



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Scheduling and Link Adaptation for Uplink SC-FDMA Systems - A LTE Case Study

Calabrese, Francesco Davide

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Calabrese, F. D. (2009). *Scheduling and Link Adaptation for Uplink SC-FDMA Systems - A LTE Case Study*.
Aalborg Universitet.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Scheduling and Link Adaptation for Uplink SC-FDMA Systems

A LTE Case Study

PhD Thesis

by

Francesco Davide Calabrese



A dissertation submitted to
the Faculty of Engineering, Science and Medicine of Aalborg University
in partial fulfillment for the degree of
Doctor of Philosophy.
Aalborg, Denmark
April 2009

Supervisor:

Preben Elgaard Mogensen, PhD,
Professor, Aalborg University, Denmark.

Co-supervisors:

Klaus Ingemann Pedersen, PhD,
Senior Wireless Network Specialist, Nokia Siemens Networks, Aalborg, Denmark.
Claudio Rosa, PhD,
Wireless Network Specialist, Nokia Siemens Networks, Aalborg, Denmark.

Assessment Committee:

Olav Tirkkonen, PhD,
Professor, Helsinki University of Technology, Finland.
Ingo Viering, PhD,
Founder and CEO at Nomor Research GmbH.
Troels Bundgaard Sørensen, PhD,
Associate Professor, Aalborg University, Denmark.

Defence Moderator:

Flemming B. Frederiksen,
Associate Professor, Aalborg University, Denmark.

To my family.

Abstract

Long Term Evolution (LTE) is a beyond 3G system conceived with the objective of providing substantial advances in terms of data rates, Quality of Service (QoS) provisioning and cost reduction for users and operators with respect to currently available 3G systems. This project focuses on a subset of Radio Resource Management (RRM) functionalities within the context of LTE up-link and specifically on Power Control (PC), Adaptive Transmission Bandwidth (ATB) and Packet Scheduler (PS). PC and ATB are link adaptation techniques whose role is to adapt the power and transmission bandwidth to the time-varying nature of the channel. The PS entity, instead, multiplexes users in time and frequency domain based on a large variety of parameters like channel conditions and QoS requirements.

In a LTE system the orthogonality of users within the same cell, realized via the Orthogonal Frequency Division Multiplexing (OFDM) scheme, eliminates the intra-cell interference leaving the PC functionality with the task of limiting the inter-cell interference - due to a frequency reuse factor of one - and the terminal power consumption. A PC algorithm therefore is proposed which, compared to the open-loop form of the standardized PC formula, introduces the dependency from the user generated interference when setting the transmitting power. The results show a considerable performance gain that can be used to improve the outage user throughput (by more than 50%) or the cell capacity (by more than 15%) by proper parameter tuning. Such performance boost is preserved only in a full (infinite) buffer traffic whereas a finite buffer traffic scenario considerably reduces the gain due to the different distribution of users in the cell and consequently the different interference patterns.

The main contribution of this research project pertains the design of the PS entity. In LTE up-link the adoption of the Single Carrier - Frequency Domain Multiple Access (SC-FDMA) scheme limits the flexibility of the allocation schemes. The problem is initially bypassed by designing a simple scheduler which assumes a Fixed Transmission Bandwidth (FTB). This gives the possibility to quantify the frequency selectivity and multi-user diversity gains obtained from a channel-aware approach compared to a channel-blind one. In a second phase a scheduling algorithm (novel to the knowledge of the author) is designed to integrate the flexibility of ATB into the scheduler. Such algorithm is shown throughout the thesis to provide flexibility in terms of inbuilt adaptation to cell load, user power limitations and QoS requirements when driven by appropriate scheduling metrics. Moreover the comparison of FTB- and ATB-based schedulers reveals interesting viewpoints regarding the principles of multi-user diversity gain.

In the last part of the thesis the design of the scheduling framework is completed by the addition of a time domain unit motivated by control channel limitations, computational complexity and QoS requirements like Guaranteed Bit Rate (GBR). The focus is therefore shifted on the design of time and frequency domain metrics which are modified as to take into account additional requirements, e.g. the GBR, and different traffic types, e.g. the mix of Constant Bit Rate (CBR) and Best Effort (BE) traffics. The proposed metrics are shown to effectively prioritize the users based on their traffic type and GBR requirement while keeping a similar cell throughput performance.

Dansk Resumé¹

Systemet "Long Term Evolution (LTE)" er et nyt system efter 3G, som er designet til at tilbyde væsentlig højere datahastighed, garanteret servicekvalitet, og til en lavere pris end tidligere 3G systemer. Dette projekt omhandler radio ressource management til LTE uplink med fokus på power kontrol og pakke schedulering med variabel båndbredde per bruger. Power kontrol og variabel båndbredde er link teknikker, som tilpasser sig det tids-varierende radiomiljø. Pakke scheduleren har til opgave at multiplexe brugerne i tid- og frekvens domænet, baseret på en række parametre og målinger. LTE er baseret på OFDMA teknikken, så der ikke er noget interference mellem brugere i samme celle, hvilket betyder, at power kontrollen primært har til opgave at kontrollere interferencen mellem celler – under antagelse af fuld frekvensgenbrug – og maksimum transmissions power per terminal. En åben loop power kontrol algoritme, som bl.a. afhænger af interferencen genereret til naboceller er derfor blevet undersøgt. Resultatet af disse undersøgelser viser stort potentiale, med mere end 50% højere datahastighed for brugere på cellekanten. Førnævnte forbedring er under antagelse af en såkaldt "Full buffer" trafik model, mens der er væsentlig mindre forbedringer, hvis en "Finite Buffer" trafik model benyttes. Hovedformålet med dette forskningsprojekt er design af pakke scheduleren. Uplink af LTE er baseret på SC-FDMA, hvilket sætter nogle begrænsninger mht. fleksibilitet, når brugere allokeres båndbredde. Som et første skridt er der blevet designet en scheduler med konstant transmissionsbåndbredde per bruger. Studierne med denne scheduler giver mulighed for at undersøge multi-bruger diversity fordele ved at allokerne brugerne der, hvor de har de bedste radiobetingelser i frekvens-domænet. I den efterfølgende fase generaliseres scheduleren til at allokerne forskellige båndbredder til brugerne. En sådan scheduler har større fleksibilitet og er bedre til at tilpasse sig under forskellige forhold som for eksempel trafikfordeling, servicekvalitets krav, osv. Sammenligningen af pakke scheduleren med konstant og variabel båndbredde viser nogle interessante effekter mht. multi-bruger diversity. I den sidste del af Ph.D. rapporten videreudvikles pakke scheduleren yderligere, så servicekvalitets krav som f.eks. garanteret datahastighed for brugeren tages med i beslutningsprocessen. Pakke scheduleren fungerer også i et scenario, hvor brugerne har forskellige trafikmodeller, med forskellig garanteret datahastighed for nogle af brugerne. De opnåede resultater viser, at den foreslæde pakke scheduler virker og bl.a. er i stand til prioritere mellem forskellige klasser af brugere og kan opfylde kravene fra brugere med minimum garanteret data hastighedskrav.

¹Translation by Klaus I. Pedersen and Jytte Larsen, Nokia Siemens Networks, Aalborg, Denmark.

Preface and Acknowledgments

This dissertation is the result of a three years research project carried out at the Radio Access Technology (RATE) section, Center for TeleInFraStruktur (CTIF), Institute of Electronic Systems, Aalborg University, Denmark, under the supervision and guidance of Professor Preben E. Mogensen (Aalborg University, Denmark), Dr. Klaus I. Pedersen (Nokia Siemens Networks, Aalborg, Denmark), and Dr. Claudio Rosa (Nokia Siemens Networks, Aalborg, Denmark). The dissertation has been completed in parallel with the mandatory course work, teaching, and project work obligations in order to obtain the PhD degree. This research project has been co-financed by Aalborg University and Nokia Networks R&D, Aalborg.

I am sincerely grateful to my supervisors Preben E. Mogensen, Klaus I. Pedersen and Claudio Rosa. Their technical knowledge and their clear ideas regarding the research direction to take has made my journey so much easier. Even more importantly, they never failed to show me their human understanding of personal issues and to give advices that felt just right even in situations that little or nothing had to do with work. I felt very privileged being supported by such a “trio” and I cannot be grateful enough for that.

I'd also like to thank my colleague and cherished friend Mohmmad Anas which carried out his PhD project in a similar period and in the same research area leading to a fruitful collaboration and several common publications. The reciprocal support we offered each other in the last three years has been very important in improving our research, moreover sharing joys and pains throughout has made the ride much more fun.

I'd also like to thank other researchers which have enabled an interesting collaboration and many interesting discussions including, in no specific order, Malek Boussif, Per-Henrik Michaelsen, Jens Steiner, Mads Brix, Guillaume Monghal, Jeroen Wigard, Daniela Laselva, Troels Kolding, Frank Frederiksen, Carlos Úbeda Castellanos.

On the administrative side the support of our beloved secretaries Lisbeth Schiønning Larsen and Jytte Larsen has made my life so simple as to forget the nightmares of bureaucracy I often experienced in Italy.

There are of course many more names, too many to distinguish individually, which have contributed to the current work. So I'd like to thank all the past and present colleagues at Radio Access Technology Section, Aalborg University and Nokia Siemens Networks, Aalborg for their support.

Sincere and heartfelt thanks go to my truly amazing friends Guillaume and Oumer. I feel blessed having met them here in Aalborg and having spent so much time talking about projects, passions and meaning of life. I hope I will have the chance to return them the kindness, the support and the patience they have always shown me.

I'd also like to express my gratitude to my lovely girlfriend Aki, who brought a new creative dimension to every aspect of my life, challenged so many of my old assumptions and made me feel very special in so many occasions.

Finally, I'd like to thank my brother Tito and my parents Maria Domenica and Pasquale for their infinite affection, patience and love: I feel I'm the luckiest brother and son in the world.

Abbreviations and Mathematical Conventions

Abbreviations and mathematical conventions used in the thesis are listed below for quick reference. The abbreviations are additionally defined at their first occurrence.

Abbreviations

- 16-QAM** 16-Quadrature Amplitude Modulation
3GPP 3rd Generation Partnership Project
64-QAM 64-Quadrature Amplitude Modulation
aGW Access Gateway
AC Admission Control
ACK Acknowledgement
AMBR Aggregate Maximum Bit Rate
AMC Adaptive Modulation and Coding
APG Average Path Gain
ARP Allocation Retention Priority
ARQ Automatic Repeat reQuest
ATB Adaptive Transmission Bandwidth
AVI Actual Value Interface
BCH Broadcast Channel
BE Best Effort
BLEP Block Error Probability
BLER BLock Error Rate
BS Base Station
BSR Buffer Status Report
CAZAC Constant Amplitude Zero AutoCorrelation
CBR Constant Bit Rate
CC Chase Combining
CDMA Code Division Multiple Access
CLPC Closed Loop Power Control
CQI Channel Quality Information
CSI Channel State Information
DL Downlink
DRX Discontinuous Reception
DL-SCH Downlink Shared Channel

ELIISE Efficient Layer II Simulator for E-UTRAN

eNode-B Evolved Node-B

EESM Exponential Effective SINR Mapping

EMG Elastic with Minimum Guarantee

EPC Evolved Packet Core

EPS Evolved Packet System

FD Frequency Domain

FDD Frequency Division Duplex

FDM Frequency Domain Multiplexing

FDPS Frequency-Domain Packet Scheduling

FFT Fast Fourier Transform

FPC Fractional Power Control

FTB Fixed Transmission Bandwidth

GBR Guaranteed Bit Rate

HARQ Hybrid Automatic Repeat reQuest

HDTV High Definition TV

HO Handover

HSDPA High-Speed Downlink Packet Access

HSPA High-Speed Packet Access

HSUPA High-Speed Uplink Packet Access

IPC Interference based Power Control

IR Incremental Redundancy

ISI Inter-Symbol Interference

ISD Interference Spectral Density

KPI Key Performance Indicator

LA Link Adaptation

LC Load Control

LMMSE Linear Minimum Mean Square Error

LTE Long Term Evolution

MAC Medium Access Control

MBMS Multimedia Broadcast Multicast Service

MCH Multicast Channel

MCS Modulation and Coding Scheme

MIESM Mutual Information Effective SINR Mapping

MIMO Multiple Input Multiple Output

MRC Maximal Ratio Combining

NACK Non-Acknowledgement

NR Noise Rise

OFDM Orthogonal Frequency Division Multiplexing

OFDMA Orthogonal Frequency Division Multiple Access

OLLA Outer Loop Link Adaptation

OLPC Open Loop Power Control

PAPR Peak-To-Average Power Ratio

PBCH Physical Broadcast Channel

PBR Prioritized Bit Rate

PC Power Control

PCFICH Physical Control Format Indicator Channel
PCH Paging Channel
PDN Packet Data Network
PDP Power Delay Profile
PDCCCH Physical Downlink Control Channel
PDP Power Delay Profile
PDSCH Physical Downlink Shared Channel
PHICH Physical Hybrid ARQ Indicator Channel
PDU Protocol Data Unit
PF Proportional Fair
PHR Power Headroom Report
PMCH Physical Multicast Channel
PRB Physical Resource Block
PS Packet Scheduler
PSD Power Spectral Density
PRACH Physical Random Access Channel
PUCCH Physical Uplink Control Channel
PUSCH Physical Uplink Shared Channel
QCI Quality Class Identifier
QoS Quality of Service
QPSK Quadrature Phase Shift Keying
RACH Random Access Channel
RAD Required Activity Detection
RAN Random
RBG Radio Bearer Group
RC Resource Chunk
RLC Radio Link Control
RNC Radio Network Controller
RR Round Robin
RRM Radio Resource Management
SC Single Carrier
SC-FDMA Single Carrier - Frequency Domain Multiple Access
SAW Stop-And-Wait
SINR Signal-to-Interference-plus-Noise Ratio
SISO Single Input Single Output
SR Scheduling Request
SRS Sounding Reference Signal
TBS Transport Block Size
TD Time Domain
TTI Transmission Time Interval
TTW Throughput To Wideband
UE User Equipment
UL Uplink
UL-SCH Uplink-Shared Channel
VoIP Voice Over Internet Protocol
WCDMA Wideband Code Division Multiple Access

Contents

1 Thesis Introduction	1
1.1 Introduction to UTRAN LTE	2
1.2 LTE Radio Interface	3
1.2.1 OFDM Transmission Technology	3
1.2.2 SC-FDMA Radio Access Scheme	3
1.3 LTE Radio Resource Management	3
1.4 General Considerations on Link Adaptation and Packet Scheduling	4
1.5 Thesis Scope and Objectives	5
1.6 Scientific Methods Employed	6
1.7 Novelty and Contributions	7
1.8 Thesis Outline	9
2 Uplink Radio Resource Management in LTE	11
2.1 Introduction	11
2.2 QoS and Associated Parameters	11
2.3 Transport and Physical Channels	12
2.4 Admission Control	14
2.5 Packet Scheduling	14
2.6 HARQ	16
2.7 Link Adaptation	17
2.7.1 Adaptive Modulation and Coding	17
2.7.2 Outer Loop Link Adaptation	18
2.7.3 Adaptive Transmission Bandwidth	20
2.7.4 Power Control	20
2.8 Uplink Signaling for Scheduling and Link Adaptation Support	21
2.8.1 Channel State Information	21
2.8.2 Buffer Status Reports	23
2.8.3 Power Headroom Reports	24
2.9 Summary	24
3 Interference based Power Control	27
3.1 Introduction	27
3.2 Open Loop Fractional Power Control	28
3.3 Interference Based Power Control Algorithm	31
3.4 Modeling Assumptions	33
3.5 Performance Evaluation	33
3.5.1 Full Buffer	34
3.5.2 Finite Buffer	38
3.6 Considerations on Inband Inter-User Interference	39

3.7 Conclusions	40
4 Fixed Transmission Bandwidth based Packet Scheduling	43
4.1 Introduction	43
4.2 Packet Scheduling Framework	44
4.3 Matrix-Based Search Algorithm	45
4.4 Modeling Assumptions	45
4.5 Performance Evaluation of Matrix-Based Search Algorithm	47
4.6 Tree-Based Search Algorithm and Results	56
4.7 Conclusions	58
5 Adaptive Transmission Bandwidth based Packet Scheduling	59
5.1 Introduction and Motivation	59
5.2 Algorithm Description	59
5.3 Modeling assumptions	61
5.4 Performance Evaluation	61
5.5 Conclusions	71
6 Scheduling for Elastic Traffic with Minimum Throughput Guarantee	73
6.1 Introduction and Motivation	73
6.2 Time and Frequency Domain PF-like Metrics	74
6.2.1 Metric Symbols	74
6.2.2 TD Metrics Definition	74
6.2.3 FD Metrics Definition	75
6.2.4 Performance Evaluation	76
6.3 Elastic Traffic with Identical Minimum Throughput Guarantee under Balanced Load	83
6.3.1 QoS-aware Metrics	83
6.3.1.1 Aggregate minimum capacity requirement lower than average sector throughput	84
6.3.1.2 Aggregate minimum capacity requirement higher than average cell throughput	86
6.4 Poisson Arrival and Admission Control	87
6.4.1 Elastic Traffic with Different Minimum Throughput Guarantees	91
6.5 Conclusions	95
7 Scheduling for Service Differentiation	97
7.1 Introduction	97
7.2 TD Metric Design	98
7.3 Modeling Assumptions	99
7.4 Performance Evaluation	100
7.4.1 Impact of scheduling metric on system performance	100
7.4.2 Impact of buffer knowledge on system performance	105
7.5 Conclusions	108
8 Overall Conclusions and Recommendations	109
8.1 Interference-based Power Control	109
8.2 Channel-Aware Scheduling	110
8.3 Scheduling for Service Differentiation	111
8.4 Topics for Future Research	112

A Semi-Static System Level Simulator Description	113
A.1 Semi-Static System Simulator	113
A.2 Link-to-System Performance Mapping	114
A.3 Traffic Models	116
A.3.1 Best Effort	116
A.3.2 Constant Bit Rate	116
A.4 Poisson process	116
A.5 Key Performance Indicators	117
A.6 Acknowledgment	118
B SC-FDMA link level performance and assumed propagation conditions	119
B.1 Introduction	119
B.2 SC-FDMA Link Level Performance	119
B.3 Propagation Conditions	121
C Statistical Significance Assessment and Convergence	123
C.1 Introduction	123
C.2 Modeling Assumptions	123
C.3 Results	123
C.4 Convergence to Steady-State: Assumptions and Results	126
Bibliography	129

Chapter 1

Thesis Introduction

Introduction

As of April 2008 the number of mobile subscribers worldwide has passed the 3 billions mark and it keeps growing at a rate of 15 new subscribers per second [1]. The mobile phone, from object of desire has become a vital communication tool playing a fundamental role in our daily life and providing a remarkable social value. Parallel to the growth of mobile subscriptions, the explosion of Internet services and amount of information readily available has shifted the attention from voice services to data services. Moreover the importance of mobile and Internet worlds has led to the conception of services which are mostly a prerogative of mobile users, e.g. location-aware services which provide information which is relevant in the context of the user physical environment.

The plethora of services available, characterized by different response patterns in terms of data rate, delays, error rates, etc. has drawn attention on the importance of providing the required Quality of Service (QoS) as well as on the need of increasing the available capacity. The Wideband Code Division Multiple Access (WCDMA) radio access technology, conceived within the 3rd Generation Partnership Project (3GPP), has been deployed starting from year 2002 to meet such needs as shown in Figure 1.1. In order to preserve competitiveness compared to other technologies, the 3GPP standardization body has proposed an evolution of the WCDMA technology by introducing the High-Speed Downlink Packet Access (HSDPA) in release 5 and High-Speed Uplink Packet Access (HSUPA) in release 6 which provide higher spectral efficiency and data rates. Such evolution, continued with release 7 which introduced technologies like beamforming and Multiple Input Multiple Output (MIMO), is generally denoted as High-Speed Packet Access (HSPA) and informally known as 3.5G and it represents a considerable improvement over the WCDMA technology. Nonetheless, the need to offer advanced solutions over a longer time frame has pushed the 3GPP into initiating a further development, also known as Long Term Evolution (LTE), which introduces significant changes in the radio access interface as well as the network architecture. Such technology, informally known as 3.9G, represents the framework within which this PhD study is carried out.

The rest of the chapter is organized as follows: In Section 1.1 the LTE targets are described. Section 1.2 contains the evolution of the radio interface. Section 1.3 gives an introduction to the Radio Resource Management (RRM) in LTE which then focuses on Link Adaptation (LA) and Packet Scheduler (PS) described in Section 1.4. Section 1.5 formulates the scope and objectives of this study. The scientific methodology used is outlined in Section 1.6. The novelty and contributions of the thesis along with the list of articles published during the PhD study is detailed in Section 1.7. Finally, Section 1.8 lays out the organization of the thesis.

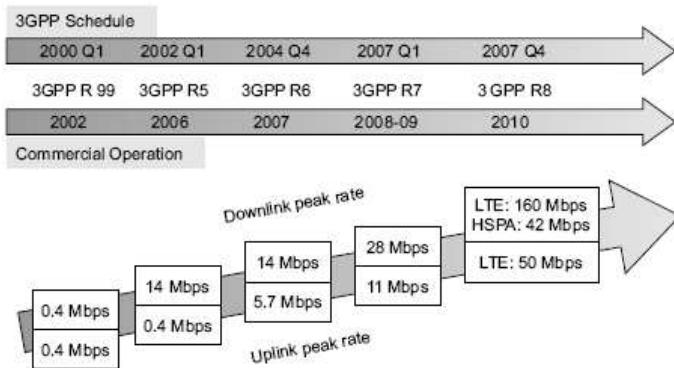


Figure 1.1: Standardization and commercial operation plan along with downlink and uplink peak data rate [2].

1.1 Introduction to UTRAN LTE

UTRAN LTE was started within 3GPP with the aim of creating a technology capable of being competitive in the long-term future by meeting increasing user demands in terms of service provisioning and cost reduction over HSPA. Compared to HSPA release 6, the number of nodes between the air interface and the backbone network has been reduced resulting also in a lower call setup time. As a result, several RRM functionalities, including Admission Control (AC) and mobility control, have been moved to the Evolved Node-B (eNode-B) which assumes a higher importance given the decentralized network architecture.

The targets agreed within 3GPP in the study item phase are enunciated in [3] and summarized as follows:

- Peak data rates exceeding 100 Mbps in the downlink and 50 Mbps in the uplink (in 20 MHz bandwidth).
- Increase in spectral efficiency by a factor of three to four times in downlink and two to three times in uplink [2].
- Significantly reduced control-plane latency as well as user plane latency (10 ms round-trip delay [4]).
- Scalable bandwidth operation up to 20 MHz, i.e., 1.4, 3, 5, 10, 15 and 20 MHz [5, 6].
- Support for packet switched domain only.
- Enhanced support for end-to-end QoS.
- Optimized performance for user speed of less than 15 kmph, and high performance for speeds up to 120 kmph, and the connection should be maintained with speeds even up to 350 kmph [7].
- Reduced cost for operator and end user.

Among the listed targets, the support for only the packet switched domain highlights the focus of LTE to enhance packet based services. The overall goal is to develop an optimized packet based and IP-based access system with high data rate and low latency. Examples of such services include High Definition TV (HDTV) broadcast, movies on demand, interactive gaming and Voice Over

Internet Protocol (VoIP) [8]. In order to achieve such goal an evolution of the network architecture and, more importantly for this research project, of the radio interface is conceived. The latter is going to be described hereafter.

1.2 LTE Radio Interface

Compared to HSPA another significant evolution is the deployment of the Orthogonal Frequency Division Multiplexing (OFDM) multi-carrier transmission scheme made possible by the availability of OFDM transceivers at feasible cost.

1.2.1 OFDM Transmission Technology

The OFDM is a Frequency Domain Multiplexing (FDM) scheme used for the modulation of multi-carrier transmissions. The information data is divided in a set of parallel data streams carried by closely spaced and orthogonal sub-carriers. Each sub-carrier is modulated with a conventional modulation scheme like Quadrature Phase Shift Keying (QPSK) or 16-Quadrature Amplitude Modulation (16-QAM). The low symbol rate makes affordable the use of guard interval between symbols which enables controlling of time-spreading and Inter-Symbol Interference (ISI). The simplified channel equalization is only one of the reasons why the OFDM was chosen as the multi-carrier transmission schemes by LTE and other standardization bodies. Other advantages include high spectral efficiency, efficient implementation via Fast Fourier Transform (FFT), inherent bandwidth scalability, flexibility of bandwidth allocation by varying the number of sub-carriers used for transmissions. The OFDM also shows some disadvantages like sensitivity to frequency synchronization and above all a high Peak-To-Average Power Ratio (PAPR). As the amplitude of the time-domain signal is dependent on hundreds of sub-carriers, large signal peaks will occasionally reach the amplifier saturation region, resulting in a non-linear distortion [9]. The last problem requires high linearity power amplifiers which operate with a large backoff from their peak power suffering from poor power efficiency.

1.2.2 SC-FDMA Radio Access Scheme

Based on the enumerated properties, among the OFDM-based multiple access techniques, the Orthogonal Frequency Division Multiple Access (OFDMA) has been selected for downlink transmission in LTE. Due to the PAPR limitation the same scheme would impose a significant burden in terms of power consumption on the mobile handset resulting in reduced battery life. For this reason the Single Carrier - Frequency Domain Multiple Access (SC-FDMA) (also known as DFT-spread OFDMA) has been selected for the uplink transmission (see Figure 1.2 for a comparison). In this case the subcarriers, due to the DFT-spread operation, are transmitted sequentially rather than in parallel thus resulting in a lower PAPR than OFDMA signals. This produces a higher ISI which the eNode-B has to cope with via frequency domain equalization. Thus the SC-FDMA, while retaining most of the benefits of OFDMA, it also offers reduced power consumption and improved coverage. On the other hand it requires the sub-carriers allocated to a single terminal to be adjacent. This constraint will prove to be very challenging when designing resource allocation schemes.

1.3 LTE Radio Resource Management

The term RRM generally refers to the set of strategies and algorithms used to control parameters like transmit power, bandwidth allocation, Modulation and Coding Scheme (MCS), etc. The aim

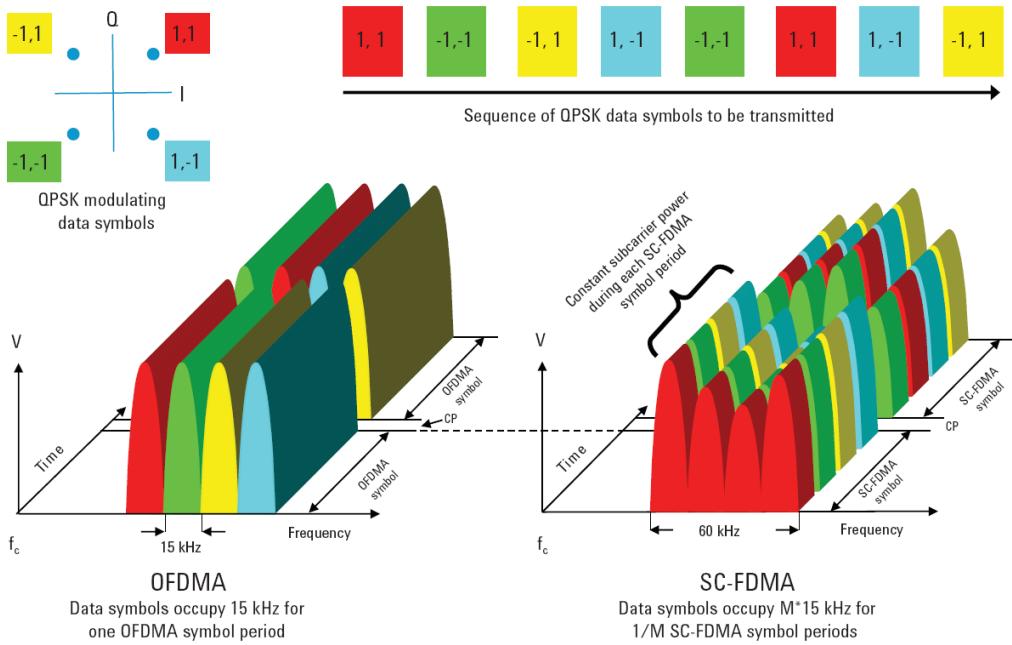


Figure 1.2: Graphical representation of OFDMA and SC-FDMA. Only four subcarriers over two symbol periods are represented. The payload is represented by QPSK data symbols [10].

is to utilize the limited radio resources available as efficiently as possible while providing the users with the required QoS.

The set of RRM functionalities in LTE has known an evolution compared to previous systems: some entities are removed while some others acquire a different role. As an example the AC and the mobility control have been moved to the eNode-B which assumes a higher importance given the decentralized network architecture compared to HSPA. The same process was started already in the HSPA evolution where the PS entity was moved from the Radio Network Controller (RNC) in the WCDMA to the Node-B in the HSPA. The advantage offered by that solution was the possibility to operate at a faster rate (every 2 ms) as consequence of being close to the radio channel, thus providing significant spectral enhancements [11].

The uplink and downlink RRM functionalities, even though they share the same general objective of efficiently utilizing the available radio resources, are mostly treated separately given the different scenarios and effects to be analyzed. The focus of this research project is on the uplink RRM. The considered functionalities include Power Control (PC), Adaptive Modulation and Coding (AMC) and Adaptive Transmission Bandwidth (ATB) (generally termed as LA techniques), scheduling and AC, though a major effort has been directed towards PC, ATB and scheduling in particular.

1.4 General Considerations on Link Adaptation and Packet Scheduling

In a real propagation environment the radio channel is affected by fast fading variations due to the scattering of multiple paths and the constructive or destructive recombination at the receiver. Such variations take place on top of slower fading variations - which change with the location - and distance-dependent path-loss - which changes with the distance of the receiver from the transmitter.

The LA refers to the techniques deployed to adapt the transmission parameters to the time-varying nature of the radio link. The transmission parameters include, for example, the MCS, the transmit power level, the transmission bandwidth [4, 12]. The spectral efficiency can for example be improved by using a more robust MCS under adverse channel conditions and vice versa when the channel conditions improve. AMC, indeed, is known to be the most efficient type of LA in improving the system performance as shown in [13, 14].

LA can take place at a fast or slow rate depending on the kind of fading the system is trying to cope with [15]. Such mechanism requires knowledge of the channel quality information. A finer frequency granularity of such information and a faster rate at which it is updated improve the precision of the LA and therefore the spectral efficiency but it also requires a larger amount of signaling, thus requiring such parameters to be chosen carefully.

The PS is the network entity responsible for the allocation of system resources to users over the shared data channel [16]. In a scenario where multiple users are sharing the same wireless media the PS can improve the spectral efficiency by utilizing the statistical behavior of the radio channel as well as that of offered traffic [17]. Similarly to the LA, the PS relies on the knowledge of the channel quality information to exploit time, space, frequency and multi-user diversity. Of particular interest in this research project is the latter which refers to independent channel fading statistics experienced by different users within the same coverage area. The multi-user diversity gives to the PS entity the possibility to schedule users which are experiencing fading peaks. Packet scheduling has been widely investigated in literature, see, for example, [18] for the Code Division Multiple Access (CDMA) and [11, 19] for the HSPA. In LTE, given the choice of supporting only data transfer, the PS plays a key role in the overall system performance. The aim is not only to improve spectral efficiency while exploiting the different forms of diversity, but also guaranteeing the QoS requirements associated with different traffic types.

Compared to HSPA systems, in LTE the PS is given high importance as reflected by the possibility of operating also in the frequency domain [20]. This opens the possibility for the exploitation of frequency diversity which refers to uncorrelated channel fading statistics for the same user on frequencies which are separated widely enough. While in Downlink (DL) the adoption of the OFDMA makes the exploitation of the frequency diversity possible, in Uplink (UL) the SC-FDMA limits such possibility but still leaves the chance to achieve multi-user diversity.

The importance of the PS in LTE is reaffirmed by the choice of a fully IP-based packet switched technology dictated by the need of providing services which exist in a wide range of data rate, delay and reliability requirements. This include also the support of voice services which will not benefit any longer of the circuit switched technology and will have to rely on the VoIP technology.

As final remark, it is worth to point out that PS and LA entities, being co-located in the eNode-B, can interact with each other in order to improve the radio resource utilization [21].

1.5 Thesis Scope and Objectives

The objective of this work is to develop algorithms for scheduling and LA in order to improve system capacity and QoS provisioning, as well as to analyze their performance in the framework of LTE Uplink.

The evolution of the radio interface, including the addition of the frequency domain and the introduction of a new multiple access scheme, the decision to support only packet based transmissions and new classes of services poses new challenges to the RRM functionalities when it comes to improving spectral efficiency and supporting the different QoS requirements. Of all the available RRM functionalities, the focus, as anticipated, is going to be on the PS and LA, though the interaction with other entities is often taken into account. Specifically, novel algorithms are derived and their performance evaluated at system level. The study is carried out within the con-

text of the LTE framework modeled according to the specifications given in [20] and is limited to the Uplink only. Other RRM functionalities like Handover (HO) and Load Control (LC) are not covered in the thesis while the AC, whose design and analysis has been carried out in a parallel study, will be deployed and be part of the default assumptions in the last part of this project.

In LTE, LA, as previously anticipated, refers to the adaptation of signal transmission parameters like power (via PC), modulation and coding (via AMC) and bandwidth (via ATB) to the varying channel conditions. The scope of such technique is to improve the spectral efficiency and the reliability of the channel [17]. The advantages of the AMC have already been proven in the mentioned works [13, 14] and the evaluation conducted in [15] confirmed that the fast AMC (performed every TTI) is able to improve the system performance over the slow AMC. For this reason it is included as a default assumption throughout the whole study. The focus is therefore restricted to PC (being the ATB considered at a later stage) and the aim is to design and evaluate the gain of an interference-based PC approach over the standard open-loop approach in a variety of deployment scenarios.

Next, the objective of the project is to design and evaluate the performance of different packet scheduling algorithms. Initially the aim is to design a channel-aware scheduling algorithm assuming a fixed-bandwidth allocation and to evaluate the gain over a channel-blind approach. Afterwards the objective is to design a more flexible algorithm which integrates the ATB functionality into the allocation algorithm. Such allocation algorithm will be assumed as default for the rest of the study.

In the last part of the thesis the scheduling framework is completed by the addition of the Time Domain (TD) unit and the attention is shifted on the design and analysis of the metrics for time and frequency domain scheduling for QoS provisioning. The goal here is to gradually expand the design of the metrics until they incorporate additional aspects (like Guaranteed Bit Rate (GBR)) and are able to serve a larger set of scenarios which differs in terms of QoS requirements and traffic types, like a mix of Constant Bit Rate (CBR) and Best Effort (BE).

