussions, we also analyzed the OT scheme under the TDMA framework. It is proved that the TDMA-based OT scheme always requires higher power than the FDMA-based OT scheme for achieving the

same QoS, but in terms of energy consumption per packet, their efficiency is equivalent.

In practice, economical and environmental design for sum energy minimization falls into the framework of various tradeoffs and constraints. A new energy efficiency metric using the sum 'power assigned per bit' is proposed, and the corresponding optimization objective is also changed for the ST and OT schemes under different QoS constraints. The advantages of using this new metric for efficiency and fairness definitions are discussed, resulting in a weighted resource allocation approach based on power efficiency and fairness tradeoffs. By adding log-normal shadowing to form stochastic fading channels, the relation among delay, outage and short-term power constraints are also analyzed. It is shown that under the specified rate requirements, both relaxing delay constraints and short-term power constraints benefit the outage performance of the two links.

We need to notice that the theoretical analysis in this chapter is based on Shannon capacity formula, which provides the upper bound of the maximum achievable data rates. Here the energy reduction problem is analyzed by inverting the Shannon equation. Although many techniques for spectral efficiency maximization can be adopted similarly from an energy efficiency perspective, their differences are obvious in a practical communication system. For example, given a fixed amount of data, using longer time to transmit will generally reduce energy consumption. However, in real communication networks, energy saving can also be obtained by deploying time-domain sleep modes, which switch off certain time period and thus use less active time for transmission. This is primarily because that in practice by gating off certain time period, we can reduce the energy used for operating other radio components such as mobile phone hardware circuits. The sleep mode scheme further highlights the difference between spectral-efficient and energy-efficient design, which can also be applied for radio base stations based on the base station energy consumption model. In Chapter 4, we will introduce the time-domain sleep mode design for radio base stations.

Chapter 4
Time-Domain Base Station Sleep Mode Optimization

4.1 Introduction

In this chapter, we aim to think about energy-efficient transmission techniques from a more holistic point of view. Starting with the analysis of base station power consumption, we know that the energy used directly in the transmission process represents only a small percentage of the total energy, while the vast majority of energy is used in hardware processes such as rectifying and cooling, digital signal processing (DSP), analogue to digital/digital to analogue conversion (ADC/DAC) and power amplification. Although reducing hardware energy consumption is not our main research focus, their effects should not be neglected when we design new radio access strategies. This further highlights the difference between spectral-efficient and energy-efficient designs. For example, in [97], both radio frequency and transceiver circuit energy consumption are taken into consideration for designing energy-efficient link adaptation strategies. In [98], training based transmission schemes are analyzed from both the spectral and energy efficiency perspectives. In [99], the bandwidth expansion concept is proposed for LTE advanced, allowing for a bandwidth-energy efficiency tradeoff. Here, we want to take advantage of our design to further reduce the overall operational energy consumption.

A straightforward idea to reduce operational energy consumption is deploying sleep modes. By switching off unused wireless resources and hardware devices, sleep modes can potentially achieve significant energy reduction. Because the users' QoS still needs to be guaranteed, reasonable assumptions of traffic loads and network constraints are essential in sleep mode design. Generally, we expect more energy reduction at low traffic loads due to the flexibility to switch off more wireless resources.

As we know, using sleep modes for energy saving is not a new concept. It has been widely applied to many wireless devices such as cell phones and computers due to the use of batteries. When there is no data to transmit or receive, user terminals are switched to non-active modes for

energy saving. Regarding implementation, sleep mode algorithms are defined in wireless communication standards and supported by user terminals. For example, LTE uses discontinuous transmission and reception techniques in the handset [100]. IEEE 802.16e also employs a sleep mode mechanism for mobile stations based on a discrete-time queueing model [101]. As these designs are based on the analysis of queueing delay and traffic profiles, they generally refer to the time-domain sleep modes [102]. According to the packet arrival rate and queue length, the mobile station can be simply switched among different modes to save energy consumption.

Besides, sleep mode design can also contribute towards energy saving in radio base stations. In [103] and [104], the proposals of geographically switching off individual base stations at low traffic loads are discussed. With some cells being turned off, radio coverage and service provisioning can be taken care of by the neighboring active cells, so as to guarantee that service is available over the whole network. Similarly, an extended discontinuous transmission mode for the base stations has been proposed for the LTE standard to handle some practical signalling issues [64]. In [105], the sleep mode design is further extended to allow gating off a number of frequency subcarriers rather than the whole base station during off peak hours. Numerical results are provided, demonstrating that under a given traffic profile and QoS objective, using the proper number of subcarriers for transmission has the potential to improve network energy efficiency.

From the network architecture point of view, sleep modes can also be deployed in heterogeneous networks. In [106], sleep modes have been used for small femtocells to form an overlay layer on the existing macrocell network. This architecture not only satisfies high data rate traffic requirements and enables breakthrough services, but also can conserve energy at low traffic loads. In [107], the operational principles as well as the transition to and from power saving mode for the femtocell base stations are further described. Also, in [108] a network scale sleep mode together with the energy-aware system selection criteria is introduced for the 2G/3G systems. By deactivating the entire system for some traffic scenarios, large energy reductions can be obtained.

Unlike the above approaches, here we focus on deploying base station sleep modes in the time-domain using ‘blank subframes’ [14]. Although the concept of delivering blank subframes during low traffic loads has been proposed for femtocells and relay networks [109], no details are provided regarding how to optimize the network and how much energy saving can be achieved. Here the baseline system we adopt is LTE, and in principle the idea can generally be applied to

all kinds of wireless systems. The energy consumption is reduced by switching off hardware processing and radio transmission functions under low traffic load conditions. This generally includes the power used by power amplifiers, ADC/DAC devices, DSP units, rectifiers and cooling systems. Meanwhile, the RF energy consumption may also be reduced according to the LTE physical layer structure. In LTE, the time frequency resources are divided into small slots, termed as physical resource blocks. Inside each PRB, some part is used for control signalling while the other part is used for data transmission. Without sleep modes, the control information is always delivered even if there is no user data for transmission, thus wasting energy at low traffic loads. With sleep modes, we can switch off some time slots and use the remaining active slots for data transmission. We can save energy by transmitting less control overhead though this will have a potential impact on channel estimation and signalling channel bandwidths.

In this chapter, Section 4.2 and Section 4.3 will describe the base station energy consumption structures and UK network profiles respectively, suggesting the potential energy saving of time-domain sleep modes. In Section 4.4, the innovations of adapting the active number of subframes in each frame depending on the traffic loads will be proposed. In Section 4.5, we will provide an in-depth technical description of sleep mode design, from the basic model assumptions to the analytical derivations and numerical results. We need to notice that the discussions in this chapter are based on the assumption of a single user SISO scenario. Also, a fixed control signalling overhead size is always used under various channel conditions for simplicity.

4.2 Base Station Energy Efficiency Model

The ambitious goal of 100x power reduction in Mobile VCE Green Radio project definition document is set by taking both embodied energy and operational energy into consideration [110]. Regarding the overall energy consumption in wireless networks, Ericsson provided a life-cycle analysis in 2004, using a model network based on the most common GSM radio base stations and radio network controllers (RNC) [111]. This life-cycle energy consumption is named as embodied energy, which includes the energy used to make any product, bring it to market, and dispose of it. Here we ignore the embodied energy consumption and only focus on the operational energy part.

Figure 4.1 shows the energy efficiency model for GSM base stations. The delivered mix of energy is recalculated into percentage of the base station consumptions, with only 2% of the

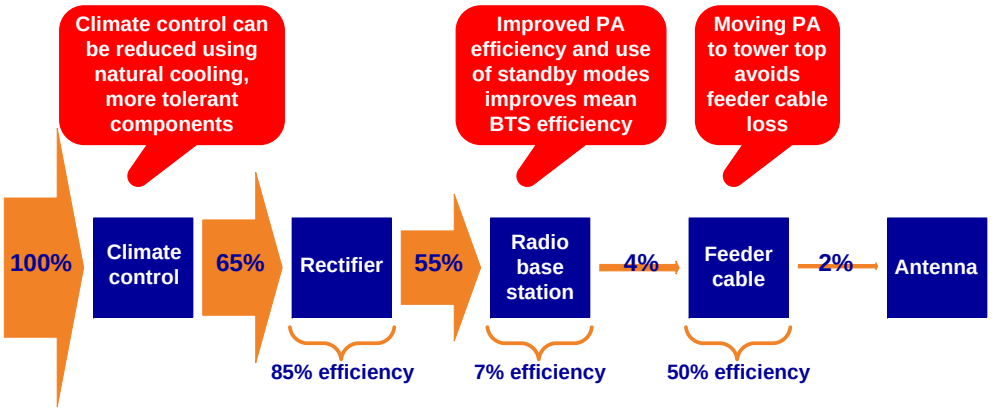

Figure 4.1: *GSM base station efficiency model [111]*

total input energy directly contributing radio antennas. From an energy efficiency perspective, this structure needs to be improved by employing intelligent radio architectures and hardware technologies. Here, the RBS energy efficiency is only 7%, which is very inefficient due to the energy loss in the ADC/DAC, baseband signal processing and the antenna power amplifier. Because one RNC typically handle many radio base station nodes, RBS nodes are identified as the most significant energy consumers in the radio access network [110]. Therefore, energy saving efforts must target RBS sites.

Similar base station power consumption measurements can also be found from [112], where the energy consumption of GSM base station cooling, power supply, power amplifier, rectifier, ADC/DAC converter, signal processing and antennas feeder loss corresponds to about 13%, 16%, 22%, 19%, 9%, 17% and 1% of total input power respectively. The energy arrives at the radio antenna for RF transmission is only about 3%. Correspondingly, the figures measured from an LTE system are listed in Table 4.1, which are based on a 3 sector macro base station with 4 antennas per sector and power amplifier efficiency of 40%. The top of cabinet (TOC) efficiency gives the ratio of the combined power output of the PAs to the power supply unit (PSU) power, and the radiated efficiency references the antenna efficiency and feeder losses [113].

Although the table suggests that only 5.3% of the total LTE base station input power is used for RF transmission, the energy consumption of most RBS components is related to this RF

Description	Power In (W)	Power Out (W)	Efficiency
Radiated power (per sector)	8	501 (27dBW)	18dBi antenna gain
Antenna and Switch	12	8	65%
Feeder	24	12	50%
PA (all sectors)	180	72	40%
Transceiver (all sectors)	180		
Free Air Cooling	40		
PSU Input	450	400	88%
TOC Efficiency			16%
Radiated Efficiency			5.3%

Table 4.1: *Estimated power consumption for base stations [113]*

power. A joint consideration of transmission protocol design with the above energy consumption model is important to optimize Green Radio technologies. In [114], the base station energy consumption is modeled as the sum of a proportional part in terms of RF transmission and a constant part. While the proportional part refers to the PA energy consumption, the constant part generally includes the energy used for signal processing. Similarly, a more complex model is proposed in [115], which further reveals that the RBS energy consumption is dependent on network traffic conditions and cell configurations. Therefore, it is expected that energy-efficient transmission strategy design can also significantly benefit energy conservation of the whole RBS.

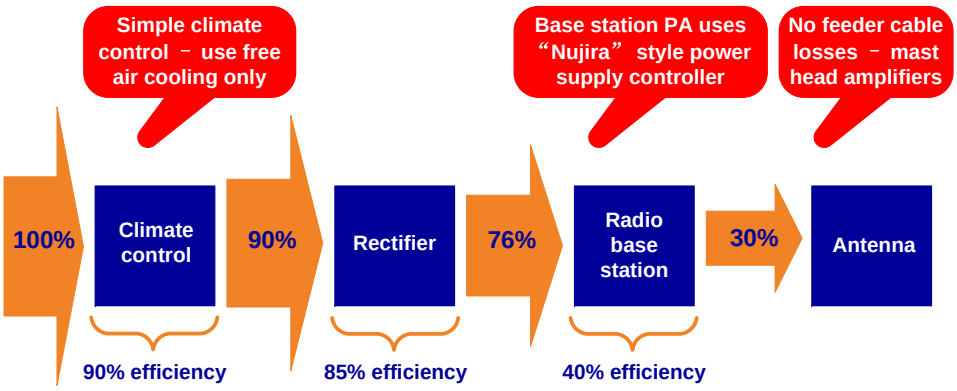

Figure 4.2: *Suggested LTE base station efficiency model*

While both academia and industry are putting their effort into reducing energy consumption, more energy-efficient radio networks and base stations can be expected. According to the presentation from Nortel [116] and the suggestions from Figure 4.1, our assumptions for future LTE base stations are updated by employing free air cooling, mast head amplifiers and energy-efficient 'Nujira' style PAs. An estimate for an improved base station designed for LTE is provided in Figure 4.2. With a large proportion of base station power contributing to antenna PAs, the energy consumption dependency on network traffic loads will further increase.

4.3 Traffic Profiles in Network

Regarding real network energy consumption figures, Vodafone in 2007 has reported its total UK base station power as 228 GWh, so with 12046 sites this gives 2.2 KW average power per site [117]. As sites share equipment, the power per site does not necessarily mean the power per base station. Also, 2.2 KW only suggests us an approximate average value. At present there is very little variation in base station transmit power with traffic. According to the hourly traffic volumes in the network, its significant variation suggests us to adapt the instantaneous network power consumption with its traffic loads.

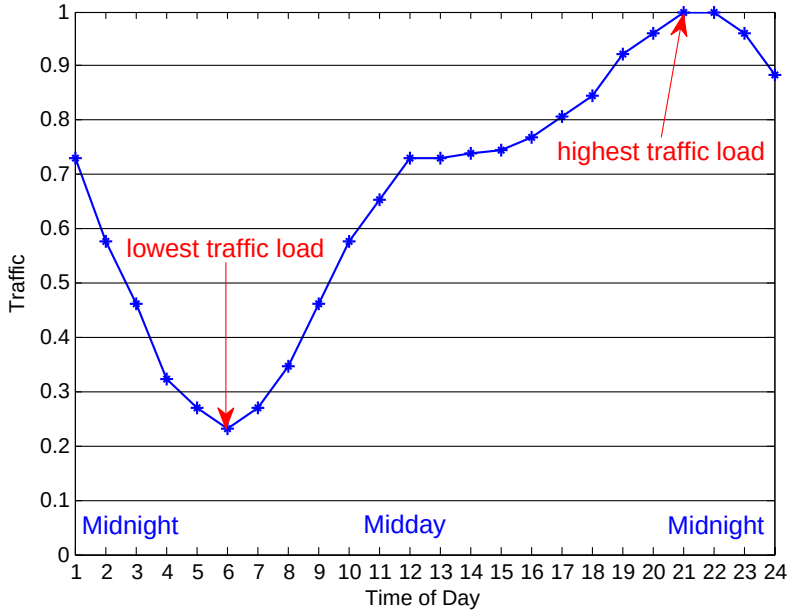

Figure 4.3: *Normalized traffic in 3G network in London [118]*

Figure 4.3 shows Vodafone results for the average traffic loads during a whole day [118]. These

data are measured based on the 3G network in London over 28 days. The obtained average hourly traffic profile is then normalized by assigning its busiest hourly traffic load to 1. In Figure 4.3, the data traffic varies significantly over time, with the smallest value obtained around 6 am and the largest value obtained around 9 pm. Also, the minimum traffic is approximately 25% of the peak traffic, and nighttime traffic loads are generally lower than daytime traffic loads.

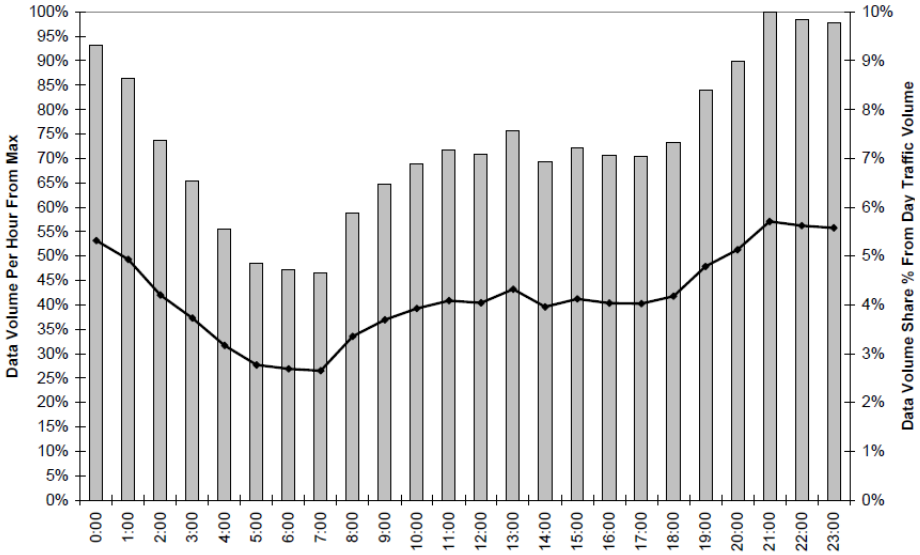

Figure 4.4: *Daily RNC level hourly data volume deviation [41]*

A similar traffic profile is shown in Figure 4.4, which is provided by Nokia depicting the data volume distribution over 24 hours for the RNC on a typical working day [41]. Compared with Figure 4.3, this figure refers to a specific scenario in high speed downlink packet access (HSDPA) network, based on the statistics of a single RNC and up to 200 radio base stations. Besides the normalized data traffic loads, the hourly traffic share percentage from the day traffic volume is also plotted. It can be seen that the single hour data volume share takes approximately 3% to 6% from the total daily traffic volume. Also, the normalized traffic load is 50% during the early morning hours and steadily increases towards the busiest hours from 9 pm to midnight. The usage increases heavily after about 6 pm which indicates that as the working day ends this is the time for the heaviest internet usage.

Comparing Figure 4.3 with Figure 4.4, it can be seen that the trends of hourly traffic changes are quite similar, although Figure 4.4 exhibits a smaller variational scale. The nighttime traffic in both figures are quite low, and most of the night hours have half the load or less relative to

the corresponding peak time. Although the network location, network type and absolute data traffic values are not provided, they can give us a pointer towards the realistic traffic distributions. Based on the above measurements, it is also expected that the analysis from broadband 3G/HSDPA networks will also be useful for the dimensioning of broadband LTE networks.

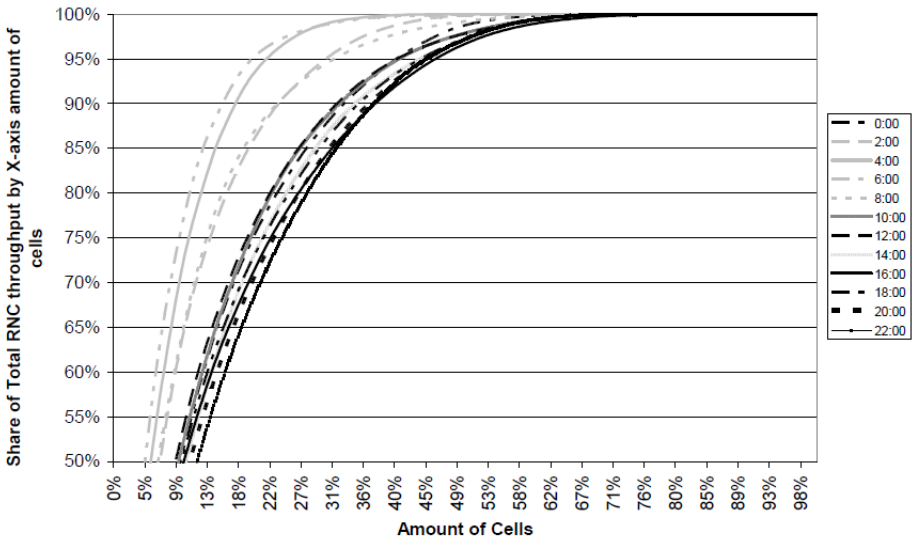

Figure 4.5: *Cells' data volume contribution to total RNC data volume [41]*

Following Figure 4.3 and Figure 4.4, another interesting problem is how much the individual base stations contribute to the hourly traffic profiles. As shown in Figure 4.5, during the night when the data volume is low and mobility is low, the traffic is heavily concentrated on certain areas (*e.g.* 14% of cells carry 90% of the traffic at 6 am), while during the day the share of cells contributing to the total data volume increases (*e.g.* 38% of cells carrying 90% of the traffic at 4 pm). This result reveals that the relatively lower nighttime traffic is also area-centralized, indicating spare wireless resources in most areas.