1.6 Scientific Methods Employed

Given the number of parameters involved and the complex interaction among the system entities the problem is too complex to be approached analytically. For this reason the preferred methodology is to construct a system model following some assumptions and guidelines, like the ones given in [20, 22], and to implement such model in a system level simulator. The results presented in this thesis have therefore been produced via extensive computer simulations using the system model developed and implemented during the course of the project.

Whenever possible the analytical method is used. The system simulator developed is a state of the art semi-static, multi-cell and multi-user simulator. It includes detailed implementation of LA based on AMC and FPC, explicit scheduling of HARQ processes including retransmissions and link-to-system mapping technique suitable for SC-FDMA. A variety of traffic is also implemented including the full infinite buffer, which is recommended by the standard, and more realistic traffic models such as finite buffer and constant bit rate.

The aim is to evaluate the system-level performance of the proposed RRM framework using the Key Performance Indicator (KPI)s listed in A.5 in order to recommend algorithms for practical implementation.

1.7 Novelty and Contributions

The main contribution of this research project is the design and analysis of a RRM framework which merges into scheduling several aspects of LA. The analysis takes into account the complex interaction of the different entities including PC, AMC and AC and is carried out in a variety of scenarios. Thus, the evaluation of the proposed algorithms required work not only in terms of conceptual design but also system modeling, software design, implementation and testing.

The first topic of research to be presented in the thesis is the design of a Interference based Power Control (IPC) algorithm which is based on the Fractional Power Control (FPC) formula agreed in 3GPP [23].

Related to PC the following contribution has been published¹:

- M. Boussif, N. Quintero, F. D. Calabrese, C. Rosa, and J. Wigard, “Interference Based Power Control Performance in LTE Uplink” in *Proceedings of the IEEE International Symposium on Wireless Communication Systems (ISWCS)*, Reykjavik, Iceland, October, 2008.

in such publication the IPC concept is proposed and the gains are evaluated in a full buffer traffic scenario. Additionally a rule for setting some of the parameters involved is derived.

Within the same topic, the following contributions have been co-authored:

- C. U. Castellanos, D. L. Villa, C. Rosa, K. I. Pedersen, F. D. Calabrese, P. H. Michaelsen, and J. Michel, “Performance of Uplink Fractional Power Control in UTRAN LTE” in *Proceedings of the 67th IEEE Vehicular Technology Conference (VTC)*, Singapore, May, 2008.

which evaluates the performance of the pure FPC strategy in different scenarios and suggests appropriate values for the related parameters.

- C. U. Castellanos, F. D. Calabrese, K. I. Pedersen, and C. Rosa, “Uplink Interference Control in UTRAN LTE Based on the Overload Indicator” in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Calgary, Canada, September, 2008.

which proposes and evaluates a method for controlling the P_0 parameter values via the exchange of the Overload Indicator message among eNode-Bs.

The second topic of research is the design of scheduling algorithms. Here, the main focus is on the potential gain achievable by the channel-aware scheduling and considerations regarding the control channel limitations are added only in a second phase. Different kinds of resource allocation algorithms and scheduling metrics are designed and evaluated and conclusions are drawn on which one provides the best performance.

As a result of such studies the following articles have been published:

- F. D. Calabrese, P. H. Michaelsen, C. Rosa, M. Anas, C. U. Castellanos, D. L. Villa, K. I. Pedersen, and P. E. Mogensen, “Search-Tree Based Uplink Channel Aware Packet Scheduling for UTRAN LTE” in *Proceedings of the 67th IEEE Vehicular Technology Conference (VTC)*, Singapore, May, 2008.

which evaluates the performance of two Fixed Transmission Bandwidth (FTB) based resource allocation strategies (matrix based and search-tree based) using the Proportional Fair (PF) metric.

¹Even though the author of this thesis does not appear as first author of the paper, the ideas presented therein have been elaborated and partly evaluated as part of this PhD project.

- F. D. Calabrese, C. Rosa, M. Anas, P. H. Michaelsen, K. I. Pedersen, and P. E. Mogensen, “Adaptive Transmission Bandwidth Based Packet Scheduling for LTE Uplink” in *Proceedings of the 68th IEEE Vehicular Technology Conference (VTC)*, Calgary, Canada, September, 2008.

which proposes an ATB based resource allocation algorithm and compares its performance and properties with the FTB based one.

- F. D. Calabrese, C. Rosa, K. I. Pedersen, P. E. Mogensen, “Performance of Proportional Fair Frequency and Time Domain Scheduling in LTE Uplink” submitted to *15th European Wireless Conference*, Aalborg, Denmark, May 2009.

which adds the time domain scheduling to the framework, analyzes different combinations of time and frequency domain metrics and identifies the one which provides the best performance.

In the last part, the QoS aspects are also taken in consideration. Specifically users with finite buffer and a GBR requirement are considered and different time domain metrics are analyzed. The performance is analyzed both with and without the AC functionality. This investigation sets the ground for the analysis of more complex scenarios in terms of traffic types and GBR requirements. Those scenarios include users with different GBR requirements as well as mix of BE and CBR traffic types.

The results of these studies have been partly published in the following conference contribution:

- M. Anas, C. Rosa, F. D. Calabrese, K. I. Pedersen, and P. Mogensen, “Combined Admission Control and Scheduling for QoS Differentiation in LTE Uplink” in *Proceedings of the 68th IEEE Vehicular Technology Conference (VTC)*, Calgary, Canada, September, 2008.

In addition the collaborative work on the design of RRM functionalities for LTE Uplink has resulted in the following publications whose results are used in the research project but are not specifically discussed or analyzed:

- C. Rosa, D. V. Lopez, C. U. Castellanos, F. D. Calabrese, P. H. Michaelsen, K. I. Pedersen, P. Skov, “Performance of Fast AMC in E-UTRAN Uplink” in *Proceedings of the IEEE International Conference on Communications (ICC)*, Beijing, China, May, 2008.
- M. Anas, C. Rosa, F. D. Calabrese, P. H. Michaelsen, K. I. Pedersen, and P. Mogensen, “QoS-Aware Single Cell Admission Control for UTRAN LTE Uplink” in *Proceedings of the 68th IEEE Vehicular Technology Conference (VTC)*, Singapore, May, 2008. [Maybe to be attached as appendix]

Finally, the following publications have been produced using a dynamic system level simulator developed in the first year of this PhD and are not included in the thesis.

- F. D. Calabrese, M. Anas, C. Rosa, K. I. Pedersen, and P. E. Mogensen, “Performance of a Radio Resource Allocation Algorithm for UTRAN LTE Uplink” in *Proceedings of the 65th IEEE Vehicular Technology Conference (VTC)*, Dublin, Ireland, April, 2007.
- M. Anas, F. D. Calabrese, P. Mogensen, C. Rosa, K. I. Pedersen, “Performance Evaluation of Received Signal Strength Based Hard Handover for UTRAN LTE” in *Proceedings of the 65th IEEE Vehicular Technology Conference (VTC)*, Dublin, Ireland, April, 2007.

- M. Anas, F. D. Calabrese, P. E. Östling, K. I. Pedersen, P. Mogensen, “Performance Analysis of Handover Measurements and Layer 3 Filtering for UTRAN LTE” in *Proceedings of the 18th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Athens, Greece, September, 2007.

A significant part of the PhD project has been devoted to the modeling, implementation and testing of the semi-static system-level simulator - from which the results presented in this thesis have been generated - as well as the dynamic system-level simulator named Efficient Layer II Simulator for E-UTRAN (ELIISE), whose results have not been presented in the thesis as they followed an entirely different line of thought than the one subsequently adopted. The main contribution to the semi-static system-level simulator has revolved around the development of the scheduling framework, with emphasis on the ATB-based algorithm, as well as the different time and frequency domain metrics (both channel-aware and QoS-aware). The development of such simulator has been carried out in collaboration with other Nokia and Nokia Siemens Networks colleagues. The contribution to the dynamic system-level simulator ELIISE involved the overall design of the software from scratch as well as the modeling and implementation of features including PC, AMC, Outer Loop Link Adaptation (OLLA), scheduling and mobility. The development of this simulator has been carried out in collaboration with PhD student Mohammad Anas.

1.8 Thesis Outline

The PhD thesis is organized as follows:

- Chapter 2: *Overview of Uplink Radio Resource Management in LTE* - This chapter presents an overview of the uplink RRM functionalities in LTE. Particular attention is given to the description of PS and LA entities.
- Chapter 3: *Interference based Power Control* - This chapter describes a new algorithm for power control based on measures of the interference produced by each user within the system. Simulation results are presented mainly assuming a full buffer best effort type of traffic. This research topic has been carried out in collaboration with Nokia Siemens Networks colleague Malek Boussif to whom I recognize 50% of the work. Specifically, the author of this thesis has contributed the concept and the algorithm modeling while the software implementation and the generation of results (via simulation) has been carried out by Malek Boussif.
- Chapter 4: *Fixed Transmission Bandwidth Based Packet Scheduling* - This chapter first describes a matrix-based resource allocation algorithm and presents results in a variety of scenarios. Afterwards the same algorithm is generalized to use a tree-based approach. The obtained results are compared with the simpler matrix-based approach.
- Chapter 5: *Adaptive Transmission Bandwidth Based Packet Scheduling* - This chapter presents a packet scheduling algorithm which integrates the ATB functionality in the resource allocation algorithm. The proposed algorithm is analyzed via simulation and the results are compared with the ones obtained under FTB and presented in the previous chapter. The suggested algorithm is shown to provide flexibility compared to the cell load.
- Chapter 6: *Scheduling for Elastic Traffic with Minimum Throughput Guarantee* - In this chapter the full scheduling framework, which includes time and frequency domain scheduling, is deployed. An investigation of different time and frequency domain metrics is carried out and some recommendations are given based on the presented results. Afterwards the

GBR is also considered and the metrics are adapted to serve such new requirement. A comparison of different metrics is shown first without AC and afterwards with AC. The AC deployed is studied by PhD student Mohammad Anas and the related published article is reprinted at the end of the thesis.

- Chapter 7: *Scheduling for Service Differentiation* - In this chapter the packet scheduling framework is tested in a traffic mix scenario which include BE users and CBR users. The results presented demonstrate the ability of the proposed metric to effectively prioritize the CBR users over the BE users which are allocated only the resources not used by the CBR users.
- Chapter 8: *Overall Conclusions and Recommendations* - This chapter provides a summary of the overall study and discusses future research issues.
- Appendix A: *Semi-Static System Level Simulator Description* - This appendix provides the detailed description of the semi-static multi-cell system level simulator including network layout, channel model, traffic model, link-to-system level mapping and definition of important KPIs.
- Appendix B: *Statistical Significance Assessment and Convergence* - This appendix presents the analysis of statistical significance of KPIs for representative simulation scenarios taken from the study.

Chapter 2

Uplink Radio Resource Management in LTE

2.1 Introduction

This chapter presents an overview of the LTE system architecture and the Uplink RRM functionalities. The entities which are not of interest for the project are not described or only briefly mentioned while the packet scheduling and the LA functionalities investigated in this research project are introduced in this chapter and then described in more details in their dedicated chapters.

Following a top-down approach, first the setting of QoS parameters at the bearer level is described in Section 2.2. Afterwards an overview of transport and physical channels, useful for a better understanding of later sections, is given in 2.3. The RRM functionalities are also described following a top-down approach: Section 2.4 gives a description of the channel-aware and QoS-aware AC algorithm developed in a parallel PhD study by Mohammad Anas [24]. The PS entity is introduced in Section 2.5 and treated in more details in Chapter 4 while Section 2.6 describes the modeling of Hybrid Automatic Repeat reQuest (HARQ) and the related assumptions. Section 2.7 describes the set of functionalities generally grouped under LA, including AMC, OLLA, ATB and PC. Finally Section 2.8 contains a description of the uplink signaling from the User Equipment (UE) to the eNode-B, specifically Channel State Information (CSI), Buffer Status Report (BSR) and Power Headroom Report (PHR). Section 2.9 contains a summary of the chapter.

2.2 QoS and Associated Parameters

An Evolved Packet System (EPS) bearer is the level of granularity for bearer level QoS control in the Evolved Packet Core (EPC)/E-UTRAN. One EPS bearer is established when the user connects to a Packet Data Network (PDN), and that remains established throughout the lifetime of the PDN connection to provide the user with always-on IP connectivity to that PDN. That bearer is referred to as the default bearer. Any additional EPS bearer that is established to the same PDN is referred to as a dedicated bearer. The initial bearer level QoS parameter values of the default bearer are assigned by the network, based on subscription data. The decision to establish or modify a dedicated bearer can only be taken by the EPC, and the bearer level QoS parameter values are always assigned by the EPC.

An EPS bearer is referred to as a GBR bearer if dedicated network resources related to a GBR value that is associated with the EPS bearer are permanently allocated (e.g. by an admission control function in the eNode-B) at bearer establishment/modification. Otherwise, an EPS bearer is referred to as a Non-GBR bearer. A dedicated bearer can either be a GBR or a Non-GBR bearer

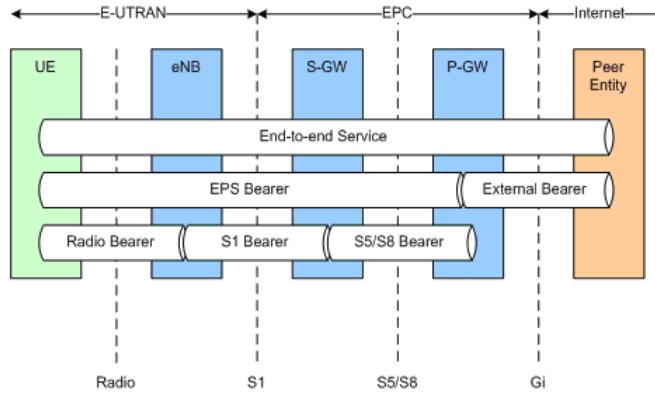


Figure 2.1: EPS Bearer Service Architecture [25].

while a default bearer shall be a Non-GBR bearer.

The EPS bearer service layered architecture is illustrated in Figure 2.1.

Each EPS bearer (GBR and Non-GBR) is associated with the following bearer level QoS parameters signaled from the Access Gateway (aGW) (where they are generated) to the eNode-B (where they are used):

- Quality Class Identifier (QCI): scalar that is used as a reference to access node-specific parameters that control bearer level packet forwarding treatment (e.g. bearer priority, packet delay budget and packet loss rate), and that have been pre-configured by the operator owning the eNode-B. A one-to-one mapping of standardized QCI values to standardized characteristics is captured in [26].
- Allocation Retention Priority (ARP): the primary purpose of ARP is to decide whether a bearer establishment / modification request can be accepted or needs to be rejected in case of resource limitations. In addition, the ARP can be used by the eNode-B to decide which bearer(s) to drop during exceptional resource limitations (e.g. at handover).

Each GBR bearer is additionally associated with the GBR bearer level QoS parameter, which is the bit rate that can be expected to be provided by a GBR bearer. The GBR denotes bit rate of traffic per bearer while Aggregate Maximum Bit Rate (AMBR) denotes a bit rate of traffic per group of bearers.

Additionally there is the Prioritized Bit Rate (PBR), which is set from the eNode-B in uplink for both GBR and non-GBR bearers in order to avoid starvation of low priority flows [25]. It should be noted that PBR is only relevant for users with multiple bearers.

2.3 Transport and Physical Channels

The mapping between the transport channels, which describe how and with what characteristics the data are transferred over the air interface, and the physical channels which corresponds to a set of resource elements carrying information originating from higher layers, takes place between Layer 2 and Layer 1[25, 27].

In downlink, four types of transport channels exist: Broadcast Channel (BCH), Downlink Shared Channel (DL-SCH), Paging Channel (PCH) and Multicast Channel (MCH). The BCH is characterized by a fixed and pre-defined transport format and is required to be broadcasted in the entire coverage area of the cell. The DL-SCH is the most flexible and is characterized by

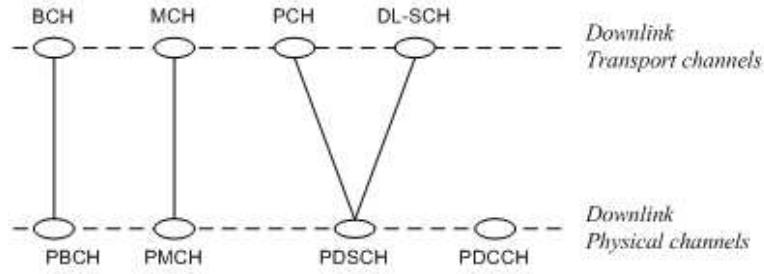


Figure 2.2: Mapping between downlink transport channel and downlink physical channel [25].

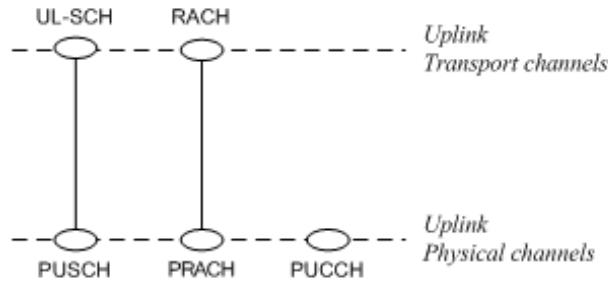


Figure 2.3: Mapping between uplink transport channel and uplink physical channel [25].

support for HARQ, support for dynamic link adaptation via variation of modulation, coding and transmit power, possibility to be broadcasted in the entire cell, support for dynamic and semi-static resource allocation and support for Discontinuous Reception (DRX) to enable UE power saving. The PCH is characterized by support for DRX to enable UE power saving and requirement to be broadcasted in the entire cell. The MCH is required to broadcast in the entire coverage of the cell and offers support for Multimedia Broadcast Multicast Service (MBMS) transmission on multiple cells. The mapping of downlink transport channels to downlink physical channels is represented in Figure 2.2 where only four of the six physical channels are indicated. They are Physical Downlink Shared Channel (PDSCH), Physical Broadcast Channel (PBCH), Physical Multicast Channel (PMCH), Physical Control Format Indicator Channel (PCFICH), Physical Downlink Control Channel (PDCCH), Physical Hybrid ARQ Indicator Channel (PHICH).

In uplink, two types of transport channels exist: Uplink-Shared Channel (UL-SCH) and Random Access Channel (RACH). Like in downlink the UL-SCH is the most flexible and is characterized by support for HARQ, dynamic link adaptation, support for dynamic and semi-static resource allocation. The RACH is used for the initial access to the system, the call setup and the exchange of limited control information. The mapping with uplink physical channels is represented in Figure 2.3. There are three types of physical channels: Physical Uplink Shared Channel (PUSCH), Physical Uplink Control Channel (PUCCH), Physical Random Access Channel (PRACH). The PUSCH carries the UL-SCH, PUCCH carries HARQ Ack/Nack's in response to downlink transmission, scheduling requests, and Channel Quality Information (CQI) reports. Due to the Single Carrier (SC) constraint a user cannot transmit at the same time on PUCCH and PUSCH.

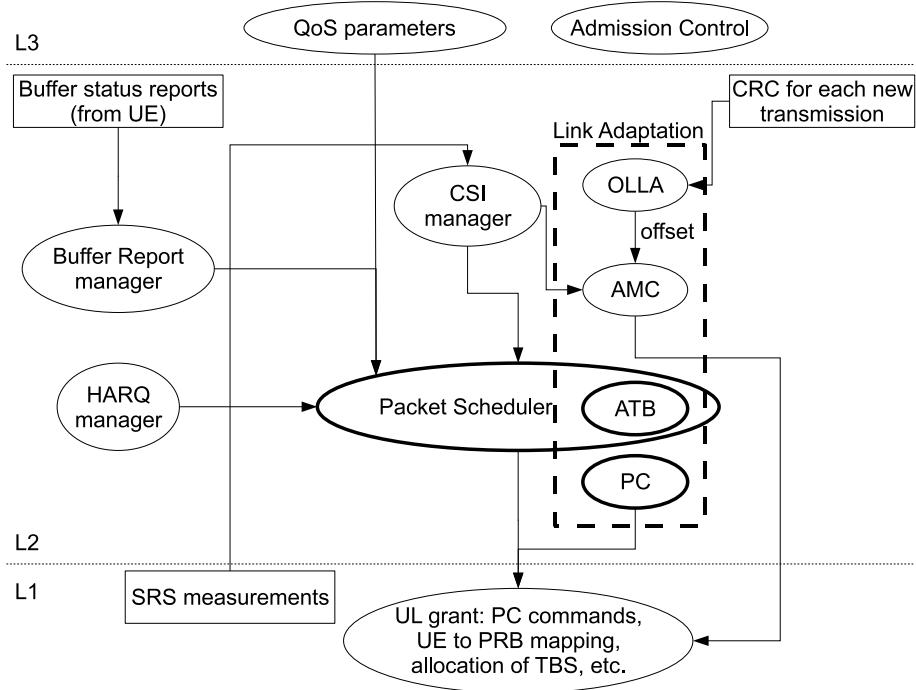


Figure 2.4: Interaction between RRM functionalities with focus on scheduling and adaptation.

2.4 Admission Control

An overview of the functionalities of main interest to this project, their interaction and location in the protocol stack is shown in Figure 2.4.

The AC is a Layer 3 (network layer) functionality located in the eNode-B whose task is to admit or reject the request of a new radio bearer or of an handover candidate. The criterion used to admit or reject a new bearer is based on ensuring an efficient utilization of the available radio resources, by admitting new bearers as long as radio resources are available while at the same time guaranteeing the QoS provisioning of ongoing sessions by rejecting requests which cannot be accommodated [25]. The AC needs therefore to be QoS aware so that a new user can be granted or denied access based on whether the QoS of such user can be fulfilled while guaranteeing the QoS of the existing users. For the same reason the AC for BE users is optional. In order to evaluate the possibility of admitting a new user the AC needs to utilize the information on the local cell load. The eNode-B could also interact with neighboring cells in order to make AC decisions based on multi-cell information. In this project, though, only local cell information is considered. Moreover it is assumed that each user has only a single bearer and only new calls are considered for admission as handover is not implemented due to the lack of mobility.

The QoS-aware AC algorithm adopted in this study uses the user radio channel conditions to make an admission decision and considers the GBR as the only QoS criterion. It is based on the FPC formula agreed within 3GPP [23] and it is described in details in the paper reprinted at the end of the thesis.

2.5 Packet Scheduling

The PS is an entity located in the Medium Access Control (MAC) sublayer which aims at utilizing efficiently the UL-SCH resources. The main role of the PS is to multiplex the users in time and

frequency domain. Such multiplexing takes place via mapping of users to the available physical resources. If the system is affected by time and frequency selective fading the PS entity can exploit the multi-user diversity by allocating the users to the portions of the bandwidth which exhibit favorable channel conditions. In this way the radio channel fading, which used to be a limitation to the performance of wireless system, is turned into an advantage.

In this study the PS has the possibility of performing a mapping of users to Physical Resource Block (PRB)s on a Transmission Time Interval (TTI) basis and is therefore referred to as *fast scheduling*. The data in the frequency domain can be multiplexed via localized or via distributed transmission. In the distributed transmission the subcarriers of one PRB are distributed over the entire frequency band with an equal distance of each other. For this reason it can exploit the *frequency diversity* and offers robustness against frequency selective fading. In the localized transmission, which is the focus of this study, the subcarriers assigned to one user are adjacent with each other. Localized transmission can potentially achieve *multi-user diversity* in the presence of frequency selective fading by assigning each user to a portion of the bandwidth where the user exhibits favorable channel conditions.

The PS is strongly related to LA (in particular ATB) and HARQ functionalities. Indeed, the scheduling decisions need to take into account a large set of factors including payloads buffered in the UE, HARQ retransmissions, estimation of CSI, UE sleep cycles, QoS parameters, etc. Most of these parameters are going to be considered in the thesis though some of them will be introduced only in the last chapters.

The problem that the PS tries to solve can be formulated as an optimization problem. The solution to the optimization problem can be very complex given the number of different parameters involved and the resulting possible combinations. Due to the hard time constraints encountered in a real system and the computational limitations, the focus is going to be on the design of practical framework and scheduling algorithms which aim at solving the problem introducing some simplifications or intelligently reducing the space of possible solutions.

Even following a practical approach, scheduling remains fundamentally an optimization problem and the preferred approach is to adopt a search algorithm which traverses the space of possible solutions in order to identify the optimum or a sub-optimum. In order to make decisions along the search path, the algorithm requires an optimization criterion, in the following referred to as *metric*, which is designed as to take into account different scheduling aspects including channel gain, QoS requirement, traffic type, etc. Assuming that such criterion (the metric) can be evaluated for every user and every PRB, the optimization problem can be formulated as:

$$\text{maximize } M_{sum} = \sum_{i \in \Omega_i, j \in \Omega_j} M_{i,j} A_{i,j} \quad (2.1)$$

subject to:

$$A_{i,j} = \begin{cases} 1 & \text{if user } i \text{ is allocated to PRB } j \\ 0 & \text{otherwise} \end{cases} \quad (2.2)$$

$$\sum_{i \in \Omega_i} A_{i,j} \leq 1, \forall j \in \Omega_j \quad (2.3)$$

$$A_{i,j} - A_{i,(j+1)} + A_{i,k} \leq 1, \text{ for } k = j + 2, j + 3, \dots, |\Omega_j| \quad (2.4)$$

where $M_{i,j}$ is the metric for user i and PRB j , Ω_i is the set of users, Ω_j is the set of PRBs, $A_{i,j}$ a selection variable and $|\Omega_j|$ the cardinality of Ω_j (that is, the total number of PRBs). The inequality (2.3) expresses the orthogonality of the users within the same cell, that is, at most one user can be allocated to a certain PRB. The inequality (2.4) expresses the requirement that the

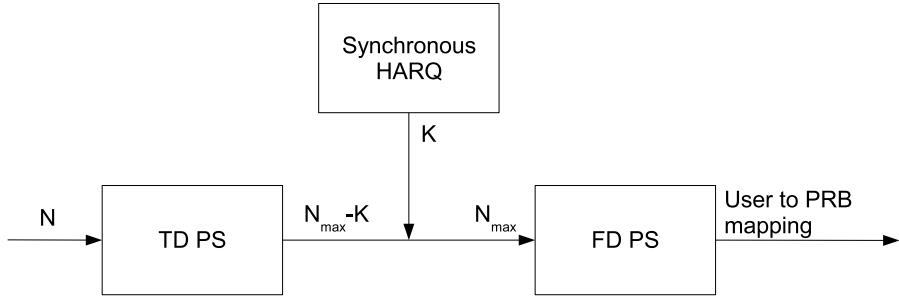


Figure 2.5: General scheduling framework.

PRBs allocated to the same user must be adjacent. If $A_{i,j} = 1$ and $A_{i,(j+1)} = 0$, then $A_{i,k} \leq 0$ for $k > j + 1$. If on the other hand both $A_{i,j} = 1$ and $A_{i,(j+1)} = 1$, the inequality requires that $A_{i,k} \leq 1$. If $A_{i,j} = 0$ then the inequality becomes redundant as it states that $A_{i,k} \leq 1 + A_{i,(j+1)}$.

The problem we are trying to solve involves two design phases: one is the *design of the allocation algorithm*, that is the algorithm that determines the user to PRB mapping; the other is the *design of the metric*, that is the optimization criterion whose value is compared by the allocation algorithm to decide which user to serve first..

The allocation algorithm, in turn, can also be split in two sub-phases which can be indicated as a phase of *user selection* - referred to as TD scheduling - and a phase of *PRB allocation* or *user to PRB mapping* - referred to as Frequency Domain (FD) scheduling (see Figure 2.5). These two steps can actually be performed in one solution by the FD scheduler and this is, indeed, the first simplified approach followed in this work to give an initial system evaluation. In the later stages also the TD scheduler is introduced to take into account limitations due to control channels and computational complexity. The user to PRB mapping phase in uplink is complicated by the presence of the SC-FDMA which limits to a great extent the flexibility of allocation. This aspect of the allocation phase will be vastly discussed in the coming chapters.

The metric has to be designed as to enclose the same principles which are at the base of the RRM, that is, efficient use of the time-frequency resources and QoS provisioning due to the variety of existing services. In case the scheduling is split in TD and FD two metrics are needed, one for the TD and one for the FD. The role of the two metrics is different and so is their design. As general rule of thumb the TD metric is assigned the role of provisioning the required QoS while the FD metric is assigned the role of efficiently using the time-frequency resources. This indication can on the other hand change significantly according to the number of users in the cell and the variety of their QoS requirements. In case of a low number of users in the cell, for example, QoS provisioning may also be needed in FD.

2.6 HARQ

In LTE both retransmission functionalities Automatic Repeat reQuest (ARQ) and HARQ are provided. ARQ provides error correction by retransmissions in acknowledged mode at the Radio Link Control (RLC) sublayer of Layer 2. HARQ is located in the MAC sublayer of Layer 2 and ensures delivery between peer entities at Layer 1 [28]. In case a data packet is not correctly received, the HARQ ensures a fast Layer 1 retransmission from the transmitter (UE). In this way the HARQ provides robustness against LA errors (due, for example, to errors in CSI estimation and reporting) and it improves the reliability of the channel.

The HARQ has the following characteristics:

- It uses a N -process Stop-And-Wait (SAW) protocol between the UE and the eNode-B. Each

process takes place on a different time channel in order to ensure continuous transmission to the UE. Each packet, if not successfully received, is transmitted for a given number of attempts before getting discarded.

- It is based on Acknowledgement (ACK)/Non-Acknowledgement (NACK) messages. This means that data packets are acknowledged after each transmission. A NACK, instead, indicates that a retransmission is requested either in form of Incremental Redundancy (IR) or to enable Chase Combining (CC). The first consists in incrementally sending additional redundancy to facilitate the decoding of the packet. The second consists in sending an identical copy of the packet so that the receiver can obtain an SNR gain by soft combining the information contained in all the transmissions.
- It is synchronous (in uplink) and adaptive, like in dynamic scheduling, or non-adaptive, like in semi-persistent scheduling. Synchronous refers to the fact that retransmissions need to occur at specific time instants while adaptive refers to the possibility of changing transmission parameters like resource allocation and MCS in the subsequent retransmissions.

In case of HARQ retransmissions failure, the ARQ in the RLC sublayer can handle further retransmissions using knowledge gained from the HARQ in the MAC sublayer. The ARQ retransmissions may not be required for example in case of VoIP traffic due to the short delay budget. Such retransmissions are not investigated in this study.

2.7 Link Adaptation

As previously anticipated, the LA is a fundamental functionality in a channel affected by fading. In the following, the mechanisms that control the adaptation of the main transmitting parameters, that is MCS, bandwidth and power, are described.

2.7.1 Adaptive Modulation and Coding

It is well known that AMC can significantly improve the spectral efficiency of a wireless system [12]. The MCS selection algorithm is based on mapping tables which return an MCS format (and hence a Transport Block Size (TBS)) after having received an SINR value and, optionally, the Block Error Rate (BLER) target at first transmission as input. In LTE uplink the supported data modulation schemes are QPSK, 16-QAM and 64-Quadrature Amplitude Modulation (64-QAM) [5].

The expected instantaneous throughput per TTI for a given MCS and SINR can be defined as:

$$T(MCS, SINR) = TBS(MCS) \cdot (1 - BLEP(MCS, SINR))$$

where the Block Error Probability (BLEP) represents the probability that the transmitted block is going to be in error. Different algorithms are possible for the selection of the MCS:

1. Select the MCS which maximizes the expected throughput, that is the throughput calculated using the expected BLEP.
2. Select the MCS which maximizes the throughput under the constraint that the estimated BLEP is smaller or equal than the BLER target at first transmission.
3. Select the MCS which minimizes the difference between the expected BLEP and the BLER target at first transmission.

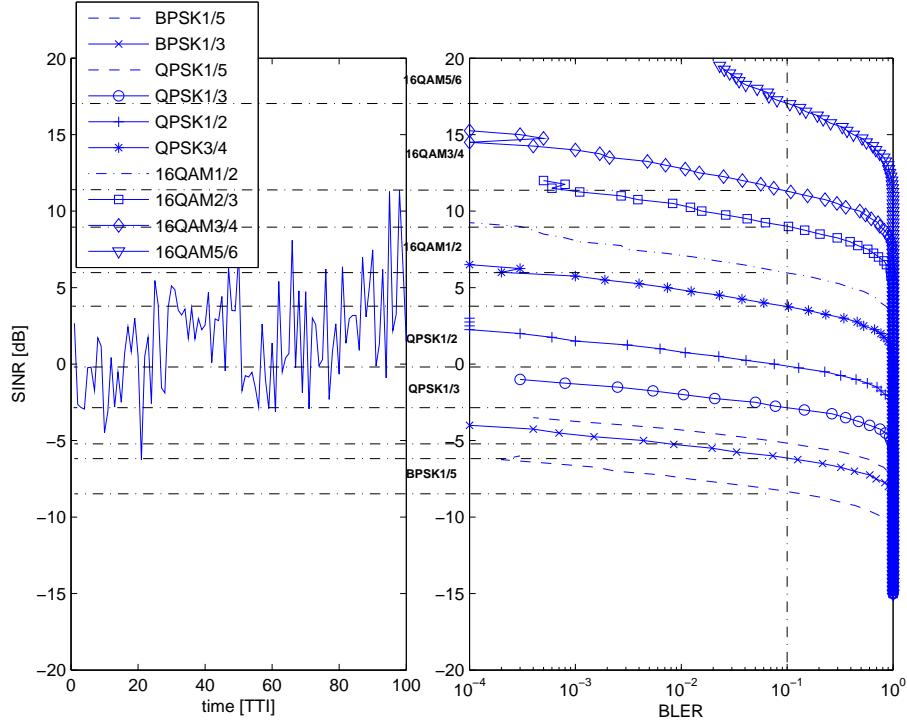


Figure 2.6: AMC mechanism: MCS selection based on estimated SINR.

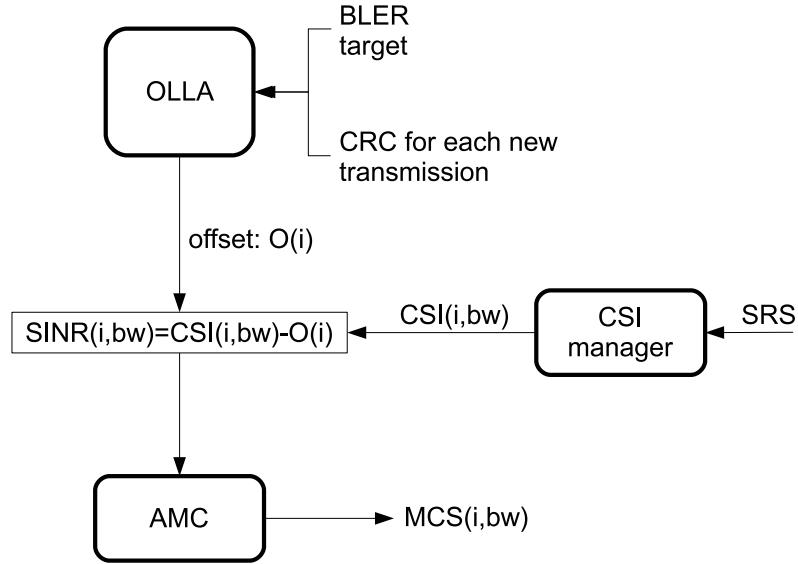
An example of the second approach is given in Figure 2.6. The first approach is the one chosen for this study. The first approach introduces an error which is going to be corrected by the OLLA as long as it falls within the OLLA range.

The AMC can be performed on a slow basis, for example with the same rate of the power control commands to exploit the slowly changing channel variations, or on a faster basis, for example every TTI, to exploit the high instantaneous SNR conditions. A detailed performance analysis of the AMC functionality is carried out in [15] where the fast AMC is shown to exhibit a gain above 20% compared to the slow AMC. For this reason it is selected as default assumption for the rest of this study.

2.7.2 Outer Loop Link Adaptation

The described link adaptation mechanism based on fast AMC is implicitly characterized by different errors. Some examples of fast AMC errors include:

- Bias when combined with channel-aware scheduling, due to the tendency to schedule on PRBs with positive measurement errors.
- Bias due to differences between measures of CSI (that is SINR estimated based on uplink Sounding Reference Signal (SRS) strength measurement) and experienced SINR on the data channel. They are due to:
 - CSI measurement errors.
 - Different ways of calculating CSI and SINR (see equations (2.7) and (2.9)).
 - Different Power Spectral Density (PSD) used for the SRS and for the transmission on PUSCH (not modeled in the present work).

**Figure 2.7:** Interaction of OLLA and AMC.**Table 2.1:** Default settings for the considered OLLA parameters.

Parameter	Setting
$BLER_T$	30 %
Step size (S)	0.5 dB
Min offset (O_{min})	-4 dB
Max offset (O_{max})	4 dB

- Different average level of interference experienced on sounding channel and on PUSCH (not modeled in the present work).

In order to maintain the BLER at first transmission as close as possible to the target an OLLA algorithm is needed to offset the CSI measurements as shown in Figure 2.7 for a user i and a bandwidth bw .

The offset $O(i)$ is adjusted following the same rules of outer loop PC in WCDMA [29]:

1. If a 1st transmission on PUSCH is correctly received, $O(i)$ is decreased by $O_D = S \cdot BLER_T$
2. If a 1st transmission on PUSCH is not correctly received, $O(i)$ is increased by $O_U = S \cdot (1 - BLER_T)$

where S represents the step size and $BLER_T$ the BLER which the algorithm will converge to if the offset $O(i)$ remains within a specified range $O_{min} \leq O(i) \leq O_{max}$.

The BLER target can be expressed as function of O_D and O_U as follows:

$$BLER_T = \frac{O_D}{O_D + O_U} = \frac{1}{1 + \frac{O_U}{O_D}}$$

Table 2.1 summarizes the settings of the OLLA parameters used throughout this thesis.

2.7.3 Adaptive Transmission Bandwidth

In a SC-FDMA system, which inherently enables bandwidth scalability, the adaptability of the transmission bandwidth represents a fundamental feature given the variety of services that an LTE system is called to provide. The ATB, therefore, becomes a necessary technique to cope with different traffic types, varying cell load and power limitation in the UE.

Some services, e.g. VoIP, require a limited amount of bandwidth while a user with BE type of traffic may receive as much bandwidth as it is available as long as there are data in the buffer and power available at the UE. The power limitations also represent a constraint which highlights the importance of the ATB: The PSD a user is required to transmit with may be as high, due to adverse channel conditions, as to limit the user to support only a limited bandwidth. Additionally a varying cell load also calls for the adaptability of the transmission bandwidth as the bandwidth a user can receive depends also on the number of other users in the system.

The ATB is ultimately a functionality which allows the allocation of different portions of bandwidths to different users and therefore offers a significant flexibility when exploited as part of the scheduling process. The integration of the two functionalities, indeed, gives the possibility to better exploit the frequency diversity by limiting the user bandwidth allocation to the set of PRBs which exhibit the largest metric value. The benefits offered by such functionality will be clarified with the results presented in Chapter 5.

2.7.4 Power Control

In a OFDM-based system like LTE, where the orthogonality removes the intra-cell interference and the near-far problem typical of CDMA systems, the role of PC is changed into providing the required Signal-to-Interference-plus-Noise Ratio (SINR) while at the same time controlling the intercell interference. The classic idea of PC in uplink is to modify the user transmit power as to receive all the users with the same SINR at the Base Station (BS). Such idea is known as full compensation of the path-loss. In 3GPP the idea of FPC has been introduced. In this scheme the users are allowed to compensate for a fraction of the path-loss so that the users with higher path-loss will operate with a lower SINR requirement and will likely generate less interference to neighboring cells.

The agreed FPC scheme to set the power on PUSCH is based on an Open Loop Power Control (OLPC) algorithm aiming at compensating for slow channel variations. In order to adapt to changes in the inter-cell interference situation or to correct the path-loss measurements and power amplifier errors, aperiodic close-loop adjustments can also be applied. The user transmit power is set according to the formula (2.5) expressed in dBm [23]:

$$P = \min\{P_{max}, P_0 + 10 \cdot \log_{10} M + \alpha \cdot L + \Delta_{mcs} + f(\Delta_i)\} \quad [dBm] \quad (2.5)$$

where P_{max} is the maximum user transmit power, P_0 is a user-specific (optionally cell-specific) parameter, M is the number of PRBs allocated to a certain user, α is the cell-specific path-loss compensation factor that can be set to 0.0 and from 0.4 to 1.0 in steps of 0.1, L is the downlink path-loss measured in the UE based on the transmit power P_{DL} of the reference symbols [20], Δ_{mcs} is a user-specific parameter (optionally cell-specific) signaled by upper-layers, Δ_i is a user-specific close-loop correction value and the $f(\cdot)$ function performs an absolute or cumulative increase depending on the value of the UE-specific parameter *Accumulation-enabled*.

If the absolute approach is used the user applies the offset given in the PC command using the latest OLPC command as reference. If the cumulative approach is used the user applies the offset given in the PC command using the latest transmission power value as reference. In the latter case Δ_i can take one of four possible values: $-1, 0, 1, 3 \text{ dB}$.

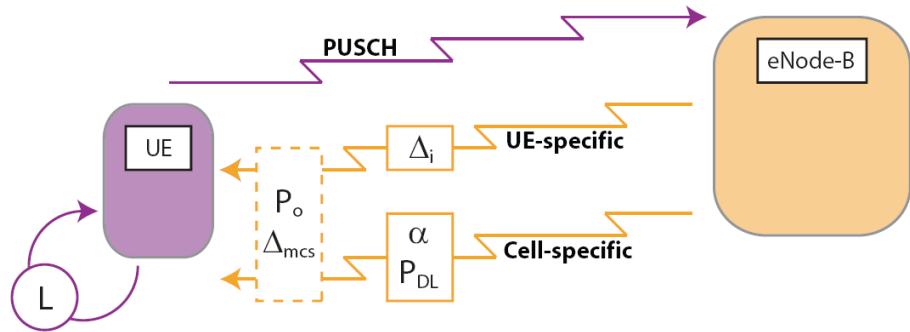


Figure 2.8: Power control signaling

In case the Closed Loop Power Control (CLPC) term is not used, the formula is simplified to include only the open loop terms as indicated in (2.6).

$$P = \min\{P_{max}, P_0 + 10 \cdot \log_{10}M + \alpha \cdot L\} \quad [dBm] \quad (2.6)$$

The exchange of the different signals related to PC is exemplified in Figure 2.8.

2.8 Uplink Signaling for Scheduling and Link Adaptation Support

The PS and LA entities rely on the CSI gathered via SRSs to perform channel-aware scheduling and AMC. Similarly, the allocation of time-frequency resources to users requires knowledge of their buffer status to avoid allocating more resources than are needed. Finally, the knowledge of how close the user is to its maximum transmit power is especially relevant for ATB operations. For this reason it is worth describing in more details the signaling needed to support such operations as simplified in Figure 2.9.

2.8.1 Channel State Information

The uplink CSI can be described as the SINR measurement of the SRS. CSI measurements are used to gain knowledge of the channel and perform fast AMC and Frequency-Domain Packet Scheduling (FDPS).

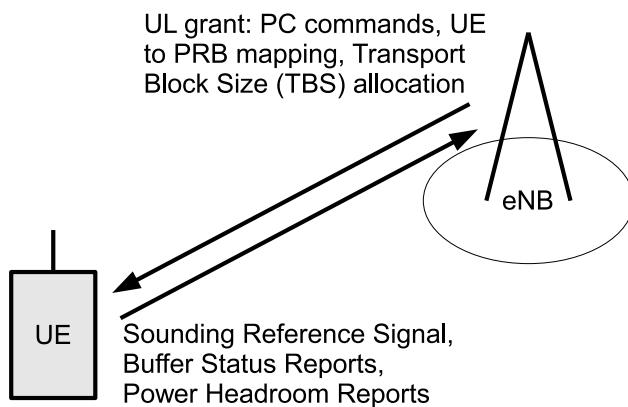


Figure 2.9: Signaling exchange between UE and eNode-B.

The SRS is transmitted over a fraction or over the full scheduling bandwidth. Users in the same cell can transmit in the same bandwidth without interfering with each other thanks to the orthogonality provided by Constant Amplitude Zero AutoCorrelation (CAZAC) sequences and the uplink synchronous transmission. In reality there exists a constraint on the number of users in one cell that can simultaneously sound the same bandwidth without interfering with each other. The PSD on the pilot channel is the same as the one used on the data channel. UE power capabilities typically impose a limit on the sounding bandwidth, or, alternatively, on the level of accuracy of the corresponding SINR measurements. However power limitations for the SRS are not taken into account, that is, the user will always send a SRS over the whole scheduling bandwidth with the same PSD used on the data channel even if that is going to violate the power limitation constraints. Additionally we assume, in this work, that CSI is available at the eNode-B every TTI, over the entire scheduling bandwidth, for all active users in the corresponding cell, and with a given resolution in the frequency domain (which we refer to as CSI granularity). Finally, due to the dynamic scheduling and the variability of the instantaneous interference conditions in uplink the interference component is averaged over a certain time window. This is shown to be beneficial for the channel estimation and consequently for an improvement of average cell throughput and outage user throughput as shown in [15].

The CSI of user i on PRB r at time instant t is modeled as:

$$CSI(i, r, t) = \sum_{m=1}^M \left(\frac{\sum_{r' \in R} S^m(i, r', t)}{\sum_{r' \in R} (\bar{I}^m(b(i), r', t) + N_{PRB})} \right) \cdot 10^{\frac{\epsilon(r, t)}{10}} \quad (2.7)$$

where:

- M is the number of receiving antennas at the eNode-B
- R is the set of simultaneously sounded PRBs to which the PRB r belongs to. The number of simultaneously sounded PRBs depends on the CSI resolution which corresponds to the size of R
- $S^m(i, r', t)$ is the SRS power received from user i at time instant t on PRB r' and antenna m
- $b(i)$ is the serving eNode-B of user i
- $\bar{I}^m(b(i), r', t)$ is the average interference signal power of eNode-B b calculated at time instant t on PRB r' and antenna m
- N_{PRB} represent the thermal noise over the bandwidth of one PRB
- $\epsilon(r, t)$ is a zero mean Gaussian distributed random variable with standard deviation σ_{CSI} introduced to model measurement errors. The random variables $\epsilon(r, t)$ and $\epsilon(r, t + m)$ are uncorrelated for $m \neq 0$. Similarly, the random variables $\epsilon(r, t)$ and $\epsilon(r + s, t)$ are uncorrelated for $s \neq 0$.

The interference component $\bar{I}^m(b(i), r', t)$ is calculated via exponential averaging as:

$$\bar{I}^m(b, r', t) = \eta \cdot I^m(b, r', t) + (1 - \eta) \cdot \bar{I}^m(b, r', t - 1) \quad (2.8)$$

In (2.8) η controls the averaging period of the interference used in CSI measurements.

Table 2.2 shows the assumptions used throughout the thesis in relation to the CSI.

Table 2.2: Default settings for the considered CSI parameters.

Parameter	Setting
CSI frequency resolution	2 PRBs
CSI delay	0 ms
Filter length (η)	100 ms (0.01)
Std. CSI error (σ_{CSI})	1 dB
PSD	Same as on data channel
Max number of simultaneously sounding users	Unlimited

The uplink SINR experienced at the eNode-B by a user u at time instant t , is calculated as:

$$SINR(i, t) = \sum_{m=1}^M \left(\frac{\sum_{r' \in R} S^m(i, r', t)}{\sum_{r' \in R} (I^m(b(i), r', t) + N_{PRB})} \right) \quad (2.9)$$

where

- M is the number of receiving antennas at the eNode-B
- R represents the set of consecutive PRBs on which the user is allocated
- $S^m(i, r', t)$ represents the signal power received from user i at time instant t on PRB r' and antenna m
- $b(i)$ is the serving eNode-B of user i
- $I^m(b(i), r', t)$ is the interference signal power of eNode-B b calculated at time instant t on PRB r' and antenna m
- N_{PRB} represent the thermal noise over the bandwidth of one PRB

The definition of the SINR together with the Actual Value Interface (AVI) represents the interface with the link level results. In downlink the SINR to be used in input to the AVI is computed by using the Exponential Effective SINR Mapping (EESM) model which exploits an estimate of the SINR per subcarrier obtained in a previous transmission or the Mutual Information Effective SINR Mapping (MIESM) [30]. In uplink the SINR per subcarrier is not directly related to the data symbol. This is because of the SC-FDM transmission, which spreads each data symbol over the whole bandwidth, so that, even though every sub-carrier experiences a different channel gain, the differences are averaged out over a sufficiently large bandwidth. Therefore the average power over the transmission bandwidth (i.e., the sum of power over the different PRBs, divided by the number of PRBs) divided by the average interference over the transmission bandwidth is a sufficiently accurate model for the SINR calculation and therefore it has been used as interface with the link level results.

2.8.2 Buffer Status Reports

The Buffer Status reporting procedure is used to provide the serving eNode-B with information about the amount of data available for transmission in the UL buffers of the UE.

A BSR can be triggered in one of three forms:

- “Regular BSR”: The UE buffer has to transmit data belonging to a radio bearer (logical channel) group with higher priority than those for which data already existed in the buffer (this include as special case the situation in which the new data arrive in an empty buffer) or in case of a serving cell change.
- “Padding BSR”: UL resources are allocated and number of padding bits is equal to or larger than the size of the BSR MAC control element.
- “Periodic BSR”: issued when the periodic BSR timer expires.

Only a regular BSR is followed by a Scheduling Request (SR) when the UE is not scheduled on PUSCH in the current TTI. When available, the SR can be transmitted using one dedicated bit on the PUCCH otherwise it is transmitted using the Random Access procedure. A BSR (of any type) is also transmitted when the UE has resources allocated on PUSCH in which case it is transmitted as a MAC control Protocol Data Unit (PDU) with only header, where the length field is omitted and replaced with buffer status information.

BSR are reported on a per Radio Bearer Group (RBG) basis as result of a compromise between the need of differentiation of data flows based on QoS requirements and the need of minimizing the resources allocated for signaling. The number of RBG is limited to 4. Each RBG groups radio bearers with similar QoS requirements.

2.8.3 Power Headroom Reports

Due to the open-loop component of the standardized PC formula (see (2.5)) the eNode-B cannot always know the PSD transmitted by the UE. Such information is important for different RRM operations including the allocation of bandwidth, modulation and coding scheme. Assuming that the eNode-B knows the user bandwidth, the transmission power can be derived from the information on the PSD. For this reason the power headroom reports have been standardized in [31].

The Power Headroom reporting procedure is used to provide the serving eNode-B with information about the difference between the nominal UE maximum transmit power and the estimated power for UL-SCH transmission. A PHR is triggered if any of the following criteria is met:

- A predefined timer expires or has expired and the path-loss has changed more than a predefined threshold since the last power headroom report when the UE has UL resources for new transmission.
- The predefined timer expires.

2.9 Summary

This chapter provides an overview of the uplink RRM in LTE Release 8. After a description of the QoS parameter setting and the mapping of transport to physical channels, the functionalities located in the Layers 3, 2 and 1 of the protocol stack are described following a top-down approach.

First the principles of AC are described. The AC used in this thesis is channel-aware and is used to preserve the QoS of the ongoing calls. Then the PS framework and its interaction with the HARQ is presented. The PS multiplexes the users in time and frequency domain trying to efficiently use the available system resources. The data is then scheduled on the Physical Uplink Shared Channel (PUSCH) based on the uplink grant sent using the Physical Downlink Control Channel (PDCCH).

The LA functionalities including Adaptive Modulation and Coding (AMC), Outer Loop Link Adaptation (OLLA), Adaptive Transmission Bandwidth (ATB) and Power Control (PC) are also

described in detail. The AMC, corrected by the OLLA decides the most suitable MCS for transmission on a per TTI basis. The ATB allows the allocation of different bandwidths to different users and plays an important role in coordination with the PS.

Both the scheduling and LA functionalities rely on the use of signaling including Channel State Information (CSI), Buffer Status Report (BSR) and Power Headroom Report (PHR) which are also described in detail.

Further considerations regarding PS and PC, which are central in this thesis, will be given in the following chapters.

Chapter 3

Interference based Power Control

3.1 Introduction

In a multi-cellular system, the aim of PC is to minimize the interference produced towards users in other cells. This is especially relevant in a system characterized by a reuse factor of 1, as in LTE. Additionally PC aims also at controlling the transmit power of the UE in order to reduce the battery consumption and to limit the dynamic range of the received signal at the base station.

In LTE, there is no a central entity in charge of coordinating the power control parameters of the cells in the system. Rather, in a distributed fashion, the eNode-Bs are each in charge of controlling such parameters (and therefore the transmit power of the users in the cell) with the aim of improving the overall system performance.

The objective of this chapter is to describe the IPC algorithm proposed during this study in [32].

The FPC standardized formula, given in (2.6) and studied in [33] in the open-loop form offers the reference against which the performance of the proposed algorithm is evaluated. The parameter P_0 is chosen as to optimize the 5% outage user throughput for a given α . α , in turn, is chosen as a compromise between cell edge and capacity performance. To facilitate the comparison, the relevant results presented in [33] will be reproduced along the way together with some new and useful considerations.

On the topic of open-loop FPC, the work presented in [34] shows promising results. The FPC algorithm is compared to the traditional approach where all the users are received with the same SINR and to the approach where the PC is absent and all the users transmit at the maximum power. The outcome of the study indicates that the FPC can provide a significant spectral efficiency increase compared to the traditional approach and, at the same time, a remarkable cell edge improvement compared to a full power approach. In addition, the same work shows that the sum of the interference produced by the first and second strongest interferers largely surpasses the sum of interference produced by all the other sectors¹ and thus accounts for most of the interference relevant for the performance.

The work presented in [35] shows an interesting approach in setting the SINR target for the users, where the concept of PC is explicitly intended as a mean for controlling the inter-cell interference. For this reason, the path-loss to the strongest neighboring cell is also taken into account by setting the target SINR for the user as a function of the difference between such a path-loss and the path-loss to the serving cell. This approach provides a significant improvement in both, outage user throughput and average cell throughput. Moreover it leads to a lower variance of the interference level which allows a more reliable channel estimation and hence a more efficient functioning

¹Please note that the terms “sector” and “cell” are used interchangeably throughout the whole thesis.

Table 3.1: Main simulations parameters

Parameter	Setting
# UEs per sector	10
# PRBs per UE	6 (fixed bandwidth)
Cell-level user distribution	Uniform
Scheduling metric	PF
BLER target at 1 st transmission	30%
Propagation scenarios	3GPP Macro case 1 (ISD of 500 m)
Traffic model	Full (infinite) buffer

of the AMC.

The rest of the chapter is organized as follows: The main results and considerations of the FPC formula together with the findings which motivate an interference based approach is described in Section 3.2. The proposed algorithm is described in Section 3.3. The modeling assumptions and the related results are described in Sections 3.4 and 3.5, respectively. Section 3.6 offers some considerations regarding the effects of intra-cell interference. The conclusions are given in Section 3.7.

3.2 Open Loop Fractional Power Control

Before entering into the details of PC, it is worth clarifying the quantity we aim at controlling. Unlike a system such as WCDMA, a user in LTE can transmit using a fraction of the available system bandwidth (in multiples of a PRB). For this reason, it is more appropriate to adjust the PSD rather than the total transmit power. The PSD is assumed to be the same on all the allocated PRBs. Given the constraint on maximum transmit power only a certain number of PRBs can be allocated before the user hits the power limit. Allocating a number of PRBs larger than the ones the user can support leads to the user transmitting at a lower PSD than the one set via the open-loop command.

The FPC formula in the simplified open-loop form is recalled here for clarity.

$$P = \min\{P_{max}, P_0 + 10 \cdot \log_{10} M + \alpha \cdot L\} \quad [dBm] \quad (3.1)$$

where P is the transmitting power of the user, P_{max} the maximum transmit power, M the number of allocated PRBs, L the path-loss to the serving station, P_0 (in $[dBm/PRB]$) a broadcast parameter later discussed and α is a broadcast parameter representing the fraction of path-loss that we aim at compensating for. In the traditional OLPC the objective is to receive all users at the same SINR level. This corresponds to having $\alpha = 1$ in (3.1). At the other extreme we have $\alpha = 0$ which corresponds to having all the users transmitting with the same power which results in poor cell edge performance. By setting a different value for the α factor we can lower the SINR of the users at the cell edge and possibly decrease the interference generated towards other cells.

The effect of P_0 and α is shown in Figure 3.1 under the assumptions listed in Table 3.1². A variation of P_0 determines a curve shift while a variation of α not only shifts the curve but also changes its variance. We can better understand the relation between P_0 and α by plotting (see Figure 3.2) the PSD vs the path-loss. The parameters have been adjusted as to keep a similar

²A more complete list of the default parameter used throughout the thesis is available in Appendix A in Table A.1.

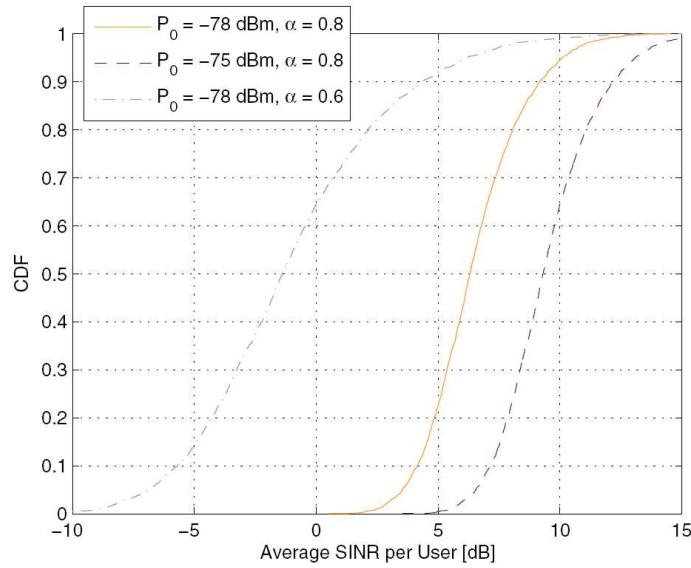


Figure 3.1: CDF of the average SINR per user for different P_0 and α . Fixed transmission bandwidth with $M = 6$ PRBs per user is assumed.

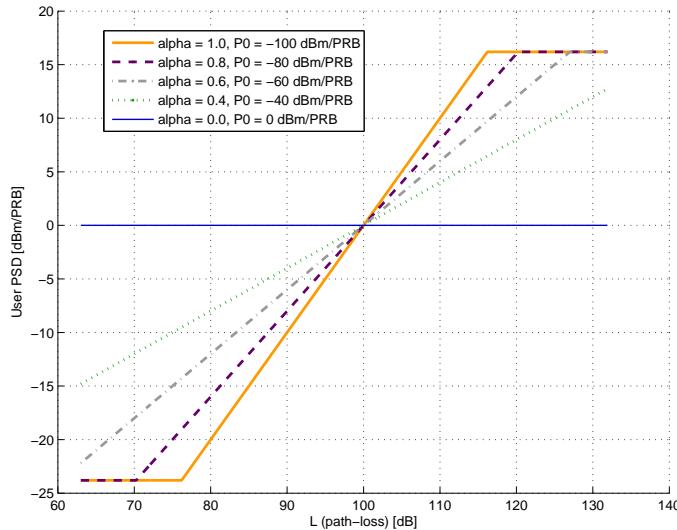


Figure 3.2: User transmission PSD [dBm/PRB] vs path-loss [dB]. Fixed transmission bandwidth with $M = 6$ PRBs per user is assumed.

distribution of the Noise Rise (NR) defined as:

$$NR = \frac{I + N}{N} \quad (3.2)$$

where N represents the noise spectral density and I is the interference spectral density. The NR indicates the level of interference in the system assuming the noise as reference. In case there is no interference in the system, the NR is equal to 1 in linear (0 dB). The transmission bandwidth (that is, the M parameter in (3.1)) is fixed, therefore the plot of the transmission power vs the path-loss can be obtained simply by translating the curves by the quantity $10 \cdot \log_{10}(M)$.

When $\alpha = 0$, P_0 represents the power per PRB that all the users in the cell would have within the path-loss range represented in the Figure (which is a typical range for a Macro 1 case). If we increase the value of α we have to decrease the value of P_0 in order to keep the same transmit

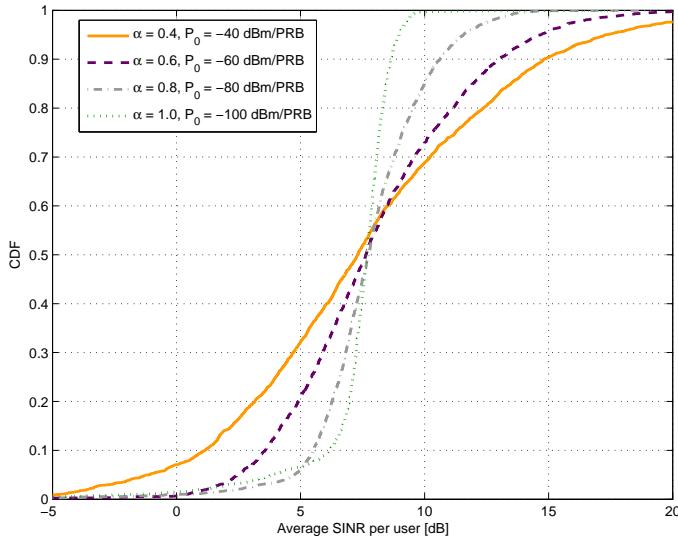


Figure 3.3: CDF of the average SINR per user obtained by varying α and accordingly P_0 as to keep a similar NR level.

power at the center of the path-loss range. In other words, in order to keep a similar NR level, we have to rotate the curve around a center determined by the mid-point of the path-loss range and the level of power we want to keep at that path-loss. For example, if we want to keep a power spectral density of approximately $0 \text{ dBm}/\text{PRB}$ at a path-loss of approximately 100 dB , assuming a full compensation approach ($\alpha = 1$) we need to set $P_0 \approx -100 \text{ dBm}/\text{PRB}$. Similarly, given α , we can derive the P_0 values which preserve a similar NR level using the relation:

$$P_0 = -100 \frac{\text{dBm}}{\text{PRB}} \cdot \alpha \quad (3.3)$$

It's important to notice that following this approach we preserve a similar NR level but not necessarily the highest 5% outage user throughput which has been the optimization criterion followed throughout most of this work. For this reason, further adjustments of P_0 are needed and the values used in the following may differ from the ones derived using (3.3). The approach still remains useful to provide an initial estimate of the P_0 value.

The experienced SINR per user can be expressed as:

$$S = P - L - NR - N \quad [\text{dB}] \quad (3.4)$$

The dependence of the SINR from the P_0 and α parameters can be easily obtained by substituting into (3.4) the transmission power expressed as in (3.1). Neglecting power limitations we obtain:

$$S_{FPC} = P_0 + 10 \cdot \log_{10} M + (\alpha - 1) \cdot L - NR - N \quad [\text{dB}] \quad (3.5)$$

where S_{FPC} indicates the SINR for the FPC approach. Changing the values of P_0 and α while keeping the same NR as explained above leads to the SINR distributions depicted in Figure 3.3 which shows, as expected, a different standard deviation depending on the value of α while keeping a similar median value.

The described approach to PC does not take explicitly into account the interference generated from each user to other cells. In other words, the assumption that users at the cell edge generate

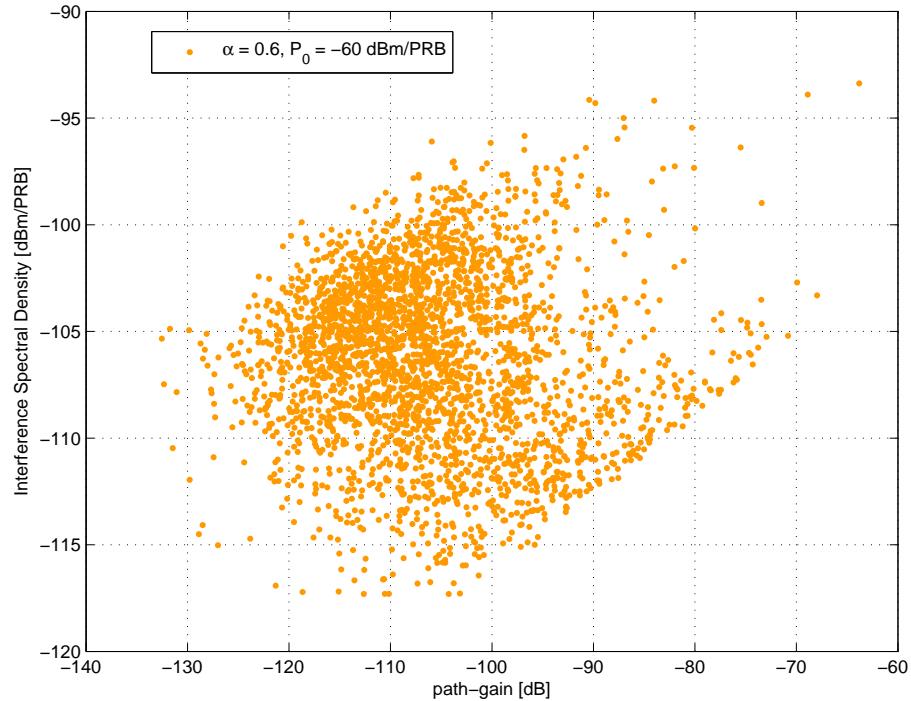


Figure 3.4: Interference spectral density vs path-gain. Every dot represents a user in the network.

more interference than users at the cell center may not be entirely true in a three-sectorized network topology (and even more so in a real scenario). In particular, users experiencing the same path-loss (or path-gain) to the serving cell may be generating a considerably different interference depending on their location in the network. This is confirmed by Figure 3.4 where the Interference Spectral Density (ISD) generated by each user (a dot in the figure) is represented against the path-gain to the serving sector. The ISD is a measure of interference normalized to the bandwidth and is defined as:

$$I_0^i = P_0^i \cdot \sum_{j \in \Omega_i} G_{i,j} \quad [mW/PRB] \quad (3.6)$$

where I_0^i and P_0^i are respectively the ISD and PSD for user i in mW/PRB , $G_{i,j}$ indicates the path-gain³ from the user i to the sector j and Ω_i indicates the set of all the sectors of the network excluded the serving sector of the user i . Figure 3.4 shows clearly how users with a low path-gain may generate a lower interference than users with a high path-gain and users experiencing the same path-gain may generate an interference that can vary by up to 20 dB.

3.3 Interference Based Power Control Algorithm

The finding from the previous section motivates a different approach to PC where also the level of interference in the system is controlled. We could therefore set a per-user limit or target on the interference generated from each user so that the power transmitted becomes a function of the user

³including distance-dependent path-gain, shadowing and antenna gain

channel conditions. Thus, based on equation (3.6), we can derive the condition (3.7):

$$I_0^i = I_0 \Leftrightarrow P_0^i = \frac{I_0}{\sum_{j \in \Omega_i} G_{i,j}} \quad (3.7)$$

where I_0 represents the ISD target, that is, the ISD that each user should ideally generate in the system (unless power limitations occur) and Ω_i , as in the previous case, the set of all the sectors in the network excluded the serving sector of the user i . Taking into account only the interference generated in the system would result in treating equally the users which are generating the same interference but experiencing different path-gains to the serving sector. For this reason, it is worth keeping the measure of the path-gain into the PC formula which thus becomes:

$$P_0^i = \frac{I_0}{G_{i,i} \cdot \sum_{j \in \Omega_i} G_{i,j}} \quad (3.8)$$

Simplifying the notation we can set $G_{i,i} = G_{own}$ and $\sum_{j \in \Omega_i} G_{i,j} = G_{other}$ and rewrite the formula more compactly as:

$$P_0^i = \frac{I_0}{G_{own} \cdot G_{other}} \quad (3.9)$$

It is important to notice that G_{other} here represent the sum of the path-gains to the other sectors but, in a practical implementation, it could be limited to the path-gain to the strongest interfering sector. Indeed this is the approach followed in [35] and, as already mentioned in Section 3.1, in [34] the authors show that the first and second strongest interfering sectors account for most of the interference.

Finally, following the same approach of the FPC, it may be beneficial to compensate only for a fraction of the path-gain or a fraction of the generated interference. For this reason, we keep the weight α already used in the open-loop form of the FPC formula to indicate a fraction of the path-gain and introduce a similar weight β to indicate a fraction of the sum of path-gains to other sectors. In this way the PC formula, neglecting power limitations, becomes:

$$P_0^i = \frac{I_0}{G_{own}^\alpha \cdot G_{other}^\beta} \quad (3.10)$$

Converting the formula to logarithmic scale and replacing the path-gain G with the path-loss L (related as $G = -L$) we obtain:

$$P_0^i = I_0 + \alpha \cdot L_{own} + \beta \cdot L_{other} \quad [dBm/PRB] \quad (3.11)$$

Recalling the open-loop formula (3.1) presented at beginning of this chapter and rewriting it by neglecting the power limitations P_{max} and the transmission bandwidth M , we obtain:

$$P_0^i = P_0 + \alpha \cdot L_{own} \quad [dBm/PRB] \quad (3.12)$$

Comparing (3.11) and (3.12) we discover that the I_0 and P_0 parameters actually play exactly the same role (and will therefore have similar values) and our proposed approach to PC simply replaces the weighted path-loss to the serving station ($\alpha \cdot L$) with a linear combination of the path-loss to the serving and to the non-serving cells ($\alpha \cdot L_{own} + \beta \cdot L_{other}$). Obviously, setting $\beta = 0$ would return the initial open-loop PC formula.