4.4 Innovations of Time-Domain Sleep Mode Design

This section proposes a time-domain approach for base station transmission sleep modes, taking the long term evolution system as a specific example. The idea of this design is to enable the base station to reduce energy consumption in low traffic conditions, while still properly supporting active user connections. As shown in Figure 4.6, the time-domain sleep mode is

expected to scale the base station power with traffic loads, further saving energy consumption.

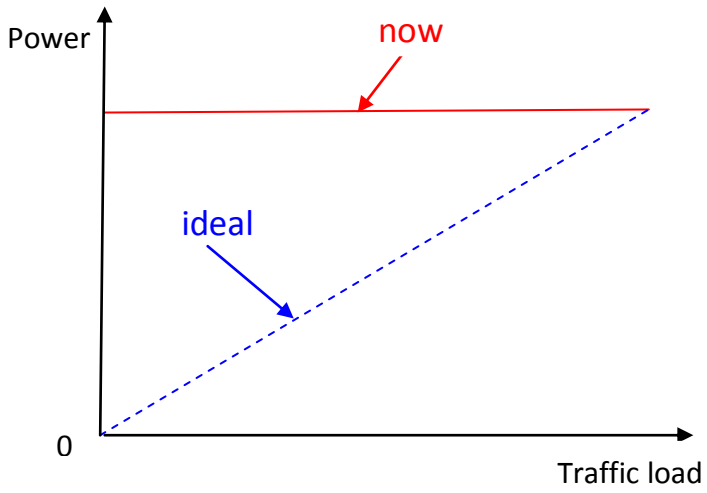

Figure 4.6: *Demonstration of traffic-dependent power control*

4.4.1 Possible Energy Reductions of Sleep Modes

Considering the base station power consumption in Figure 4.2 and the network traffic profiles in Figure 4.3 and Figure 4.4, we expect that employing sleep modes in LTE can significantly reduce energy consumption at low traffic loads. Generally, this can be achieved by gating off the base station periodically during off peak times, thus reducing the operational energy consumption. In real applications, since time-domain switching may introduce significant delay for data transmission, the maximum non-active time needs to be constrained according to any active user's QoS requirements.

Firstly, with sleep modes RF energy saving can be achieved by transmitting less signalling overhead. The LTE system downlink transmission is based on a frame transmission structure of duration 10 ms. With frequency division duplex (FDD), each frame consists of 10 subframes of 1 ms. Each subframe is further divided into two slots, each of 0.5 ms in duration. Each slot consists of 7 OFDM symbols if the short cyclic prefix is employed. This time-domain structure is shown in Figure 4.7.

In each subframe, except the user data intended for transmission, we need to deliver the corresponding control information, which includes a fixed amount of pilot signals, some broadcast

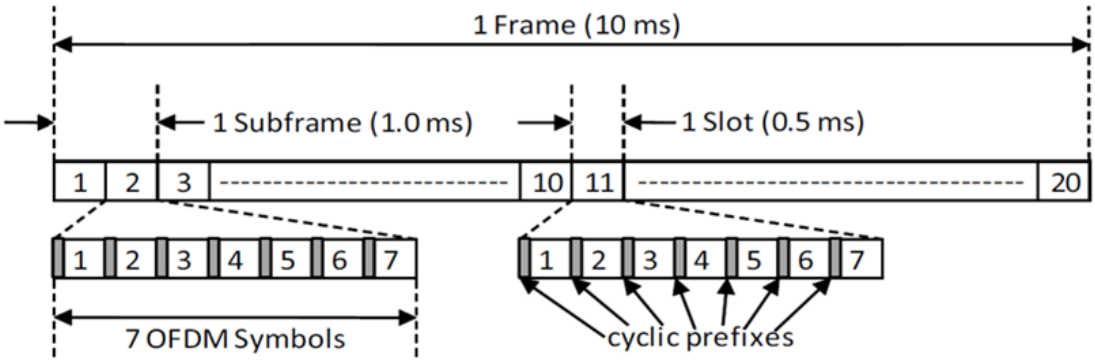

Figure 4.7: *LTE general frame structure [119]*

channel and synchronization signals, and a scalable amount of user-specific signalling. This control signalling overhead takes approximately 5-30% of the total subframe [120]. The remaining 70-95% of the subframe is available for supporting user data transmission as shown in Figure 4.8.

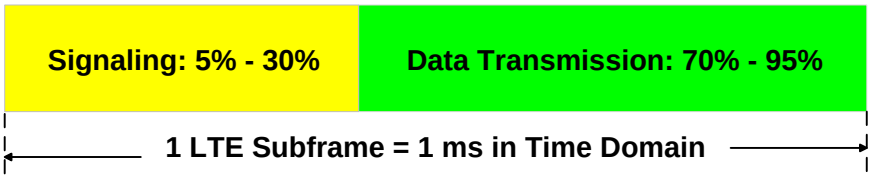

Figure 4.8: *LTE frame structure assumption for control signalling and data transmission*

In the first release of LTE, the control signalling is active in all subframes. Even if there is no user data for transmission during the current frame, the RBS still needs to deliver the control information of all its subframes, thus wasting energy. The idea of sleep modes for energy saving is to switch off some subframes and only let the active subframes carry user and signalling information. Since the control overhead in non-active subframes is not transmitted anymore, the energy consumption due to signalling will be reduced. In practice, if there are no active users in the cell, or the base station is serving users in discontinuous reception mode, the base station could decide to switch off the entire frames to further reduce energy consumption.

Generally, the more subframes we switch off, the less signalling bits we need to transmit, and the more energy we can save in the signalling portion. For example, in Figure 4.9, we use yellow colour to denote the resources used for control signalling, and green colour to mark the

resources used for data transmission. We assume that the duty cycle of the time-domain sleep mode is 50%. The load factor (LF) is defined as the quotient of serviced user data rate divided by the maximum allowed spectral efficiency under system power constraint. As shown in the figure, for the first case where LF=0%, we can save half of the energy consumption.

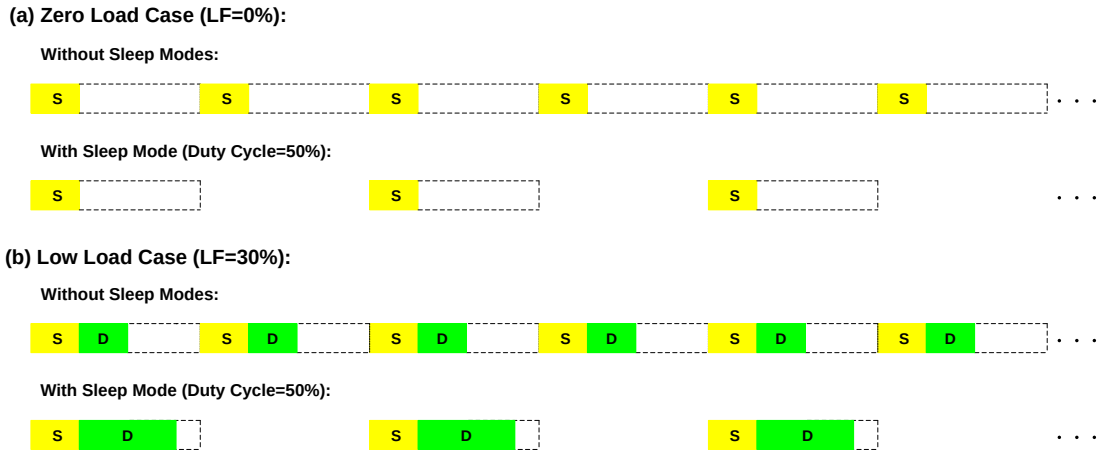

Figure 4.9: *Energy gains of time-domain base station sleep modes*

When there are higher traffic loads, such as LF=30% in Figure 4.9(b), it is still possible to reduce the energy consumption by switching off certain subframes and transmit all the required user data in the active subframes. Assuming a fixed amount of user data needs to be transmitted in each frame, there will be a lower bound on the active time per frame, which is mainly decided by the instantaneous channel quality, the system peak power constraint and the traffic load. Also, from implementation aspects, active user connections need to be maintained, and the maximum allowed delay is constrained in order to meet user's service requirements. According to information theory, with more subframes been switched off, the energy consumption in the data portion will increase due to the higher data rate per active packet. Within active subframes, a linear increase in data volume in the sleep mode case will lead to an exponential increase of power consumption due to Shannon equation. Therefore, the essence of optimizing the total RF energy saving by sleep modes lies on how to select the proper number of active/sleep subframes.

Besides the RF energy, the RBS operational energy consumption shown in Figure 4.2 typically refers to the energy used by the PA, DSP, radio and power supply. With the PA consuming a significant proportion of the total RBS power, we expect to reduce it by deploying sleep modes. Obviously, by switching off some subframes, the base station PA can also be gated off during

the inactive time. It will then transmit fuller data packets during the active time [134]. For DSP processing, if no signal is transmitted, the DSP processing for a given subframe can also be reduced significantly. Meanwhile, the user terminal can be informed that it can switch off the receiver DSP as well, further reducing power consumption. For the RBS radio, the base station DAC, user terminal ADC and RF receiver circuitry can also be switched off correspondingly, similar to the DRX mode already defined in LTE [100]. Here, we notice that although the power supply should be active all the time, and the energy consumption due to cooling systems and rectifier circuits might not be directly reduced through gating off the PA and DSP, knock-on energy savings can also be achieved by reducing RF energy consumption.

4.4.2 Sleep Modes with Resource Management

As we discussed above, if we aim to minimize the RF energy consumption, the optimum duty cycle of sleep mode can be determined by trading off the energy used by control signalling and user data transmission. According to the LTE frame structure shown in Figure 4.8, this can be achieved by switching off 1 to 10 subframes in each frame. We need to notice that since this sleep mode design is in time-domain, the idea can be applied to an LTE system with any bandwidth, although the size of the control signalling is bandwidth-dependent.

As we know, the LTE system is designed to meet carrier needs for high-speed data and media transport as well as high-capacity voice support well into the next decade. Its physical layer is a highly efficient means of conveying both data and control information between an enhanced base station and a user terminal [119]. In LTE, the minimum unit for resource allocation is the physical resource block, which is 1 subframe (1 ms) in time and 12 subcarriers (180 KHz) in frequency. Within each PRB, there are 14 OFDM symbols in time and 12 subcarriers in frequency. One subcarrier across 1 OFDM symbol is termed as 1 resource element (RE). In the frequency-domain, orthogonal frequency division multiple access is selected as an excellent multiplexing scheme in terms of efficiency and latency for LTE downlink transmission [119]. Users are allocated a specific number of physical resource blocks depending on the scheduling decision at the 3GPP base station.

For the sake of energy efficiency, we assume that full system bandwidth is always used. With each subcarrier bandwidth being 15 KHz, the total number of available subcarriers depends on the overall system bandwidth. The LTE specifications define parameters for system bandwidth from 1.25 MHz to 20 MHz, which correspond to 72 to 1200 subcarriers or 6 to 100 PRBs across

Parameters	Values
Carrier frequency f_{c}	5GHz
Cell diameter D	800m
Total bandwidth B	10MHz
Peak power P_{m}	40W
Operation temperature T_0	290 K

Table 4.2: *Model parameters for LTE downlink transmission*

the frequency domain. For a single user scenario, we simply allocate all the subcarriers to that user, thus the optimization is only in the time-domain by deciding the sleep mode duty cycle. For the multi-user case, different PRBs are assigned to different users according to certain time-frequency patterns depending on different scheduling strategies. Multi-user diversity can be obtained in the frequency-domain especially when frequency-selective fading is included in the analysis.

4.5 Basic Sleep Mode Modeling for SISO System

Assuming a simple single user SISO scenario, this section describes the analytical solution of sleep mode duty cycle optimization. Under different load factors and transceiver distances, numerical results suggest significant energy reduction gains.

4.5.1 Model Assumptions

To demonstrate sleep mode design, here we only consider a single user SISO case in an urban macro-cell environment with the parameters shown in Table 4.2. Using d to denote the distance between the base station and the user terminal, the path-loss L can be represented according to the WINNER II urban macro-cell channel model [43] as

$$L(\mathrm{dB}) = 26\log_{10}(d) + 39 + 20\log_{10}(f_{\mathrm{c}}/5). \tag{4.1}$$

And the corresponding channel gain due to path-loss is calculated as

$$G = 10^{\frac{-L}{10}} = 10^{\frac{-26\log_{10}(d)-39-20\log_{10}(f_{\mathrm{c}}/5)}{10}}. \tag{4.2}$$

For the specified operation temperature T_0, the noise density in the cell could also be obtained according to

$$N_0(\mathrm{W/Hz}) = KT_0 = 1.38 \times 10^{-23} \times 290 = 4.0 \times 10^{-21}. \tag{4.3}$$

We further assume a slow fading scenario with channel states changing every frame, where the fading follows the Rayleigh distribution with unit variance. By denoting the gain due to fading as h, the total channel gain can be represented as

$$g = Gh^2. \tag{4.4}$$

Regarding the signalling overhead, it basically includes reference signals for channel estimation and user-specific resource assignment. Other signals carrying system information are also need to be configured. Here the system and user-specific information can only be carried by the resource elements located in the control plane, which usually sit on the first 1 to 3 OFDM symbols in each subframe. For the SISO case, reference signals are inserted as shown in Figure 4.10.

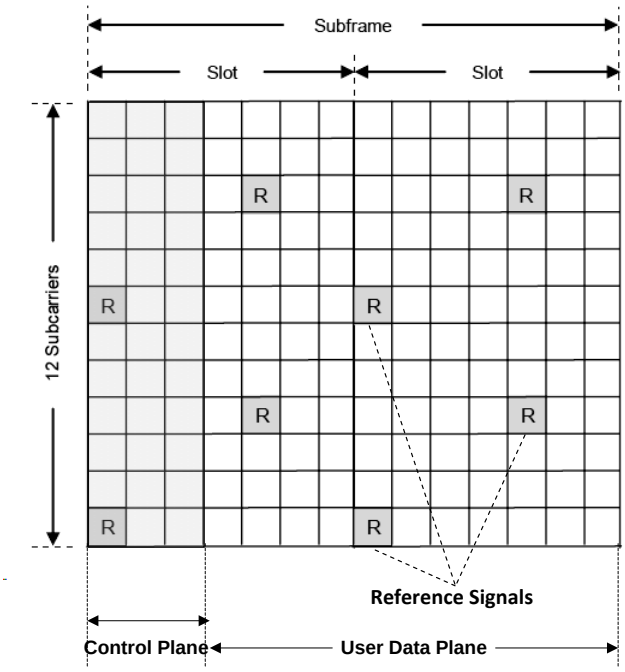

Figure 4.10: *DL control signalling assumptions [119]*

Using the above signalling pattern as an example, a fixed proportion of the control overhead can be calculated as $3/14 + 6/(11 \times 12) = 21.75\%$. However, we need to notice that this overhead estimation only provides us an upper bound. In practice, not all the REs in the control plane are used for signaling. Some of these resources remain blank. The size of the control overhead actually been delivered is decided by the number of scheduled users, number of transceiver antennas and instantaneous channel conditions. For the single user SISO case, here we assume a 10% control overhead, while the other 90% resources can be used for user data transmission.

4.5.2 On/Off Time Optimization

With full bandwidth transmission, we use $x_{\mathrm{s}} = 0.1$ to denote the proportion of time/frequency resources used for control overhead signalling in each subframe. We also assume that this signalling overhead is transmitted using the system peak power level P_{m} to guarantee reliable decoding over the cell area. Meanwhile, we would like to deliver a fixed amount of data in each frame which comprises 10 subframes. Without sleep modes, since all subframes within the same frame are assumed to have the same channel gain, they are assigned an equal proportion of user data for transmission.

According to the Shannon equation, if R_{d} bits/frame/Hz is delivered in the current frame using time t_{d} and power P_{d}, we can represent this as

$$R_{\mathrm{d}}(\mathrm{bits/frame/Hz}) = t_{\mathrm{d}} \log_2(1 + \frac{P_{\mathrm{d}} g}{N_0 B}). \tag{4.5}$$

Alternatively the required power for transmitting at data rate R_{d} can be written as

$$P_{\mathrm{d}}(\mathrm{W}) = \frac{(2^{\frac{R_{\mathrm{d}}}{t_{\mathrm{d}}}} - 1) N_0 B}{g}, \tag{4.6}$$

where $t_{\mathrm{d}} = (1 - x_{\mathrm{s}}) t_{\mathrm{f}}$ and $t_{\mathrm{f}} = 0.01$ s is defined as the frame length. Obviously, the time used for control signaling is $t_{\mathrm{s}} = x_{\mathrm{s}} t_{\mathrm{f}}$, and the RF energy consumed in each frame can be written as

$$E(\mathrm{J}) = P_{\mathrm{m}} t_{\mathrm{s}} + P_{\mathrm{d}} t_{\mathrm{d}}. \tag{4.7}$$

With sleep modes, we can optimally select the number of active subframes in each frame to reduce energy consumption. If there is no user data for transmission in that frame, we still

transmit one subframe of the control signalling to inform the receiver, which means that the number of active subframes per frame $N_{\rm act}$ can be selected between 1 and 10. Here, to transmit $R_{\rm d}$ bits/frame/Hz in each frame, the required power for data transmission in the active subframes becomes

$$P_{\rm ds}({\rm W}) = \frac{(2^{\frac{R_{\rm d}}{t_{\rm ds}}} - 1)N_0 B}{g}, \tag{4.8}$$

where $t_{\rm ds} = (1 - x_{\rm s})t_{\rm f}N_{\rm act}/10$ and $t_{\rm ss} = x_{\rm s}t_{\rm f}N_{\rm act}/10$. The RF energy consumption per frame can be calculated as

$$E_{\rm s}({\rm J}) = P_{\rm m}t_{\rm ss} + P_{\rm ds}t_{\rm ds}. \tag{4.9}$$

In order to minimize the RF energy consumption, we want to find the optimum number of active subframes $N_{\rm act} \in \{1, \cdots, 10\}$ in each radio frame. To do this we relax the above problem by searching for a continuous valued $N_{\rm act}$. We take the first and second order derivatives of $E_{\rm s}$ in terms of $N_{\rm act}$ as

$$\begin{cases} \frac{\partial E_{\rm s}}{\partial N_{\rm act}} = \frac{P_{\rm m}x_{\rm s}t_{\rm f}}{10} + \frac{(2^{\frac{R_{\rm d}}{N_{\rm act}(1-x_{\rm s})t_{\rm f}/10}} - 1)N_0 B t_{\rm f}(1-x_{\rm s})}{10g} - \frac{2^{\frac{R_{\rm d}}{N_{\rm act}(1-x_{\rm s})t_{\rm f}/10}} \ln 2 R_{\rm d} N_0 B}{N_{\rm act} g}, \\ \frac{\partial^2 E_{\rm s}}{\partial N_{\rm act}^2} = 2^{\frac{R_{\rm d}}{N_{\rm act}(1-x_{\rm s})t_{\rm f}/10}} \frac{10(\ln 2)^2 N_0 B R_{\rm d}^2}{(1-x_{\rm s})t_{\rm f} g N_{\rm act}^3} \geq 0. \end{cases} \tag{4.10}$$

This means $E_{\rm s}$ is a convex function in terms of $N_{\rm act}$, and there is a unique $N_{\rm opt}$ to minimize the energy consumption $E_{\rm s}$. By setting $\partial E_{\rm s}/\partial N_{\rm act} = 0$, the optimum number of active subframes per frame $N_{\rm opt}$ can be computed as

$$N_{\rm opt} = \frac{10 \ln 2 R_{\rm d}}{(1 - x_{\rm s})t_{\rm f}\left(1 + \mathbb{LW}\left[\frac{N_0 B x_{\rm s} - N_0 B + P_{\rm m} x_{\rm s} g}{N_0 B (1-x_{\rm s}) e}\right]\right)}, \tag{4.11}$$

where the lambert W-function $\mathbb{LW}[\ \ ]$ is defined as the inverse function of $\mathbb{F}[W] = We^W$. For sleep mode design, because $N_{\rm act}$ represents the number of active subframes per frame transmission, it should be an integer between 1 and 10. Regarding this, the convexity of the $E_{\rm s}$ function reveals that the practical number of active subframes per frame will be a rounding of $N_{\rm opt}$ either up or down. So evaluating the two nearest integer values of $N_{\rm act}$ to $N_{\rm opt}$ allows the minimum energy solution to be identified.