Table 3.2: Main simulations parameters

Parameter	Setting
# UEs per sector	10
# PRBs per UE	6 (fixed bandwidth)
FD scheduling	PF
Cell-level user distribution	Uniform
BLER target at 1st transmission	30%
Propagation scenario	3GPP Macro case 1 (ISD of 500 m)
α (open-loop PC)	from 1.0 to 0.0 in steps of 0.1
P_0 (open-loop PC)	See Table 3.3 for $\beta = 0$
α (closed-loop PC)	from 1.0 to 0.0 in steps of 0.1
β (closed-loop PC)	from 1.0 to 0.0 in steps of 0.1
I_0 (ISD target)	See Table 3.3
Traffic model	Full and finite buffer with balanced load
Buffer size for finite buffer	1 Mb

3.4 Modeling Assumptions

The performance evaluation of the proposed algorithm is carried out using a detailed multi-cell system level simulator which follows the guidelines given in [20] and is described in Appendix A.1. The simulation parameters and assumptions relevant for the results presented in this chapter are listed in Table 3.2 (a full list of the default simulation parameters is presented in Table A.1). As this topic has been investigated in parallel with the initial studies on scheduling which are discussed afterwards, some of the simulation assumptions used here will be explained in details only in later chapters though are not deemed important for the understanding of the results here presented. As an example, the scheduling used to generate such results assumes a fixed bandwidth allocation and the use of the PF metric. Two traffic models are considered: A full buffer scheme, where the users always have data to transmit from the buffer considered infinite. A finite buffer scheme, where the users have a limited amount of data in the buffer so that when a user empties its buffer, it is replaced by another user in the same sector but in a generally different (random) location.

The I_0 values (or equivalently P_0) which optimize the 5% outage user throughput are indicated in Table 3.3. Such values are obtained using a simpler and faster offline tool which searches the value which optimizes the 5% outage use throughput around an initial value chosen according to the criterion given in (3.3). Additionally, some selected values have been verified with the simulator used in this research project.

It is worth pointing out that the secondary diagonal of the matrix in Table 3.3 and in general all the points for which $\alpha + \beta = c$ (c being a constant between 0 and 1), show very similar or identical I_0 values.

3.5 Performance Evaluation

In the following the performance evaluation of the proposed scheme is given. First the full buffer scenario will be evaluated. Afterwards the same scheme will be evaluated in a finite buffer scenario. The studies on the PC were carried out in parallel with the studies on scheduling but for ease of organization and presentation the results related only to the topic of PC will be all listed in this chapter.

Table 3.3: I_0 (or equivalently P_0) values in dBm/PRB that maximizes the 5% outage user throughput for each combination of α and β .

α β	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.0	16	2	-8	-21	-32	-44	-56	-68	-80	-92	-103
0.1	-4	-11	-21	-32	-44	-55	-68	-79	-92	-103	-114
0.2	-7	-20	-32	-44	-56	-68	-80	-92	-103	-114	-123
0.3	-21	-33	-45	-57	-69	-81	-92	-104	-114	-123	-132
0.4	-33	-45	-57	-69	-81	-92	-104	-115	-123	-132	-141
0.5	-45	-56	-68	-79	-93	-103	-114	-123	-135	-142	-153
0.6	-56	-68	-80	-93	-104	-113	-124	-135	-146	-155	-162
0.7	-69	-80	-93	-104	-113	-124	-135	-144	-155	-164	-171
0.8	-80	-93	-104	-113	-124	-133	-146	-155	-164	-171	-181
0.9	-91	-102	-113	-124	-133	-145	-155	-165	-172	-181	-192
1.0	-102	-113	-123	-134	-145	-156	-165	-174	-184	-190	-201

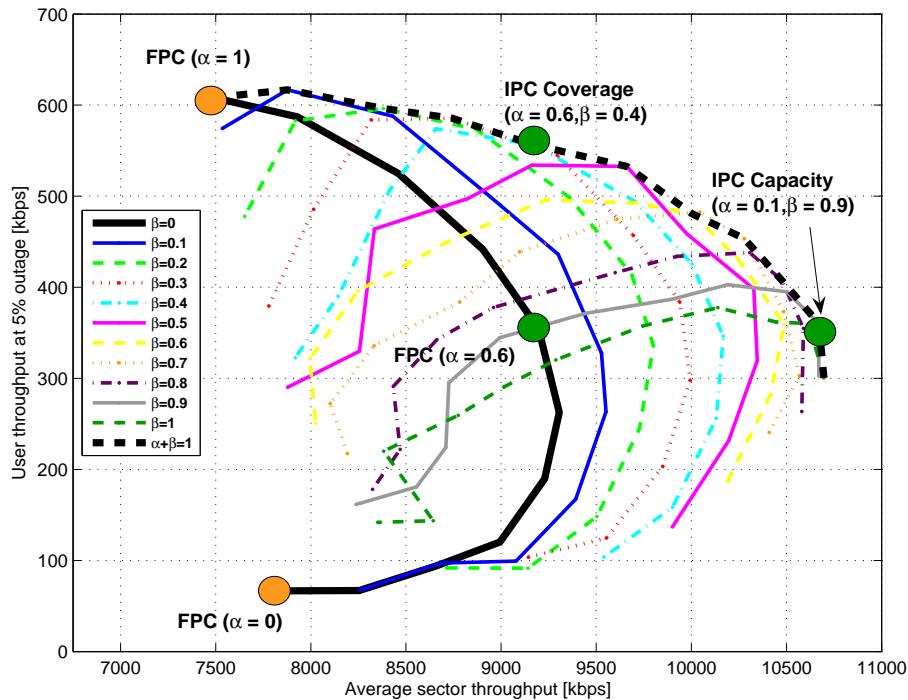


Figure 3.5: Outage user throughput at 5% vs average sector throughput. Full buffer traffic.

3.5.1 Full Buffer

The overall performance of the proposed IPC is summarized in Figure 3.5. Each curve corresponds to a value of β and is obtained by varying α from 1.0 to 0.0 in steps of 0.1. Each point of each curve corresponds to a combination of values of α and β (and the corresponding I_0 value from Table 3.3). The curve obtained for $\beta = 0$ corresponds to the FPC case. Connecting the points which offer the highest outage user throughput and average cell throughput, we obtain an envelope which we could consider as the collection of points offering the best performance.

Interestingly enough, such curve is obtained under the case of $\alpha + \beta = 1$. This characteristic will be analyzed from a different point of view later on when comparing the obtained results with

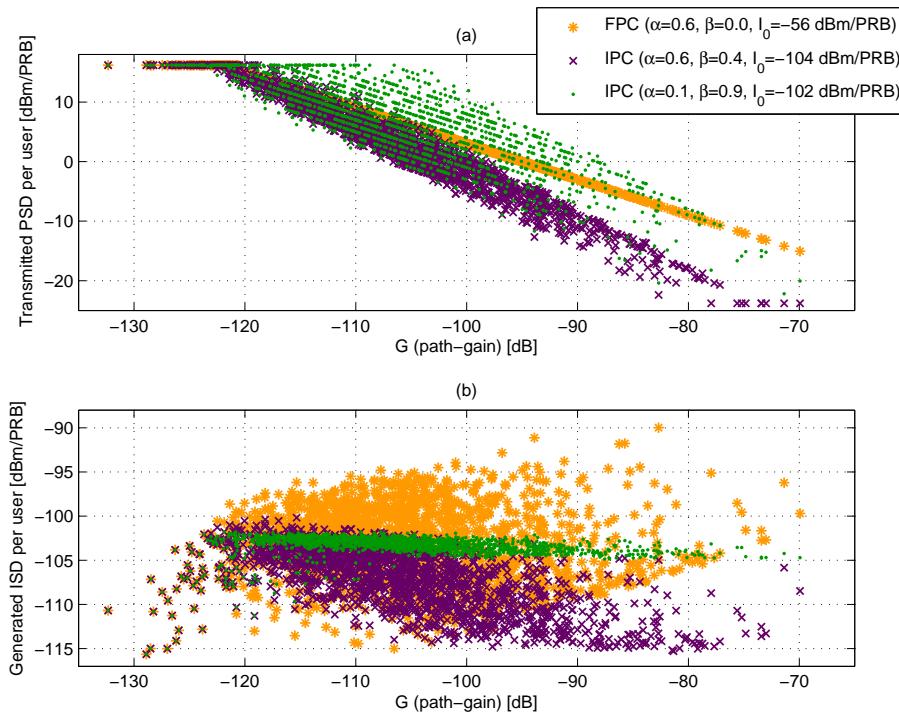


Figure 3.6: (a) User transmission PSD in dBm/PRB vs path-gain. (b) User generated ISD in dBm/PRB vs path-gain. Every dot corresponds to a single user. Full buffer case.

the findings of [35].

In the open-loop FPC study conducted in [33], the recommended settings for a good trade-off between 5% outage user throughput and average cell throughput were $\alpha = 0.6$ and $P_0 = -58 \left[\frac{dBm}{PRB} \right]$. The performance under such settings is marked as “FPC ($\alpha = 0.6$)” in Figure 3.5. In order to compare the reference case of the open-loop FPC with the IPC, we choose two points from the IPC for which $\alpha + \beta = 1$, that is, the point having the same average sector throughput but a higher outage user throughput and the point having the same outage user throughput but a higher average sector throughput. The first point, indicated as “IPC Coverage” in Figure 3.5, corresponds to $\alpha = 0.6$ and $\beta = 1 - \alpha = 0.4$, while the second point, indicated as “IPC Capacity” in Figure 3.5, corresponds to $\alpha = 0.1$ and $\beta = 1 - \alpha = 0.9$. The results show a considerable performance increase obtained by taking into account the interference generated in the system. Specifically the “IPC Coverage” point shows a gain in outage user throughput of more than 50% over the reference while the “IPC Capacity” point shows a gain in average sector throughput of ca. 16% over the reference. The following figures provide some information which explain the behavior of the algorithm and help identifying the reasons of such performance boost.

Figure 3.6 shows the user transmit PSD for the three considered cases: the reference FPC case (obtained for $\beta = 0$ and $\alpha = 0.6$) and the two IPC cases, that is, IPC Coverage and IPC Capacity. The reference FPC case shows a linear dependence only from the path-gain (apart from power limitations) expressed by the factor α . As we introduce the dependence also from the interference, by changing β to a value greater than 0, we obtain a cloud of points or, more specifically, a set of points for each value of path-gain. This happens because some users generate a lower interference and are therefore allowed to transmit with a higher power than others and vice versa. In other words, α changes the slope of the curve whereas β controls its spreading so that larger values of β determines the transmission power values obtained for a certain path-gain to be more dispersed over the axis.

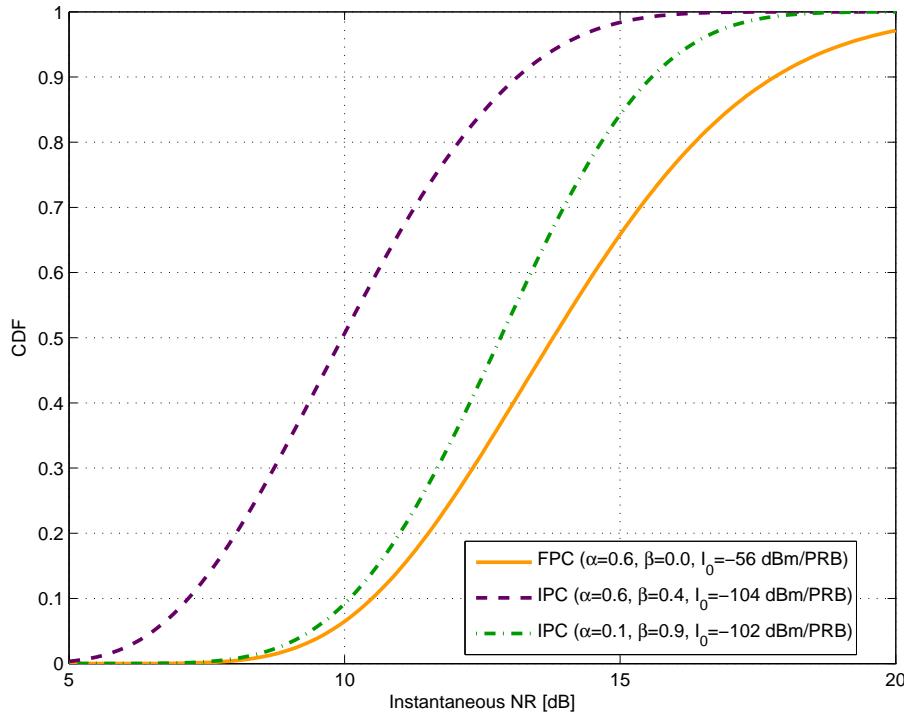


Figure 3.7: Distribution of the instantaneous NR for the reference FPC case and the chosen IPC cases.

Looking at Figure 3.6(b), that is, the received ISD (in dBm/PRB) we notice the opposite trend. The FPC reference case, already shown in this chapter as motivation for a different approach to PC, shows a cloud of points which are very dispersed for a given value of path-gain. At the other extreme, the case with the largest value of β and therefore with the highest dependence from the generated interference (referred to as IPC Capacity and having $\alpha = 0.1$ and $\beta = 0.9$), produces a cloud of points with the least dispersed values of ISD independently of the path-gain.

This is reflected in the distribution of the interference which, in Figure 3.7, is expressed as NR. The distribution exhibits a lower variance as the β factor, expressing the dependence from the path-loss to non-serving cells G_{other} , is increased. This, in turn, results in a more reliable channel estimation and MCS selection as shown by the distributions of the CSI and OLLA error in Figure 3.8.

The CSI error is defined in dB as:

$$CSI_{error} = SINR_{experienced} - SINR_{estimated}$$

where $SINR_{experienced}$ is the SINR experienced by the UE at the eNode-B and $SINR_{estimated}$ is the SINR estimated at the eNode-B by processing the UE Sounding Reference Signal (SRS) and is affected by Gaussian measurement error. A smaller variance of the CSI error distribution indicates a more precise estimation of the SINR as a result of a more stable interference scenario.

The OLLA error is defined in dB as:

$$OLLA_{error} = SINR_{experienced} - SINR_{input_{AMC}}$$

where $SINR_{input_{AMC}}$ is the SINR given in input to the AMC function defined as:

$$SINR_{input_{AMC}} = SINR_{estimated} - OLLA_{offset}$$

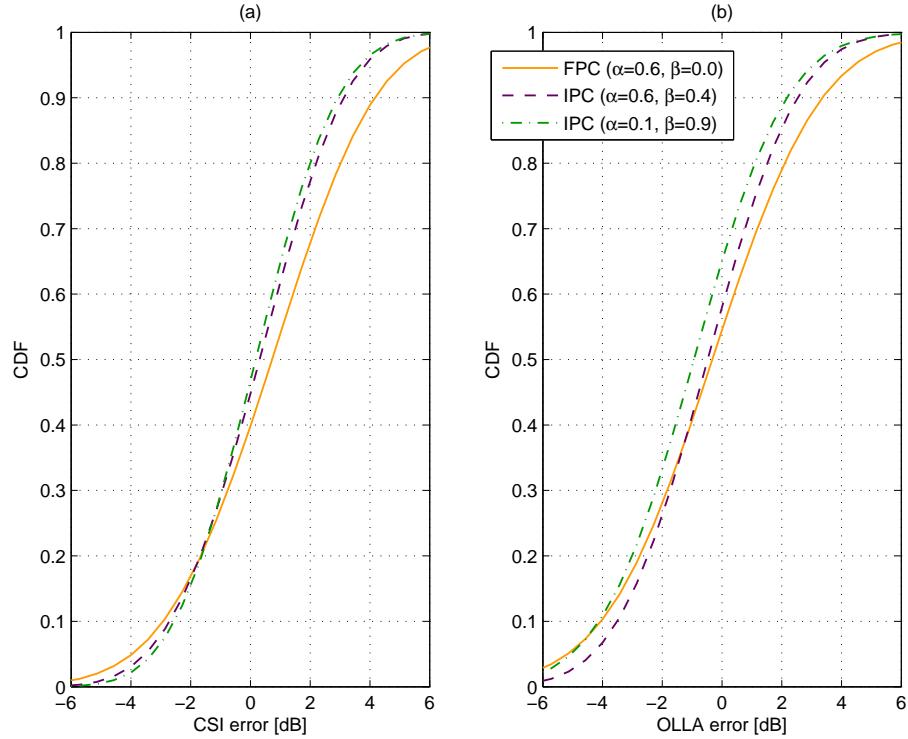


Figure 3.8: Distributions of CSI and OLLA errors.

where OLL_{offset} is the correction made by the OLLA before the SINR is given in input to the AMC. A OLLA error distribution with a smaller variance indicates a more appropriate MCS selection, that is, a more robust selection in case of positive error or a more aggressive selection in case of negative error. This translates into a higher throughput which increases the coverage or the capacity depending on the settings.

It is worth, at this point, trying to relate our approach with the one followed in the reference paper [35] which, interestingly, shows very similar results to the ones here obtained under the condition $\alpha + \beta = 1$. In the paper the author proposes to set the target SINR according to the formula:

$$S = S_{edge} + (\alpha - 1) \cdot \Delta L \quad [dB] \quad (3.13)$$

where $\Delta L = L_{own} - L_{strongest-interferer}$ and S_{edge} is the target SINR for the users which experience $\Delta L = 0$ (supposedly the users at the cell edge).

Similarly, the target SINR under open-loop FPC, which was previously derived as:

$$S_{FPC} = P_0 + 10 \cdot \log_{10}(M) + (\alpha - 1) \cdot L - IoT - N$$

can be rewritten as:

$$S_{FPC} = S_0 + (\alpha - 1) \cdot L \quad (3.14)$$

where

$$S_0 = S \mid_{L=0} = P_0 + 10 \cdot \log_{10}(M) - IoT - N$$

In the same way, the target SINR under IPC can be written as:

$$S_{IPC} = S_0 + (\alpha - 1) \cdot L_{own} + \beta \cdot L_{other} \stackrel{(\alpha+\beta=1)}{=} S_0 + (\alpha - 1) \cdot \Delta L \quad (3.15)$$

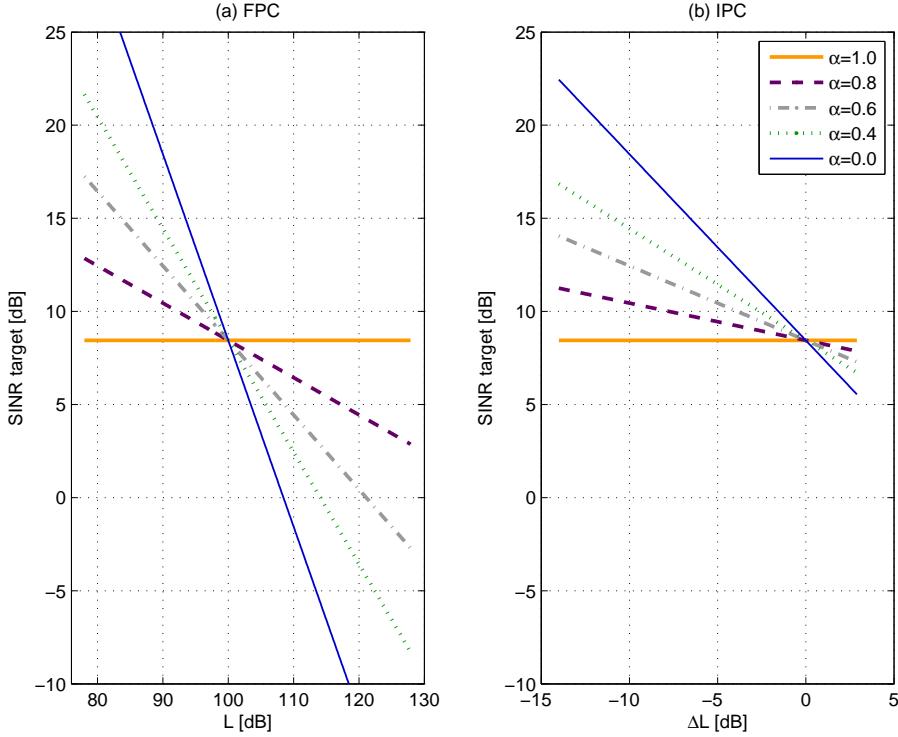


Figure 3.9: (a) SINR target calculated using equation (3.14) for a typical range of path-loss. (b) SINR target calculated using equation (3.15) for a typical range of ΔL . Macro 1 case.

with $\Delta L = L_{own} - L_{other}$. The formula (3.15) is equivalent to the formula (3.13) once we replace the sum of path-loss to all non serving sectors (L_{other}) with the path-loss to the strongest interferer only ($L_{strongest-interferer}$). Following two different approaches, one based on the modification of the FPC formula and the other based on setting the SINR target, we have therefore come to the same results and have shown that in the considered scenario it is beneficial to take into account also the path-loss to the non serving sector.

The equations (3.14) and (3.15) are represented in Figure 3.9 (a) and (b) respectively for a typical range of L and ΔL . It is important to highlight that in equation (3.15) S_0 does not depend on α because it is obtained under the assumption of $\alpha + \beta = 1$ for which P_0 is approximately constant. In equation (3.14), on the other hand, S_0 has to be tuned according to α because in the FPC formula a different α requires a different P_0 .

3.5.2 Finite Buffer

The very promising results obtained in the full buffer scenario do not hold any longer when considering a finite buffer scenario. The curve obtained under $\beta = 0$ (pure FPC), and the one obtained under $\alpha + \beta = 1$ (combination of best IPC cases), are represented for both full buffer and finite buffer traffics in Figure 3.10. The finite buffer curves show a considerable performance degradation for both FPC and IPC cases compared to the curves obtained under full buffer. More importantly under the finite buffer case, which is closer to a real scenario than the full buffer case, the FPC and the IPC curves show a very similar trend and the gain of the IPC over the FPC is reduced considerably.

This is assumed to be a consequence of the distribution of the users over the cell: the finite buffer causes the users to finish their session faster for users close to the serving station than to the cell edge. As the users finish their sessions and are replaced by new users randomly generated in

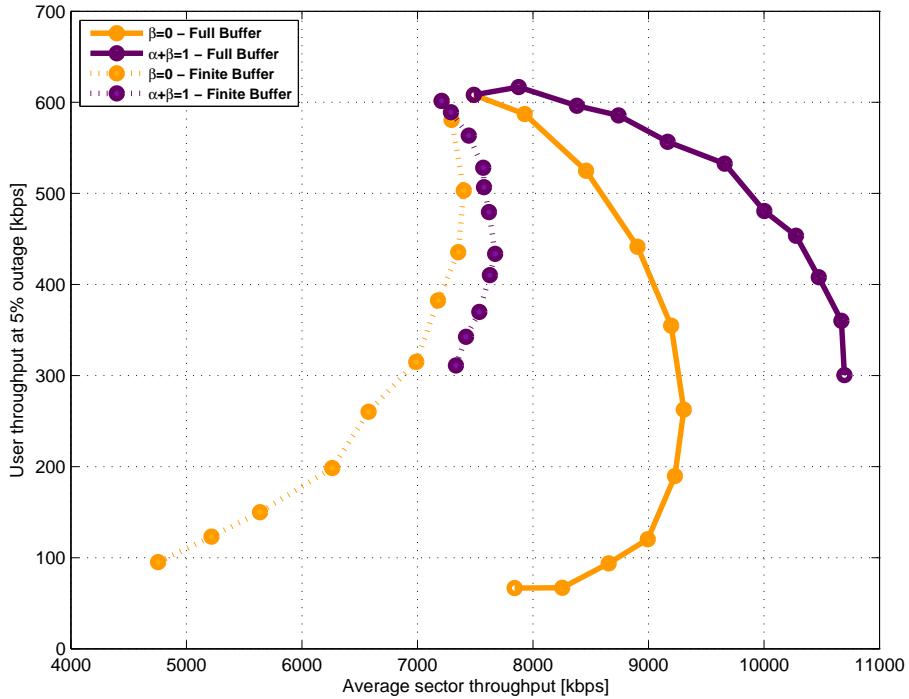


Figure 3.10: User throughput at 5% outage vs average sector throughput. Full and finite buffer cases.

the cell, a larger number of them gathers at the cell edge thus affecting the path-loss distribution and the gain provided by the IPC. It is also possible that the I_0 parameter, whose values have been left going from the full buffer to the finite buffer case, should be adjusted to fit the different scenario.

The clarification of this aspect is left for future studies and it does not affect the results of the following chapters as they have been mostly obtained before the investigations on IPC and under the assumption of pure FPC (with $\alpha = 0.6$).

3.6 Considerations on Inband Inter-User Interference

In ideal conditions the SC-FDMA transmission maintains the user orthogonality. However, the presence of hardware imperfections in the RF transceiver, such as frequency offset and phase noise, destroys the subcarrier orthogonality [36], and the user signal leakage start to interfere with other users. In practice, the received signals from different users experience various frequency offset/phase noise and lead to the inter-user interference. The PSD of the received signals from different users will also vary and the frequency offset/phase noise and signal leakage from the adjacent users with higher PSD can lead to a significant interference level and degrade the system performance.

As the average received interference, due to dynamic scheduling, is the same for all the users, we can refer to the distribution of the received SINR, which is shown in Figure 3.11 for the Macro 1 case, rather than the distribution of the received signal. We can see that the distribution is contained in a range of approximately 22 dB in case of instantaneous SINR and 15 dB in case of average SINR. As shown in [37], the SNR loss for the interferer UEs with a frequency offset of 200 Hz and PSD offset of 10 dB is very small, less than 0.1 dB. Increasing the PSD offset by another 10 dB results in an SNR loss of 0.3 dB. Moreover the probability of two users located in the uppermost and lowermost part of the distribution being allocated next to each other is very low.

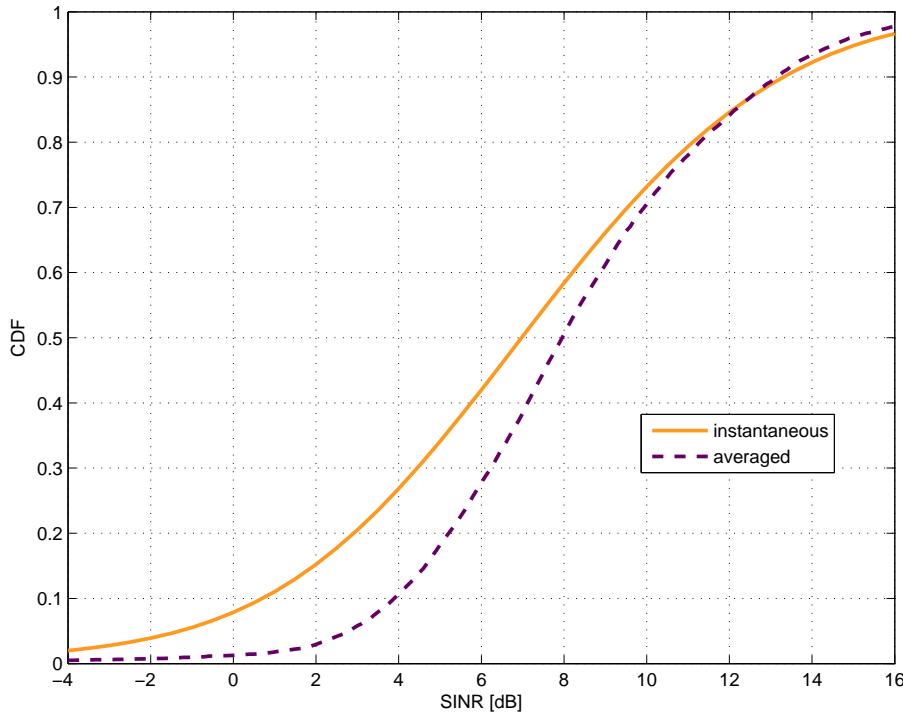


Figure 3.11: Distribution of instantaneous SINR and time-averaged SINR per user. 16 users per cell and 6 PRBs per user.

The SNR degradation can be more significant in case the user with lower PSD has also a small bandwidth allocated. In case further investigations show a significant SNR degradation, not only proper frequency synchronization mechanisms are needed but also allocation algorithms which take such aspects into account could be devised.

3.7 Conclusions

In this chapter an interference-based PC algorithm is introduced. The main idea of the algorithm is to set to user transmit power not only according to the path-loss to the serving cell but also according to the interference generated in the system which is expressed by the sum of all the path-loss to the non-serving cells.

Even though the approach to the problem is independent from other recent work in the open literature, the development of the algorithm and the careful evaluation of the results lead to the same conclusions presented in one of the mentioned references.

The performance of the algorithm is evaluated in terms of comparison with the open-loop FPC algorithm. The main performance indicators used are the outage user throughput and the average cell throughput but other quantities are also taken into account to understand the behavior of the algorithm and gain a better knowledge of the system. Compared to the FPC approach for $\alpha = 0.6$ the IPC algorithm can provide a gain of above 50% in outage or approximately 16% in average cell throughput depending on the parameter settings. Such gains are the result of a more appropriate SINR setting, which takes into account also the user generated interference, as well as a more stable interference in the system.

The deployment of a finite buffer scenario produces a different path-loss distribution of the users in the network and, according to the current results, a considerable reduction of the gain of the proposed strategy over the open-loop FPC. The reason for such performance drop could also

be attributed to the improper tuning of the P_0 parameter (or, equivalently I_0) and the related further investigation is left for future studies.

Throughout the rest of the thesis the open-loop FPC formula will be taken as default assumption.

Chapter 4

Fixed Transmission Bandwidth based Packet Scheduling

4.1 Introduction

This chapter is devoted to the performance analysis of an uplink channel aware scheduling algorithm elaborated during this PhD project and described in [38]. The proposed algorithm is evaluated in both Macro 1 and Macro 3 cases and under a varied set of parameter settings in order to understand its potential as well as the interaction with other system entities.

The problem of channel aware scheduling has already been investigated in downlink for example in [39] and [40] for an OFDMA-based system.

In [39] a scheduling algorithm based on a generalized PF metric is presented. The definition of the PF metric, adopted for example in [18] for a TD based system, is extended to the FD and analyzed in terms of fairness and system throughput. Compared to a system without FD scheduling, the proposed approach increases the system throughput and yields an improved fairness with respect to allocated resources and achieved data-rate per user.

In [40] the potential of FD channel aware scheduling is investigated for the downlink of LTE via system-level simulation. Also in this case, the PF metric is assumed and it is shown to have gains in the order 40% in average system capacity and cell-edge data rates compared to a frequency-blind and time-opportunistic only scheduling.

At the time of this study fewer publications were available regarding the uplink of SC-FDMA based systems. The constraint in resource allocation for the uplink channel in LTE systems makes channel dependent scheduling a more challenging task due to the fact that some resources may have to be allocated to satisfy the constraint rather than the channel condition. In [41], for example, the SC constraint is not taken into consideration in the algorithm design. This means that the PRBs are allocated to users which exhibit the highest marginal utility regardless of the location of other PRBs that are already allocated.

More recently Al-Rawi *et al.* [42] approached the FD scheduling, in a single-cell scenario, as an integer programming problem which takes into account also the SC constraint. The solution to the integer programming problem, being computationally intensive, is only used to set the upper bound to the achievable performance. A more efficient algorithm is therefore derived which, in case of perfect channel conditions, achieves a performance 10% below the optimal one and 80% above the performance of a channel blind scheduler.

In this thesis, given the complexity introduced by the SC constraint, we first bypass the problem by considering an FTB-based approach to derive an initial performance evaluation. In the next chapter, which is a natural continuation of this chapter, we integrate the flexibility of the ATB into the scheduling.

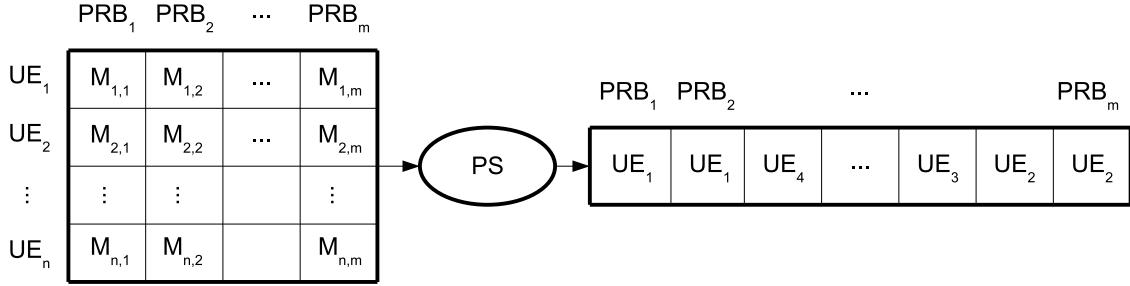


Figure 4.1: User to PRB mapping realized by the allocation algorithm based on the metrics received as input.

Different heuristic scheduling algorithms are proposed along the way and are evaluated by means of extensive system level simulations.

The chapter is organized as follows: The PS framework is described in Section 4.2. Section 4.3 describes the algorithm. The performance evaluation of the algorithm is given in Section 4.5 after the modeling assumptions listed in Section 4.4. Section 4.6 describes a generalization of the scheduling algorithm together with additional results. Finally, conclusions are presented in Section 4.7.

4.2 Packet Scheduling Framework

In this section, the scheduling framework already described in Section 2.5 is recapitulated with the addition of few details useful for the understanding of the proposed algorithm.

The task of the PS is to multiplex in time and frequency domain the active users (which are able to transmit in the next TTI and are not in DRX mode) by dynamically allocating the available time-frequency physical resource units under a defined set of constraints. Such objective can be formulated as:

$$\text{maximize } M_{sum} = \sum_{i \in \Omega_i, j \in \Omega_j} M_{i,j} A_{i,j} \quad (4.1)$$

subject to the constraints (2.2), (2.3) and (2.4), where $M_{i,j}$ is the metric for user i and PRB j , Ω_i is the set of users and Ω_j is the set of PRBs.

In the following, M_{sum} will be referred to as global metric. The metric defines the optimization criterion to be used by the PS, that is, how the different aspects of fairness, channel awareness and QoS requirements should be taken into account. Decoupling the allocation algorithm from the choice of the metric greatly simplifies the overall design without compromising the performance. Following this approach, indeed, the PS algorithm can be designed with the single objective in mind of maximizing the global metric while the strategy to be taken for the different users in terms of fairness, channel awareness and QoS requirements is solely reflected by the metric expression.

The PS takes a set of metrics (which can be represented in form of matrix) as input and outputs an allocation table, that is a user to PRB mapping as shown in Figure 4.1. The allocation takes place under a set of hard constraints including the scheduling of users which have to undergo a retransmissions or the allocation of adjacent PRBs to the same user due to the adoption of the SC-FDMA scheme.

Throughout this study the PS is assumed to be dynamic, that is, it produces a new allocation table every TTI in order to exploit the time selectivity of the mobile propagation channel. Given the computational complexity of the algorithm and the limited time available for completion (1 TTI),

	RC_1	RC_2	RC_3
UE_1	$M_{1,1} = 380$	$M_{1,2} = 670$	$M_{1,3} = 1530$
UE_2	$M_{2,1} = 300$	$M_{2,2} = 730$	$M_{2,3} = 1390$
UE_3	$M_{3,1} = 650$	$M_{3,2} = 810$	$M_{3,3} = 1280$

Figure 4.2: Metric value for each UE and each RC in form of matrix. The circles indicate the elements chosen by the algorithm. The hatching covers the elements discarded after a user to PRB selection has been made.

it is proposed to split the algorithm in two phases indicated as TD scheduling and FD scheduling (as previously shown in Figure 2.5). In the first phase the TD scheduler filters out the users which are not likely to be scheduled in the next TTI because their expected performance or their priority is, according to some criteria, largely below the one of the other users. The reduced set of users is handed over to the FD scheduler which performs the most computationally intensive operations and outputs the user to PRB mapping.

4.3 Matrix-Based Search Algorithm

Given the complexity of the PS algorithm, mainly due to the constraint imposed by the SC-FDMA technology, in the first phase of this work the approach is simplified by assuming the bandwidth to be fixed and equal in size for all the users. Such bandwidth is constituted by a set of consecutive PRBs and is indicated in the following as Resource Chunk (RC). The size of the RC is chosen to be a sub-multiple of the system bandwidth so that an integer number of users can be accommodated without creating bandwidth fragmentation. The metric is calculated over the bandwidth of the RC rather than the bandwidth of a single PRB.

Assuming that each user can be allocated at most one RC, we can adopt a greedy optimization strategy based on arranging users and bandwidth chunks in a matrix containing the metric value for each user and each RC, as exemplified in Figure 4.2. Such matrix is then fed as input to the FDPS which performs the following steps:

1. Find the UE and RC with the highest metric (e.g. UE_1 at RC_2 in Figure 4.2)
2. Allocate the RC to the UE
3. Remove UE (matrix row) and RC (matrix column)
4. Repeat from step 1 using the resulting sub-matrix

4.4 Modeling Assumptions

In this initial phase of the study the scheduling framework is simplified as all the users are handed over to the FD scheduler making the TD scheduler irrelevant for the final allocation. In FD the metrics considered are Random (RAN) and PF.

When the RAN metric is used, a random value extracted by a uniform random variable in the range between 0 and 1 is assigned to each UE and each RC every TTI. It is similar to the Round Robin (RR) metric in that it gives the same priority to the users over a sufficiently long period of time and it is blind to the channel conditions. On the other hand it is preferred to the RR metric because it guarantees frequency and interference diversity.

The PF metric for a UE i , on PRB r , at TTI t , is defined as (see [39] and [40]):

$$M_{PF}(i, r, t) = \frac{\hat{T}(i, r, t)}{\bar{T}(i, t)}$$

where $\hat{T}(i, r, t)$ is the estimated Layer 1 achievable throughput (estimated according to AVI tables and BLER target using CSI) for user i , on PRB r , at scheduling instant t and $\bar{T}(i, t)$ is the past averaged acknowledged¹ Layer 1 throughput for user i , at scheduling instant t calculated using an exponential averaging filter as:

$$\bar{T}(i, t) = (1 - \frac{1}{N_{TTI}})\bar{T}(i, t-1) + \frac{1}{N_{TTI}}T(i, t)$$

where $T(i, t)$ is the Layer 1 acknowledged throughput for user i , at scheduling instant t and N_{TTI} defines the filter length or filter memory, which in this study is assumed to be the same for all the users.

Table 4.1: Main simulations parameters

Parameter	Setting
# UEs per sector	from 4 to 60
# UEs handed to the FD scheduler	from 4 to 60
# PRBs per UE (or RC size)	2, 4, 6, 8, 12
Cell-level user distribution	Uniform
TD scheduling	Not considered
FD scheduling	RAN or PF
BLER target at 1 st transmission	30%
Propagation scenarios	3GPP Macro case 1 and 3
PF filter length (N_{TTI})	100 ms
Initial throughput value ($T(0)$)	500 kb
α (for PC)	0.6
P_0 (for PC)	-58dBm/PRB for Macro case 1; -64dBm/PRB for Macro case 3
Traffic model	Full buffer with balanced load

Table 4.1 shows the simulation parameters which are relevant for the results presented in this chapter. The HARQ is assumed to be synchronous and adaptive with Chase Combining. This means that retransmissions have to be scheduled in a specific TTI but can take place anywhere within the bandwidth. No specific optimization in this respect is deployed in FD. The AMC selects the most appropriate MCS based on the SINR estimated over the selected RC. Such estimation is obtained from the SRS transmitted by the UE and processed at the eNode-B to extract near-instantaneous frequency selective CSI as described in Section 2.8.1. The PC functionality is implemented according to the open-loop formula given by equation (3.1), that is, no closed-loop adjustments are considered.

¹“Acknowledged” throughput is meant to be the successfully delivered throughput and is therefore affected by the experienced BLER

4.5 Performance Evaluation of Matrix-Based Search Algorithm

The performance of the proposed algorithm is evaluated in terms of SINR and user throughput distributions, average cell throughput and 5% outage user throughput. Along the way, in order to better understand the system and the interaction of the scheduler with other entities a set of additional statistics is presented including average BLER per user, average transmit power per user, NR and OLLA error distributions.

Figure 4.3 (a) shows the distribution of the scheduled or experienced SINR per user under the RAN and PF metrics. The PF metric provides an improvement over the whole SINR range quantified in approximately 1 dB at the median value. This translates into a corresponding improvement in the throughput distribution as shown in Figure 4.3 (b).

Figure 4.4 (a) shows that approximately 80% of the users experience an average BLER which matches the BLER target of 0.3 under both RAN and PF metrics. In form of verification Figure 4.4 (b) and 4.4 (c) show respectively the instantaneous NR and the average user power distributions. As expected, they overlap with each other because the power is set according to the OLPC formula which is independent on the scheduling metric deployed. The PC settings specified in Table 4.1 lead to a NR median of approximately 13 dB and to approximately 5% of the users in maximum transmit power because this is the situation in which the 5% outage user throughput is maximized.

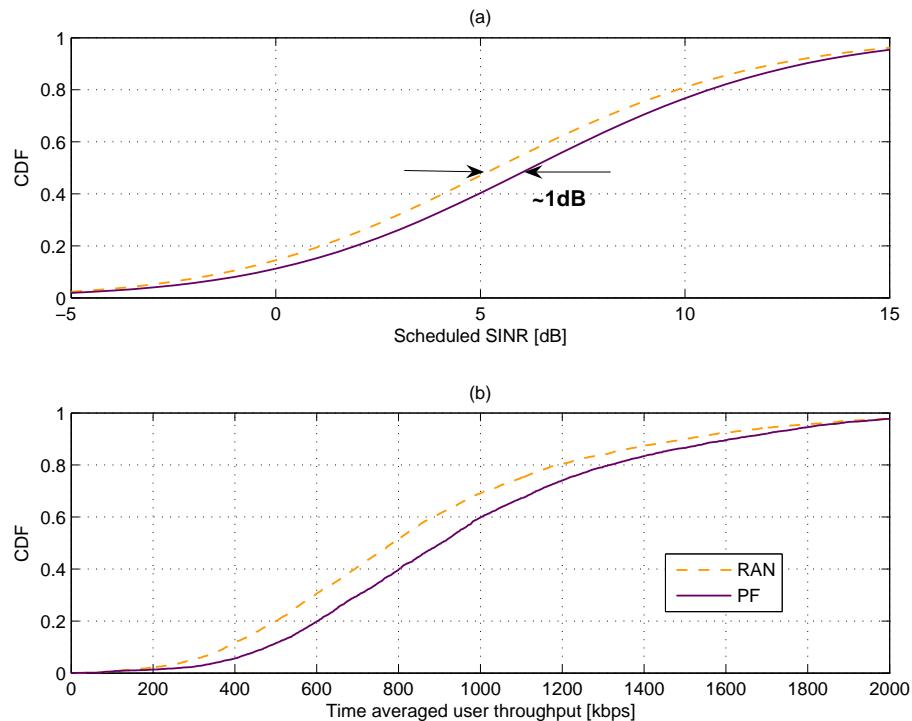


Figure 4.3: Scheduled SINR and time averaged throughput per user using RAN and PF metrics. 8 users per sector and 6 PRBs per user.

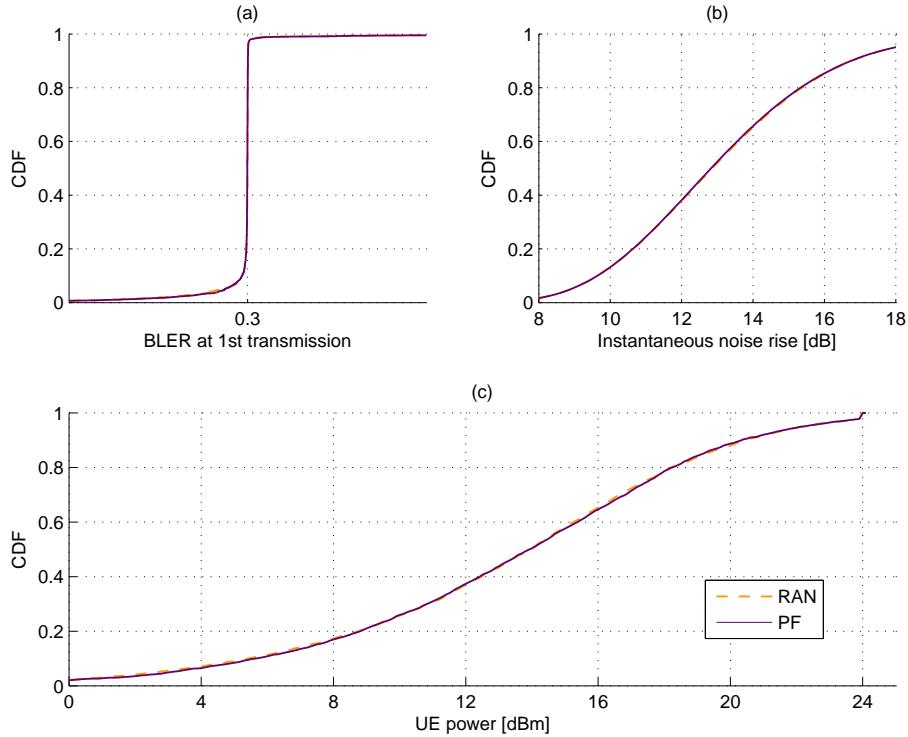


Figure 4.4: BLER at 1st transmission, instantaneous NR and power per user using RAN and PF metrics for 8 users per sector.

Figure 4.5 shows an interesting statistic regarding the time correlation introduced in the resource allocation by the PF and RAN metrics. More specifically, it shows the distribution of the number of consecutive² re-allocations of the same RC to the same user. Under the PF metric there is a higher probability of allocating the same RC to the same user over consecutive TTIs. Specifically, the probability of this happening for at least one TTI is equal to 0.27 (1-0.73, referring to the Figure) in case of PF metric while it is equal to 0.125 or 1/8 (1-0.875, referring to the Figure) in case of RAN metric³. On the other hand this effect is not significant enough as to stabilize the inter-cell interference and reduce the standard deviation of the OLLA error distribution. The only visible effect is that the PF metric tends to schedule users characterized by a positive error on the estimated SINR thus producing a shift of the OLLA error distribution to the right by approximately 0.5 dB.

²By “consecutive” it is meant over subsequent TTIs

³1/8 is simply the probability that a user occupies a specific RC out of the 8 available.

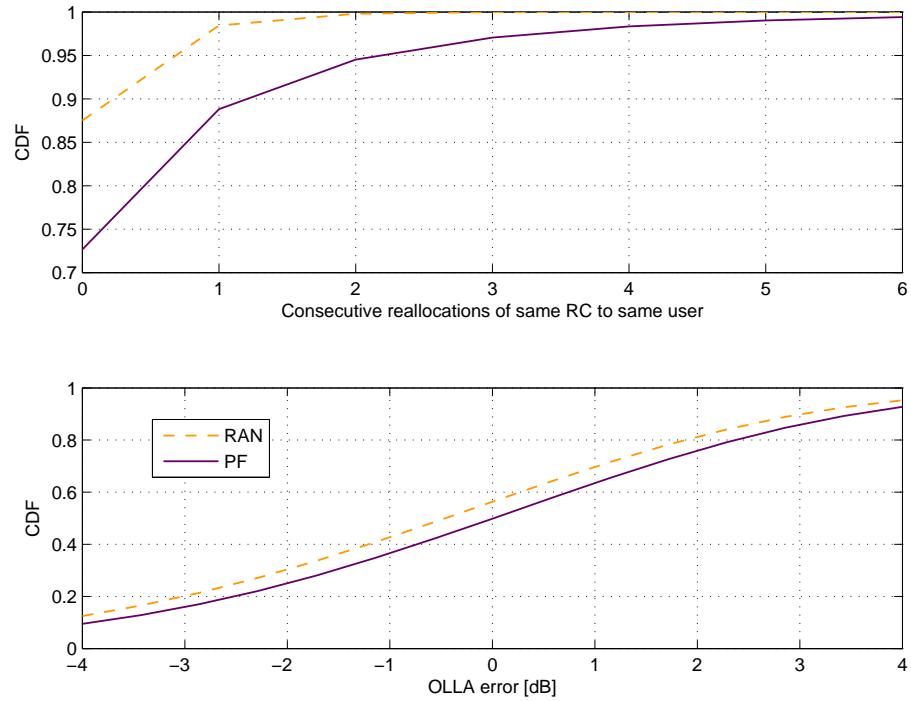


Figure 4.5: (a): time-consecutive re-allocations (in time) of the same RC to the same user. (b): OLLA error. RAN and PF metrics, 24 users per sector, 6 PRBs per user.

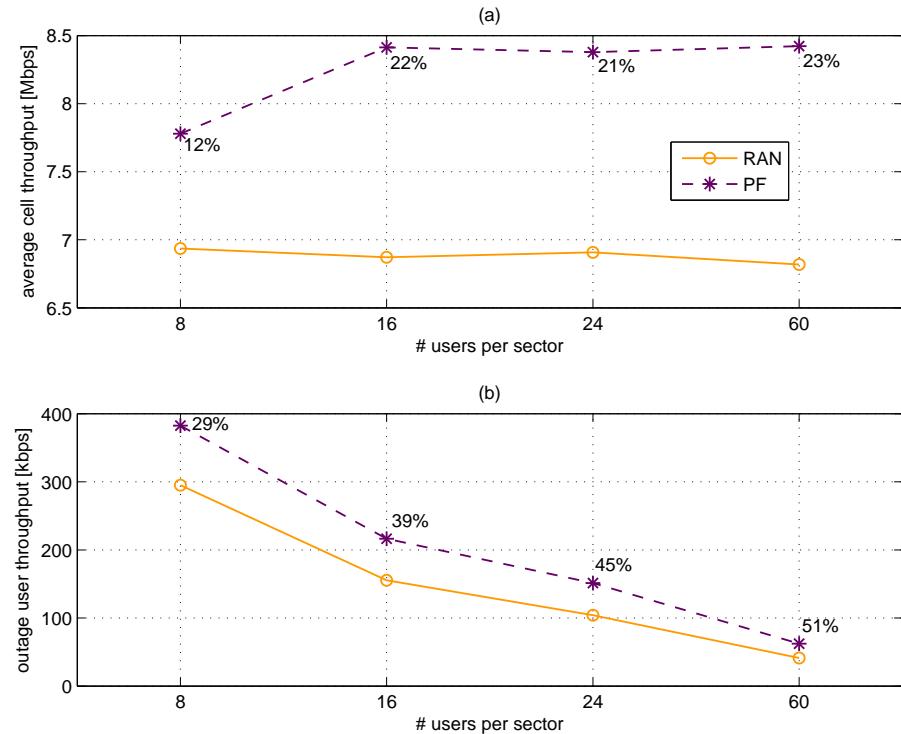


Figure 4.6: Average cell throughput and 5% outage user throughput under RAN and PF metrics for 8, 16, 24 and 60 users per sector. Macro 1 case. 6 PRBs per user.

Figure 4.6 shows the absolute values of the average cell throughput and outage user throughput

as well as the gains of the PF over the RAN metric for different number of users and a fixed bandwidth of 6 PRBs.

The first observation is that the gain in outage user throughput is higher than the gain in average cell throughput for all the considered cases. Indeed, the gain at 5% outage of the user throughput is generally higher than the gain in average sector throughput due to the non-linear mapping of SINR to throughput. Figure 4.7 explains this more explicitly by showing the average user throughput vs the Average Path Gain (APG)⁴ for the cases of 8 and 16 users under RAN and PF metrics. It is noticeable how the absolute increase of throughput of the PF over the RAN is similar over the whole range but the increase in percentage becomes smaller simply because the absolute throughput value used as reference grows higher.

The second observation is that the gain increases by approximately 10% in both average cell throughput and outage user throughput when the number of users per sector is increased from 8 to 16 while the number of scheduled user is kept to 8.

To explain this effect we can refer to Figure 4.7 where the average user throughput is shown vs the APG. In Figure 4.7(a) with 8 users per cell, the best users reach the same maximum throughput under PF and RAN and the curves overlap in the upper range. The reason is that after the users reach the highest MCS and their BLER starts decreasing, the expected throughput \hat{T} , and consequently the PF metric, becomes the same over all the PRBs (and all the RCs) making the allocation of such users similar to the one performed by the RAN metric. Similarly, the lowest throughput value in Figure 4.7(a) is the same under the PF and RAN metrics because the selected MCS is the lowest (as can be expected given the higher BLER) and the expected throughput \hat{T} as well as the PF metric become the same over all the RCs.

Figure 4.7(b), which shows the same variables for the case of 16 users, verifies this interpretation. In this case, the gain under the PF metric is visible over the full range of throughput values including the minimum and maximum.

The higher minimum throughput value of PF compared to RAN is explained by looking at Figure 4.8(d) where the scheduling activity for the PF is higher than for the RAN. This is due to the higher experienced BLER, which not only increases the scheduling activity of both RAN and PF by increasing the number of retransmissions, but also lowers down the past average throughput \bar{T} thus increasing the PF priority and consequently the scheduling activity. This does not happen in the case of 8 users per cell because the scheduling activity is always 100% for all the users as shown in Figure 4.8(c). The higher maximum throughput is also explained by the higher scheduling activity under the PF metric with respect to the RAN metric.

With 24 and 60 users no further gain in average cell throughput is shown in Figure 4.6 while there is an improvement in outage user throughput due to different scheduling activity.

⁴By average path-gain we mean the product (in linear) of distance-dependent path-gain, shadowing and antenna gain. It is sometimes simply referred to as path-gain and is the inverse (in linear) of the path-loss.

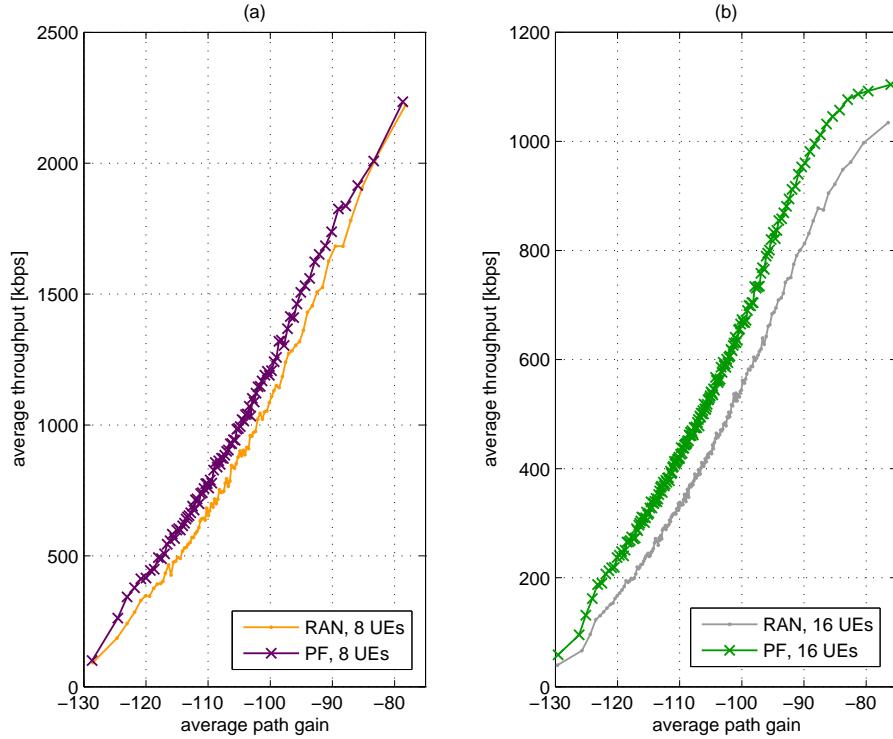


Figure 4.7: Average throughput vs APG using RAN and PF metrics for 8 and 16 UEs per sector. Each point represents the mean of 50 values consecutive with respect to the APG.

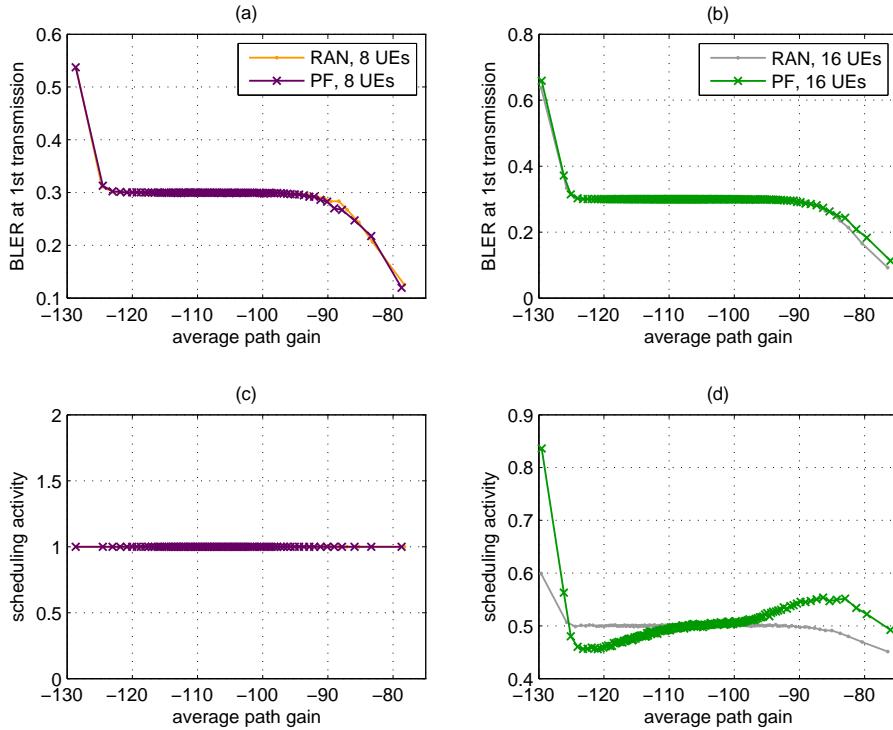


Figure 4.8: (a) and (b): BLER at 1st transmission vs average path gain. (c) and (d): scheduling activity vs average path gain. RAN and PF metric. Macro 1 case. 8 and 16 users per sector. 6 PRBs per user. Each point represents the mean of 50 consecutive values.

Figure 4.9 shows for the Macro 3 case the same performance indicators of average cell throughput and outage user throughput already shown for the Macro 1 case. Unlike the Macro 1 case, we can observe that with 8 users per cell there is no gain in outage cell throughput. This occurs because users at 5% outage have a BLER higher than the target (unlike in Macro 1 where only few users well below the outage have a BLER higher than the target) and therefore also a PF metric which behaves like a RAN metric without differentiating among RCs or PRBs. This explains not only the same minimum and maximum average throughput for PF and RAN metrics shown in Figure 4.10(a) but also the overlapping of the two curves in the lower throughput range characterized by a BLER higher than the target as shown in Figure 4.11(a). The outage gain for the case of 16 users per cell is due to those users having a higher scheduling activity (see Figure 4.11(d)) because of a BLER higher than the target (see Figure 4.11(c)). So while in the Macro 1 case the gain in outage for 16 UEs is actually a channel gain due to the PF metric preserving aspects of channel awareness via the expected throughput \bar{T} , in the Macro 3 case the gain in outage is a scheduling activity gain due to the PF metric having a lower past average throughput \bar{T} which is caused by an experienced BLER higher than the target.

Figure 4.10 compared to Figure 4.7 shows another difference between the Macro 3 and the Macro 1 cases. The average throughput versus the APG decreases more rapidly below -133 dB of average path gain because the users incur into power limitations. Such value can be easily derived from the OLPC formula where the number of PRBs is set to 6⁵. Moreover, comparing Figure 4.10(a) and (b) we notice a different trend in the lowest throughput range which can be explained again considering the different scheduling activity respectively under 8 and 16 users per cell.

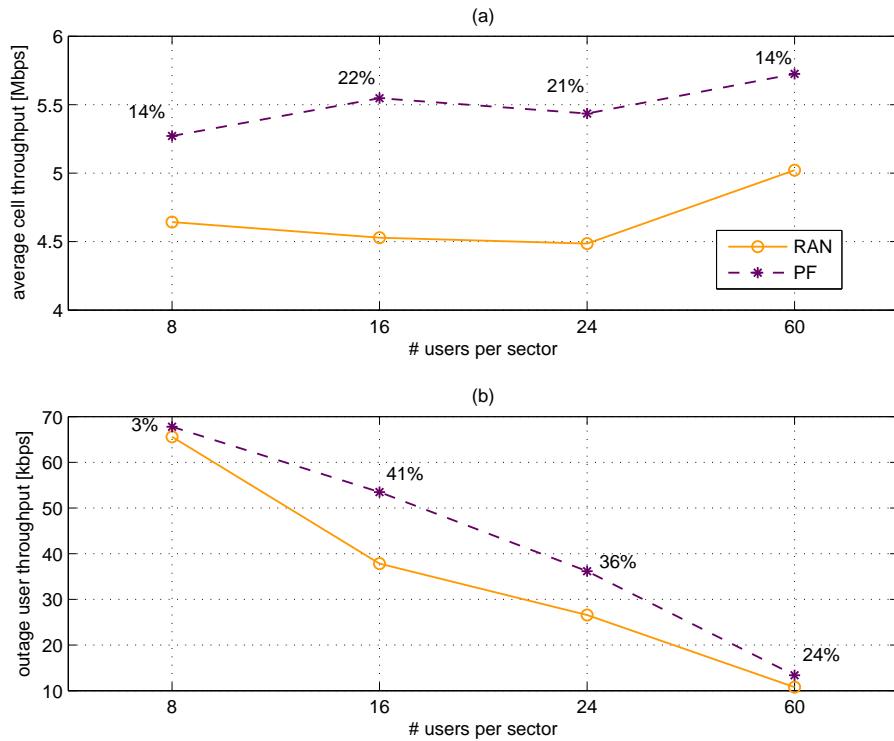


Figure 4.9: Average cell throughput and 5% outage user throughput under RAN and PF metrics for 8, 16, 24 and 60 users per sector. Macro 3 case. 6 PRBs per user.

⁵At maximum transmit power we have $P_{max} = P_0 + 10 \cdot \log_{10}(M) + \alpha \cdot L \Rightarrow L = \frac{P_{max} - P_0 - 10 \cdot \log_{10}(M)}{0.6} = \frac{24dBm + 64dBm/PRB - 7.7dB}{0.6} \approx 133dB$

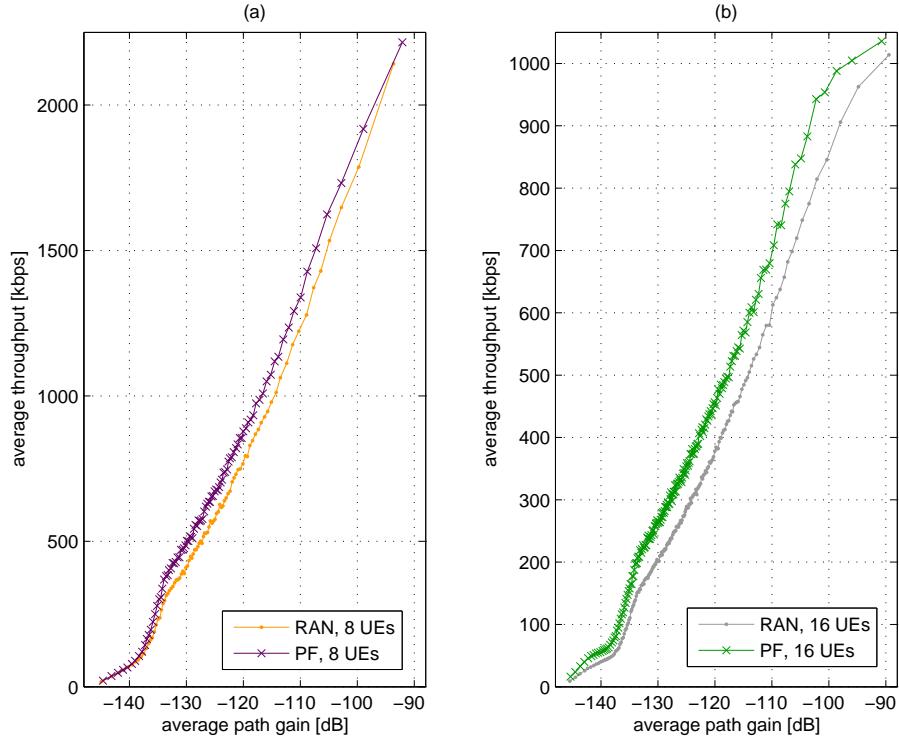


Figure 4.10: Average throughput vs APG using RAN and PF metrics for 8 and 16 users per sector. Each point represents the mean of 50 consecutive values. Macro 3 case.

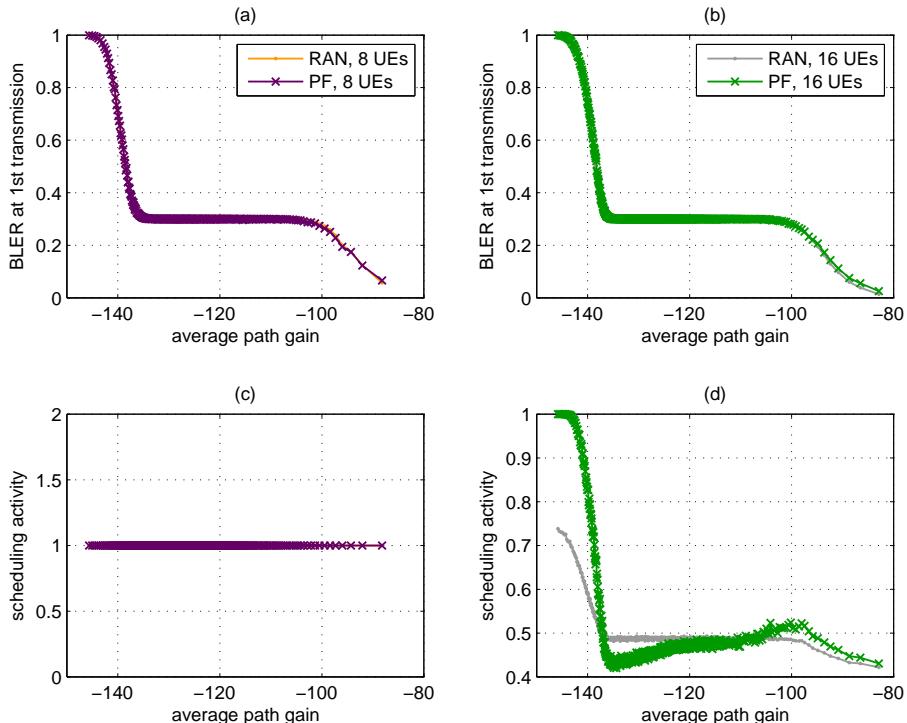


Figure 4.11: (a) and (b): BLER at 1st transmission vs average path gain. (c) and (d): scheduling activity vs average path gain. RAN vs PF metric. Macro 3 case. 8 and 16 users per sector. 6 PRBs per user. Each point represents the mean of 50 consecutive values.

Another interesting analysis pertains the gain achievable by modifying the bandwidth of the RC allocated to each UE as shown in Figure 4.12. With a RC of 2 PRBs we have a frequency selectivity equal to the granularity of the CSI reports and similar to the coherence bandwidth of the channel. Therefore, for the same number of UEs, we achieve the highest gain in average cell throughput and outage user throughput. Larger bandwidth values lead to a lower throughput under the PF metric because of lower frequency diversity and to a slightly larger throughput under the RAN metric because of higher coding gain (included in the AVI tables) consequently reducing the gain of FDPS.

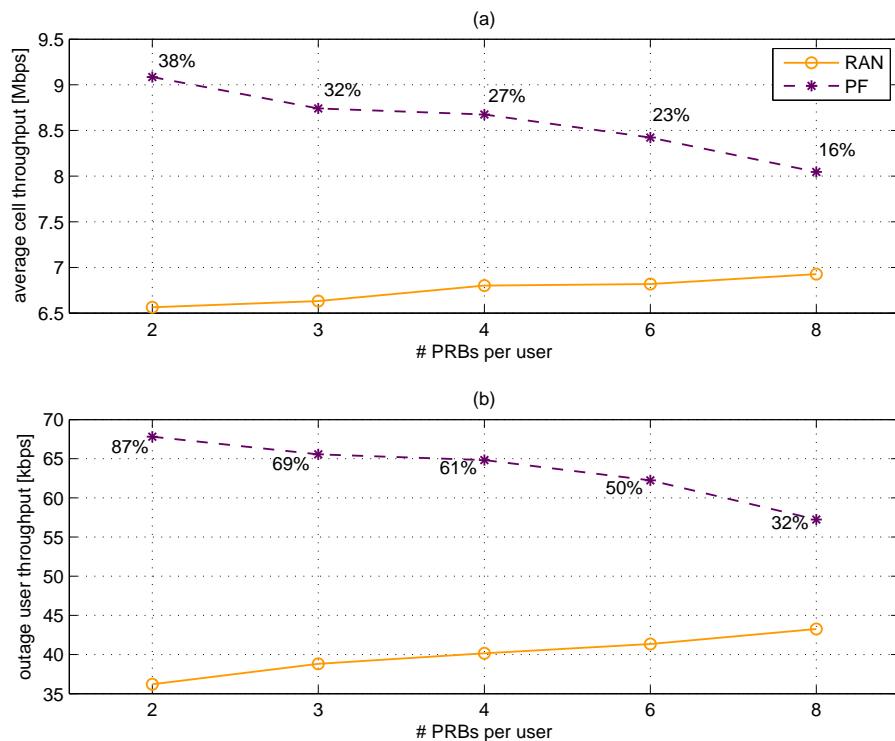


Figure 4.12: Average cell throughput and 5% outage user throughput under RAN and PF metrics for 2, 3, 4, 6 and 8 PRBs per user. 60 users per sector. Macro 1 case.