We need to note that the theoretical value of $N_{\rm opt}$ and the convexity of $E_{\rm s}$ only give us a pointer for selecting the number of active subframes in each frame. In practice, there are more constraints that need to be satisfied. Firstly, the system peak power constraint provides an upper limit on subframe capacity, indicating a lower bound $N_{\rm min}$ on $N_{\rm act}$ to transmit at rate $R_{\rm d}$. In

bad channel conditions, this may lead to $N_{\min} > 10$, thus we could not successfully deliver R_{d} within the current frame, causing the desired QoS to be violated. Secondly, the power level used for data transmission could not go beyond system peak power constraint. In good channel conditions, the optimum theoretical value N_{opt} may result in a higher P_{ds} than P_{m}.

4.5.3 Performance Analysis of SISO Sleep Modes

To investigate the performance of sleep modes for energy saving, we define the instantaneous data capacity as the maximum amount of user information that could possibly be delivered in the current frame, which is decided by the system peak power constraint and the channel gain as

$$R_{\mathrm{m}}(\mathrm{bits/frame/Hz}) = t_{\mathrm{f}} \log_2(1 + \frac{P_{\mathrm{m}} g}{N_0 B}). \tag{4.12}$$

This means that the instantaneous data capacity for one subframe is $R_{\mathrm{m}}/10$, and any data rate $R_{\mathrm{d}} > R_{\mathrm{m}}$ could not be effectively transmitted in the current frame. Also, the instantaneous data load factor can be interpreted as the proportion of required user information R_{d} divided by instantaneous data capacity R_{m}. According to the parameters in Table 4.2, we select a user with a distance of $d = 400$ m from the base station as an example, and plot its RF energy consumption E_{s} varying with the number of active subframes N_{act} under different load factors. In Figure 4.11, the frame we investigated is the sample selected from a Rayleigh process with $h = 1$ and $g = G = 10^{-\frac{L}{10}}$. Here we need to notice that all the parameters in the above equation take instantaneous values. Although the wireless channel varies over time, the following trends hold for all the channel realizations.

As the figure suggests, firstly, there is a minimum value for the number of possible active subframes under the specified load factor. This lower bound $N_{\min}$ increases with LF due to the limited subframe data capacity. In the case of LF=100% ($R_{\mathrm{d}} = R_{\mathrm{m}}$), R_{d} can only be transmitted using peak power P_{m} for all of the 10 subframes due to the lack of spare resources. Secondly, we notice that the RF energy consumption E_{s} is linearly increasing in terms of N_{act} when LF=0%. This is consistent with our signalling assumptions, where a fixed proportion of time x_{s} is assigned for control signaling, and there is no power adaptation for overhead transmission. Thirdly, the convex energy function E_{s} suggests that minimizing N_{act} does not necessarily minimize E_{s}, and increasing N_{act} can allow more energy-efficient coding and modulation schemes to be used. Although the energy consumed by control signalling increases with N_{act}, the energy used for delivering R_{d} decreases because of the relaxed delay constraints. Generally, the

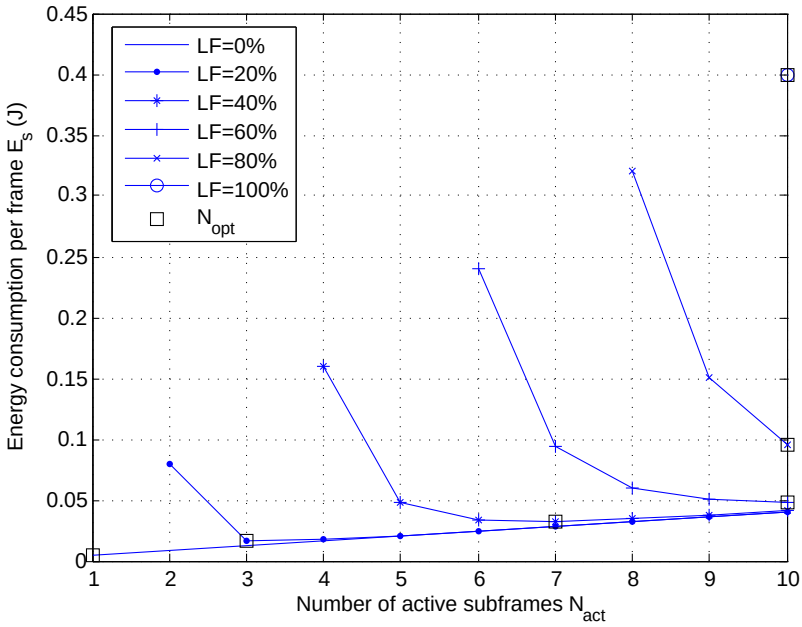

Figure 4.11: *Energy consumption for different numbers of active subframes*

optimum value of N_{act} can be rounded up or down from N_{opt} in the range of $1 \leq N_{\text{act}} \leq 10$.

Under different load factors and channel conditions, sleep mode protocols can select the optimum integer values of N_{act} between 1 and 10 to reduce RF energy consumption. For a long-term process, we further define the average load factor (ALF) by scaling the ergodic data capacity under a system peak power constraint, thus ALF=100% corresponds to the ergodic data capacity in Equation (4.13).

$$R_{\text{ergodic}}(P_{\text{m}}, g[\tau], B)(\text{bits/frame/Hz}) = \mathbb{E}\left[t_{\text{f}} \log_2(1 + \frac{P_{\text{m}} g[\tau]}{N_0 B})\right]. \tag{4.13}$$

.

In Figure 4.12, we compare the energy consumption of non-sleep mode and sleep mode for a cell-edge user with $d = 400$ m under different load factors. In this case, we need to notice that the required data rate is fixed in each frame, thus the idea applies to the applications with constant rate transmission.

With the increase of ALF, the required energy consumption increases for both non-sleep mode and sleep mode cases. Regarding sleep mode, since we assume at least one active subframe for signalling purposes, an 90% energy saving can be achieved at ALF=0%, but this energy

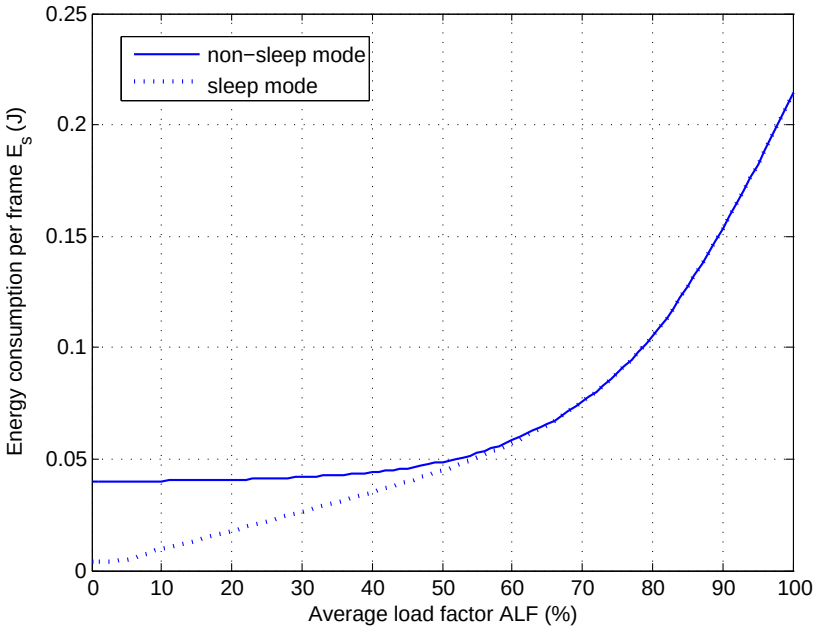

Figure 4.12: *Non-sleep/sleep mode energy consumption under different ALFs*

reduction gain decreases with the increased ALF. When the load factor goes beyond 60%, there are no obvious energy gains. This means that increasing data rates increases the required data transmit power, reducing the energy reduction gain.

To further illustrate the energy saving by sleep modes, the definition of ERG is applied as

$$\mathrm{ERG} = \frac{E - E_{\mathrm{s}}}{E} \times 100\%, \tag{4.14}$$

where E denotes the energy consumption of non-sleep mode, and E_{s} represents the energy consumption of sleep mode. According to the traffic profiles in Chapter 4.3, since the majority of cells carry very low or even no traffic especially during night time, the proposed time-domain sleep mode has great potential to reduce energy consumption in realistic scenarios. For example, Figure 4.5 suggests that at 6 am, 14% of the cells contribute to 90% of the total traffic, while 86% of the cells only contribute to 10% of the traffic. Thus on average each cell is 21.2% loaded. Even at the busiest time 4 pm, 38% of the cells carry 90% of the traffic, an average load factor of 40.4% can be estimated. Further, the figure indicates that at least 30% of the cells are usually empty, which allows us to shut down the base station entirely to save energy.

In Figure 4.13, under fixed average load factors, the properties of the energy reduction gain are investigated in terms of user's distances. Here, the required data rate R_{d} decreases with

the increased user distances, as ALF=100% always corresponds to the maximum achievable ergodic capacities under the peak power constraint P_{m}.

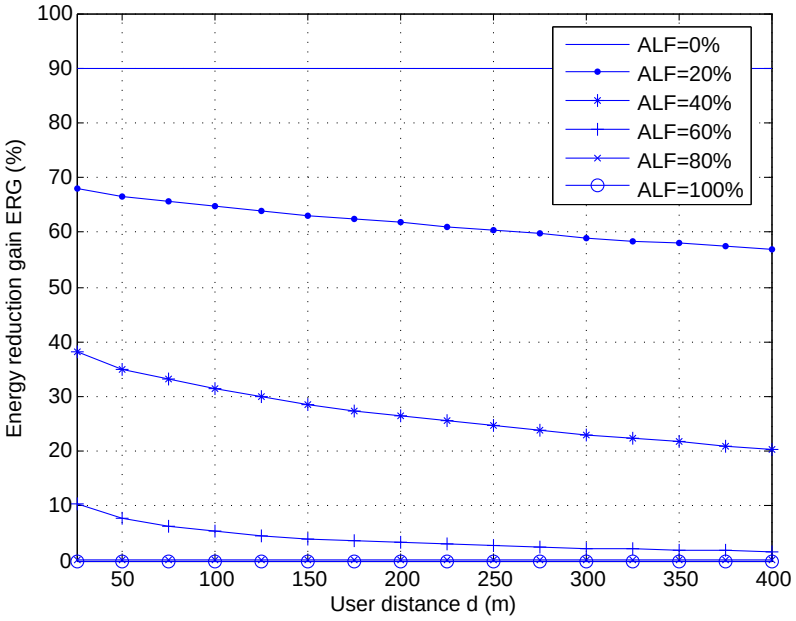

Figure 4.13: *Energy consumption under various user's locations with fixed ALF.*

From the above results, we can see that the time-domain sleep mode could effectively reduce RF energy consumption under low load conditions. Firstly, for any specified user's location, the energy reduction gain decreases with the increased ALF. The largest energy reduction gain is 90%, which corresponds to the zero-load case. Meanwhile, the energy saving performance of sleep modes is sensitive to the transceiver distances. While the data capacity generally decreases due to the degraded channel conditions, the energy reduction gain ERG decreases with d under any fixed ALF. Also, with ALF$\leq$ 60%, the energy reduction gain is obvious. Comparing these results with Figure 4.3, where one third of the time (2 am to 10 am) during a day is under 60% traffic load, the time-domain sleep modes are expected to achieve significant energy reduction in practice.

4.6 Summary

In this chapter, we have proposed an energy-efficient time-domain sleep mode design based on the LTE physical layer structure. According to the traffic profile, we adaptively selected

the on/off duty cycle in each frame. Compared to non-sleep modes, all user information needs to be transmitted during the limited active time, and the energy consumption can be reduced by discarding the signalling overhead during the inactive period. Besides, the energy used for ADC/DAC, DSP, rectifiers and cooling systems *etc* is expected to be conserved due to switching off RF transmission and base station hardware operation.

It is shown that the traffic load plays an important role in sleep mode design. Based on our model assumptions, up to 90% energy reduction can be obtained at low loads. While this gain decreases with the increased traffic, no significant energy can be saved when the load factor goes beyond 60%. Similarly, when the user moves from cell-center to cell-edge under the same data loads, the energy reduction gain of sleep modes decreases because of the reduced data capacity.

We need to note that increasing the sleeping time will not necessarily reduce the energy consumption. Although the RF energy of the signalling overhead can be reduced by this gating-off procedure, the available time resources for data transmission are also reduced, leading to an increased energy consumption for delivering user information. Thus, how to select the switching-off pattern is based on a tradeoff of these two factors. By taking the first order derivative of the total energy, we can find the optimum number of active/sleep subframes in each frame, which is primarily decided by the data rate requirements and instantaneous channel gains.

Although this chapter provides in-depth analysis of the time-domain sleep mode design, suggesting significant energy reduction gains, the examples we adopted for simulations are based on a simple SISO model with fixed control overhead size. In practice, multiple antennas are widely deployed, and the control overhead can be adapted according to the wireless channel variations. Also, there are more signalling issues we need to consider before switching off the base stations. This includes some system information and synchronization signals *etc*, which are essential for the other cells to be configured or operated. Therefore, which subframes are possible to be switched off needs to be evaluated carefully according to the current LTE standards. In Chapter 5, we will further discuss these LTE control signals and extend our current design to the MIMO scenarios.

Chapter 5
Sleep Modes with MIMO Systems

5.1 Introduction

In the previous chapter, we assumed a fixed size of control signalling overhead. In this chapter we further refine the model assumptions according to the LTE standards [125]. In practice, the size of control signalling overhead varies with system bandwidth, number of scheduled users, number of transmit antennas and instantaneous channel state information. Generally, the overhead size increases with the number of transmit antennas due to the extra pilots for channel estimation, and decreases with improved channel conditions since less control channel redundancy is required. How to adapt the control signalling overhead given the above parameters also depends on the network operator's decisions. In this chapter, we assume a single user case and provide a general estimation of signalling overhead size by summing up all the system signals and control channels.

While MIMO techniques are becoming widely deployed, it is also necessary to examine the performance of sleep modes in a MIMO environment. As extra circuit power and control signalling overhead are required in MIMO systems, switching off individual antenna ports can also reduce energy consumption. Regarding this, [126] has introduced the idea of smart MIMO for IEEE 802.11 wireless local area networks (WLAN), which demonstrates that different MIMO transmission schemes can be adopted according to the channel conditions to improve energy efficiency. In [127], a mechanism to switch between MIMO and SIMO is proposed to conserve mobile terminals' energy as well. Apart from these, the documents from Huawei [128] and the presentation from Vodafone [129] also suggest that gating off antennas could be an effective solution for Green Communications.

Here we still focus on optimizing RF energy consumption. According to the LTE physical layer structure, MIMO requires a larger control overhead due to the increased number of spatial channels, thus it will consume more energy for control signalling. However, MIMO spatial multiplexing potentially saves energy for user data transmission. A spatial-domain optimization

problem is introduced based on this tradeoff. Combining with the sleep modes in the time-domain, it is also useful to find out the most energy-efficient strategy by gating off certain number of subframes and antenna ports together [15].

In this chapter, Section 5.2 will provide in-depth analysis for properly estimating the LTE control overhead size under different channel conditions and antenna configurations for single user system. After that, Section 5.3 will introduce the idea of switching off a number of transmit antennas in the spatial-domain to save energy consumption. Based on the overhead estimation in the previous section, simulation results will be provided to verify the energy reduction of the spatial-domain sleep mode and spatial-time sleep mode. In Section 5.4, some practical issues of sleep modes implementation will be further discussed. When multiple users are scheduled at the same time, the energy reduction gain also depends on the adopted scheduling strategies. In the LTE downlink, frequency-domain scheduling is widely accepted based on OFDMA [99]. Since the control signalling overhead contributes to a large portion of the total RF energy consumption, considering the energy consumption including both control signalling and user data transmission together is very important in designing energy-efficient scheduling schemes [17]. New optimization objectives may be proposed for building scheduling decision functions. Here, we will also investigate how to incorporate sleep modes into frequency-domain scheduling schemes, and the possibility of fitting sleep modes into voice over IP (VoIP) protocol and legacy handsets.

5.2 LTE Downlink Transmission

In order to design the spatial-domain sleep mode based on the energy tradeoff between control signalling and user data transmission, in this section we aim to further investigate the LTE control signalling process to provide an accurate estimation of control overhead size in different scenarios. Assuming only one user been scheduled, the two main factors which dominate this are related to the number of transmit antennas and channel conditions. Starting with analyzing the LTE physical layer structure, firstly we introduce and describe various control and data channels. Then a reasonable assumption of control signalling overhead is described according to the documented signaling process. The factors affecting this overhead size generally include system bandwidth, number of scheduled users, number of transmit antennas, instantaneous channel state information *etc*. User data transmission and other issues regarding resource management are also addressed according to LTE standards.

5.2.1 LTE Physical Layer Structure

In LTE systems, the physical resources used for control signalling are not fixed. Unlike Figure 4.8, the size of control signalling overhead x_{s} depends on the system bandwidth, number of transmit antennas, number of scheduled users, instantaneous channel conditions *etc*. In order to evaluate x_{s}, firstly we de-construct the LTE control channels within 1 radio subframe. Taking a 1.25 MHz system as an example, Figure 5.1 provides the general control signalling structure of the LTE physical layer with 6 PRBs.

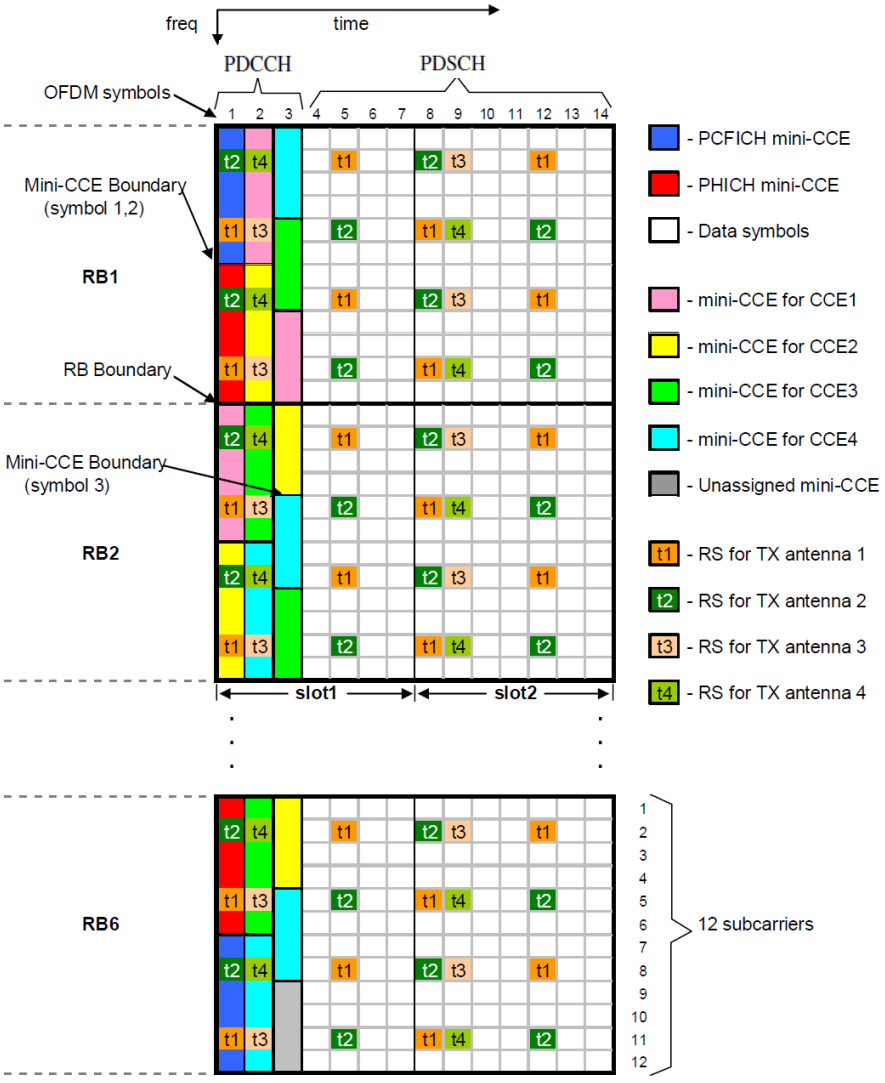

Figure 5.1: *LTE downlink control signalling structure [122]*

As the figure suggests, each resource block contains 14 OFDM symbols (1 ms) in the time-

domain and 12 subcarriers (180 KHz) in the frequency-domain. Thus there are 168 resource elements in each resource block. Within each PRB, the main components are the RSs, the physical downlink control channel (PDCCH) and the physical downlink shared data channel (PDSCH). Here we assume that there are 4 antennas equipped at the transmitter side. We try to schedule the maximum number of users at the same time.