The Macro 3 case depicted in Figure 4.13 shows some differences compared to the Macro 1 which can be summarized in a much larger outage gain for a bandwidth of 2 PRBs compared to the Macro 1 case and in a much larger drop of such gain as soon as the bandwidth becomes larger than 2 PRBs.

The first point is easily explicable considering again the logarithmic SINR to throughput mapping. Indeed the absolute increase of throughput under Macro 3 is actually smaller than the one obtained under Macro 1 (the user is under power limitations and close to having a BLER higher than the target with a consequent lower variability of the PF metric over the different RCs) but such increase is related to a smaller value of throughput given that in Macro 3 the user at 5% outage experiences a much lower average path gain. As a result the percentage gain is significantly higher.

The second point is explained considering that a larger bandwidth (e.g. 4 PRBs) actually brings the UEs into more serious power limitations leading to a BLER increase and therefore to a gain of the PF metric over the RAN only due to higher scheduling activity. Figure 4.14 shows that the scheduling activity at 5% outage is the same under PF and RAN metrics for the case of 2 PRBs while it's higher for PF than RAN for the case of 4 PRBs indicating that the gain obtained for 2

PRBs is channel gain while the gain obtained for 4 PRBs is gain due to larger scheduling activity.

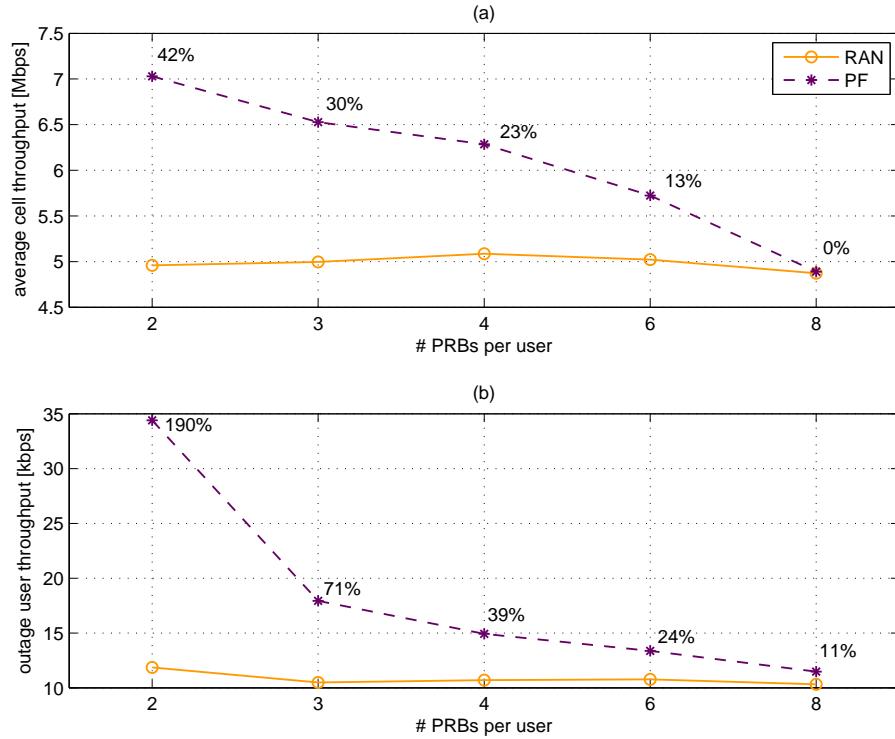


Figure 4.13: Average cell throughput and 5% outage user throughput under RAN and PF metrics for 2, 3, 4, 6 and 8 PRBs per user. 60 users per sector. Macro 3 case.

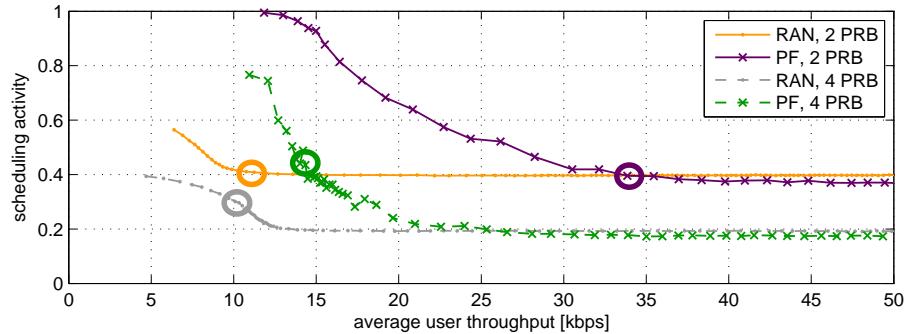


Figure 4.14: Scheduling activity vs average throughput for 2 and 4 PRBs and for PF and RAN metrics for 60 UEs per sector. Macro 3 case. The circles indicate the UEs at 5% outage.

Finally Figure 4.15 shows the average cell throughput performance under the PF metric for different combinations of RCs and number of users per sector. The first interesting aspect is that there are several cases where, for the same number of users, a larger bandwidth gives a larger throughput than a smaller one. See, for example, the case of 12 users combined with bandwidths of 4 or 6 PRBs. As general rule, it is preferable to have a larger bandwidth and more users than it is possible to schedule rather than a narrower bandwidth and as many users as it is possible to schedule. Another visible trend is that the multi-user diversity gain saturates faster (with respect to the number of users) for a larger bandwidth than for a smaller one.

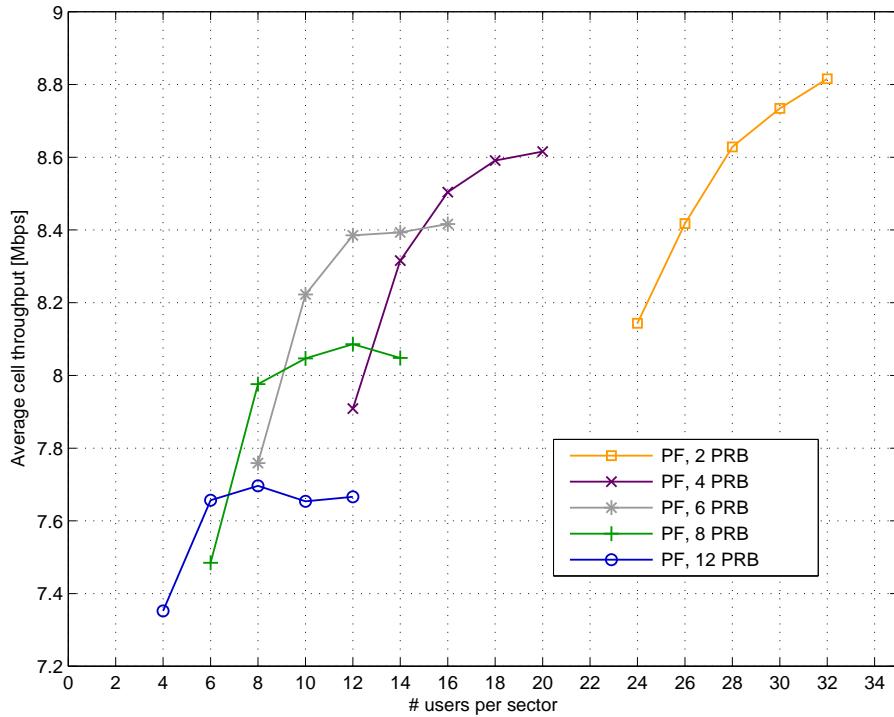


Figure 4.15: Average sector throughput for different number of PRBs and different number of users under the PF metric. Macro 1 case.

4.6 Tree-Based Search Algorithm and Results

In this section we are going to propose a more general scheduling algorithm which includes the matrix-based approach as special case. We have shown that the matrix-based approach provides a significant gain over a blind scheduling (like the one obtained under the RAN metric), but does not achieve the global optimum. As an example let's consider a simple case with two UEs and two RCs with the fictitious metrics in Figure 4.16.

	RC ₁	RC ₂
UE ₁	$M_{1,1}=960$	$M_{1,2}=980$
UE ₂	$M_{2,1}=870$	$M_{2,2}=970$

Figure 4.16: Simple scenario with two UE and two RCs: the algorithm fails to identify the optimum.

If we apply the algorithm used so far, we would end up with RC₂ allocated to UE₁ and RC₁ allocated to UE₂. The resulting global metric, which was defined as the sum of the metrics of the allocated users, would be $M_{sum} = 1850$. Performing the opposite allocation (RC₁ to UE₁ and RC₂ to UE₂) would provide the maximal global metric $M_{sum} = 1930$. This offers a hint on how the algorithm could be improved to perform a more exhaustive search. In every step of the search algorithm, rather than considering only the best metric, we also consider what is, *globally*, the second best metric. In other words, in every derived sub-matrix we pick the two best metrics

and build, in this way, a binary search tree where the best mapping of UEs to RCs is given by the path with the highest sum of metrics. The number of branches departing from a node of the tree is called out-degree and is indicated in the following as N_{out} . For a binary tree $N_{out} = 2$, but the approach can easily be generalized to any value of N_{out} . As anticipated at the beginning of this paragraph the matrix-based algorithm is equivalent to the tree-based algorithm for an out-degree of 1 ($N_{out} = 1$). This means that the search-tree algorithm is a generalization of the greedy matrix-based algorithm which considers more and more possibilities as the value of N_{out} increases. However, the number of combinations increases dramatically with N_{out} for a reasonable number of users and this represents a serious limitation considering the real time constraints. For this reason N_{out} should not be higher than 2 also considering that the gain provided from increasing such a parameter very soon saturates. In any case this algorithm could be thought of as a novel application of a well known search algorithm to the problem of scheduling.

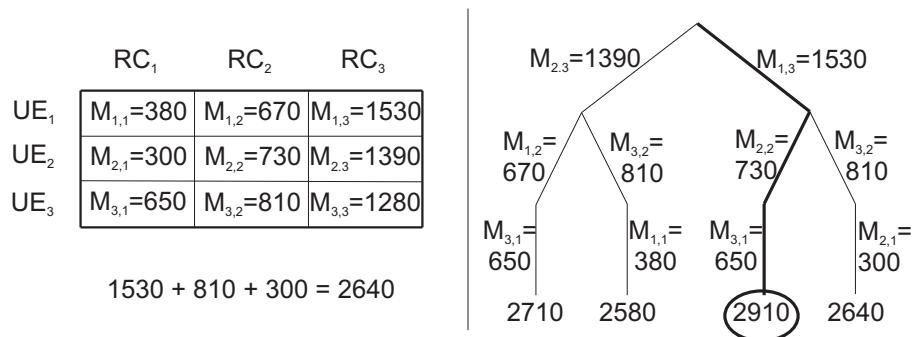


Figure 4.17: Scheduling example with three UEs and three RCs. To the left the circles indicate the allocation performed using the matrix algorithm. To the right the thick line indicates the allocation performed using the tree algorithm with $N_{out} = 2$.

Fig.4.17 provides an example of a binary tree comparing the two algorithms for the case of three UEs and three RCs. On the left the circles indicate the metrics chosen by the matrix algorithm. On the right the thick line indicates the path in the binary tree which leads to the best allocation of RCs to UEs for the considered combinations. For the fictitious metric values considered, the tree algorithm is able to provide two allocations (corresponding to two paths) whose global metric is higher than the one provided by the matrix algorithm.

The gains will of course depend on the nature of the metric adopted. Interestingly enough, in the case of the PF metric, the throughput improvement provided by the search-tree algorithm is very limited and not as significant as to justify the increased complexity. This means that, by applying the most efficient scheduling strategy, we are also able to exploit most of the throughput gain.

Figure 4.18 shows the KPIs for a few selected combinations of number of users and PRBs and N_{out} degree of 1 and 2. The gains are approximately 2% in average cell throughput and 6% in outage user throughput for the cases in which there are as many users as it is possible to schedule (8 UEs and 6 PRBs or 12 UEs and 4 PRBs) while are even lower - 1% in average cell throughput and 3% in outage user throughput - in the other cases.

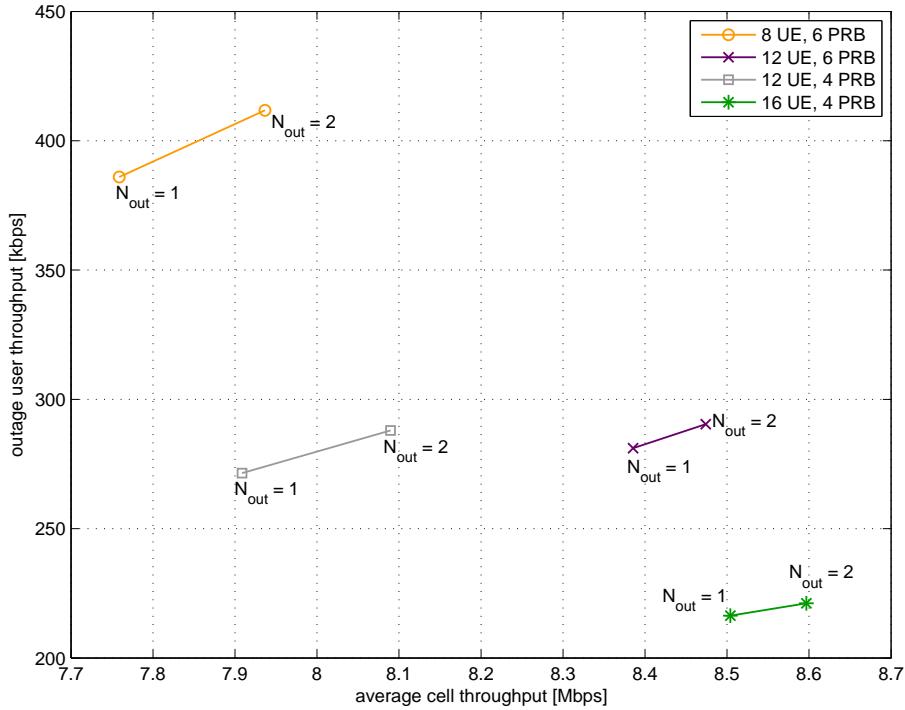


Figure 4.18: Average sector throughput and outage UE throughput for $N_{out} = 1$ and $N_{out} = 2$. Macro 1 case.

4.7 Conclusions

In this chapter we have evaluated, in different scenarios, the performance of channel-aware scheduling by comparing Proportional Fair (PF) and Random (RAN) metrics. A greedy matrix-based search algorithm is used for the allocation.

The results show a variable gain of the PF metric depending on the number of users and the size of the Resource Chunk (RC). For a Macro 1 case the average cell throughput gain can range from 4% in case of 4 users per cell and a RC of 12 PRBs (per user) to 38% in case of 60 users and a RC of 2 PRBs (per user). The Macro 3 case shows a much larger range of gains, especially in outage user throughput, due to power limitations which impact the performance especially for a RC size larger than 2 PRBs.

In the last part a generalization of the matrix-based search referred to as tree-based search is proposed as allocation algorithm. Such algorithm is characterized by a higher computational complexity due to the larger search space but does not achieve a significant gain. For this reason it is set aside and the simpler yet effective matrix-based algorithm is retained and referred to as the FTB-based algorithm.

It is worth noting that the gains shown don't take into account the constraints of control channels and computational complexity which will limit the number of users which can be scheduled in one TTI. In this sense the case of 60 users per cell, for example, is used only to find the limits of user diversity gain. These and other aspects will be further discussed in the following chapters.

Chapter 5

Adaptive Transmission Bandwidth based Packet Scheduling

5.1 Introduction and Motivation

This chapter is closely related to the previous one in that it deals with channel aware scheduling in the frequency domain. The main novelty here is represented by the introduction of the ATB feature which is merged into the scheduling algorithm. The ATB is a functionality needed to accommodate for different traffic types (e.g. VoIP, which requires a limited bandwidth and BE which would take as much bandwidth as it is left), different data rate requirements and different user power capabilities. Integrating the ATB into the PS would result in a RRM framework simpler than it would otherwise be if each functionality had its own algorithm as well as more flexible. We have seen in the previous chapter how the selection of the most appropriate bandwidth depends on the number of scheduled users. The advantage of the ATB-based approach is that no additional algorithm is required to tune the bandwidth. In other words, the capability of coping with varying traffic loads and power limitations¹ is inbuilt in the algorithm in the sense that the ATB-based PS guarantees automatic adaptation to variations in the cell load. Moreover QoS-aware schedulers, required to cope, for example, with different data rate requirements, can be easily combined with this kind of channel-aware scheduling by weighting the frequency-selective scheduling metrics and allowing, in this way, a different distribution of the bandwidth to different users based on their requirements.

The chapter is therefore devoted to the performance analysis of such ATB-based scheduling algorithm proposed during this PhD and published in [43]. As the main addition is the exploitation of the ATB functionality when performing the allocation of PRBs, it is natural to compare, under the same conditions and scenarios, its performance with the one obtained under the FTB-based algorithm previously described.

This chapter is organized as follows: The ATB-based PS is described in Section 5.2. The modeling assumptions are listed in Section 5.3 while the performance evaluation is carried out in Section 5.4. The conclusions are presented in Section 5.5.

5.2 Algorithm Description

The main challenge to face in attempting the introduction of the ATB functionality in the scheduling is represented, as already highlighted, by the SC constraint. As soon as a PRB is allocated to

¹With the generic term of power limitation we refer to the situation where users use the maximum transmit power because of their low average path gain

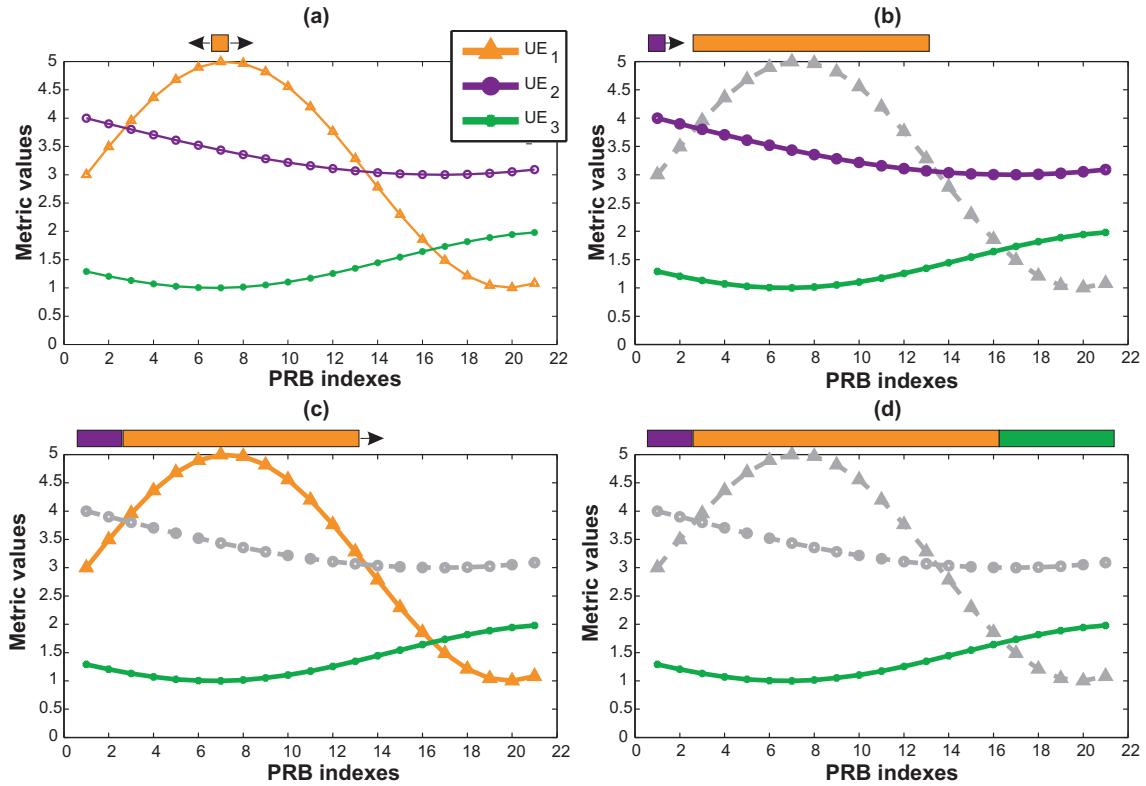


Figure 5.1: Algorithm description with 3 UEs and 21 PRBs. The gray dashed curves indicate UEs (and associated metrics) which have been temporarily or permanently excluded.

a user, different constraints are created. First, for another PRB to be allocated to the same user, it will have to be adjacent to the first one. Secondly, another user to be allocated will find the available bandwidth to be already fragmented so that its bandwidth will be located either entirely to the right of the first PRB or entirely to the left.

For this reason the proposed algorithm aims at containing the bandwidth fragmentation by picking the user with the highest metric and then expanding its bandwidth as long as its metric is highest. The algorithm is heuristic but, on the other hand, a theoretical approach would be impractical because of computational complexity as also shown in [42].

The steps of the algorithm, for first transmission users, are exemplified in Figure 5.1 and described as follows:

1. Find, within the matrix of metric values, the UE i and the PRB j with the highest metric value² and allocate PRB j to UE i .
2. Expand the bandwidth of UE i until one of the following conditions is met:
 - (a) another user has a higher metric on the adjacent PRB (Figure 5.1(a));
 - (b) the expansion has reached physical constraints on one side (a bandwidth edge or another user already allocated) and condition (a) on the other side;
 - (c) the expansion has reached physical constraints on both sides;
 - (d) the estimated transmit power is above the maximum.

²For a user partially allocated only the adjacent PRBs are considered.

3. Temporarily exclude UE i and its metric values if conditions 2(a) or 2(b) are verified, otherwise permanently exclude the user.
4. Repeat steps 1, 2 and 3 considering the reduced set of users and metrics (Figure 5.1(b)).
5. If any, readmit temporarily excluded users as further expansion may be possible because of the exclusion of other users and relative metrics (Figure 5.1(c)).
6. Repeat the steps from 1 to 5 until all the users have reached a permanent stopping condition (Figure 5.1(d)).

The retransmissions, when they occur, are placed starting from the left edge of the bandwidth. This is possible because of the adaptive HARQ and is done to avoid bandwidth fragmentation. In this way the algorithm can be applied to first transmission users within the remaining portion of the bandwidth.

5.3 Modeling assumptions

Table 5.1: Main simulations parameters

Parameter	Setting
# UEs per sector	from 4 to 60
# UEs handed to the FD	from 4 to 60
# PRBs per UE	[1 to 24] and [2 to 24]
TD scheduling	Not considered
FD scheduling	RAN or PF
BLER target at 1st transmission	30%
PF filter length (N_{TTI})	100 ms
Initial throughput value ($T(0)$)	500 kb
α (for PC)	0.6
P_0 (for PC)	-58 dBm/PRB for Macro 1; -64 dBm/PRB or -62 dBm/PRB for Macro 3
Traffic model	Full buffer with balanced load

The modeling assumptions followed for the performance evaluation of the ATB-based algorithm are like the ones described in Section 4.4 for the FTB-based algorithm and are reported in Table 5.1. It is worth reminding the use of the full (infinite) buffer traffic model as well as the deployment of only the FD scheduler. First some statistics related to the ATB functionality are presented. Afterwards, the performance evaluation is presented in form of comparison with the FTB-based algorithm.

5.4 Performance Evaluation

Unlike the FTB-based algorithm, the ATB-based scheduling allocates a variable number of UEs in FD depending on a set of parameters including the number of UEs handed over to the FD, the metric deployed in FD and the power limitations of the UEs.

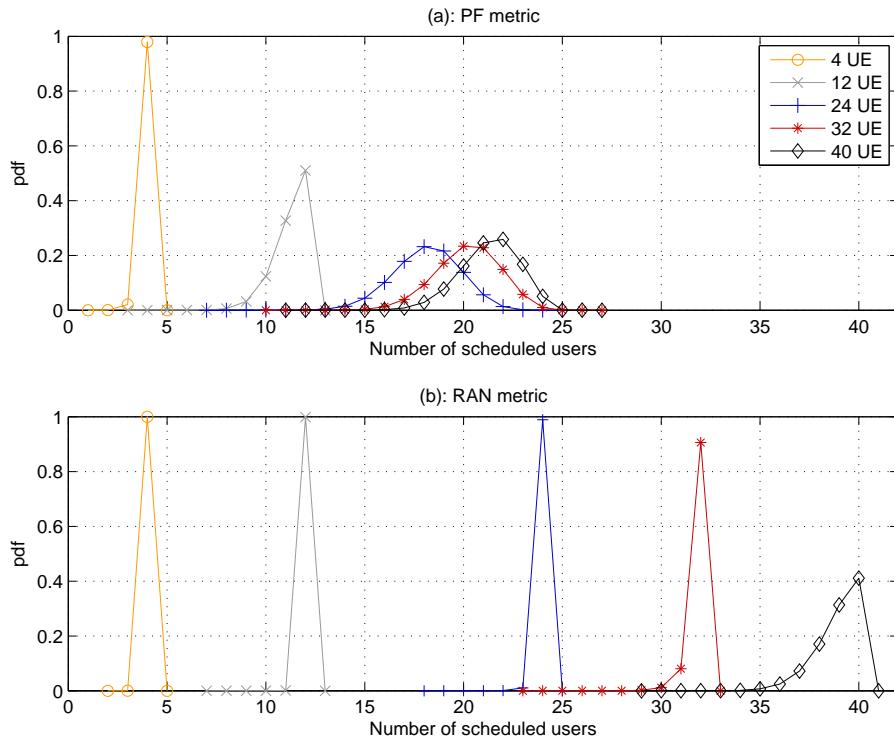


Figure 5.2: Distribution of the number of scheduled users per TTI for a different number of users per sector. RAN and PF metrics.

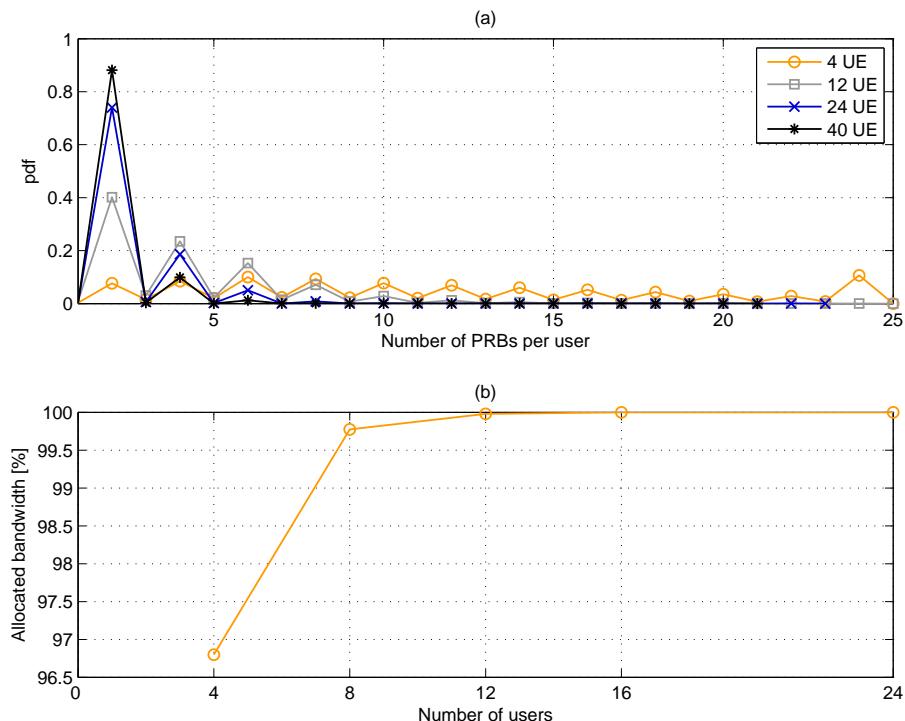


Figure 5.3: Number of scheduled PRBs per user per TTI and total bandwidth utilization for a different number of users. Macro 1 case. PF metric.

Figure 5.2 shows an example of such distributions for a different number of UEs under the PF and RAN metrics, both introduced in Chapter 4. The PF metric, being channel aware, is affected by the coherence of the channel over the frequency and therefore, as the number of users grows, it tends to allocate a lower number of users per TTI compared to the RAN metric which, instead, does not have any correlation.

Related to the distribution of the number of scheduled users is the number of scheduled PRBs per user per TTI and the total bandwidth utilization which are shown in Figure 5.3. The higher probability of allocation of an even number of PRBs seen in Figure 5.3(a) is a consequence of the frequency resolution of the CSI set to two PRBs. An odd number of PRBs is allocated when at least one user is affected by power limitations. If this was not the case the expansion of the user would occur in steps of 2 PRBs.

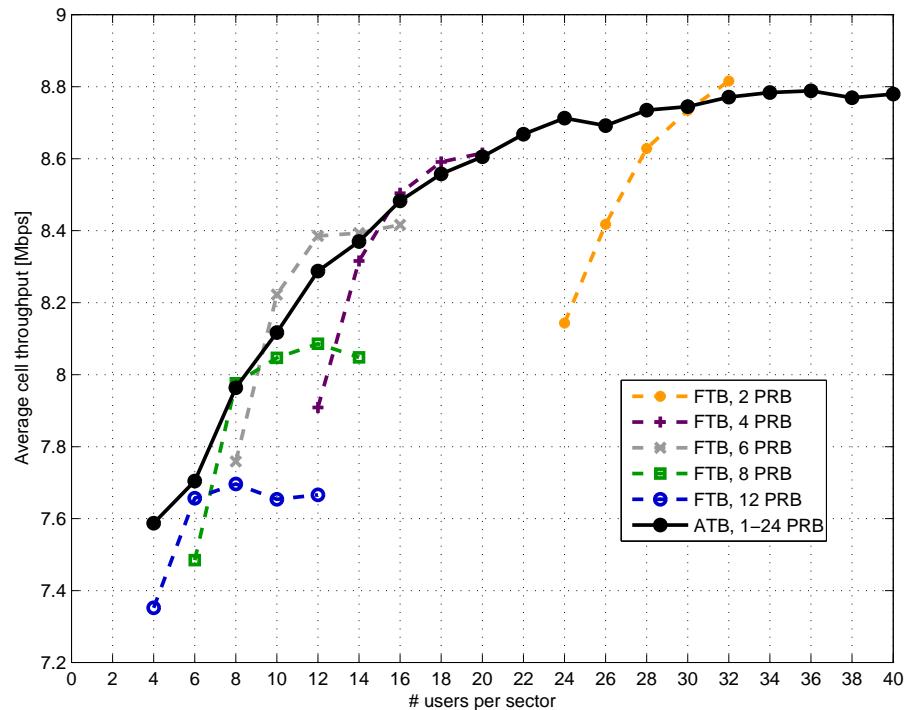


Figure 5.4: Average cell throughput under FTB and ATB for increasing number of users per sector. Macro 1 case. PF metric.

The main benefit provided by the ATB is the inbuilt bandwidth adaptation which enables the exploitation of multi-user diversity. Figure 5.4 shows the performance of the ATB algorithm against the performance of the FTB algorithm obtained for different number of users and different bandwidths. The ATB is able to exploit the multi-user diversity without incurring into gain saturation like it happens for the cases of fixed bandwidth. Even though not perfectly, the curve obtained under ATB tends to follow the envelope of the curves obtained for different fixed bandwidths. Like the AMC, which adapts to the SINR by selecting the most appropriate MCS, the ATB adapts to different cell loads by selecting the most appropriate bandwidth. Some combinations of bandwidth and number of users show a slightly better performance in case of the FTB algorithm because we are using two different algorithms for FTB and ATB but the parallelism between AMC and ATB helps understanding the flexibility introduced by the latter. The loss in performance visible in some cases occurs because the ATB, for simplicity of design, deals in a simplified manner with the HARQ retransmissions compared to the FTB. While the FTB based algorithm treats equally first transmissions and retransmissions so that both benefit from multi-user diversity, the ATB, in

order to avoid bandwidth fragmentation, simply schedules the retransmissions in sequence at the beginning of the system bandwidth so that they don't benefit from multi-user diversity. Moreover the first transmissions have a more limited frequency diversity, and therefore a lower gain, because their allocation is optimized within a comparatively smaller bandwidth (in average 70% of the system bandwidth given the 30% BLER target).

Figure 5.5(a) summarizes the behavior of the ATB anticipated in the previous figures: as the number of users in the cell increases, the number of scheduled users per TTI increases as well (though not linearly) while the average bandwidth per user decreases. In case of full bandwidth utilization the two quantities are related by the basic formula:

$$\overline{PRB}_{UE_{sched}} = \frac{N_{PRB}}{\overline{UE}_{sched}} \quad (5.1)$$

where N_{PRB} indicates the number of PRBs available for scheduling in the system bandwidth, \overline{UE}_{sched} indicates the average number of scheduled users per TTI and $\overline{PRB}_{UE_{sched}}$ indicates the average number of allocated PRBs per scheduled user.

Figure 5.5(b) compares the bandwidth per user under ATB with the bandwidth per user under FTB. The dashed line indicates the combinations of number of users and number of PRBs which give the best average cell throughput performance for the FTB algorithm. For a given number of users, there is only one specific bandwidth (or RC size) under which the FTB can achieve a performance similar to the ATB and such bandwidth is generally larger than the average one allocated by the ATB (because, as already shown, multi-user diversity under FTB is preferable to the frequency selectivity). For all the other bandwidths the FTB performs worse than the ATB. Specifically, if we consider the same number of users and the same number of PRB (fixed or in average) the ATB offers a significantly higher performance compared to the FTB as shown in Figure 5.4.

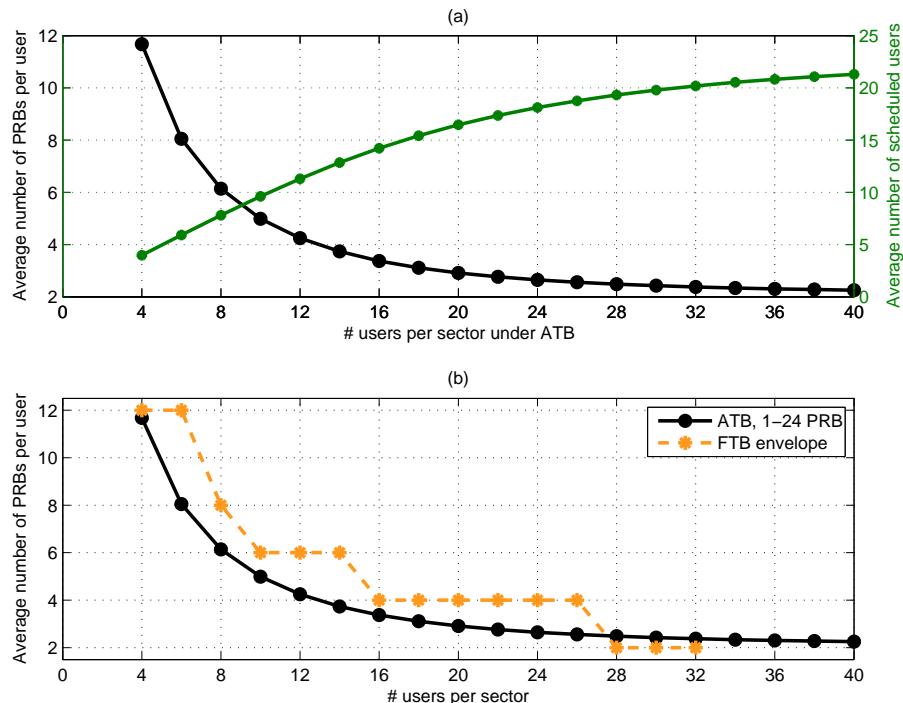


Figure 5.5: (a) Average number of scheduled PRBs and scheduled users for increasing number of users per sector under ATB. (b) Average number of scheduled PRBs for increasing number of users per sector under ATB and FTB.

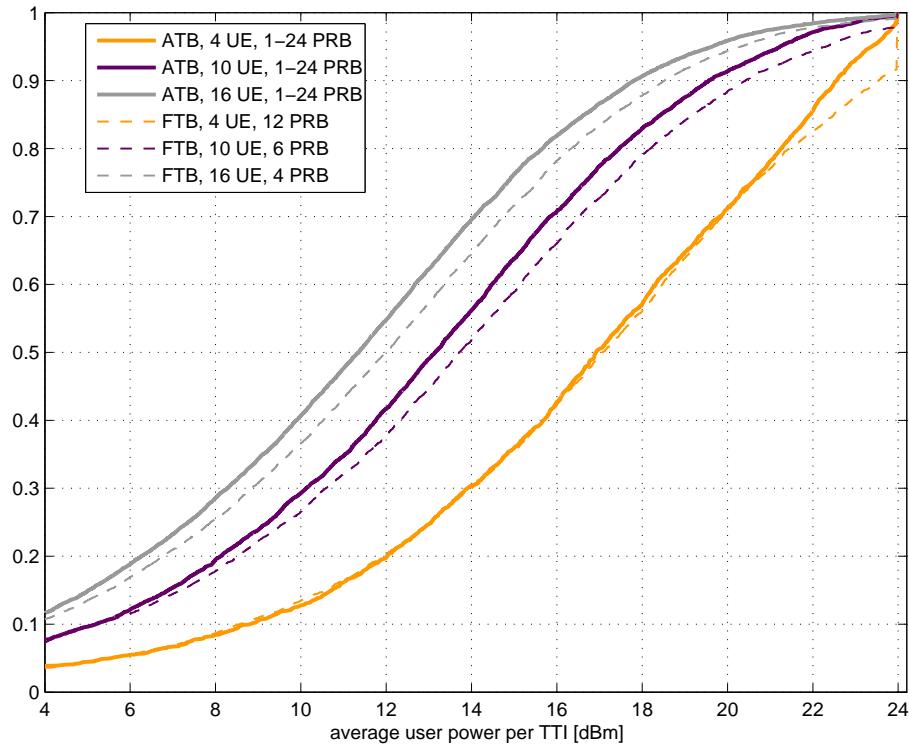


Figure 5.6: Average user transmit power scaled to scheduling activity for three different cases of number of users per sector under FTB and ATB. The bandwidths chosen for the FTB under different number of users per sector are the ones providing the largest cell throughput. Macro 1 case. PF metric.

The difference in bandwidth is reflected also in the distribution of the power per user normalized to the scheduling activity of the user as shown in Figure 5.6.

For the same number of users the ATB shows a lower transmit power than the FTB counterpart because the average transmit bandwidth is generally lower. Of the three cases depicted the case of 4 users has the same transmit bandwidth for both ATB and FTB but shows overlapping curves only up to approximately 21 dBm. After such threshold the transmit power under FTB becomes higher than ATB because the ATB, by design, stops expanding the user bandwidth when the maximum power limitation is reached. In the Macro 1 case this results in having no users in power limitations under the ATB algorithm and a certain fraction of users (between 1% and 10% depending on the bandwidth chosen) in power limitations under FTB. Having a certain percentage of users in power limitation is beneficial for the performance of the corresponding percentile outage user throughput. For this reason looking only at the 5% outage user throughput would show a performance for the ATB lower than for the FTB, because the PC parameter is optimized for the 5% outage under the FTB algorithm with 6 PRBs. A more complete picture which considers also a smaller percentile value (e.g. 1%) is depicted in Figure 5.7. The set of curves for the 1% outage user throughput show that the ATB is always on top of the other curves with the exception of the curve obtained for 4 PRBs. The reason is that such a curve has indeed 1% of the users in power limitations and therefore will have the highest outage throughput at the first percentile and it will slightly outperform the ATB which doesn't have any user under max transmit power.

Another interesting aspect we can derive from the outage curves is that the FTB is more sensible to the power settings than the ATB. On the one hand this means that the FTB can be finely tuned as to maximize even a specific percentile of the outage user throughput. On the other hand this also means that such a setting is quite likely to be not optimal if we consider a different bandwidth or a different percentile. It's the case for the curve obtained for 8 PRBs which has a good

performance at the 5% outage but a quite poor one at the 1% outage as it's the case for the 5% outage curves obtained for 6 PRBs and 2 PRBs which are respectively performing better and worse than the ATB. Unlike the FTB, the ATB is more robust to the different power settings (specifically P_0) and therefore it shows a reasonably good performance for the different number of users and for the different percentiles with the 1st percentile being the best performing. What makes the ATB more robust to different power settings is the capability of reacting to different power settings by limiting the minimum bandwidth in a way that the FTB cannot.

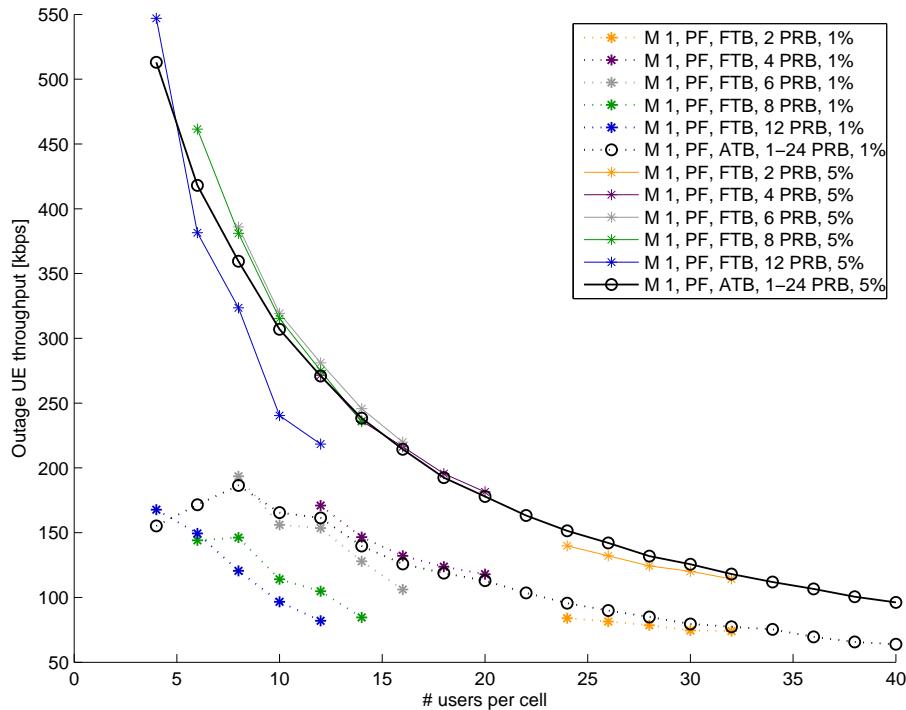


Figure 5.7: 1% and 5% outage user throughput under FTB and ATB for an increasing number of users per sector. Macro 1 case. PF metric.

In the Macro 3 case the scenario changes significantly because power limitations become even more constraining at the point of increasing the experienced BLER beyond the target for a significant percentage of users. This is shown in Figure 5.8 where the transmission power per user and the corresponding BLER at 1st transmission are shown for the cases of 4, 10 and 20 users. As the number of users increases and the allocated bandwidth per user decreases the power distribution curves tend to look more alike and so is the BLER performance. The difference in power limitation is reflected also in the lower range of the instantaneous SINR distribution shown in Figure 5.9.

Figure 5.10 and 5.11 show respectively the average cell throughput and the 5th percentile of the outage user throughput. For the average cell throughput we can see a gain in all the considered cases. As for the outage user throughput we can see that the FTB algorithm offers a better performance but the performance of the ATB can be improved by increasing to 2 the minimum number of PRBs allocated per user.

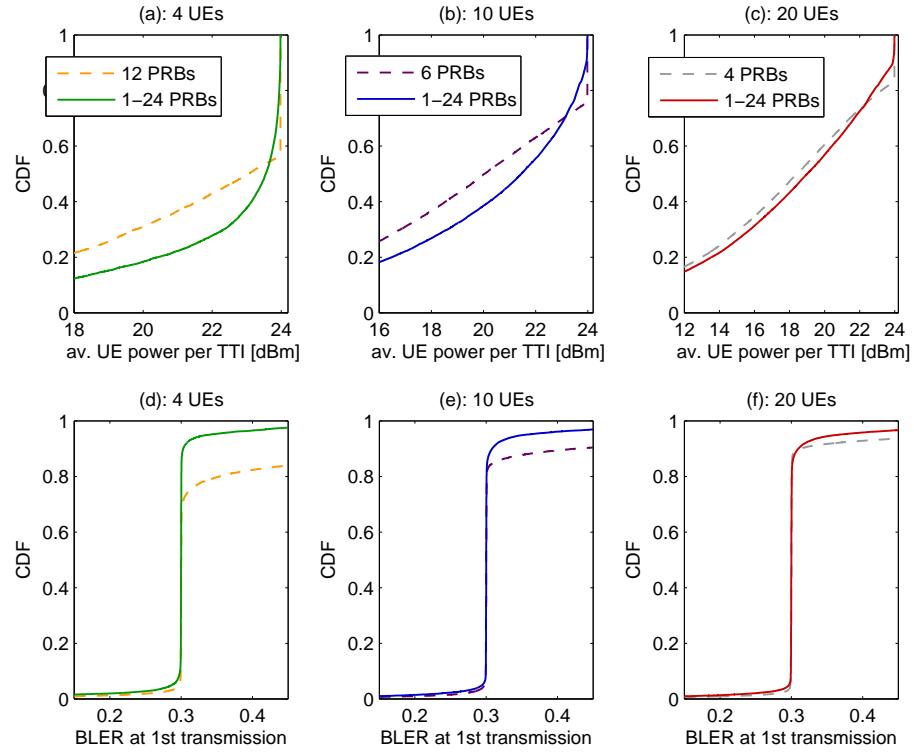


Figure 5.8: Average user power per TTI (that is, normalized to scheduling activity) and BLER at 1st transmission under FTB and ATB for 3 different numbers of users per sector. -64 dBm/PRB and -62 dBm/PRB respectively for FTB and ATB. Macro 3 case.

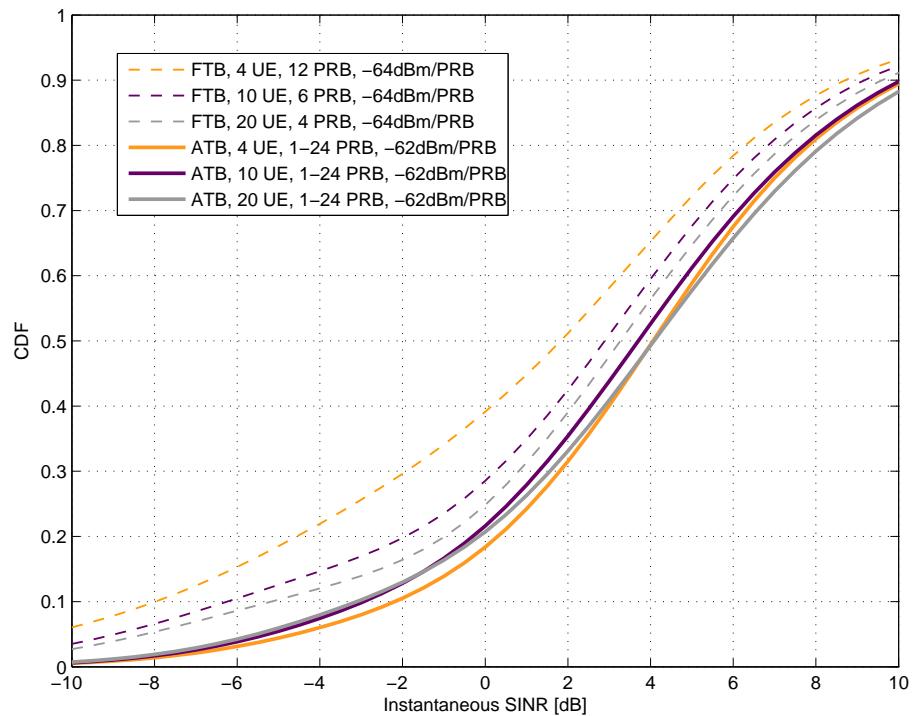


Figure 5.9: SINR distribution under FTB and ATB for 3 different numbers of users per sector.

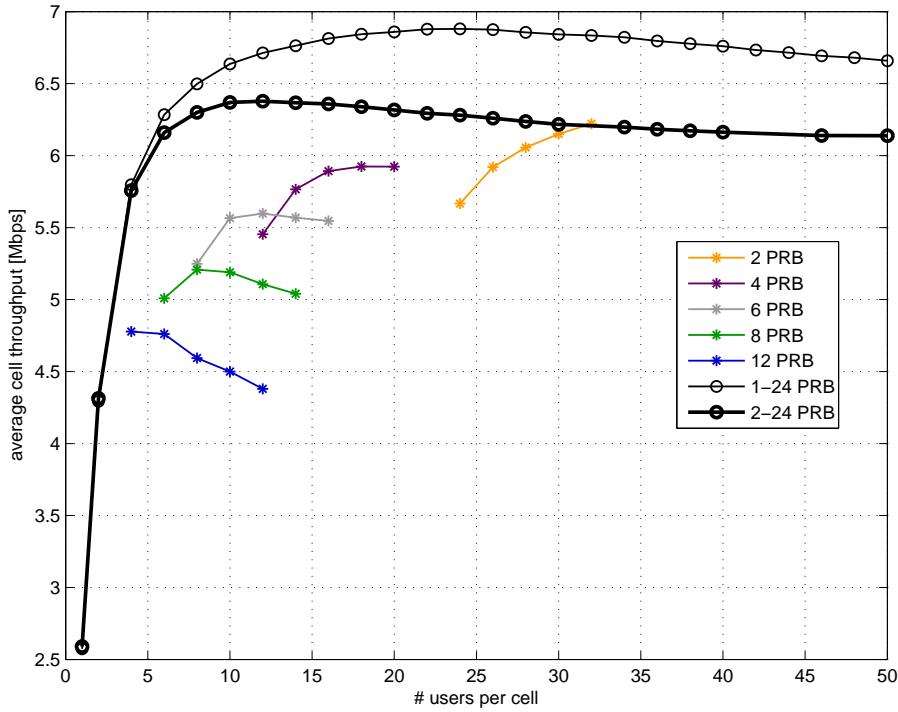


Figure 5.10: Average cell throughput under FTB and ATB for increasing UDO. Macro 3 case.

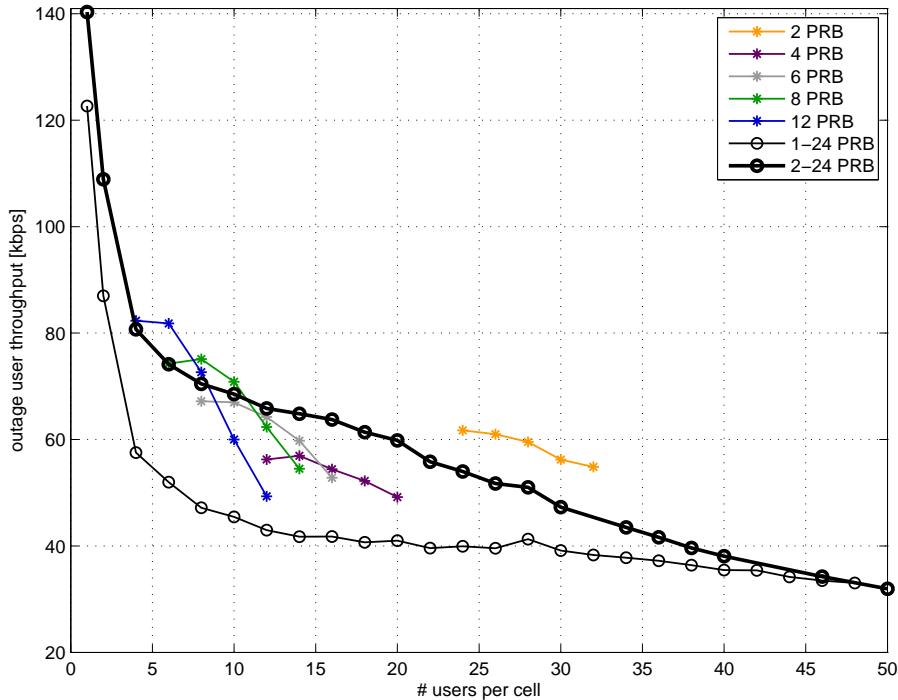


Figure 5.11: 5% and 10% outage user throughput under FTB and ATB for an increasing UDO. Macro 3 case.

Another situation where the ATB and the FTB show rather divergent behavior is an unbalanced load scenario. An unbalanced load scenario is different from the balanced load scenario in that the users are distributed uniformly over the network area rather than the sector. The number of users in the system is therefore the same as in the balanced load case but the number of users per sector changes from sector to sector. In this situation the ATB shows, especially for a low number of users, a considerably higher average cell throughput, shown in Figure 5.12(a), as a result of a higher bandwidth utilization, shown in Figure 5.12(b).

On the other hand the outage user throughput performance is characterized by an opposite trend. Figure 5.13(a) shows a considerably higher outage performance under the FTB than the ATB. This is the result of lower NR values under FTB compared to ATB, as shown in Figure 5.13(b), due to partial bandwidth utilization.

The situation changes in a Macro 3 case where the interference has a much lower impact (being a noise limited scenario). In this case the gain is visible, especially for low loads (intended as number of users), in both average cell throughput and outage user throughput as shown Figure 5.14(a) and (b).

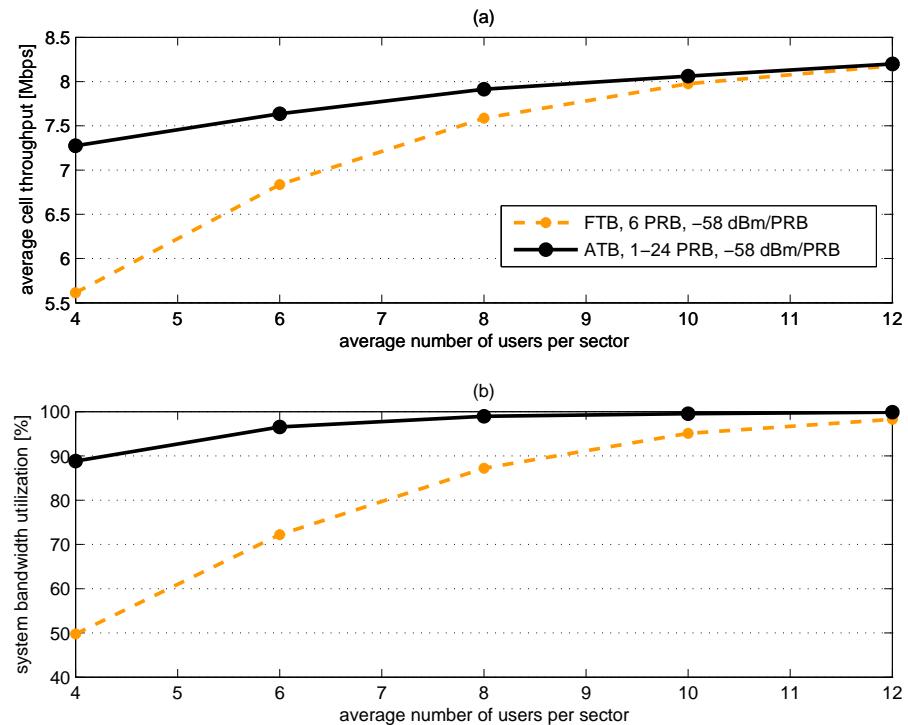


Figure 5.12: Average cell throughput and bandwidth utilization under FTB and ATB in an unbalanced load Macro 1 scenario.

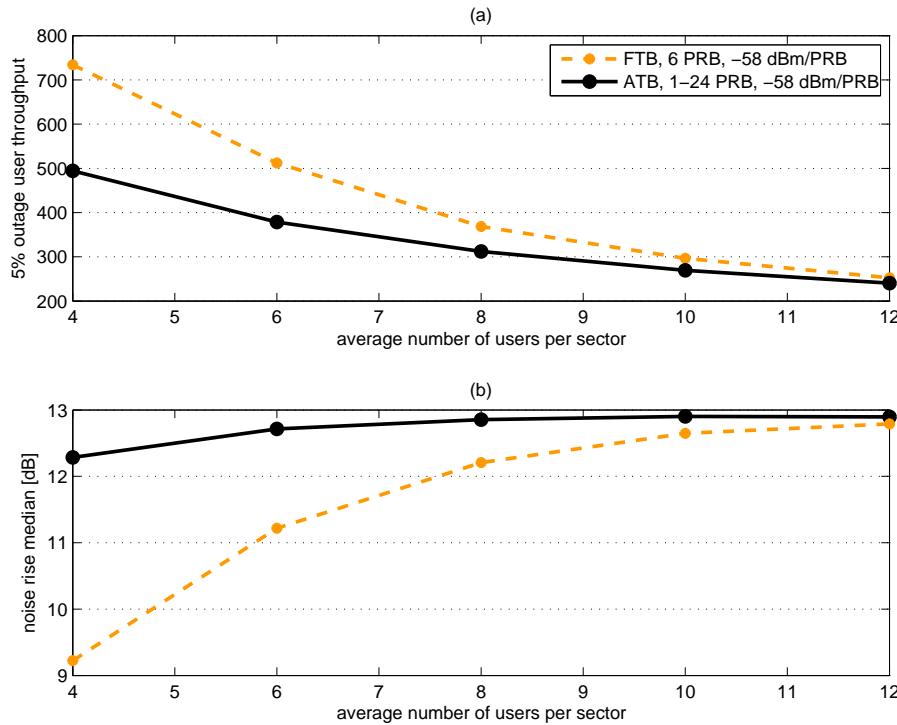


Figure 5.13: Outage user throughput and NR median under FTB and ATB in an unbalanced load Macro 1 scenario.

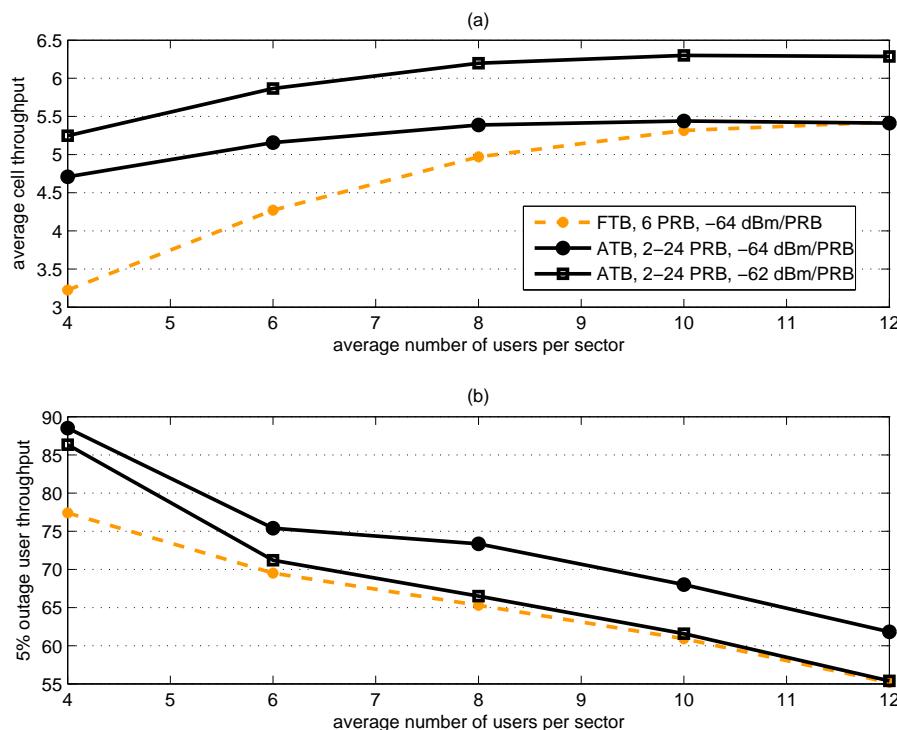


Figure 5.14: Average cell throughput and 5% outage user throughput under FTB and ATB in an unbalanced load Macro 3 scenario.

5.5 Conclusions

In this chapter a new scheduling algorithm is proposed which makes use of the ATB functionality in performing the allocation and is therefore referred to as ATB-based PS. The performance of the algorithm is evaluated in a variety of scenarios and compared to the FTB-based PS.

The main advantage provided by such scheduling algorithm is the flexibility to the varying cell load. Specifically, the scheduler automatically tunes the transmission bandwidth according to the cell load offering a performance similar to the FTB-based PS whose bandwidth has been chosen to provide the highest cell capacity and outage user throughput.

In the Macro 3 case, thanks to the capability of modifying the transmission bandwidth according to the power limitations of the users, the proposed algorithm does not only provide flexibility but also a performance gain.

In an unbalanced load scenario the ATB-based PS guarantees a high bandwidth utilization which results in a higher cell throughput than FTB but lower outage user throughput because of increased noise rise. In a Macro 3 scenario, where the interference is much more limited, the ATB is able to provide a gain in both average cell throughput and outage user throughput in a measure that depends on the power settings.

Chapter 6

Scheduling for Elastic Traffic with Minimum Throughput Guarantee

6.1 Introduction and Motivation

The performance of channel aware scheduling in the previous chapters has been assessed via extensive system level simulations assuming as metric the ratio of achievable instantaneous data rate over average received data rate [18]. The same metric was already investigated for a OFDMA-based system in [39] and applied to the DL of UTRA LTE in [40]. The performance of different TD and FD scheduling metrics for the DL of UTRA LTE has been further investigated in [44] where the authors introduced an SINR-based PF metric which, despite its good performance, is hardly applicable in the downlink of LTE due to the format of the feedback returned by the user. On the other hand, an SINR-based PF metric could potentially be derived for the uplink where frequency-selective CSI is directly available at the eNode-B based on SRS measurements and does not need to be quantized and fed back to the eNode-B. Therefore, the focus throughout the chapter is going to be on the different PF metrics used by the TD and FD schedulers, their design and interactions as well as their impact on system performance in terms of throughput, fairness and QoS provisioning. Even though the performance is evaluated in the context of uplink LTE the conclusions can be generalized to other SC-FDMA-based systems as well.

Unlike Chapters 4 and 5, where the scheduler was represented only by the FD entity, from now on the full scheduling framework, as described in Section 2.5, is going to be fully utilized by introducing also the TD entity. The main motivation for introducing the TD scheduler is represented by the necessity of meeting control channel limitations as well as the computational complexity of the FD scheduler by filtering the users according to their metrics. As shown afterwards, the TD metrics will be designed as to embody the user QoS requirements, therefore the TD scheduler will play an important role in QoS provisioning.

Additionally, a new set of assumptions, which aim at resembling a more realistic system, are going to be progressively introduced. They include the adoption of a finite buffer traffic model, the introduction of the GBR as the only QoS parameter to enable user differentiation, the deployment of the AC functionality paired with a user arrival rate modeled by a Poisson process.

The chapter is organized as follows: Section 6.2 presents a comparison of different PF-like metrics with emphasis on the FD and proposes a pair of metrics to be used as default for the TD and FD schedulers based on an analysis of the results obtained. Section 6.3 introduces the GBR as QoS requirement and proposes a modification of the TD metric to accommodate for it assuming the system has enough capacity to serve the users. The system instability which occurs in case of too high capacity requirement motivates the introduction of the AC functionality paired with a Poisson arrival process in Section 6.4. Furthermore, results are shown in Section 6.4.1 for the case

of two categories of users having each a different GBR requirement. Time and frequency domain metrics are modified in order to cope with the requirements of the new scenario. The conclusions are presented in Section 6.5.

6.2 Time and Frequency Domain PF-like Metrics

In this section we are going to analyze a set of different channel aware metrics which all share the principle of giving priority to users which are in their best relative conditions. In this sense they are said to be PF-like.

6.2.1 Metric Symbols

Table 6.1 lists the symbols which are going to be used throughout this chapter together with their meaning.

Table 6.1: Metrics related symbols.

Symbol	Meaning
$T(i, t)$	Acknowledged throughput for user i , at scheduling instant t
$\bar{T}(i, t)$	Past averaged throughput for user i , at scheduling instant t
$\bar{T}_{sch}(i, t)$	Past averaged throughput for user i , at scheduling instant t , updated only in the scheduling instants in which the user is allocated some resources
$\hat{T}(i, r, t)$	Estimated achievable throughput for user i , on PRB r , at scheduling instant t
$\hat{T}_w(i, t)$	Estimated wideband achievable throughput for user i , at scheduling instant t
$SINR_{CSI}(i, r, t)$	CSI (or SINR of the SRS) measured at the eNode-B for user i , on PRB r , at scheduling instant t
$SINR_{CSI,w}(i, t)$	Wideband CSI (or SINR of the SRS) measured at the eNode-B for user i , at scheduling instant t

6.2.2 TD Metrics Definition

In TD we are going to consider two metrics, PF and RR.

The PF metric in TD is defined as:

$$M_{PF_w}^{TD}(i, t) = \frac{\hat{T}_w(i, t)}{\bar{T}(i, t)} \quad (6.1)$$

In the definition of such metric we have to resolve to using the wideband expected throughput $\hat{T}_w(i, t)$ because we do not know, in the TD phase, which PRB the user is going to occupy. The PF metric in TD is the closest expression to the original definition of PF metric which is largely discussed, among others, in [18], [45], [46] and [47].

The RR metric is defined as:

$$M_{RR}^{TD}(i, t) = t - t_s(i) \quad (6.2)$$

where $t_s(i)$ indicates the TTI in which the user i has been scheduled last time.

6.2.3 FD Metrics Definition

The PF metric used in previous chapters is reported again here for completeness. It is defined as

$$M_{PF}^{FD}(i, r, t) = \frac{\hat{T}(i, r, t)}{\bar{T}(i, t)} \quad (6.3)$$

where \bar{T} is updated using an IIR filter defined as

$$\bar{T}(i, t) = (1 - \frac{1}{N_{TTI}})\bar{T}(i, t - 1) + \frac{1}{N_{TTI}}T(i, t)$$

where N_{TTI} defines the filter length or filter memory, which in this study is assumed to be the same for all the users. The first \bar{T} value is also the same for all users unless a GBR is specified, in which case \bar{T} is initialized to the GBR value. \bar{T} , and therefore the scheduling priority, is updated only in case the user has data in the buffer to avoid increasing the priority of users who have nothing to transmit. It's also important to highlight that $T(i, t)$ represents the acknowledged throughput (that is, the successfully delivered throughput) and not the scheduled throughput. This means that when a transmission fails the metric is updated with a value of 0. In other words the past averaged throughput \bar{T} in this way embeds also the effect of the experienced BLER. If the experienced BLER increases above the target the priority of the metric increases as well. The presence at the denominator of a value updated via an IIR filter raises some issues regarding the time needed for the metric to converge to its steady state. This depends on the choice of the filter length as well as the initial value of \bar{T} and will be taken into account when analyzing the results. $\hat{T}(i, r, t)$ is the estimated achievable Layer 1 throughput (estimated based on the AVI tables and the BLER target). So the relation between $\bar{T}(i, t)$ and $\hat{T}(i, r, t)$ is that the first depends on the experienced BLER while the second depends on the BLER target resulting in a different scheduling activity compared to other users when the experienced BLER deviates from the target BLER. Another relevant characteristic of this metric is its dependence on past scheduling decisions which is embedded in the throughput averaging at the denominator.

The second metric which we are going to analyze will be an SINR based metric and will be indicated as PF-SINR. Within the LTE context, it is first proposed in [44] and it is defined as:

$$M_{PF-SINR}^{FD}(i, r, t) = \frac{SINR_{CSI}(i, r, t)}{SINR_{CSI,w}(i, t)} \quad (6.4)$$

Unlike the PF metric, the PF-SINR is ideally always on steady state, that is, does not need to converge to it because the wideband SINR is assumed to be known on a TTI basis. In reality, also in this case we have an initialization and convergence problem as users typically do not send a wideband SRS, therefore the wideband SINR must be obtained by averaging (in time) the CSI obtained from several narrowband SRS measurements. Such level of detail is not included in our results, that is, only wideband sounding is assumed. Unlike the PF metric, the PF-SINR is independent of its past scheduling decisions.

The third metric considered is another throughput based metric (like the first) and it will be indicated as PF-Throughput To Wideband (TTW). It is first proposed in [48] and it is defined as:

$$M_{PF-TTW}^{FD}(i, r, t) = \frac{\hat{T}(i, r, t)}{\hat{T}_w(i, t)} \quad (6.5)$$

Like the PF-SINR, such metric is ideally always on steady state because the wideband SINR which is mapped into the wideband throughput is assumed to be known on a TTI basis. The PF-TTW, like the PF-SINR, is also independent of past scheduling decisions.