Regarding the reference signals for this MIMO system, there are 24 REs equally distributed in each PRB. They are used as pilot signals for channel estimation. The RSs exist in all PRBs, occupying $24/(14\times12) = 14.4\%$ physical resources [120]. Regarding the MIMO systems with fewer transmit antennas, this figure decreases because there are fewer spatial channels between different transceivers for channel estimation. When two antennas are used for transmission, the RS overhead becomes 16 REs evenly located in each PRB. For SISO system, there are $8/(14 \times 12) = 4.8\%$ resources used on reference signals [119].

The PDCCH stands for the physical downlink control channel. It is used for carrying user-specific resource assignment, thus each scheduled user is allocated 1 PDCCH [122]. The PDCCH can only be located on the free time-frequency resources within the control plane, which refer to the REs which are not used for reference signals and other control channels within the first 1 to 3 OFDM symbols across all the subcarriers in each subframe [41]. Using either 1, 2 or 3 OFDM symbols for the control plane is defined by the physical control format indication channel (PCFICH). In Figure 5.1, we assign 3 OFDM symbols as the PDCCH region. In this region, every 4 REs are grouped as resource element group (REG), and every control channel element (CCE) consists of 9 REGs. Each PDCCH can be composed from 1, 2, 4 or 8 CCEs. We need to notice that the limited control plane provides an upper bound on the maximum number of users being scheduled. This is related to the size of each user's PDCCH, which further depends on the instantaneous channel conditions.

For the remaining 11 to 13 OFDM symbols, the available resources are used for delivering user data. This is called the physical downlink shared data channel (PDSCH). For data transmission, QPSK, 16-QAM and 64-QAM are the supported modulation formats in LTE [41]. Together with time-frequency resource allocation, power control is used to transmit the coded and modulated data symbols for achieving specified user QoS requirements.

Apart from the three components discussed above, the LTE physical layer structure also defines resources for synchronization signals, the physical broadcast channel (PBCH), the physical

control format indicator channel (PCFICH), the physical HARQ indicator channel (PHICH), the physical multicast channel (PMCH) *etc* [120]. Instead of existing in each PRB, these signals or channels are only mapped to certain OFDM symbols or subcarriers according to a predetermined pattern. Therefore, to estimate the control overhead size x_{s}, we need to add their contribution and average over the whole system bandwidth and time. In general, the control signalling overhead includes the following components [41]:

- Reference signal: used for channel estimation
- PDCCH: carries user-specific resource assignment
- PCFICH: defines how many OFDM symbols (1, 2 or 3) are used for control channel in each subframe
- PHICH: carries HARQ ACK/NACK to indicate whether the eNodeB has correctly received a transmission on uplink
- PBCH: carries basic system information which allows the other channels in the cell to be configured and operated
- Synchronization signal: cell search and synchronization.

Although all the above signals and channels are necessary for signalling purposes, the main part of the control overhead is still contributed by RSs and PDCCHs, which are delivered using QPSK modulation. The PMCH is used for radio multicast, which sits on the user data plane together with the PDSCH. Here we omit the details about these physical channels but focus on calculating a reasonable value of x_{s} for more accurate sleep mode design.

5.2.2 LTE Control Signalling Process

In order to estimate LTE control signalling overhead, we need to sum up the resources occupied by all the control channels together. As discussed above, the size of reference signals depends on the number of transmit antennas. As the PCFICH, PHICH, PBCH and synchronization signals only exist in certain subcarriers, their size depends on the total system bandwidth. The contribution of PDCCH mainly comes from:

- Number of PDCCHs: depends on how many users being currently scheduled

- Individual PDCCH size: depends on the instantaneous channel conditions.

Here we assume a fixed $B = 10$ MHz system with only 1 user being scheduled. We aim to study how the PDCCH size is adapted to the instantaneous channel conditions. For the number of transceiver antennas, we assume SISO, 2×2 MIMO and 4×4 MIMO are available for comparison. As we know, Shannon capacity always provides the upper-bound for the information rate at which we can transmit. According to the LTE standards, the Alamouti scheme is used for precoding the control signalling, thus the instantaneous channel capacity of the signalling channels for SISO and 2×2 MIMO can be written as [123]

$$C_{N\times M} = \log_2\left(1 + \frac{\bar{\rho}}{M}\sum_{m=1}^{M}\sum_{n=1}^{N}|h_{nm}|^2\right), \tag{5.1}$$

where M denotes the number of transmit antennas and N denotes the number of receiver antennas. $\bar{\rho}$ is the average receiver side SINR with $\bar{\rho} = P_{\mathrm{m}}G/(N_0B + I_{\mathrm{int}})$ for full power transmission. The scalar h_{nm} is the channel gain due to Rayleigh fading, which is a random variable with unit variance.

For the 4×4 MIMO case, the precoding format is slightly changed. Taking a block of modulated symbols $\{x_1, x_2, x_3, x_4\}$ from the layer mapper at time instances τ and $\tau + 2$, two consecutive symbols are transmitted in parallel in two symbol periods using four antennas with the mapping pattern shown in Figure 5.2, where $[\quad]^*$ denotes the complex conjugate operation. Correspondingly, the instantaneous capacity for the signalling channels can be written as

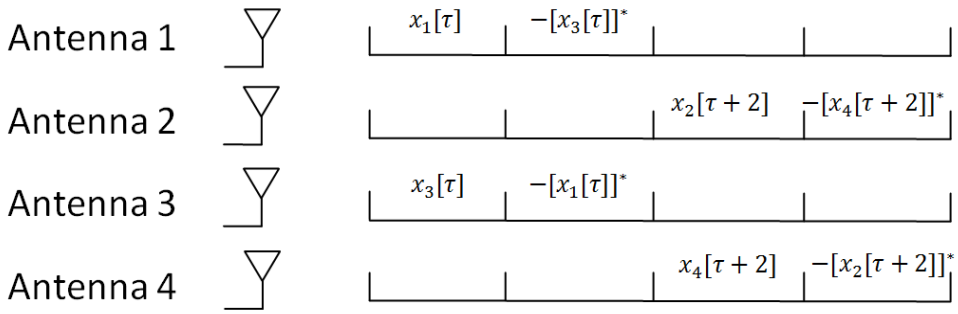

Figure 5.2: *Antenna port precoding for* 4×4 *MIMO [124]*

$$C_{4\times4} = \left[\log_2\left(1 + \frac{\bar{\rho}}{4}\sum_{m=1}^{M=2}\sum_{n=1}^{N=4}|h_{nm}|^2\right) + \log_2(1 + \frac{\bar{\rho}}{4}\sum_{m=3}^{M=4}\sum_{n=1}^{N=4}|h_{nm}|^2)\right]/2. \tag{5.2}$$

Regarding the information bits which the PDCCH actually carries, it is called the downlink control information (DCI) [125]. The DCI size depends on the selected DCI format, antenna configuration, system bandwidth *etc*. In general, DCI format 1 is used for the scheduling of single PDSCH codeword, and DCI format 2 is used for scheduling PDSCH to the UEs configured in the closed-loop spatial multiplexing mode.

For example, in a single user SISO system with $B = 10$ MHz and DCI format 1, 47 bits are required to represent the downlink control information. These 47 bits further decide how much resources we need to allocate to that PDCCH. In LTE, the so-called CCE aggregation level α is defined, representing the control channel coding redundancy. Generally, with bad channel conditions higher α values are required for the sake of reliable delivery of the control information. The value of α can be selected as 1,2,4 or 8, which means each PDCCH can contain 1, 2, 4 or 8 control channel elements, while each CCE consists of 36 REs [130]. Choosing the PDCCH size according to the channel condition is called CCE aggregation. Using QPSK with convolutional coding to deliver control signalling, we can determine the channel coding rate as follows:

$$\theta = \frac{\pi}{\alpha\eta \log_2 4} = \begin{cases} \frac{47}{1\times 36\times \log_2 4} \approx \frac{2}{3} & \text{if } \alpha = 1, \\ \frac{47}{2\times 36\times \log_2 4} \approx \frac{1}{3} & \text{if } \alpha = 2, \\ \frac{47}{4\times 36\times \log_2 4} \approx \frac{1}{6} & \text{if } \alpha = 4, \\ \frac{47}{8\times 36\times \log_2 4} \approx \frac{1}{12} & \text{if } \alpha = 8. \end{cases} \tag{5.3}$$

In Equation (5.3), π represents the number of DCI bits, α is the CCE aggregation level, and $\eta = 36$ means each CCE has 36 REs. Because each RE carries $\log_2 4 = 2$ bits according to QPSK modulation, each CCE contains 36×2 bits. Therefore, the coding rate θ can be calculated as the quotient of the information bits π divided by the total bits transmitted using the available time-frequency resources.

In the following context, we take the coding rate set $\{2/3, 1/3, 1/6, 1/12\}$ derived from Equation (5.3) as an example to illustrate the control signalling process of the 10 MHz SISO system with DCI format 1. Using the Shannon equation, we can determine the CCE aggregation levels based on these coding rates, since the successfully transmitted information bits are always

smaller than the theoretical values provided by the Shannon capacity

$$C_{N\times M} = \log_2\left(1+\frac{\bar{\rho}}{M}\sum_{m=1}^{M}\sum_{n=1}^{N}|h_{nm}|^2\right) > \begin{cases} \frac{\log_2 4\times 2/3}{\delta T\delta f} = \frac{56}{45} & \alpha = 1,2,4,8, \\ \frac{\log_2 4\times 1/3}{\delta T\delta f} = \frac{28}{45} & \alpha = 2,4,8, \\ \frac{\log_2 4\times 1/6}{\delta T\delta f} = \frac{14}{45} & \alpha = 4,8, \\ \frac{\log_2 4\times 1/12}{\delta T\delta f} = \frac{7}{45} & \alpha = 8. \end{cases} \tag{5.4}$$

Here the unit of $C_{N\times M}$ is bps/Hz, while $\delta T = \frac{1}{14}\times 10^{-3}$ s denotes the time span of 1 OFDM symbol and $\delta f = 15$ KHz denotes the bandwidth of 1 subcarrier. In order to use the PDCCH efficiently, we can adaptively select the lowest CCE aggregation level α satisfying the Shannon capacity in each channel state.

For simulations, we further assume that the user's location follows a uniform distribution between 40 and 400 m from the base station, while the interference comes from its neighboring cell with peak power transmission on the same bandwidth. According to Equation (5.1) and Equation (5.2), we can calculate the capacity of the signalling channels for each SISO/MIMO case. Figure 5.3 (a) plots the ergodic capacity over 1000 runs at each user's location. Firstly, the channel capacity increases with SINR for all SISO/MIMO cases. Secondly, MIMO outperforms SISO, while 4×4 MIMO achieves slightly better performance than 2×2 MIMO.

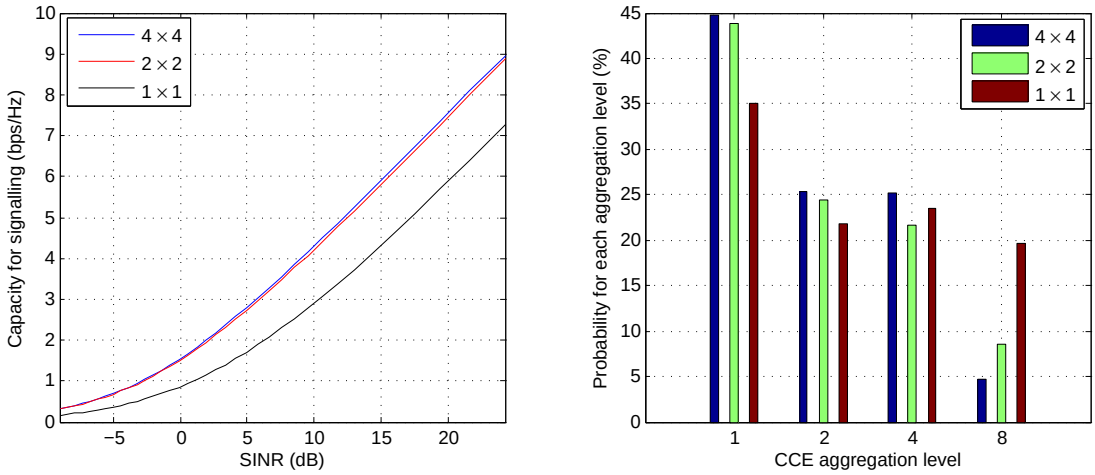

Figure 5.3: *CCE aggregation for PDCCH (a) Ergodic capacity of LTE control signalling channels (b) Probabilities of different CCE aggregation levels for PDCCH*

In Figure 5.3 (b), the probabilities of selecting different CCE aggregation levels under different channel conditions in a long-term process are plotted. In general, the figure suggests that SISO requires a higher CCE aggregation level than MIMO. This is as expected according to our CCE aggregation process in Equation (5.4). Because MIMO achieves higher capacity than SISO with the Alamouti scheme shown in Figure 5.2, the required SINR for supporting any convolutional coding rate $\{2/3, 1/3, 1/6, 1/12\}$ is lower for the MIMO case. Therefore, more channel realizations in Figure 5.3 (b) can achieve relatively lower CCE aggregation levels. Although MIMO requires more resources for reference signals, the required redundancy for coding the PDCCH is generally smaller compared to SISO.

5.2.3 LTE Control Overhead Resizing

Based on the above discussions, we can see that the size of control signalling overhead varies with different antenna configurations and channel conditions. Assuming a 10 MHz system with single user being scheduled, we can estimate the size of control overhead by summing up the resources occupied by different control signals and control channels [131].

- Reference signal:
  SISO system: 8 REs per PRB contributing to $\frac{8}{12\times14} = 4.8\%$ overhead;
  2×2 MIMO: 16 REs per PRB contributing to $\frac{16}{12\times14} = 9.6\%$ overhead;
  4×4 MIMO: 24 REs per PRB contributing to $\frac{24}{12\times14} = 14.4\%$ overhead.

- PDCCH:
  $\alpha = 1$: 36 REs per subframe contributing to $\frac{36}{12\times14\times50} = 0.43\%$ overhead;
  $\alpha = 2$: 72 REs per subframe contributing to $\frac{72}{12\times14\times50} = 0.86\%$ overhead;
  $\alpha = 4$: 144 REs per subframe contributing to $\frac{144}{12\times14\times50} = 1.72\%$ overhead;
  $\alpha = 8$: 288 REs per subframe contributing to $\frac{288}{12\times14\times50} = 3.43\%$ overhead.

- PCFICH:
  16 REs mapped to the first OFDM symbol over full system bandwidth in each subframe, which contributes to $\frac{16}{12\times14\times50} = 0.19\%$ overhead.

- PHICH:
  With the smallest scaling factor $1/6 \in \{1/6, 1/2, 1, 2\}$ for controlling the number of PHICH groups, 2 PHICH groups containing $2 \times 8 \times 1 = 16$ REs will contribute to $\frac{2\times8\times1}{12\times14\times50} = 0.19\%$ control overhead [132].

Parameters	$\alpha = 1$	$\alpha = 2$	$\alpha = 4$	$\alpha = 8$
1×1	6.20%	6.73%	7.48%	9.20%
2×2	11.00%	11.53%	12.28%	14.00%
4×4	15.80%	16.33%	17.08%	18.80%

Table 5.1: *Estimated size of control signalling overhead for single user case*

- PBCH:
  40 REs per PRB mapped to the central 72 subcarriers (6 PRBs) during the first subframe per frame, which contributes to $\frac{40\times6}{12\times14\times50\times10} = 0.29\%$ overhead.
- Synchronization signals:
  Primary and secondary synchronization signals are mapped to the central 62 subcarriers during the 6th and 7th OFDM symbols in subframe 1 and subframe 6 per frame, which contributes to $\frac{62\times2\times2}{12\times14\times50\times10} = 0.3\%$ overhead.

Therefore, the estimated size of control signalling overhead for the single user case is summarized in Table 5.1.

In Figure 5.4, we simulated the average control overhead size changing with transceiver distance over 5 million runs according to the assumptions in Section 4.5.1. As expected, the average control signalling overhead increases with the degraded channel conditions, due to the higher probabilities of requiring large CCE aggregation levels. Meanwhile, the major contributor of control overhead is reference signals, thus using more transmit antennas will in general require larger control signalling.

Further, we need to notice that the control signalling overhead estimates are dependent on the total system bandwidth. With different bandwidth, we still need to transmit the same amount of PCFICH, PBCH and synchronization signals [41]. However, the total number of REs in the system varies. Using N_{PRB} to denote the number of PRBs in each subframe, we can calculate the overhead size of PCFICH, PHICH, PBCH and synchronization signals as

$$2 \times \frac{16}{12 \times 14 \times N_{\mathrm{PRB}}} + \frac{40 \times 6}{12 \times 14 \times N_{\mathrm{PRB}} \times 10} + \frac{62 \times 2 \times 2}{12 \times 14 \times N_{\mathrm{PRB}} \times 10} = \frac{0.4810}{N_{\mathrm{PRB}}}. \quad (5.5)$$

Similarly, the overhead size due to one PDCCH can be represented as

$$\frac{36 \times \alpha}{12 \times 14 \times N_{\mathrm{PRB}}} = \frac{0.2143 \times \alpha}{N_{\mathrm{PRB}}}. \quad (5.6)$$

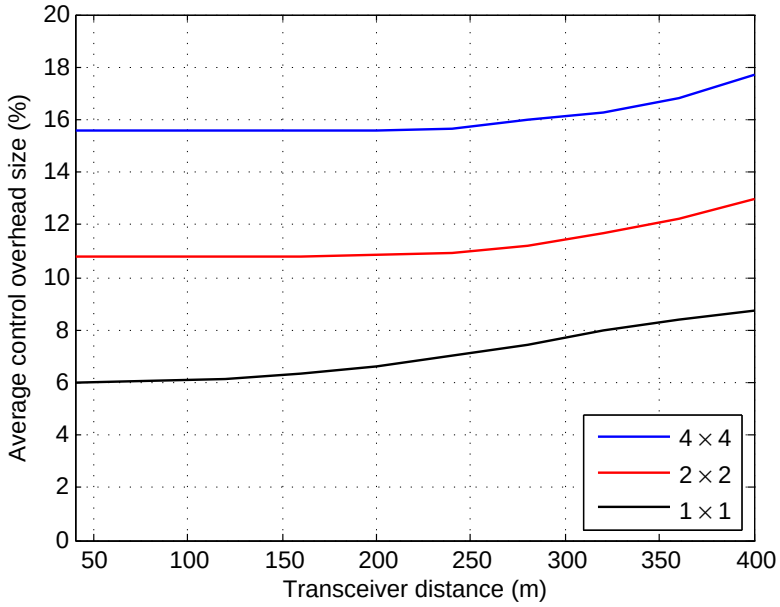

Figure 5.4: *Averaged control overhead size under specific user distance*

For a single user SISO system, we can use Equation (5.5) and Equation (5.6) to evaluate how the control overhead size changes with system bandwidth in Figure 5.5. As the figure suggests, the percentage of control signalling decreases with expanded bandwidth for each specific CCE aggregation level. This trend still holds for the MIMO systems, since the amount of their extra reference signals are proportional to the number of increased physical resource blocks.

Besides system bandwidth, another major factor which may affect the estimated overhead size shown in Table 5.1 is the number of users being scheduled. However, this is dependent on the scheduler's decisions, which is out of scope of our discussion. As mentioned in Section 5.2.1, the maximum number of users being scheduled is limited by the PDCCH control region, which is 1 to 3 OFDM symbols per subframe indicated by PCFICH [125].

Other issues which may also affect the estimated values of control signalling are listed as follows. However, these effects are generally minor. The values shown in Table 5.1 can be treated as a reasonable estimate for the single user case in a $B = 10$ MHz system.

- PHICH: PHICH is used for delivering Hybrid-ARQ acknowledgements in response to uplink transmission. The number of required PHICHs is decided by the uplink traffic.
- CCE aggregation for PDCCH: The calculated convolutional coding rates in Section 5.3.2 are based on a SISO system with 10 MHz bandwidth in DCI format 1. For MIMO, the

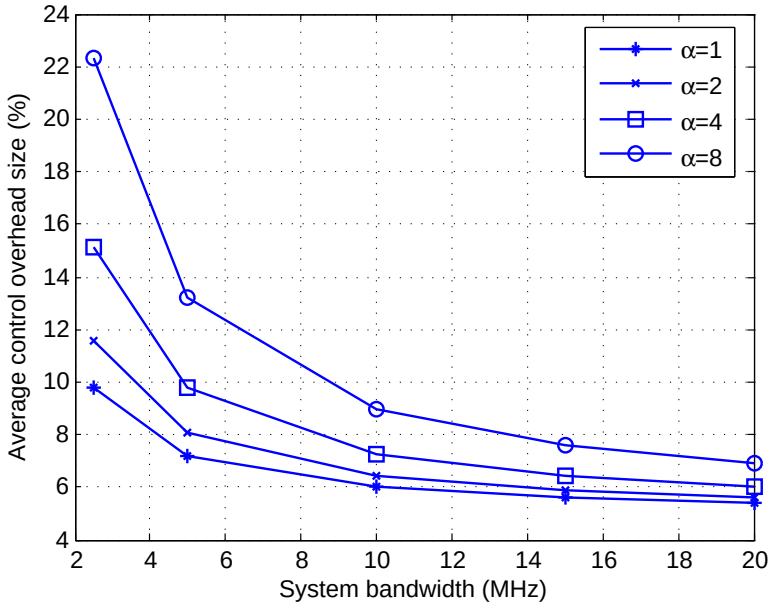

Figure 5.5: *Control overhead size under different system bandwidth for SISO*

DCI size is generally larger, which may lead to different CCE aggregation levels.

5.3 Sleep Modes in a MIMO Environment

In this section, we further extend the sleep mode design in Chapter 4 to a MIMO configuration. Besides the time-domain optimization, we can switch off a number of antennas to save energy in the spatial-domain. Although using more antenna ports can deliver user data more efficiently based on MIMO spatial multiplexing, a larger control signalling overhead is required which will consume more RF energy.

5.3.1 Optimization Issues for MIMO Sleep Modes

Section 5.2 not only provides more details for control signalling in LTE, but also considers the impacts of using MIMO. Nowadays, MIMO techniques are widely used in wireless communication systems, since they can effectively improve system spectral efficiency based on spatial multiplexing [133].

To start with, we ignore the control signalling issues and only focus on user data transmission. The achieved capacity of SISO, 2×2 MIMO and 4×4 MIMO under the same sum power

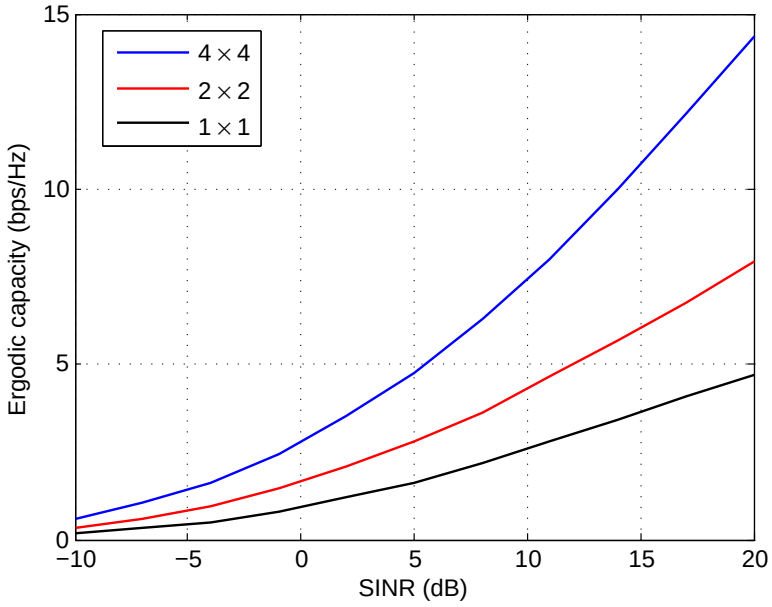

Figure 5.6: *Spectral efficiency of MIMO spatial multiplexing*

constraint can be represented as

$$C_{\text{ergodic},N\times M}(\bar{\rho},\mathbf{H}[\tau]) = \mathbb{E}\left[\log_2 \det\left(\mathbf{I}_N + \frac{\bar{\rho}}{M}\mathbf{H}\mathbf{H}^*\right)\right], \tag{5.7}$$

where N and M denote the total number of receiver and transmit antennas. The function $\mathbb{E}[\ \ ]$ denotes the expectation operation. The scalar $\bar{\rho}$ is the average receiver SINR with $\bar{\rho} = P_{\text{m}}G/(N_0B + I_{\text{int}})$, and $\mathbf{I}_N$ is the $N \times N$ identity matrix. The notation $\mathbf{H}$ is an $N \times M$ matrix with each of its element h_{nm} representing the channel gain amplitude due to Rayleigh fading from transmit antenna m to receiver antenna n as

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1M} \\ h_{21} & h_{22} & \dots & h_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ h_{N1} & h_{N2} & \dots & h_{NM} \end{bmatrix}. \tag{5.8}$$

According to the parameters shown in Table 3.1, simulated results are obtained in Figure 5.6. Firstly, the achieved capacity increases with system SINR for all antenna configurations, suggesting higher spectral efficiency can be obtained under better channel conditions. Secondly, for each specific SINR value, using more MIMO antenna ports can achieve better performance.

For example, when SINR=5 dB, we can use the system peak power 40 W to deliver about 2 bps/Hz data in SISO system, 3 bps/Hz data in 2×2 MIMO system and 4.8 bps/Hz data in 4×4 MIMO system. This also indicates that in order to achieve a data rate target, using more antenna ports is more energy-efficient. For instance, when SINR=5 dB, if we aim to deliver 4 bps/Hz we can only use 4×4 MIMO. If we aim to transmit 2 bps/Hz, 40 W power is consumed for the SISO case while much less power is required for MIMO.

Generally, the above discussions suggest that MIMO can save energy consumption for delivering a fixed amount of user data compared to SISO. However, with more transmit antennas, more energy is required for the signalling purposes due to the larger control overhead size. Figure 5.7 illustrates this tradeoff from an energy efficiency point of view.

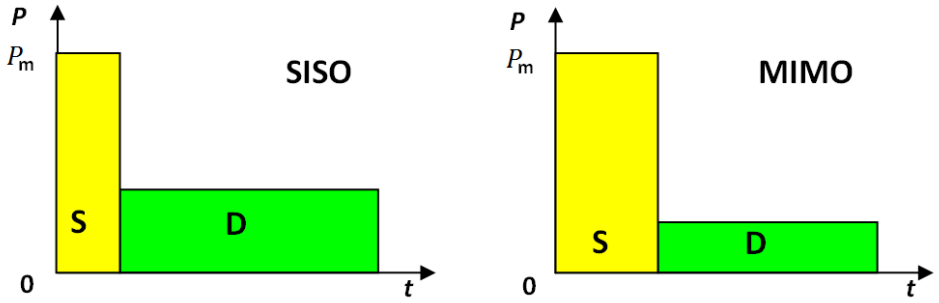

Figure 5.7: *Comparisons of energy consumption between SISO and MIMO*

For both SISO and MIMO, we use the same peak power to transmit the control overhead. With a larger overhead size in MIMO, more energy is consumed for delivering these extra reference signals. Regarding the user data transmission, less power is required in the MIMO case for achieving a fixed data rate requirement. The energy consumption for delivering user data is reduced as well. Therefore, a tradeoff exists on how to select the most energy-efficient antenna configurations, if the total energy reduction including both control signalling and user data is targeted.

If we ignore the hardware constraints on radio antennas and signal processing components, the above tradeoff suggests that in a MIMO system switching off individual radio antennas could improve energy efficiency. Thus the sleep mode design in MIMO environments can include two dimensions: one is gating off radio subframes in the time-domain, the other is switching off individual radio antennas in the spatial-domain.

5.3.2 Model Assumptions

In order to design MIMO sleep modes, the following assumptions are made based on the model described in Section 4.5.1.

- Available antenna configurations include SISO, 2×2 MIMO and 4×4 MIMO.
- Use Alamouti space-time block coding to deliver MIMO control signalling.
- Use spatial multiplexing to transmit MIMO user data.
- The peak power constraint applies to the sum of transmit antenna energy consumptions.
- Hardware constraints on antenna PAs and signal processing components are ignored.
- Adapt control overhead size to instantaneous channel conditions with lowest coding redundancy according to Equation (5.4). Select the proper antenna configuration and CCE aggregation level according to Table 5.1.
- Maximum spectral efficiency is constrained by the available LTE coding and modulation schemes for each SISO/MIMO case.

Here we need to notice that with MIMO spatial multiplexing, the waterfilling algorithm is used to allocate power across different MIMO antennas in each channel state. Instead of obtaining an analytical solution of N_{opt} as described in Section 4.5.2, we use numerical methods to search for the optimum number of active subframes in each frame for the time-domain optimization.

Regarding the spatial-domain, we aim to identify the boundary where equal energy is consumed under different antenna configurations. For instance, if we use SISO and 2×2 MIMO as examples, 2×2 MIMO will consume more energy for delivering its larger control overhead, but save energy for transmitting the same amount of user data. In mathematical notation, this equality can be written as

$$E_{\mathrm{x},2\times2} - E_{\mathrm{x},1\times1} = E_{\mathrm{d},1\times1} - E_{\mathrm{d},2\times2}, \tag{5.9}$$

where $E_{\mathrm{x},1\times1}$ and $E_{\mathrm{x},2\times2}$ denote the control signalling energy consumption for the SISO and 2×2 MIMO cases respectively; $E_{\mathrm{d},1\times1}$ and $E_{\mathrm{d},2\times2}$ denote the user data energy consumption for the SISO and 2×2 MIMO cases respectively.

Under specified channel conditions, the left side of Equation (5.9) is a constant, since the size of the control signalling can be uniquely determined according to our assumptions. This control signalling is always transmitted using system peak power, thus leading to a fixed amount of energy consumption. Regarding the right side of the equation, the energy consumption can be calculated numerically according to waterfilling power allocation scheme if a specific data rate is targeted. Alternatively, we can say that the energy consumption is uniquely decided by the user data rate requirements under the current channel conditions.

Therefore, the problem becomes identifying the data rate $R_{\mathrm{eq},11/22}$ bps/Hz with which SISO and 2×2 MIMO consume the same amount of energy for the current channel condition. This can be written as

$$\begin{cases} P_{\mathrm{m}} t_{\mathrm{f}} (x_{\mathrm{s},2\times2} - x_{\mathrm{s},1\times1}) = t_{\mathrm{f}} \left[\frac{(2^{R_{\mathrm{eq},11/22}} - 1)(N_0 B + I_{\mathrm{int}})(1 - x_{\mathrm{s},1\times1})}{h_{1\times1}^2 G} - P_{\mathrm{d},2\times2}(1 - x_{\mathrm{s},2\times2}) \right] \\ P_{\mathrm{d},2\times2} = \mathbb{F}\left[R_{\mathrm{eq},11/22} \right], \end{cases} \tag{5.10}$$

where $\mathbb{F}[\ \ ]$ represents that the power consumption of the 2×2 MIMO case is a function of the required sum data rate $R_{\mathrm{eq},11/22}$ according to MIMO waterfilling power allocation algorithm. The scalar $h_{1\times1}^2$ is the channel gain due to fading for the SISO case following Rayleigh distribution with unit variance. Since the average signal to interference plus noise ratio $\bar{\rho} = P_{\mathrm{m}} G / (N_0 B + I_{\mathrm{int}})$, with spatial independent MIMO channels Equation (5.10) can be simplified as

$$\begin{cases} \bar{\rho}(x_{\mathrm{s},2\times2} - x_{\mathrm{s},1\times1}) = \frac{(2^{R_{\mathrm{eq},11/22}} - 1)}{h_{1\times1}^2}(1 - x_{\mathrm{s},1\times1}) - \sum_{m=1}^{M=2} \frac{(2^{r_{\mathrm{eq},m}} - 1)}{h_{2\times2,m}^2}(1 - x_{\mathrm{s},2\times2}) \\ \sum_{m=1}^{M=2} r_{\mathrm{eq},m} = R_{\mathrm{eq},11/22}, \end{cases} \tag{5.11}$$

where $h_{2\times2,m}^2$ $(m = \{1, 2\})$ are the eigenvalues of the 2×2 MIMO signal covariance matrix $\mathbf{H}\mathbf{H}^*$. Similarly, $r_{\mathrm{eq},m}$ is the data rate delivered by the spatial channel m. Here, how much energy 2×2 MIMO saves for transmitting the same amount of data as the SISO system is dependent on these variables. Providing specific R_{d} and $\bar{\rho}$, the time-varying channel coefficients $h_{2\times2,m}$ $(m = \{1, 2\})$ could determine the spatial structure of the MIMO channels. Since the energy reduction achieved by MIMO derives from MIMO SM, the relative quality of different MIMO channels could decide the performance of the waterfilling algorithm for spatial multiplexing. Thus the energy reduction obtained by the MIMO system is time-dependent.

However, by taking the expectation of Equation (5.11), simulation experience suggests that the average sum energy consumption of MIMO channels converges over time, which can be represented as

$$\begin{cases} \mathbb{E}\left[\bar{\rho}(x_{\mathrm{s},2\times2}[\tau] - x_{\mathrm{s},1\times1}[\tau])\right] = \mathbb{E}\left[\frac{(2^{R_{\mathrm{eq},11/22}}-1)(1-x_{\mathrm{s},1\times1}[\tau])}{h_{1\times1}[\tau]^2} - \sum_{m=1}^{M=2}\frac{(2^{r_{\mathrm{eq},m}}[\tau]-1)(1-x_{\mathrm{s},2\times2}[\tau])}{h_{2\times2,m}[\tau]^2}\right] \\ \sum_{m=1}^{M=2} r_{\mathrm{eq},m}[\tau] = R_{\mathrm{eq},11/22}. \end{cases} \tag{5.12}$$

Therefore, with a known average channel condition $\bar{\rho}$ or transceiver distance d, we can find the data rate $R_{\mathrm{eq},11/22}$ under which 2×2 MIMO consumes the same amount of energy as that of the SISO case. If the required data rate $R_{\mathrm{d}} > R_{\mathrm{eq},11/22}$, 2×2 MIMO is more energy-efficient, since the energy saving on user data transmission is more than enough to compensate the energy used for sending the extra reference signals. In contrast, if $R_{\mathrm{d}} < R_{\mathrm{eq},11/22}$, SISO achieves better energy performance than 2×2 MIMO. Similar methods apply for comparing the energy efficiency of 2×2 MIMO with 4×4 MIMO.

5.3.3 Antenna Selection for Non-Sleep Modes

Based on the above analysis, we try to optimize the spatial-domain and time-domain separately for MIMO sleep modes. Firstly, we ignore the time-domain design in Chapter 4 and only focus on how to select the number of antenna ports for energy-efficient transmission.

Because QPSK, 16-QAM and 64-QAM are supported modulation formats in LTE, there is an upper bound of the maximum achievable data rate. According to the LTE standard [41], this is obtained by 64-QAM modulation with rate 948/1024 Turbo coding, corresponding to $R_{\max}$=5.18 bps/Hz for SISO, 10.37 bps/Hz for 2×2 MIMO and 20.74 bps/Hz for 4×4 MIMO. With the provided average channel conditions, we can also calculate the ergodic capacity R_{ergodic} for any antenna configuration. Thus the feasible data rate target can be obtained as $\{0, \min(C_{\mathrm{ergodic}}, R_{\max})\}$.

In Figure 5.8, we compare the energy consumption of SISO, 2×2 MIMO and 4×4 MIMO for delivering the same amount of data rate under given channel conditions. If we use the X axis to denote the average channel condition $\bar{\rho}$, the Y axis to represent the data rate requirement R_{d} and the Z axis to represent the total energy consumption E, this figure can be interpreted as the most energy-efficient operating region obtained under different antenna configurations projected on the X-Y plane.

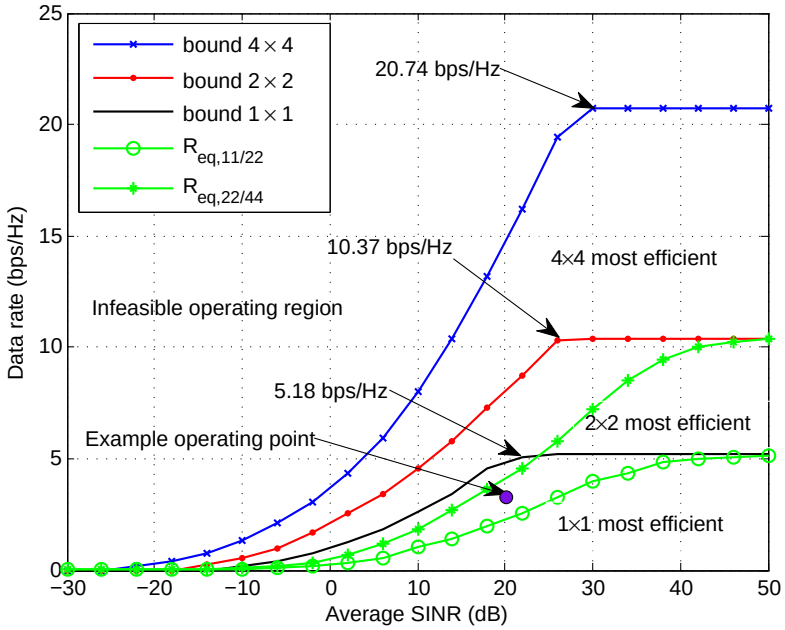

Figure 5.8: *Energy efficiency regions under different antenna configurations*

Firstly, $\{0, \min(C_{\text{ergodic}}, R_{\max})\}$ has determined the operating region. This is shown as 'bound 4×4', 'bound 2×2' and 'bound 1×1' in Figure 5.8 for 4×4 MIMO, 2×2 MIMO and SISO respectively. The data rate requirement R_{d} which falls outside 'bound 4×4' can not be delivered using our assumed antenna configurations under the current channel conditions, which is labeled as 'Infeasible operating region' in Figure 5.8.

Secondly, within the operating region, we can use the figure to identify which antenna configuration is feasible for delivering a required data rate R_{d}. In a communication system, we assume the following parameters are already known:

- Averaged channel conditions: can be obtained from either transceiver distance or channel feed back over a given time period.
- Data rate requirements: fixed data rate depending on user's application.

We can locate the operating point on the SINR-data rate grid in Figure 5.8 according to the above values. If the operating point falls between 'bound 2×2' and 'bound 4×4', we can only use 4×4 MIMO to transmit due to its high spectral efficiency. Similarly, if the operating point falls between 'bound 1×1' and 'bound 2×2', we can use either 2×2 or 4×4 MIMO for transmission. If the operating point falls into 'bound 1×1', all the three antenna configuration

are feasible.

Thirdly, Figure 5.8 can be used to identify the most energy-efficient antenna configuration. This is achieved by superimposing another two curves '$R_{\mathrm{eq},11/22}$' and '$R_{\mathrm{eq},22/44}$', which represent the data rates under which SISO and 2×2 MIMO or 2×2 MIMO and 4×4 MIMO consume the same amount of energy, and are plotted according to Equation (5.10). As we discussed in Section 5.3.2, if the operating point falls in the upper left side of the curve, a higher data rate is required than R_{eq}, which means user data transmission plays an important role in total energy consumption, thus using more antenna ports is preferred. In contrast, if the operating point falls in the lower right side of the curve, using fewer antenna ports is preferred.

Here we take the marked purple operating point in Figure 5.8 as an example. In principle, since the operating point is under the 'bound 1×1' curve, we can use SISO, 2×2 MIMO or 4×4 MIMO to deliver R_{d}=3.5 bps/Hz under an average SINR=20 dB channel. Also, the point is above '$R_{\mathrm{eq},11/22}$' curve and below '$R_{\mathrm{eq},22/44}$' curve, which means 2×2 MIMO is more energy-efficient than SISO and 4×4 MIMO.

5.3.4 Performance Analysis of MIMO Sleep Modes

Section 5.3.3 only discussed the spatial-domain optimization issues. Combining the time-domain optimization as well, MIMO sleep modes are simulated in Figure 5.9 and Figure 5.10. Here the time-domain optimization is obtained by searching the number of active subframes per frame N_{opt} within $\{1, 2, \cdots, 10\}$ in each channel state. The plotted energy consumption are time-averaged values over 20,000 channel realizations. We can also interpret these two figures as two vertical slices of Figure 5.8 at SINR=40 dB and SINR=20 dB.

In general, the energy consumption increases with R_{d} for all SISO/MIMO cases due to the higher energy used on user data transmission. Meanwhile, when comparing Figure 5.10 to Figure 5.9, because of the degraded channel conditions, higher energy consumption is required.

If we focus on the non-sleep mode cases, as suggested in Figure 5.9, using less antenna ports are always more efficient for all feasible values of R_{d}. This is consistent with Figure 5.8, since the '$R_{\mathrm{eq},11/22}$' and '$R_{\mathrm{eq},22/44}$' curves at SINR=40 dB are approaching 'bound 1×1' and 'bound 2×2' respectively. In Figure 5.10, the achieved peak data rate for 2×2 MIMO case outperforms SISO by about 3 bps/Hz under system sum power constraint P_{m}, while the 4×4 MIMO case outperforms 2×2 MIMO by about 6.2 bps/Hz. This is similar to the values of R_{eq}

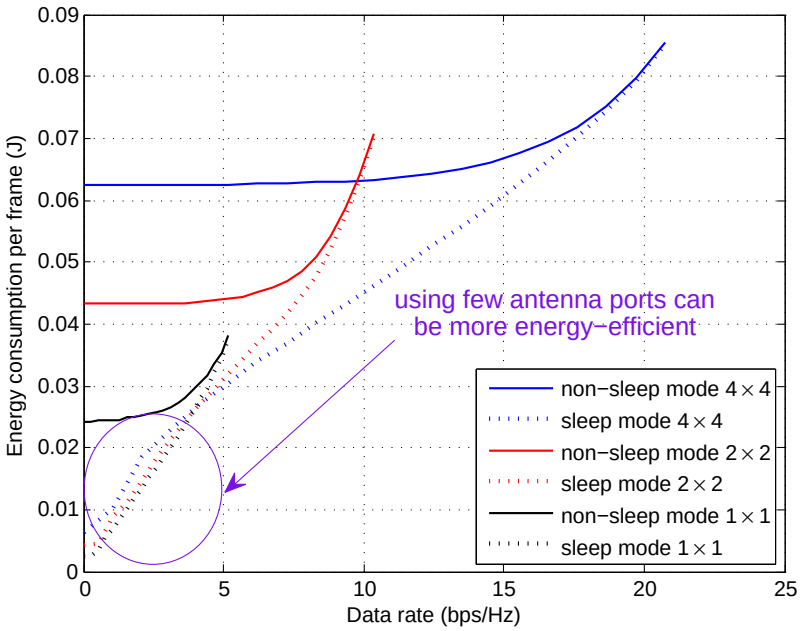

Figure 5.9: *Energy consumption of MIMO sleep modes with 40 dB SINR*

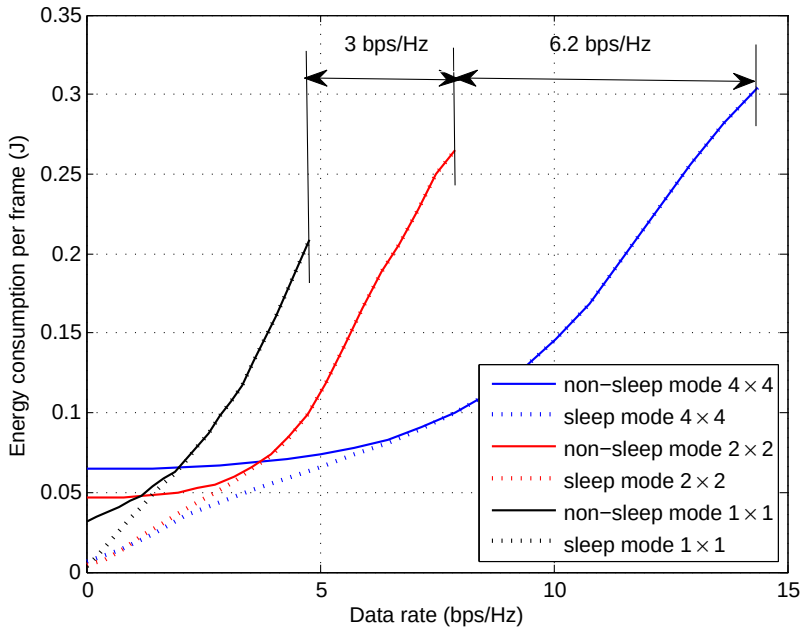

Figure 5.10: *Energy consumption of MIMO sleep modes with 20 dB SINR*

at SINR=20 dB in Figure 5.8 as well. In generally, without time-domain optimization, using fewer antenna ports can potentially save energy consumption. An antenna selection criterion can be obtained according to Figure 5.8.

With time-domain sleep modes, simulation results suggest that using more antennas is generally a good choice. This is because the extra reference signal overhead can be overcome. Although Figure 5.9 reveals that using fewer antenna ports can be more energy-efficient at low R_{d}, this energy saving is trivial. Also we notice that with lower SINR values in Figure 5.10, the advantage of using fewer antennas is also reduced. This trend continues when we further decrease the SINR values. Thus it is suggested to use all available antennas for transmission when the time-domain sleep modes are deployed.

5.3.5 Practical Issues of Spatial-Domain Antenna Sleep Modes

As suggested in Section 5.3.3 and Section 5.3.4, without time-domain sleep modes, using fewer antennas for transmission could improve system energy efficiency; with time-domain sleep modes, using all available antennas is generally a good choice. Thus in this section, we ignore the time-domain optimization problem and further analyze the impacts of antenna configuration on energy-efficient design in the spatial-domain.

In practice, apart from the SISO, 2×2 MIMO and 4×4 MIMO systems already considered, we can also use different number of antennas at the transceiver sides. For example. if two antennas are available at both the transmitter and receiver sides, there are four feasible antenna configurations: SISO, 1×2 SIMO, 2×1 MISO and 2×2 MIMO.

Regarding their energy efficiency performance, we still analyze the energy used for control signalling and data transmission separately. For control signalling, extra energy is used for delivering a larger overhead size with the increased number of transmit antenna ports, since extra REs are needed for reference signals. This means that 2×1 MISO or 2×2 MIMO will consume more energy than SISO or 1×2 SIMO on the control plane. For data transmission, higher diversity gains can be obtained with the increased number of receiver antennas. This means that 2×2 MIMO is more efficient than 2×1 MISO; 1×2 SIMO is more efficient than SISO. Combining these two observations together, we can see that potential energy saving can be achieved by switching off transmit antennas rather than receiver antennas. For the 2×2 MIMO system, it is more useful to identify the relative efficiency regions between 1×2 SIMO and 2×2 MIMO. This is shown in Figure 5.11 as '$R_{\mathrm{eq},12/22}$'. The '$R_{\mathrm{eq},11/22}$' curve in Figure 5.8 is not needed since 1×2 SIMO is always more efficient than SISO. If the spatial-domain sleep mode is enabled in the 2×2 MIMO system, we only need to consider the circumstance for switching between 1×2 SIMO and 2×2 MIMO.

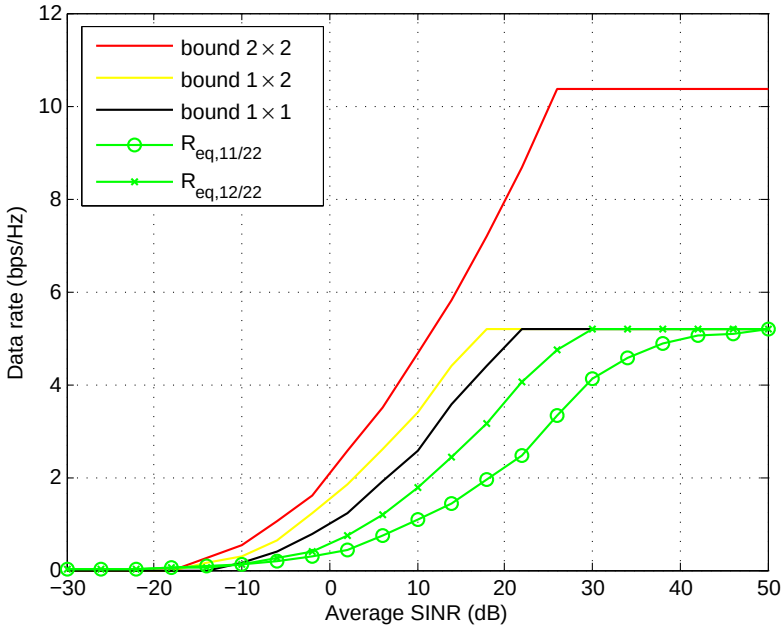

Figure 5.11: *Antenna sleep patterns within* 2×2 *MIMO system*

Similarly, if four antennas are equipped at both transmitter and receiver sides, we can identify the relevant efficiency regions in Figure 5.12. Compared with Figure 5.8, using 2×4 MIMO always achieves higher spectral efficiency than using 2×2 MIMO. The region where 2×4 MIMO is more energy-efficient than 4×4 MIMO is also larger than the region where 2×2 MIMO is more efficient than 4×4 MIMO. This suggests that when 4 antennas are equipped at both transmitter and receiver sides, no receiver antennas are required to be switched off. If the operating point falls within the curve '$R_{\mathrm{eq},14/24}$', switching off three transmit antennas is suggested. If the operating point falls between the curves '$R_{\mathrm{eq},14/24}$' and '$R_{\mathrm{eq},24/44}$', gating off two transmit antennas is the most energy-efficient. If the operating point falls between the curves '$R_{\mathrm{eq},24/44}$' and 'bound 4×4', all antennas should be activated.

Here we need to notice that although the spatial-domain antenna sleep modes can potentially improve energy efficiency, hardware constraints can not be ignored in practice. While our model assumes a sum system power constraint of 40 W across all transmit antennas, constraints may exist on individual power amplifiers depending on the antenna architecture. While different antennas are usually equipped with separate PAs, a peak power constraint per antenna may apply as well.

From a system point of view, since most energy is consumed in base station hardware rather

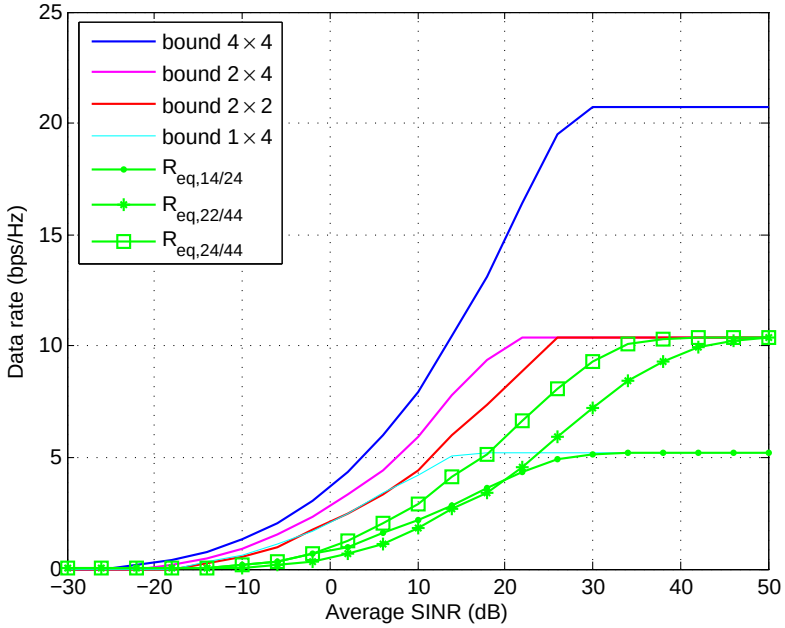

Figure 5.12: *Antenna sleep patterns within* 4×4 *MIMO system*

than the RF energy, it is important to check that implementing MIMO sleep modes will not increase the overall operational energy consumption. This hardware operational energy generally includes the energy used for cooling the base station, rectifier, ADC/DAC, DSP, PA *etc*. With specified data rate requirements, we generally expect significant energy reduction for operating power amplifiers by switching off individual antennas in the spatial-domain [134]. Since the PA energy consumption typically contributes a large portion of RBS energy consumption, the benefit of deploying spatial-domain sleep mode is significant. Here, we need to note that although the spatial-domain approach is positive from the energy efficiency point of view, noticeable downlink performance degradation may happen. Firstly, decreasing the number of transmit antenna ports may decrease the cell coverage. Secondly, since the number of active antenna ports is related to the decoding process of the broadcast channels, incorrect decoding may happen by gating off a number of radio antennas [135].

5.4 Other Discussions

Since all the above discussions are based on the single user assumption, in this section we extend the sleep mode design to a multi-user environment. The issues of combining a sleep mode option into frequency-domain schedulers are briefly introduced.

5.4.1 Multi-User Sleep Modes with Frequency-Domain Scheduling

To implement MIMO sleep modes, it is important to analyze their performance in multi-user scenarios together with the assumed scheduler. As the base station sleep modes are designed for low-traffic scenarios, we assume that the number of users requiring service is less than the maximum number of users supported by LTE control signalling. In multi-cell multi-user environments, since the energy reduction gains of MIMO sleep modes are dependent on the specific scheduling decisions, here we aim to address the issues of how to combine sleep modes with frequency-domain scheduling [136] rather than analyzing any specific example.

With more than 1 users being scheduled at the same time, the overhead size would be larger than the values summarized in Table 5.1. This is because more PDCCHs are transmitted, carrying user-specific resource assignments. Since different users have different channel conditions, the control overhead size can be estimated according to Section 5.2.3 if both the number of users and each user's channel state information are known. In Figure 5.13, the size of control signalling is plotted assuming that all the scheduled users have the same CCE aggregation level and 3 OFDM symbols are used in each subframe for the control plane. The figure generally suggests that the control overhead size linearly increases with the number of serviced users. There is a maximum value for the number of users which can be scheduled in each case. Also, the 4×4 MIMO case has a larger control overhead than the SISO case with the same CCE aggregation level and number of serviced users.

For user data transmission, OFDMA is used in the LTE downlink [41]. A PRB with one subframe in time and 12 subcarriers in frequency is the minimum unit for resource assignment. With the scheduler's decisions informed by PDCCHs, different users only decode the data carried on their allocated subcarriers. Assuming the full system bandwidth is always used in each cell, multi-user diversity can be obtained by dynamically allocating system bandwidth in each time slot. Meanwhile, with frequency-selective fading, assigning the subcarriers to the corresponding users with the best SINR can further reduce system energy consumption. Compared with the single-user scenario with the required data rate R_{d} bps/Hz transmitted over full bandwidth, if there are I users with the same data rate requirements and average channel conditions, less energy is needed based on the diversity achieved by scheduling [45].

Based on the above reasons, the energy consumption of control signalling can not be ignored when designing energy-efficient multi-user resource allocation schemes. By defining new en-

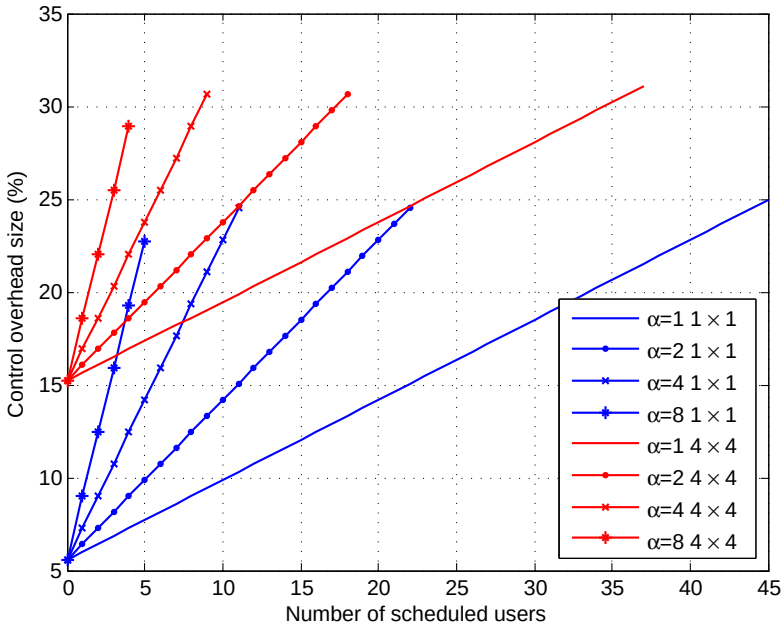

Figure 5.13: *SISO and* 4×4 *MIMO control overhead size for multi-user scenarios*

ergy efficiency metrics for scheduling functions, different energy efficient-scheduling strategies can be proposed. Besides, the time-domain base station sleep modes can be incorporated to achieve further energy gains. With a frequency-domain scheduler, the scheduling process can be described in Table 5.2.

Regarding the time-domain optimization, we need to notice that certain subframes can not be gated off. For example, the primary synchronization signals are carried by the first subframe in each radio frame on the central 62 subcarriers. Gating off these PRBs will cause failures of time synchronization during cell search. Another example is the physical HARQ indicator

Algorithm:

1. Define the peak power per subcarrier for control signalling.
2. For each subcarrier, calculate the number of active subframes per frame N_{opt} in the current radio frame, record the optimum energy consumption on each subcarrier with time-domain optimization.
3. Repeat 2. for each user required to be scheduled.
4. Allocate each subcarrier to the user with minimum energy consumption while satisfying certain fairness functions [99].
5. Adjust the allocated subcarriers to an integer number of PRBs for each user.
6. For the next radio frame, repeat 2. to 5.

Table 5.2: *Multi-user scheduling incorporating sleep modes*

channels. Since they are grouped together to map to certain resource elements for delivering ACK/NACK information, switching off PHICHs may cause the drop of uplink information. Also, the PBCH with system information is essential for the other channels to be configured and operated. With N_{opt}, we need to be careful when we decide which specific subframes to be switched off. Apart from these system operating reasons, it is also necessary to make sure that most subframes are switched off across the full system bandwidth. This will help to save hardware operating energy by gating off the entire base station, which can be done by arranging the active subframes for each subcarrier maximally overlapped on the same time slots.

5.4.2 Sleep Modes in VoIP

The basic idea of time-domain sleep modes is to save energy consumption by delivering less control signalling overhead. In principle it can be applied to any application. Because the optimization is carried out for each 10 ms LTE frame separately, in principle it can be applied to delay-limited applications as well. For VoIP, since the maximum allowed delay is 20 ms, we might use a less aggressive sleep mode with VoIP calls to minimize delay. In addition, the idea of time-domain sleep modes can be extended to other systems besides LTE. Although different models may apply, the potential energy saving due to signalling less control overhead still holds.

5.4.3 Legacy Handsets

Currently, there is a large proportion of legacy handsets still in use. When there is no user-specific traffic, the time-domain sleep mode can be activated to save handset energy consumption. Then the handsets may be waked up by a special signal when the traffic arrives. With low network traffic, the legacy network can be turned off periodically as well. For example, a GSM base station may have very low traffic loads during nighttime, sleep modes energy reduction can be obtained by delivering less signalling overheads. Although the sleep mode idea can apply to any communication system, we need to notice that for the old systems which are dominated by voice traffic, less aggressive sleep modes are generally suggested.

5.5 Summary

In this chapter, the idea of sleep mode design is further examined according to the LTE standards. Firstly, the control signalling overhead is adapted to the instantaneous channel state information, which is called CCE aggregation. Secondly, we extended the design to MIMO scenarios. Regarding the first issue, we provided a detailed estimation of control signalling overhead size by evaluating the contribution of all the system signals and control channels. While the PBCH, PCFICH, PHICH and synchronization signals occupy a fixed signalling overhead, the overall control overhead size varies with the number of transmit antennas and instantaneous channel conditions as shown in Table 5.1. Regarding MIMO antenna configurations, energy-efficient spatial-domain optimization is introduced. Although MIMO has larger control overhead due to the extra reference signals compared to SISO, it requires less energy for delivering the same amount of user data based on spatial multiplexing. For this spatial-domain optimization, we combined time-domain non-sleep mode and sleep mode for analysis separately. Without time-domain sleep modes, we identified the regions achieved by the most efficient antenna configurations. This provides guidance for selecting the number of transceiver antennas knowing the average channel SINR and data rate requirements. For the zero-load case, switching off antennas from 4×4 MIMO to 4×1 MISO can save 75% RF energy consumption. With time-domain sleep modes, simulation results suggest that using all available antennas is generally the most energy-efficient choice.

We further analyzed the performance of MIMO non-sleep/sleep modes with unequal number of transceiver antennas. The results suggest that energy gains can be obtained by switching off transmit antennas rather than receiver antennas for the non-sleep mode case. In addition, since the analysis is based on single user scenario, we briefly discussed the issues of implementing sleep modes in multi-user scenarios with frequency-domain scheduling. The proposed base station sleep modes can be potentially applied to any applications.

Chapter 6
Conclusions

In this chapter, the contributions of this thesis will be summarized. By properly allocating frequency subcarriers to different users, interference can be migrated depending on users' locations. By adapting the number of active subframes and antenna ports to the network traffic, we can transmit less control signalling overhead. Significant energy reduction gains have been obtained and the limitations as well as some possible future research topics will be pointed out.

6.1 Contributions

Today the energy consumption of mobile networks increases rapidly, which significantly exceeds the speed of revenue growth. At the same time, excessive CO_2 emissions will have negative impacts on our environment. Based on these economic and environmental considerations, reducing energy consumption has become an important research topic. The Mobile VCE Green Radio project aims to achieve 100x energy reduction by developing new network architectures and resource management solutions. In this thesis we focused on designing energy-efficient resource assignment strategies, which are different from most of the previous resource allocation approaches which aim to improve users' quality of service. Our new schemes took LTE downlink transmission as an example, and identified the possible energy reduction from the frequency, time and spatial domains respectively. In general, these energy gains derive from deploying the intelligent multiple access techniques to reduce interference, and novel base station sleep mode techniques to reduce LTE control signalling.

To analyze energy-efficient resource allocation from an information-theoretic point of view, Chapter 3 has inverted the Shannon equation by assuming fixed data rates and calculated energy consumption. To simplify the analysis, we started by comparing the energy consumed by the non-orthogonal and orthogonal techniques based on a two-link scenario, where the two mobile users can move on the line between the two base stations in their own cell. The non-orthogonal technique which is also termed as simultaneous transmission, refers to using full system bandwidth to maximize re-use efficiency, where co-channel interference exists between the two

links. The orthogonal technique refers to allocating orthogonal system bandwidth or time resources to different users in order to avoid interference, which is named as FDMA/TDMA-based orthogonal transmission. With specific link data rate constraints, it is proved that TDMA-based OT requires higher instantaneous power than FDMA-based OT. However, their average energy consumption over a specified period of time will be the same. Meanwhile, the non-orthogonal transmission generally outperforms the orthogonal transmission when the mobile users are located near cell-center, while the orthogonal transmission is more energy-efficient when the users are approaching cell-edge. This indicates that soft frequency reuse can be employed for energy-efficient network architecture design. Further, instead of specifying the data rate on both links, we have simulated the scenarios where the network sum data rate constraints are provided. We fixed the location of one user and move the other user from the cell-center to the cell-edge. An interesting result is obtained by plotting the geographic boundaries where OT outperforms ST under different sum rate constraints. The simulation results suggest that these boundaries generally depend on the network sum data rates no matter what rate proportions are shared between these two links. Apart from these, we have defined a new energy efficiency matric 'energy consumption rating' with its corresponding power fairness index metric. With new optimization objectives, the optimum bandwidth allocation for the OT scheme is changed. Also, based on the efficiency-fairness tradeoff, a weighted approach is proposed by combining energy efficiency and fairness into a new matric with different weighing factors. Finally, we extended our two-link model into a multi-link scenario. Simulation results suggest that the conclusions obtained from the two-link scenario also hold for the multi-link case. This suggests how to design location-dependent soft frequency reuse schemes given network sum data rate constraints.

While Chapter 3 bridged the connections of resource allocation strategy design from the spectral efficiency perspective to the energy efficiency perspective based on information theory, Chapter 4 and Chapter 5 took more practical issues into consideration. This is because the radio base station consumes a large proportion of energy on operating its hardware such as power supply units, antenna power amplifiers, baseband signal processing, ADC/DAC devices, RBS cooling *etc*. The energy actually used for RF transmission is currently below 6% of the total. Thus it will be good if the new resource allocation techniques can also contribute to the reduction of the total hardware energy consumption. The sleep mode has been identified as a potential technique to achieve this goal. In Chapter 4, a time-domain sleep mode based on single user SISO system has been described. By adapting the active number of subframes in

each radio frame depending on the current network traffic and channel conditions, RF energy reduction gain can be achieved due to not transmitting redundant control signalling overhead. Meanwhile, since the transmission is switched off during certain period of time, the energy consumed by the power amplifiers, RBS cooling, baseband signal processing *etc* will decrease. A closed-form expression for calculating the optimum sleep mode duty cycle has been derived, taking into account the user's data rate requirements and current channel state information. We also defined the average load factor as the quotient of the required data rate divided by the ergodic data capacity to depict the time-averaged network traffic situation. For constant rate transmission, simulation results suggest that when the average load factor exceeds 60%, there is no obvious ERG by deploying the time-domain sleep mode optimization. During the low traffic conditions, which usually happen at night time, the proposed sleep mode design can achieve an up to 90% RF energy reduction gain. This energy gain is generally decided by the average load factors, while slightly decreases with the increased transceiver distance.

Similar ideas can be applied to the spatial-domain optimization by gating off individual transmit antennas. In Chapter 5, we firstly provided a detailed estimation of LTE control signalling overhead under different antenna configurations, which generally occupies $5\% - 30\%$ physical resources in each PRB. This overhead size increases with the increased number of transmit antennas, since more reference signals are required for channel estimation. Meanwhile, depending on the channel conditions, different CCE aggregation levels can be selected for the PDCCH. Basically, a higher CCE aggregation level is typically used for poor channel conditions to enhance the channel coding redundancy. Among the various control signals and channels, reference signals contribute the largest portion of the total overhead, thus reducing the MIMO order is expected to reduce control signalling energy consumption. Also, the energy used by power amplifiers will decrease correspondingly. Based on these analysis, we modeled the energy consumption of 4×4 MIMO, 2×2 MIMO and SISO scenarios under given user data rates. Simulation results suggest that with the same user data requirements, using less antennas can be more energy-efficient. The boundaries which separate different optimum antenna configuration regions from an energy efficiency perspective are superimposed on the SNR-data rate grid. In general, with low data rate requirements, using fewer antenna ports for transmission is suggested. With high data rate requirements, using more antenna ports will benefit energy reduction. Here we also need to notice that the data rates we delivered are upper-bounded by the system configurations, including the channel state information and the number of antenna ports. The improved spectral efficiency of spatial multiplexing indicates that high data rates can

only be delivered with more antennas under given channel conditions. The feasible operating regions, which are within the minimum of the channel capacity and the maximum supported coding and modulation formats, are plotted in the SNR-data rate grid as well. Further, we have extended our discussions to include various of MISO and SIMO scenarios, and suggested how to extend the current work to multi-user scenarios with frequency-domain resource allocation. Finally, we combined the spatial-domain sleep mode with the time-domain sleep mode described in Chapter 4. Simulation results illustrate that without time-domain optimization, significant energy gains can be obtained by switching off transmit antennas; with time-domain optimization it is generally more efficient by using all available radio antennas.

6.2 Future Work

This thesis has identified promising techniques for designing energy-efficient communication systems. With the support of analytical results and numerical simulations, both the location-based soft frequency reuse and time/spatial-domain sleep modes have demonstrated the ability to significantly reduce network energy consumption for LTE. However, the proposed techniques have certain limitations due to the initial modeling assumptions. More efforts can be made in the future to generalize the analysis, or to design practical schemes when targeting at certain scenarios.

Regarding the frequency-domain approach documented in Chapter 3, the two-link scenario analysis has suggested how the energy reduction gain is obtained, followed by the extension of multi-link scenario solutions. Taking the three-link case as an example, we can always use the energy efficiency based subcarrier allocation algorithm to orthogonally assign each subcarrier to the appropriate user. Thus the optimum frequency allocation can be achieved for the OT scheme. When there are more than three links, the BABS algorithm applies similarly. In Chapter 3, it is also argued that when the network sum data rate is known, the ST/OT boundary can be determined without knowing the rate proportions among different links. Thus in principle the users which are located within the boundary in a cell can use full frequency bands. However, this conclusion may depend on the specific models. Firstly, the three-link model still constrains the three users moving on the line towards the junction point J. There may be more users in practice who have been scheduled at the same time, and the users can be located anywhere in the cell. We need to verify whether our conclusions still hold for the generalized case. A system level simulation is suggested to investigate this problem. Secondly, we have ignored

the scenario where several users are all very close to their base station in the same cell. In LTE, since the orthogonality should be maintained among them, further studies are required on how to expand their bandwidth. Obviously, the simplest approach is to let only one cell-center user to reuse all the subcarriers from other cells based on the BABS algorithm. Alternatively, we can further study how to assign these 'other cell subcarriers' orthogonally for sharing among several cell-center users.

For the time-domain sleep mode design in Chapter 4, the energy saving derived from delivering less control signalling overhead. However, we need to be careful when we switching off the subframes, since some control signals are essential for the system operation. For example, the PHICH carrying HARQ ACK/NACK can not be switched off when there is uplink transmission. Switching off PBCH will affect the other channels in the cell to be properly configured and operated. Synchronization signals should also remain active for cell search and synchronization purposes. Since all of these signals and channels are defined with a specific pattern in the LTE physical layer, which subframes to switch off should be carefully selected when the time-domain sleep mode is deployed. Also, sometimes it is better to adopt the sleep mode less aggressively, in the case of delay-sensitive applications such as VoIP or when there are legacy handsets within the cell. Apart from these signalling issues, the current model is based on single user SISO system. Future work can further study how to deploy the time-domain sleep mode together with frequency-domain scheduling, or how to decide the sleep mode duty cycle when multiple antennas are used. Also, the current simulation assumes a constant data rate to transmit based on the Shannon theory, more practical design with adaptive modulation and coding schemes can be applied similarly.

In Chapter 5, the potentials of energy reduction by gating off a number of transmit antennas are discussed. Similar to the time-domain sleep mode, extra efforts on how to solve the signalling issues when the number of antennas are changed need to be discussed. Principally, the antenna number change can be detected by the UE from reading PBCH information. The dedicated antenna configuration or transmission mode have to be changed at the same time to keep the ongoing service continuity. Therefore, for the users with constant data rate requirements and fixed location, the optimum number of active antennas can be decided at the beginning. When the data rate adaptation or the user mobility is enabled, a proper LTE control signalling design need to be further addressed. Finally, the current work indicates that the sleep mode design will impact the hardware energy consumption for operating LTE base stations. In the spatial-

domain, it is expected that the PA energy consumption will reduce with the decreased RF energy directly. For future studies, other operational energy or even embodied energy can be taken into consideration for optimization. This will affect the solutions of energy-efficient resource allocation depending on our constructed energy model.

Appendix A
Appendix: Publication List

Journal

1. R. Wang, J. Thompson, H. Haas, P.M. Grant, Sleep Mode Design for Green Base Stations, submitted to *IET Communications: Special Issue on Green Technologies for Wireless Communications and Mobile Computing*.

Conference

1. R. Wang, J. Thompson, H. Haas, A New Framework for Designing Power-Efficient Resource Allocation under Rate Constraints, *IEEE 70th Vehicular Technology Conference Fall*, Anchorage, USA, 2009.

2. R. Wang, J. Thompson, H. Haas, A Novel Time-Domain Sleep Mode Design for Energy-Efficient LTE, *IEEE 4th International Symposium on Communications, Control and Signal Processing*, Limassol, Cyprus, 2010.

3. S. McLaughlin, P.M. Grant, J. Thompson, H. Haas, D. Laurenson, C. Khirallah, Y. Hou and R. Wang, Techniques for Improving Radio Base Station Efficiency, submitted to *IEEE International Conference on Wireless Communications*, Vancouver, Canada, 2011.

Patent

1. S. Videv, H. Haas, R. Wang, J. Thompson, Energy Efficient Scheduling, *UK Patent 1013771.9*, filed on 2nd Sep 2010.

References

[1] 2020 Vision: Enabling the Digital Future, *Mobile VCE Vision Group*, 2007. URL: http://www.mobilevce.com.

[2] Y. Chen, S. Zhang, S. Xu and G.Li, Fundamental Tradeoffs on Green Wireless Networks, accepted by *IEEE Communications Magazine*, 2011.

[3] H. Leem, S.Y. Baek and D.K. Sung, The Effects of Cell Size on Energy Saving, System Capacity, and Per-Energy Capacity, *IEEE Wireless Communications and Networking Conference*, Page 1–6, Sydney, Australia, 2010.

[4] T. Svensson, T. Frank, T. Eriksson, D. Aronsson, M. Sternad and A. Klein, Block Interleaved Frequency Division Multiple Access for Power Efficiency, Robustness, Flexibility and Scalability, *EURASIP Journal on Wireless Communications and Networking*, Volume 2009, Page 1–18, 2009.

[5] X. Wang and G.B. Giannakis, Power Efficient Resource Allocation for Time Division Multiple Access over Fading Channels, *IEEE Transactions on Information Theory*, Volume 54, Issue 3, Page 1225–1240, 2008.

[6] Y. Yao, X. Cai and G.B. Giannakis, On Energy Efficiency and Optimum Resource Allocation of Relay Transmissions in the Low-Power Regime, *IEEE Transactions on Wireless Communications*, Volume 4, Issue 6, Page 2917–2927, 2005.

[7] V. Pitaval, R.A. Blostein, S. Riihonen and T. Wichman, Green Cooperative Communication using Threshold-Based Relay Selection Prootocols, *International Conference on Green Circuits and Systems*, Page 521–526, Shanghai, China, 2010.

[8] D. Grace, J. Chen, T. Jiang and P.D. Mitchell, Using Cognitive Radio to Deliver Green Communications, *Cognitive Radio Oriented Wireless Networks and Communications*, Page 1–6, Hanover, Germany, 2009.

[9] F. Meshkati, H. V. Poor, S.C. Schwartz and R.V. Balan, Energy Efficient Resource Allocation in Wireless Networks with Quality of Service Constraints, *IEEE Transactions on Communications*, Volume 57, Issue 11, Page 3406–3414, 2009.

[10] X. Mao, A. Maaref and K.H. Teo, Adaptive Soft Frequency Reuse for Inter-Cell Interference Coordination in SC-FDMA Based 3GPP LTE Uplinks, *IEEE Global Telecommunications Conference*, Page 1–6, New Orleans, USA, 2008.

[11] K. Lei, G. Wong and D. Tsang, Performance Study and System Optimization on Sleep Mode Operation in IEEE 802.16e, *IEEE Transactions of Wireless Communications*, Volume 8, Issue 9, Page 4518–4528, 2009.

[12] K.W. Wong, Q. Zhang and H.K. Tsang, Switching Cost Minimization in the IEEE 802.16e Mobile WiMAX Sleep Mode Operation, *Wireless Communications and Mobile Computing*, Volume 10, Issue 12, Page 1576–1588, 2010.

[13] R. Wang, J. Thompson, H. Haas, A New Framework for Designing Power-Efficient Resource Allocation under Rate Constraints, *IEEE 70th Vehicular Technology Conference Fall*, Anchorage, USA, 2009.

[14] R. Wang, J. Thompson, H. Haas, A Novel Time-Domain Sleep Mode Design for Energy-Efficient LTE, *IEEE 4th International Symposium on Communications, Control and Signal Processing*, Limassol, Cyprus, 2010.

[15] R. Wang, J. Thompson, H. Haas, P.M. Grant, Sleep Mode Design for Green Base Stations, submitted to *IET Communications: Special Issue on Green Technologies for Wireless Communications and Mobile Computing*.

[16] S. McLaughlin, P.M. Grant, J. Thompson, H. Haas, D. Laurenson, C. Khirallah, Y. Hou and R. Wang, Techniques for Improving Radio Base Station Efficiency, submitted to *IEEE International Conference on Wireless Communications*, Vancouver, Canada, 2011.

[17] S. Videv, H. Haas, R. Wang, J. Thompson, Energy Efficient Scheduling, *UK Patent 1013771.9*, filed on 2nd Sep 2010.

[18] J. Gozalvez, Green Radio Technologies, *IEEE Vehicular Technology Magazine*, Volume 5, Issue 1, Page 9–14, 2010.

[19] J. Hildersley, High Efficiency, Broadband Multimode PAs for 4G systems, *Nujira*, 2010. URL: https://ktn.innovateuk.org/c/document_library/get_file?p_l_id=737699 &folderId=870487&name=DLFE-7186.pdf.

[20] A. Dua, Power controlled random access, *IEEE International Conference on Communications*, Volume 6, Page 3514–3518, Paris, 2004.

[21] Long Term Evolution (LTE): an introduction, *Ericsson*, 2007. URL: http://www.ericsson.com/technology/whitepapers/lte_overview.pdf.

[22] Long Term Evolution: the vision beyond 3G, *Nortel*, 2008. URL: http://4g-wirelessevolution.tmcnet.com/topics/4g-wirelessevolution/articles/Nortel

[23] Green Telecom-Calling for a Better Future, IDATE News 453 (05/01/2009), URL: http://www.idate.fr/pages/index.php?rubrique=etudeidr=16idp=334idl=7.

[24] F. Meywerk, The Mobile Broadband Vision-How to Make LTE a Success, *LTE World Summit*, London, UK, 2008.

[25] T. O'Farrell, Metrics and Network Architectures for Green Radio, *Presentation from Mobile VCE TSG*, Bristol, UK, 2009.

[26] B. Badic, T. O'Farrell, P. Loskot and J. He, Energy Efficient Radio Access Architectures for Green Radio: Large versus Small Cell Size Deployment, *IEEE Vehicular Technology Conference Fall*, Page 1–5, Anchorage, USA, 2009.

[27] P. Kolios, V. Friderikos and K. Papadaki, Inter-Cell Interference Reduction via Store Carry and Forward Relaying, *IEEE Vehicular Technology Conference Fall*, Page 1–5, Ottawa, Canada, 2010.

[28] I. Ku, C. Wang and P. M. Grant, Impact of Receiver Interference Cancellation Techniques on Base Station Transmission Energy in MIMO Systems, *IEEE Global Communication Conference*, Miami, USA, 2010.

[29] S.Videv and H. Haas, Energy-Efficient Scheduling and Bandwidth-Energy Efficiency Trade-Off with Low Load, *IEEE International Conference on Communications*, Kyoto, Japan, 2011.

[30] O. Arnold, F. Richter, G. Fettweis and O. Blume, Power Consumption Modelling of Different Base Station Types in Heterogeneous Cellular Networks, *Future Network and MobileSummit 2010 Conference Proceedings*, 2010.

[31] K. Mimis, K.A. Morris, J.P. McGeehan, A 2GHz GaN Class-J Power Amplifier for Base Station Applications, *IEEE Radio and Wireless Symposium*, Phoenix, USA, 2011.

[32] T. Origuchi, Contribution to D3 (deliverable 3) Network Energy Efficiency Metrics and Related Requirements, *ITU Focus Group on ICT and Climate Change, Deutsche Telekom AG*, Germany, 2009.

[33] Y. Song, C. Zhang and Y. Fang, Minimum Energy Scheduling in Multi-Hop Wireless Networks with Retransmissions, *IEEE Transactions on Wireless Communications*, Volume 9, Issue 1, Page 348–355, 2010.

[34] J. Kim and D. Cho, Resource Allocation Scheme for Minimizing Power Consumption in OFDMA Cellular Systems, *IEEE Vehicular Technology Conference Fall*, Page 1862–1866, Baltimore, USA, 2007.

[35] F. Chen, J. Cao, Z. Tan and L. Xu, Combined resource allocation for multi-user OFDM systems, *IET 2nd International Conference on Wireless, Mobile and Multimedia Networks*, Beijing, China, 2008.

[36] F. Meshkati, H. V. Poor and R. V. Balan, Energy-Efficient Power and Rate Control with QoS Constraints: A Game-Theoretic Approach, *International Conference on Wireless Communications and Mobile Computing*, USA, 2006.

[37] P. Agyapong, H. Haas, A. Tyrrell and G. Auer, Interference Tolerance Signaling Using TDD Busy Tone Concept, *IEEE Vehicular Technology Conference Spring*, Page 2850–2854, Dublin, Ireland, 2007.

[38] D.J. Goodman, Trends in Cellular and Cordless communications, *IEEE Communication Magzine*, Page 31–40, 1991.

[39] W.C.Y. Lee, Mobile Cellular Telecommunications, Analog and Digital Systems, *New York McGraw-Hill*, 1995.

[40] A. Brand and H. Aghvami, Multiple Access Protocals for Mobile Communications, *John Wiley & Sons Ltd.*, 2002.

[41] H. Holma and A.Toskala, LTE for UMTS, OFDMA and SC-FDMA Based Radio Access, *John Wiley & Sons Ltd*, 2009.

[42] R. Struzak, Radio Propagation Basics, *The Abdus Salam International Centre for Theoretical Physics ICTP*, Trieste, Italy, 2005. URL: http://wireless.ictp.trieste.it/school_2005/lectures/struzak/R_Propg_Basics.pdf.

[43] WINNER II Channel Models, Part I: Channel Models, *WINNER Deliverable D1.1.2*, 2007. URL: http://www.ist-winner.org

[44] LTE_Channel_ITU (ITU Downlink EVM Channel Model), *Agilent Technologies*, URL: http://cp.literature.agilent.com/litweb/pdf/ads2008/3gpplte/ads2008/LTE_Channel_ITU_(ITU_Downlink_EVM_Channel_Model).html.

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

[46] X. Liu and A. J. Goldsmith, Optimal Power Allocation over Fading Channels with Stringent Delay Constraints, *IEEE Internation conference on Communications 2002*, Volume 3, Page 1413–1418, 2002.

[47] A. J. Goldsmith and P. P. Varaiya, Capacity of Fading Channels with Channel Side Information, *IEEE Transactions of Information Theory*, Volume 43, Issue 6, Page 1986–1992, 1997.

[48] J. M. Wozencraft and B. Reiffen, Sequential Decoding, *Cambridge MA: MIT Press*, 1961.

[49] P. Elias, Coding for Noisy Channels, *IRE Convention Record*, 1955.

[50] G. Ungerboeck, Channel Coding with Multilevel/Phase Signals, *IEEE Transactions on Information Theory*, Volume 28, Issue 1, Page 55–56, 1982. URL: http://engrwww.usask.ca/classes/EE/814/Papers/ungerboeck82.pdf.

[51] H. Imai and S. Hirakawa, A New Multilevel Coding Method using Error Correcting Codes, *IEEE Transactions of Information Theory*, Volume 23, Issue 2, Page 371–377, 1977.

[52] E. Zehavi, 8-Psk Trellis Codes for a Rayleigh Channel, *IEEE Transactions on Communications*, Volume 40, Issue 5, Page 873-C884, 1992.

[53] C. Berrou, A. Glavieux, and P. Thitimajshima, Near Shannon Limit Error Correcting Coding and Decoding: Turbo Codes, *Proceedings of IEEE Internatioanl Conference on Communications*, Page 1064–1070, Geneva, Switzerland, 1993.

[54] R. G. Gallager, Low Density Parity Check Codes, *Cambridge MA: MIT Press*, 1962.

[55] 3GPP TR 25.858 Physical Layer Aspects of UTRA High Speed Downlink Packet Access, *3GPP Specification series*, 2001. URL: http://www.3gpp.org/ftp/Specs/archive/25_series/25.848.

[56] MIMO OFDM Testbed for future mobile networks, *Institute for Applied Radio System Technology*, URL: http://iaf-bs.de/projects/3GLTE-mimo-ofdm-testbed.en.html.

[57] P. Viswanath, D. Tse and R. Laroia, Opportunistic Beamforming Using Dumb Antennas, *IEEE Transactions of Information Theory*, Volume 48, Issue 6, Page 1277–1294, 2002.

[58] K. Humblet, Information Capacity and Power Control in Single-Cell Multiuser Communications, *IEEE International Conference on Communications*, Volume 1, Page 331–335, Seattle, USA, 1995.

[59] E. Biglieri, J. Proakis and and S. Shamai, Fading Channels: Information-Theoretic and Communications Aspects, *IEEE Transactions of Information Theory*, Volume 44, Issue 6, Page 2619-2692, 1998.

[60] T. Zhang, Z. Zeng, C. Feng, J. Zheng and D. Ma, Utility Fair Resource Allocation Based on Game Theory in OFDM System, *IEEE International Conference on Computer Communications and Networks*, Honolulu, USA, 2007.

[61] H. Boche, M. Schubert, N. Vucic and S. Naik, Non-Symmetric Nash Bargaining Solution for Resource Allocation in Wireless Networks and Connection to Interference Calculus, *Proceedings of the 15th European Signnal Processing Conference*, Poznan, Poland, 2007.

[62] Z. Han, Z. Ji, and K. J. R. Liu, Fair Multiuser Channel Allocation for OFDMA Networks Using Nash Bargaining Solutions and Coalitions, *IEEE Transactions on Communications*, Volume 53, Page 1366–1376, 2005.

[63] R. Jain, D. Chiu and W. Hawe, A quantitative measure of fairness and discrimination for resource allocation in shared computer systems, *DEC Research Report TR-301: Networking and Internet Architecture*, 1984. URL: http://www.cs.wustl.edu/ jai n/papers/ftp/fairness.pdf.

[64] Extended Cell DTX for Enhanced Energy-Efficient Network Operation, *Ericsson*, Korea, 2009. URL: http://ftp.3gpp.org/specs/html-info/TDocExMtg–R1-59–27294.htm.

[65] Y. Akita and K. Sega, Test strategies for 2 by 2 MIMO in 802.11n systems, *Tektronix*, 2007. URL: http://www.eetimes.com/design/microwave-rf-design/4012917/Test-strategies-for-2-by-2-MIMO-in-802-11n-systems.

[66] Spatial Multiplexing Techniques and Antenna Diversity for Multiuser MIMO Capacity Improvement, *Andromida*, 2011. URL: http://hubpages.com/hub/spatial-multiplexing-spatial-diversity.

[67] B. Muquet, E. Biglieri, A. Goldsmith and H. Sari, MIMO Techniques for Mobile WiMAX Systems, *Sequans Communications*, URL: http://www.sequans.com/wp-conte nt/uploads/mimo_whitepaper.pdf.

[68] H. Bertoni, Radio Propagation for Modern Wireless Systems, *Prentice Hall Professional Technical Reference*, 1999.

[69] H. Zhang and H. Dai, Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks, *EURASIP Journal on Wireless Communications and Networking*, Volume 2004, Issue 2, Page 222–235, 2004.

[70] J. Li, B. Khaled and Z. Cao, Co-channel interference cancellation for space-time coded OFDM systems, *IEEE Transactions on Wireless Communications*, Volume 2, Issue 1, Page 41–49, 2003.

[71] Mapping the Global Future, *US National Intelligence Council*, 2004. URL: http://www.foia.cia.gov/2020/2020.pdf.

[72] T. Edler and S. Lundberg, Energy Efficiency Enhancements in Radio Access Networks, *Ericsson Review Issue no. 01*, 2004. URL: http://www.ericsson.com/ericsson/corpinfo/publications/review/2004_01/f iles/2004015.pdf

[73] D. O'Neill, A. Goldsmith and S. Boyd, Cross-layer design with adaptive modulation: delay, rate and energy tradeoffs, *IEEE Global Communication Conference*, New Orleans, 2008.

[74] Sustainable Energy Use in Mobile Communications: White Paper, *Ericsson Research Group EAB-07:021801 Uen Rev C*, 2007. URL: http://www.ericsson.com/technology/whitepapers/sustainable_energy.pdf

[75] K. Werbach, Radio Revolution: The Coming Age of Unlicensed Wireless, *New America Foundation: Public Knowledge*, 2004. URL: http://werbach.com/docs/R adioRevolution.pdf.

[76] T. View, Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms, *IEEE Wireless Communications (see also IEEE Personal Communications)*, Volume 12, Issue 4, Page 50–58, 2005.

[77] H. Poor, Signal processing across the layers in wireless networks, *European Signal Processing Conference*, Firenze, 2006.

[78] A. Gjendemsjø, D. Gesbert, G. Øien and S. Kiani, Optimal power allocation and scheduling for two-cell capacity maximization, *International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, Volume 2, Page 1–6, Boston, 2006.

[79] D. Gesbert, S. Kiani, A. GjendemsjØ and G. Øien, Adaptation, coordination, and distributed resource allocation in interference-limited wireless networks, *In Proceedings of IEEE*, Volume 95, Issue 12, Page2393–2409, 2007.

[80] C. Zhou, M. Honig and S. Jordan, Utility-based power control for a two-cell CDMA data network, *IEEE Trasactions on Wireless Communications*, Volume 4, Issue 6, Page 2764–2776, 2005.

[81] F. Meshkati, H. Poor and S. Schwartz, Energy-efficient resource allocation in wireless networks: an overview of game-theoretic approaches, *IEEE Signal Processing Magazine*, Volume 24, Page 58–68, 2007.

[82] S. Sinanovic, N. Serafimovski, H. Haas and G. Auer, System spectral efficiency analysis of a 2-link ad hoc network, *In Proceedings of the IEEE Global Telecommunications Conference*, Page 3684–3688, Washington, D.C., 2007.

[83] P. Agyapong, H. Haas, A. Tyrrell and G. Auer, Interference tolerance signaling using TDD busy tone concept, *IEEE Vehicular Technology Conference*, Page 2850–2854, Dublin, 2007.

[84] R. Etkin, A. Parekh and D. Tse, Spectrum sharing for unlicensed bands, *IEEE Journal on Selected Areas in Communications*, Volume 25, Page 517–528, 2007.

[85] G. Miao, N. Himayat, Y. Li and D. Bormann, Energy efficient design in wireless OFDMA, *IEEE International Conference on Communications*, Page 3307-3312, Beijing, 2008.

[86] J. Chuang, N. Sollenberger, B. Res and R. Bank, Uplink power control for TDMA portable radio channels, *IEEE Transactions on Vehicular Technology*, Volume 43, Issue 1, Page 33–39, 1994.

[87] Z. Beyaztas, A. Pandharipande and D. Gesbert, Optimum power allocation in a hierarchical spectrum sharing scheme, *IEEE International Conference on Communications*, Page 91–101, Beijing, 2008.

[88] R. Knopp and P. Humblet, On coding for block fading channels, *IEEE Transactions on Information Theory*, Volume 46, Issue 1, Page 189–205, 2000.

[89] 3GPP TSG-RAN WG1 Further Analysis of Soft Frequency Reuse Scheme, *Huawei*, UK, 2005. URL: http://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_42/Docs.

[90] Y. Choi, C. Kim and S. Bahk, Flexible Design of Frequency Reuse Factor in OFDMA Cellular Networks, *IEEE International Conference on Communications*, Page 1784–1788, Istanbul, 2006.

[91] M. Piscgella and J. C. Belfiore, Achieving a Frequency Reuse Factor of 1 in OFDMA Cellular Network with Cooperative Communications, *IEEE 67th Vehicular Technology Conference Spring*, Page 653–657, Singapore, 2008.

[92] 3GPP TSG-RAN R1-061374 Downlink Inter-Cell Interference Coordination/Avoidance of Frequency Reuse, *Ericsson*, China, 2006. URL: http://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_45/Docs.

[93] X.Fan, S. Chen and X. Zhang, An Inter-cell Interference Coordination Technique Based on Users Ratio and Multi-Level Frequency Allocations, *IEEE International Conference on Wireless Communications, Networking and Mobile Computing*, Page 799–802, China, 2007.

[94] B. Fan, Y. Qian, K. Zheng and W. Wang, A Dynamic Resource Allocation Scheme Based on Soft Frequency Reuse for OFDMA Systems, *IEEE International Symposium on Microwave, Antenna Propagation, and EMC Technologies for Wireless Communications*, Page 1–4, Hangzhou, 2007.

[95] J. Heo, I. Cha and K. Chang, A Novel Transmit Power Allocation Algorithm Combined with Dynamic Channel Allocation in Reuse Partitioning-based OFDMA/FDD System, *IEEE International Conference on Communications*, Page 5654–5659, Istanbul, 2006.

[96] Z. Li, G. Zhu, W. Wang and J. Song, Improved Algorithm of Multiuse Dynamic Subcarrier Allocation in OFDM System, *IEEE International Conference on Communication Technology Proceedings*, Volume 2, Page 1144–1147, 2003.

[97] G. Miao, N. Himayat and G.Y. Li, Energy-Efficient Link Adaptation in Frequency-Selective Channels, *IEEE Transactions on Communications*, Volume 58, Issue 2, Page 545–554, 2010.

[98] M.C. Gursoy, On the Capacity and Energy Efficiency of Training-Based Transmissions over Fading Channels, *IEEE Transactions on Information Theory*, Volume 55, Issue 10, Page 4543–4567, 2009.

[99] S. Videv, H. Haas, J. Thompson and P. Grant, Energy-Efficient Scheduling and Bandwidth-Energy Efficiency Tradeoff with Low Load, accepted by *IEEE International Conference on Communications*, Kyoto, Japan, 2011.

[100] C. Bontu and E. Illidge, DRX Mechanism for Power Saving in LTE, *IEEE Communications Magazine*, Volume 47, Issue 6, Page 48–55, 2009.

[101] S. Tsao , Y. Chen, Energy-efficient packet scheduling algorithms for real-time communications in a mobile WiMAX system, *Computer Communications*, Volume 31, Issue 10, Page 2350–2359, 2008.

[102] F. Zhu, Y. Wu and Z. Niu, Queueing Delay and Energy Efficiency Analyses of Sleep Based Power Saving Mechanism, *IEICE Transactions on Communications*, Volume E93–B, Issue 4, Page 1069–1072, 2010.

[103] M.A. Marsan, L. Chiaraviglio, D. Ciullo and M. Meo, Optimal Energy Saving in Cellular Access Networks, *IMDEA Networks Spain*, 2009. URL: http://www.networks.imdea.org/Portals/8/Downloads/Publications/Optimal-energy-savings-in-cellular-access-networks-2009-en.pdf.

[104] J. Gong, S. Zhou, Z. Niu and P. Yang, Traffic-Aware Base Station Sleeping in Dense Cellular Networks, *18th International Workshop on Quality of Service*, Page 1–2, 2010.

[105] 3GPP TSG-SA5 Additional Information for the Rationale, *Vodafone*, Ireland, 2009. URL: ftp://ftp.3gpp.org/tsg_sa/WG5_TM/TSGS5_66/Docs.

[106] I. Ashraf, F. Boccardi and L. Ho, Power Savings in Small Cell Deployments via Sleep Mode Techniques, *IEEE 21st International Symposium on Personal, Indoor and Mobile Radio Communications Workshops*, Page 307–311, Istanbul, 2010.

[107] A. Mader, J. Zhou, L. Fan, M. Einhaus and T. Ikeda, Power Saving Mode for Femtocell Base Stations, *IEEE 802.16 Broadband Wireless Access Working Group, NEC*, 2008.

[108] L. Saker, S.E. Elayoubi and H.O. Scheck, System Selection and Sleep Mode for Energy Saving in Cooperative 2G/3G Networks, *IEEE 70th Vehicular Technology Conference Fall*, Anchorage, 2009.

[109] 3GPP TSG-RAN WG1: Benefits of Blind Subframes in Rel-9, *Qualcomm*, China, 2009. URL: http://ftp.3gpp.org/Specs/html-info/TDocExMtg–R1-58–27293.htm.

[110] Green Radio Project Definition Document, *Mobile VCE*, 2009. URL: http://www.mobilevce.com/dloads-mem/C5_PDD_GreenRadio.pdf.

[111] T. Edler and S. Lundberg, Energy Efficiency Enhancements in Radio Access Networks, 2004. URL: http://www.ericsson.com/ericsson/corpinfo/publications/review/2004_01/files/2004015.pdf.

[112] H. Karl, An Overview of Energy-Efficiency Techniques for Mobile Communication Systems, *Telecommunication Networks Group, Technical University Berlin, Tech. Rep.*, 2003.

[113] G. Fischer,Next-Generation Base Station Radio Frequency Architecture,*Bell Labs Technical Journal*, Volume 12(2), Page 3–18, 2007.

[114] F. Richter, A.J. Fehske, and G.P. Fettweis, Energy Efficiency Aspects of Base Station Deployment Strategies for Cellular Networks, *IEEE 70th. Vehicular Technology Conference Fall*, Anchorage, 2009.

[115] O. Arnold, F. Richter, G. Fettweis and O. Blume, Power Consumption Modeling of Different Base Station Types in Heterogeneous Cellular Networks, *Future Network and Mobile Summit 2010 Conference Proceedings*, 2010. URL: http://www.vodafone-chair.com/publications/2010/Arnold_O_ICTsummit_10.pdf.

[116] A. Urquhart, Macrocellular Base Station Basics, *Nortel Research*, MVCE Researcher Education Day, 2009.

[117] Vodafone UK Corporate Responsibility 2008, *Vodafone*, 2008. URL: http://online.vodafone.co.uk/dispatch/Portal/SimpleGetFileServlet?dDocName=VF009883&revisionSelectionMethod=latestReleased&inline=0.

[118] T. Gill, 3G Traffic, *Email from Vodafone Group*, 2 July, 2009.

[119] J. Zyren and W. McCoy, Overview of the 3GPP Long Term Evolution Physical Layer, *Freescale Semiconductor*, 2007. URL: http://www.freescale.com/files/wireless_comm/doc/white_paper/3GPPEVOLUTIONWP.pdf.

[120] P. Mogensen, W. Na, I.Z. Kovács, F. Frederiksen, A. Pokhariyal, KI. Pedersen, T. Kolding, K. Hugl and M. Kuusela, LTE Capacity Compared to the Shannon Bound, *IEEE 65th. Vehicular Technology Conference*, Dublin, Ireland, 2007.

[121] F. Mullany, OPERA-Net: Optimising Power Efficiency in Mobile Radio Networks, *Mobile VCE Education Day Presentation*, 2009.

[122] R. Love, R. Kuchibhotla, A. Ghosh, R. Ratasuk, B. Classon and Y. Blankenship, Downlink Control Channel Design for 3GPP LTE, *IEEE Wireless Communications and Networking Conference*, Las Vegas, USA, 2008.

[123] C. Papadias, On the Spectral Efficiency of Space-Time Spreading Schemes for Multiple Antenna CDMA Systems, *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers*, Volume 1, Page 639–643, 1999.

[124] Antenna Port Precoding in LTE, *Steepest Ascent Ltd.*, 2010. URL: http://www.steepestascent.com/content/mediaassets/html/LTE/Help/PDCCH.html.

[125] S. Sesia, I. Toufik and M. Baker, LTE- The UMTS Long Term Evolution: From Theory to Practice, *John Wiley & Sons Ltd*, 2009.

[126] B. Bougard, G. Lenoir, A. Dejonghe, L. Perre, F. Catthoor and W. Dehaene, Smart MIMO: An Energy Aware Adaptive MIMO-OFDM Radio Link control for next generation wireless local area networks, *EURASIP Journal on Wireless Communications and Networking*, Volume 2007, Issue 3, Page 1–15, 2007.

[127] H. Kim, C.B. Chae, G. Veciana and R.W. Heath, A Cross-Layer Approach to Energy Efficiency for Adaptive MIMO Systems Exploiting Spare Capacity, *IEEE Transactions on Wireless Communications*, Volume 8, Issue 8, Page 4264–4275, 2009.

[128] Energy Saving Techniques to Support Low Load Scenarios, *Huawei*, USA, 2010. URL: http://ftp.3gpp.org/specs/html-info/TDocExMtg–R1-60–28038.htm.

[129] P. Bhat, 3GPP Energy Efficiency considerations, *Vodafone*, URL: http://www.mobilevce.com/membersonly/index.html, 2010.

[130] D. Laselva , F. Capozzi, F. Frederiksen, K.I. Pedersen, J. Wigard and I.Z. Kovacs, On the Impact of Realistic Control Channel Constraints on QoS Provisioning in UTRAN LTE, *IEEE 70th Vehicular Technology Conference Fall*, Anchorage, 2009.

[131] E. Dahlman, S. Parkvall, J. Skold and P. Beming, 3G Evolution (Second edition): HSPA and LTE for Mobile Broadband, *Acamedic Press*, 2008.

[132] Hybrid Automatic Repeat Request Indicator Channel, *Steepest Ascent Ltd.*, 2010. URL: http://www.steepestascent.com/content/mediaassets/html/LTE/Help/PHICH.html

[133] F.R. Farrokhi, G.J. Foschini, A. Lozano and R.A. Valenzuela, Link-Optimal Space-Time Processing with Multiple Transmit and Receive Antennas, *IEEE Communications Letters*, Volume 5, Issue 3, Page 85–87, 2001.

[134] P. Wright, J. Lees, J. Benedikt, P.J. Tasker and S.C.Cripps, A Methodology for Realizing High Efficiency Class-J in a Linear and Broadband PA *IEEE Transactions on Microwave Theory and Techniques*, Volume 57, Issue 12, Page 3196–3204, 2009.

[135] LS on Intra-eNB Energy Saving Solutions, *Ericsson*, Spain, 2010. URL: http://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_62/Docs.

[136] M. Schellmann, L. Thiele, V. Jungnickel and T. Haustein, A Fair Score-Based Scheduler for Spatial Transmission Mode Selection, *Conference Record of the Forty-First Asilomar Conference on Signals, Systems, and Computers*, Volume 1, Page 1961–1966, 2007.