Table 6.3 summarizes the described metrics together with the main characteristics while Table 6.5 lists the parameter settings and assumptions. It is worth highlighting the deployment of TD scheduler as well as the use of finite buffer traffic model.

Table 6.3: FD metrics: definition and main characteristics.

Acronym	Definition	Main characteristics
PF	$\frac{\hat{T}(i, r, t)}{\bar{T}(i, t)}$	Dependent on past scheduling decisions (memory dependent). Convergence to steady state dependent on $\bar{T}(i, 0)$ (the initial throughput) and N_{TTI} (the memory constant)
PF-SINR	$\frac{SINR_{CSI}(i, r, t)}{SINR_{CSI,w}(i, t)}$	Independent of past scheduling decisions (memory free). Always on steady state.
PF-TTW	$\frac{\hat{T}(i, r, t)}{\hat{T}_w(i, t)}$	Independent of past scheduling decisions (memory free). Always on steady state.

Table 6.5: Main simulations parameters

Parameter	Setting
# UEs per sector in TD	30
# UEs per sector in FD	8, 10, 30
# PRBs per UE	1 to 24
TD scheduling	RR, PF _w
FD scheduling	PF, PF-SINR, PF-TTW
BLER target at 1st transmission	30%
PF filter length (N_{TTI})	100 ms
PF initial throughput value ($T(i, 0)$)	500 kb
α (for PC)	0.6
P_0 (for PC)	-58 dBm/PRB for Macro case 1 -64 dBm/PRB for Macro case 3
Traffic model	Finite buffer with balanced load

6.2.4 Performance Evaluation

The initial focus is given to the effects of FD scheduling. For this reason we assume the RR metric in TD and limit the analysis to the Macro 1 case. Moreover the attention is on how the selected metrics differentiate users characterized by different channel conditions or data rate requirements.

For this reason the behavior of several performance indicators of interest is going to be analyzed with respect to the APG (which includes distance-dependent path-gain, shadowing and antenna gain).

Figure 6.1 shows the CDF of the average user throughput for the proposed PF-like metrics in FD. PF and PF-SINR exhibit a very similar distribution unlike the PF-TTW which shows a larger percentage of users in the lower throughput range.

Figure 6.2(a), (b) and (c) show respectively the behavior in frequency (that is, the number of allocated PRBs when the user is scheduled), in time (that is, how often the user is scheduled) and the combined time-frequency behavior (that is, the average number of allocated PRBs per TTI), for the three different metrics introduced.

The PF-SINR metric shows the most fair distribution of resources among the users both in time and in frequency. The limitation in number of PRBs in the lower range of the APG is due to the user power limitations. The fairness in terms of resource allocation of the PF-SINR metric is due to the fact that the wideband SINR, at the denominator of the metric, can be assumed to be an average estimation of the channel quality of the user. Therefore the metric can be considered an estimate of the fast fading of each user on each PRB relative to the average channel quality. Being the fast fading independently and identically distributed among users, it results in an equal probability of the users being scheduled over time and frequency.

The PF metric shows a more complex behavior which involves several factors. Particularly, the fairness of allocation is affected mostly by the denominator through a series of factors including:

- the deviation from the BLER target, which results in a higher time scheduling activity in the lower APG range and a lower time scheduling activity in the upper APG range
- the lower number of PRBs allocated in the lower APG range, which also results in a higher time scheduling activity in the lower APG range
- the reduced metric diversity in the APG ranges where the deviations from the BLER target turns the PF metric into a channel-blind one.

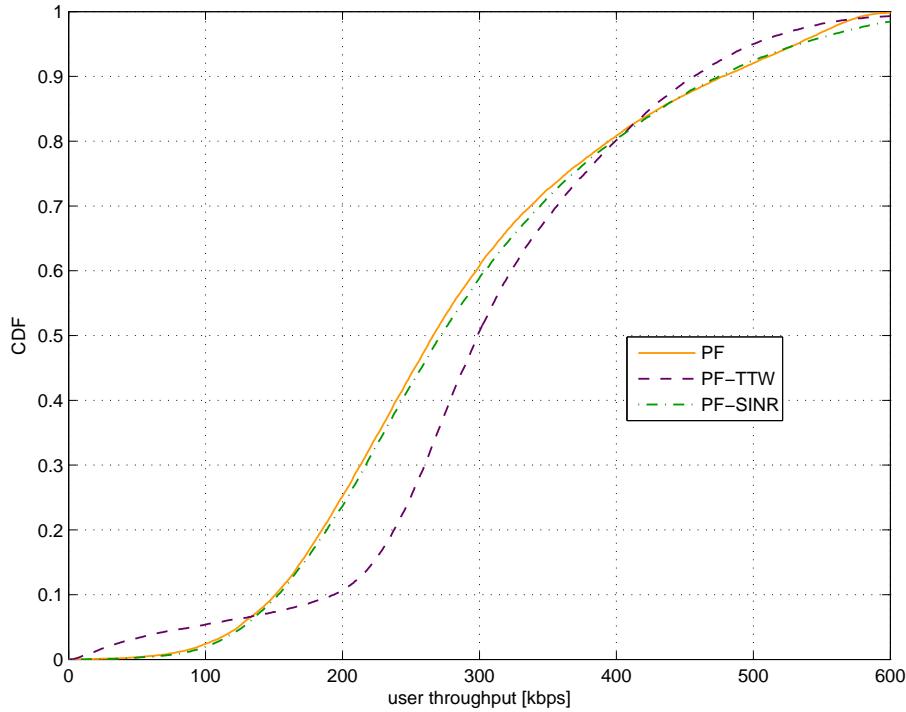


Figure 6.1: CDF of average user throughput. Macro 1 case. 30 users (per sector) in TD and all of them handed over to FD.

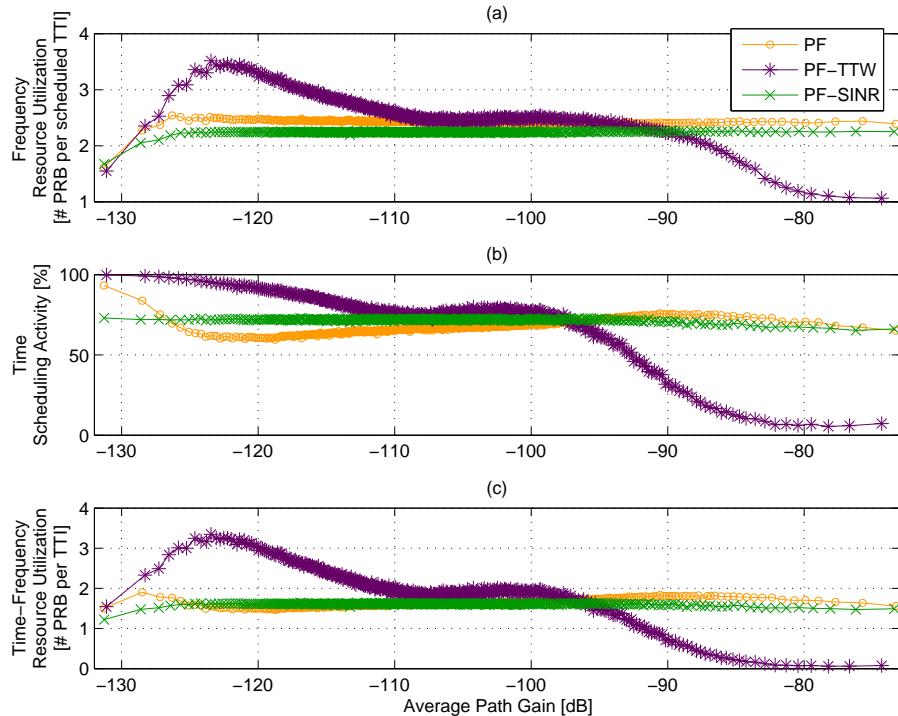


Figure 6.2: (a): Average number of allocated PRBs (when the user is scheduled by the FD). (b): Percentage of time the user is scheduled. (c): Average number of allocated PRBs per TTI. Every star represents the average over 50 consecutive users (with respect to APG). Macro 1 case. 30 users (per sector) in TD and all of them handed over to FD.

The metrics differ also in the number of users scheduled per TTI by the FD scheduler as shown in Figure 6.3. The PF and PF-TTW metrics assign in average higher priorities to some of the users than to others (as shown in the previous graphs) while the PF-SINR metric tends to assign in average the same priority to all the users (over a sufficiently long period of time). A higher average priority translates into a larger bandwidth when the user is allocated and therefore into a lower number of scheduled users compared to the PF-SINR metric.

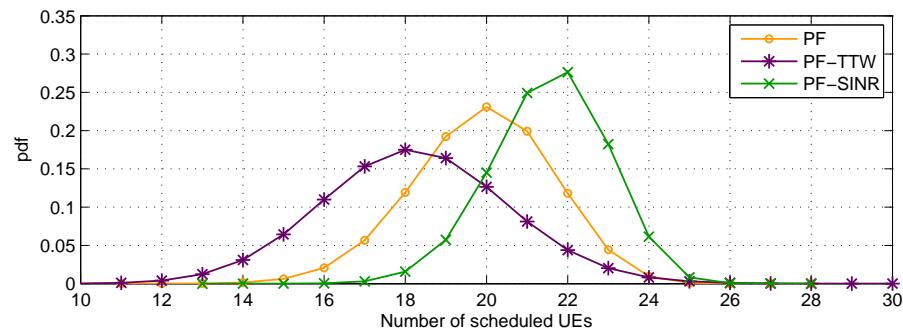


Figure 6.3: Distribution of the number of scheduled users. Macro 1 case. 30 users (per sector) in TD and all of them handed over to FD.

The PF-TTW metric is by far the one that shows the greatest variability in fairness of allocation with users in poor conditions being scheduled more often and with more PRBs than users in good conditions. The reason is that the wideband throughput (the denominator of the metric) is lower bounded by the lowest MCS (MCS_{low}) with the result that $M_{TTW} \geq C$ for users in poor conditions, where $C = \frac{TBS(MCS_{low}, 1 PRB)}{TBS(MCS_{low}, 48 PRB)}$. Similarly the wideband throughput is upper bounded by the highest MCS (MCS_{high}) so that $M_{TTW} \leq C$ for users in good conditions, where $C = \frac{TBS(MCS_{high}, 1 PRB)}{TBS(MCS_{high}, 48 PRB)}$.

The average cell throughput and outage user throughput for the different metrics are shown in Figure 6.4(a) and (b). Even though the PF-TTW metric allocates a large amount of resources to the cell edge users, the outage performance is lower than the one obtained under PF and PF-SINR. The reason is to be found in Figure 6.4(c) where we can see that the users in outage are actually the ones with the best channel conditions as they are allocated a very low amount of time-frequency resources. The PF and PF-SINR achieve similar performance even though their behavior is different in many respects.

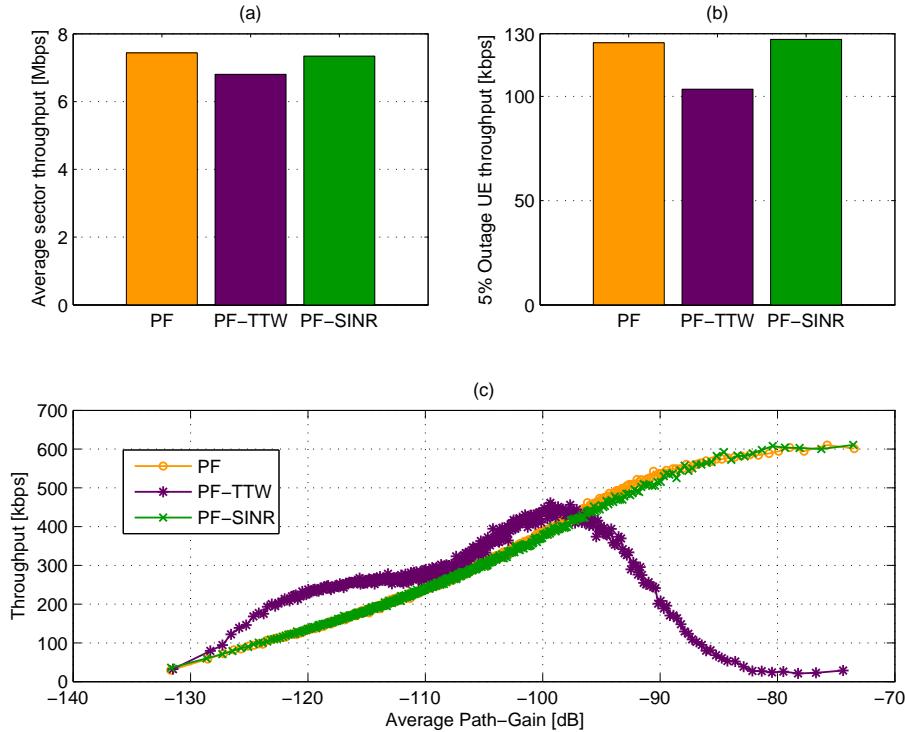


Figure 6.4: (a): Average sector throughput. (b) Outage user throughput. (c) Average user throughput vs APG. Every star represents the average over 50 consecutive users (with respect to APG). Macro 1 case. 30 users (per sector) in TD and all of them handed over to FD.

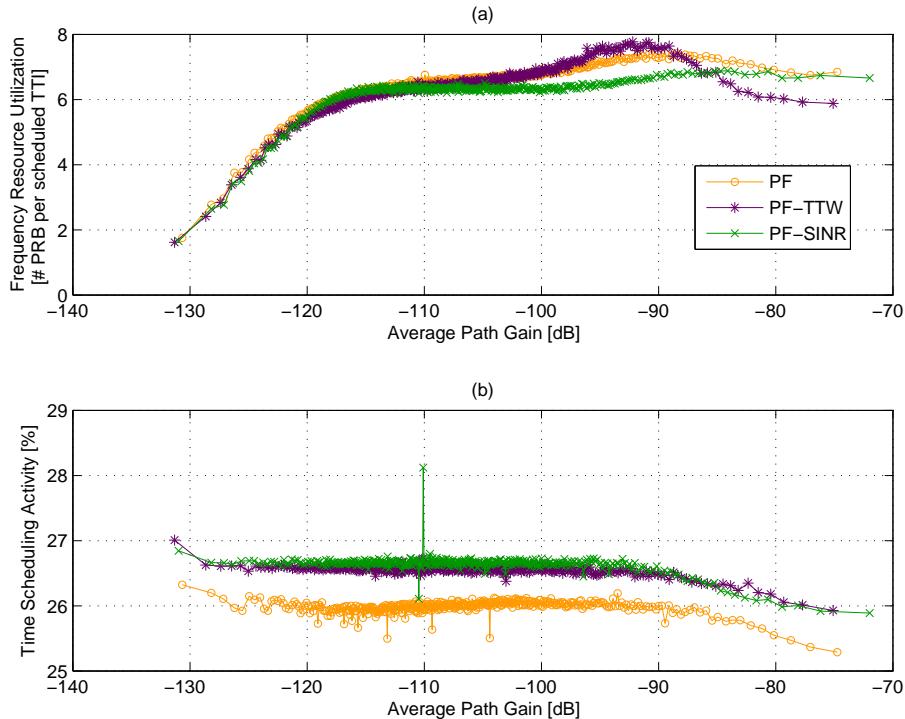


Figure 6.5: Frequency and time resources utilization vs APG. Every dot represents the average of 50 consecutive users. Macro 1 case. 30 users (per sector) in TD and 8 of them handed over to FD scheduler.

The highlighted differences among the three metrics tend to be evened out if we introduce the TD scheduler driven by a RR metric and limit the number of users handed over to the FD scheduler (like it happens in a real case scenario). In this way the RR metric guarantees a certain fairness among the groups of users passed to the FD domain while in FD the single carrier constraint, combined with a lower number of users, evens out the difference among the different users over a larger allocated bandwidth resulting in a fair distribution of resources also in FD in terms of allocated PRBs. This leads to a similar behavior and performance of the different metrics as shown in Figures 6.5. Average sector throughput and outage user throughput figures are given in Table 6.6: even though the metrics show a very similar behavior, PF-TTW is still under-performing compared to PF and PF-SINR.

Table 6.6: Average sector throughput and outage user throughput. Macro 1 case. 30 users (per sector) in TD. 8 users handed over to FD scheduler.

Metric	Average sector throughput	5% Outage user throughput
PF	6.48 Mbps	100 kbps
PF-TTW	6.31 Mbps	99 kbps
PF-SINR	6.42 Mbps	95 kbps

The results so far presented are ideal given the considered assumptions. In particular the limitations introduced by the PDCCH are to be taken into account if the aim is to resemble more closely a real system. There are different ways in which the PDCCH can have an impact on scheduling. PDCCH is error-prone and time and frequency resource consuming. Additionally it is subject to resource limitations and terminal decoding constraints. For these reasons some users may be blocked from being scheduled in a given subframe, an event referred to as PDCCH blocking. As briefly hinted, the limitations introduced by PDCCH are several but they are not explicitly modeled in this project. That is, the only way they are taken into account is by limiting the number of simultaneously scheduled users. According to [49], a realistic assumption for the number of users to be scheduled per TTI is between 8 and 10. It is therefore worth to repeat the analysis of the system performance under such constraint and to find a combination of TD and FD metrics which, in such case, can improve the performance of the system.

From the previous analysis it emerged that the PF metric schedules more or less frequently a user compared to others when the experienced BLER (which affects the throughput at the denominator of the metric) deviates from the target BLER (which is used in the estimation of the throughput at the numerator of the metric). This happens in the lower and upper regions of the SINR due, for example, to power limitations of the mobile terminal and the lack of lower order MCSs or higher order MCSs compared to the available ones. More specifically a user may experience an SINR as low as to estimate the same expected throughput (given from the lowest MCS) independently of the PRB considered. This is equivalent to turning the channel aware metric into a blind one. Same for a user experiencing a very high SINR. Therefore in the low and high SINR regions a change in scheduling activity comes together with a loss of channel awareness.

The PF-SINR metric, instead, is an estimation of the fast fading relative to the average channel quality and it therefore remains channel aware over the full SINR range, even where the PF metric turns into a blind metric. Therefore we can expect a better throughput performance in such regions as expressed in the qualitative Figure 6.6.

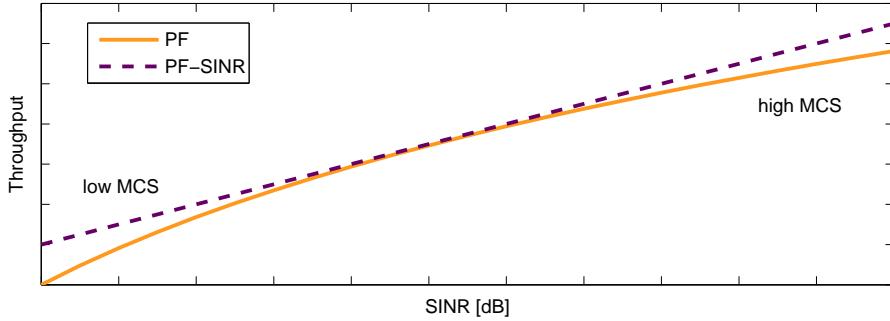


Figure 6.6: Qualitative trend of the SINR to throughput mapping for the PF and PF-SINR metrics.

As shown, the two metrics provide a similar performance even though they are governed by different mechanisms of exploiting the channel. It is therefore reasonable to expect a further gain by deploying the PF metric in TD to introduce the dependency from the throughput (and therefore increase the scheduling activity of users in disadvantaged conditions) and the PF-SINR metric in FD to exploit the channel variations over the full SINR range.

The results of such strategy are represented in Figure 6.7. In the Macro 1 case, shown in Figure 6.7(a) and (b), the gain is rather limited because only a small percentage of users deviates from the BLER target and therefore only few users can benefit from the combination of such metrics. In the Macro 3 case, on the other hand, the gain is significant as shown in Figure 6.7(c), which shows a 14% gain in average cell throughput, and 6.7(d), which shows a 30% gain in outage user throughput. In this case a larger number of users experiences a BLER larger than the target due to more severe power limitations and therefore benefit from the adoption of the proposed strategy.

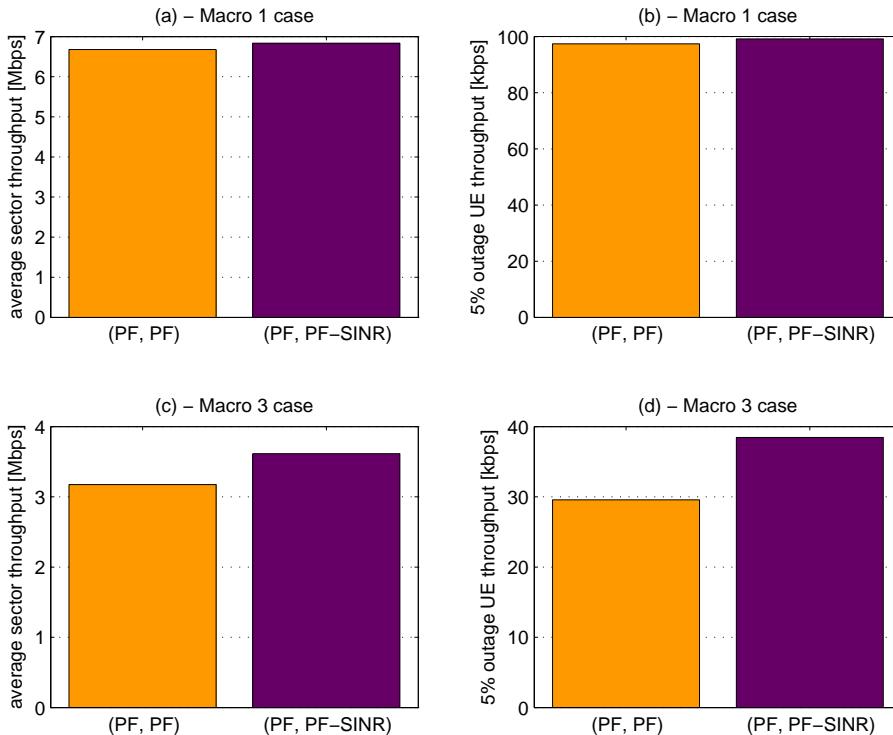


Figure 6.7: (a): Average sector throughput for Macro 1 case. (b) Outage user throughput for Macro 1 case. (c): Average sector throughput for Macro 3 case. (d) Outage user throughput for Macro 3 case. 30 users (per sector) in TD and 10 of them handed over to FD. PF metric in TD and PF vs PF-SINR in FD.

The results shown highlight the benefits of a solution where the TD scheduling select the users based on their requirements and the FD scheduling performs channel exploitation.

6.3 Elastic Traffic with Identical Minimum Throughput Guarantee under Balanced Load

In this section we are going to introduce a more general category of traffic which includes the BE traffic as special case. Such traffic is referred to as Elastic with Minimum Guarantee (EMG). The users characterized by such traffic require minimum average throughput guarantee and more if possible. The minimum average throughput guarantee is expressed by the GBR requirement which is the only QoS parameter considered throughout this project. The BE traffic so far considered represents a special case of the EMG traffic where the GBR requirement is set to 0.

Some of the more realistic assumptions introduced in the previous section are also used in the rest of the chapter. They include the use of a finite buffer model as well as limiting the number of users handed over to the FD scheduler. Moreover, as highlighted by the previous results, it is worth to perform channel exploitation in FD especially in presence of users prone to experience a BLER higher than the target. For this reason the PF-SINR metric is going to be a default assumption for the FD scheduler.

Two schemes are going to be considered to model the users arrival rate: First a scenario characterized by balanced load (without AC) is considered (a user is killed and replaced by a new user once its buffer is empty). Afterwards, the AC functionality (paired with a Poisson arrival scheme) is introduced as a mean to prevent system instability.

The AC algorithm used was developed in a parallel PhD project [24]. The main idea of the algorithm is to estimate the required number of PRBs per TTI of the new users (N_{new}) and sum them to the required number of PRBs per TTI of the existing users (N_i). If such sum is lower or equal to the total number of PRBs in the system bandwidth (N_{tot}) as indicated in 6.6, then the user is admitted.

$$\sum_{i=1}^K N_i + N_{new} \leq N_{tot} \quad (6.6)$$

The estimation of N_{new} takes into account the pathloss and the GBR of the incoming user. More details can be found in [50, 51, 24].

The adoption of new traffic models and functionalities require taking into account new statistics. Specifically, the introduction of the GBR requirement calls for a definition of the outage (or degree of unsatisfaction of users in the network) as the percentage of users which are not able to achieve the GBR. For the same average cell throughput, the lower such percentage, the better the QoS provisioning. Similarly, the introduction of the AC requires taking into account the blocking probability in the analysis of results. Such quantity will then be combined with the outage probability to derive a measure of the unsatisfaction (or satisfaction) of the users in the network.

6.3.1 QoS-aware Metrics

In the first part of this section we are going to present results in a context where the aggregate minimum capacity requirement, simply defined as $N_{UE} \cdot GBR$, is well below the average cell throughput. In the second part we are going to present the results for a scenario where the minimum capacity requirement is higher than the system capacity.

The proposed metrics applied in TD are defined by equations (6.7), (6.8), (6.9), (6.10). For comparison the RR metric and the wideband PF metric in TD will also be included in the analysis.

$$M_{GBR}^{TD}(i, t) = \frac{GBR(i)}{\bar{T}(i, t)} \quad (6.7)$$

$$M_{GBRmax}^{TD}(i, t) = \max(1.0, \frac{GBR(i)}{\bar{T}(i, t)}) \quad (6.8)$$

$$M_{PF-GBR}^{TD}(i, t) = \frac{\hat{T}_w(i)}{\bar{T}(i, t)} \cdot \frac{GBR(i)}{\bar{T}(i, t)} \quad (6.9)$$

$$M_{PF-GBRmax}^{TD}(i, t) = \frac{\hat{T}_w(i)}{\bar{T}(i, t)} \cdot \max(1.0, \frac{GBR(i)}{\bar{T}(i, t)}) \quad (6.10)$$

Equation (6.7) prioritizes the user selection based on the ratio between the GBR and the average throughput. Equation (6.8) modifies the priority only for the users whose throughput is below the GBR. Equations (6.9) and (6.10) modify the previous two equations adding the wideband PF metric so that users are prioritized according to QoS and channel aware aspects.

6.3.1.1 Aggregate minimum capacity requirement lower than average sector throughput

The results presented in this section are obtained under the parameters listed in Table 6.8.

Table 6.8: Main simulations parameters under balanced load and EMG traffic (one GBR).

Parameter	Setting
# UEs (per sector) in TD	16 (low aggregate capacity requirement); 32 (high aggregate capacity requirement)
# UEs handed to FD	10
TD scheduling	RR, PF, GBRmax, GBR, PF-GBRmax, PF-GBR
FD scheduling	PF-SINR
Propagation scenario	3GPP Macro case 1 (ISD of 500 m)
Traffic model	Finite buffer with balanced load and identical GBR requirement for all users
GBR requirement	256 kbps
Packet size call	1 Mbit

Figure 6.8 shows how the proposed QoS-aware metrics differ in terms of time scheduling activity. The corresponding behavior in frequency, that is the frequency resources utilization in terms of number of allocated PRB, is not shown as it is the same for all the considered cases. As reference two non QoS-aware metrics, specifically RR and PF, are added.

As expected, the RR metric, which is blind to both channel and QoS requirements, shows the same activity for all the users independently on their average path gain. The PF metric, which is channel aware but not QoS aware, and the GBR metric, which is QoS aware but not channel aware, draw the boundaries (in terms of scheduling activity) within which all the others metrics are contained as they are a weighted product of the two. As the weight of QoS aware component becomes more important the curve shows that a larger and larger time scheduling fraction is given to the disadvantaged users while, as consequence, a smaller and smaller time scheduling fraction is given to the best users. The order of metrics in terms of greater importance of the QoS term is PF, PF-GBR_{max}, PF-GBR, GBR_{max}, GBR.

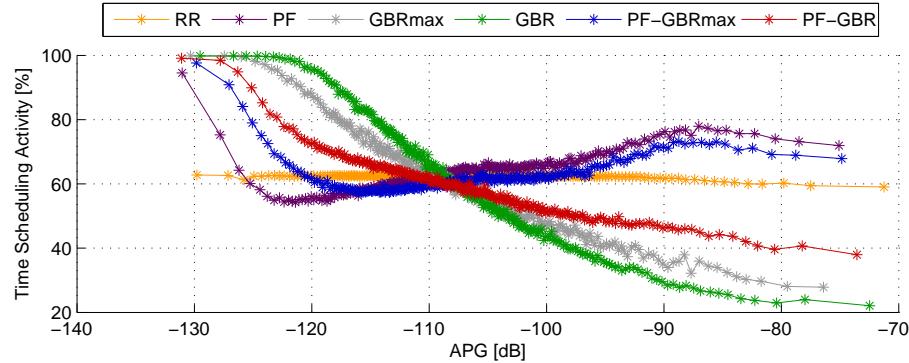


Figure 6.8: Average cell throughput and percentage of unsatisfied users using QoS aware metrics in TD and PF or PF-SINR in FD.

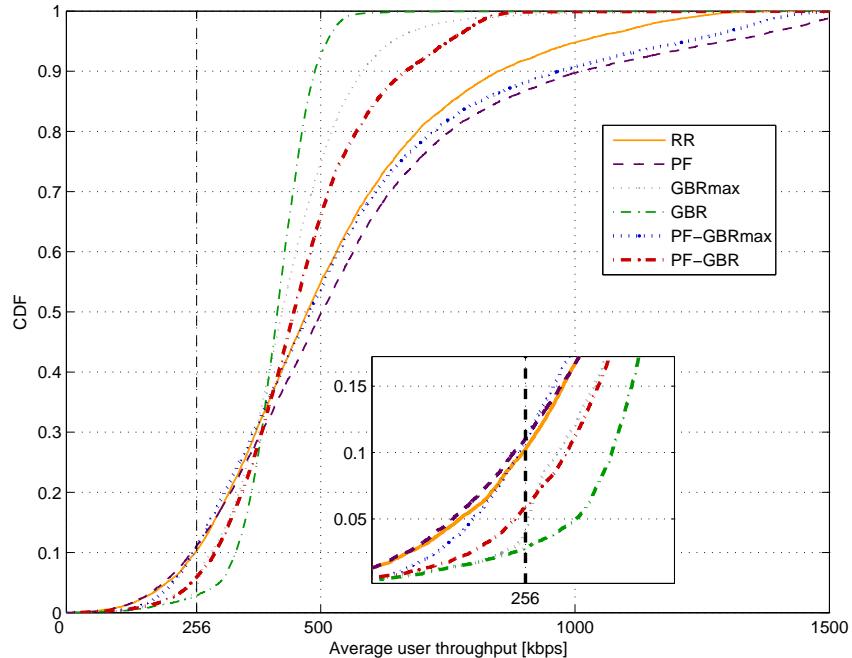


Figure 6.9: Average user throughput for different channel-aware and/or QoS-aware metrics in TD.

The effect of such metrics on the time activity is also reflected in the throughput distribution shown in Figure 6.9.

The application of a QoS aware metric in TD has the effect of reducing the average cell throughput but also the percentage of users below the GBR requirement. This is shown in Figure 6.10 where the GBR metric (purely QoS aware) reduces the average cell throughput by 4-5% and the percentage of users in outage by 7-8% compared to a RR metric. As the target, in this case, is more to provide the minimum throughput guarantee rather than increasing the average cell throughput, the GBR metric is probably to be preferred. Also the GBR_{max} metric shows an interesting behavior. The reason for introducing such a metric is to increase the priority only for the users below their GBR requirement while leaving the priority of users above the requirement untouched. The behavior is clearly noticeable in the CDF of the throughput shown in Figure 6.9 where the user throughput shows an increase right after the GBR target. Such metric is able to

provide a small improvement in cell capacity but, on the other hand, the outage is also increased.

The PF-GBR metric is also interesting as it can preserve the same average cell throughput of the RR while decreasing the number of users in outage by ca. 5%. In comparison with the GBR metric, the introduction of the channel aware aspects increase, as expected, the average cell throughput but at the same time worsen the outage performance.

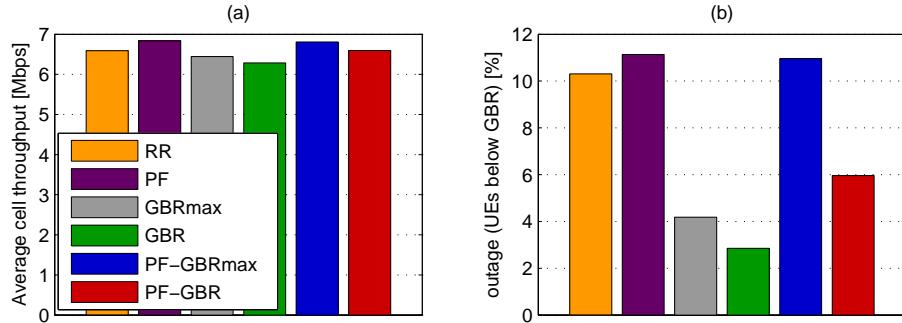


Figure 6.10: Average sector throughput and outage user throughput (that is, percentage of users below GBR requirement) using different metrics in TD and PF-SINR in FD.

6.3.1.2 Aggregate minimum capacity requirement higher than average cell throughput

In this section we are going to consider a number of users double than the one considered so far while keeping the same GBR requirement for all of them. This accounts for a minimum capacity requirement much larger than what the system can handle.

Figure 6.11 shows the CDF of the user throughput. GBR and GBR_{max} behave equally because no user is able to achieve the required GBR. On the other hand the same metrics are the most fair in terms of throughput distribution. The PF-GBR and PF-GBR_{max} show a similar trend up to the required GBR after which the PF-GBR_{max} shows a better user throughput. This has a small impact on the performance of the users below the requirement because the users above the requirement are a relatively small percentage. The introduction of a channel dependent component makes it possible to satisfy a part of the users (the ones in the best channel conditions) but it leads to a worse performance of the already unsatisfied users therefore to a larger unfairness in the throughput distribution. Finally the PF, which does not take into account the GBR requirement, shows the lowest percentage of users in outage but also the highest unfairness.

The GBR metric is the most fair in terms of throughput distribution among the different users and is also the one delivering the largest user satisfaction in case the system can accommodate the minimum data rate requested from the users. The same metric is still the most fair in terms of throughput distribution among the different users but also the one giving the lowest user satisfaction when the user requirements become excessive for what the system can handle.

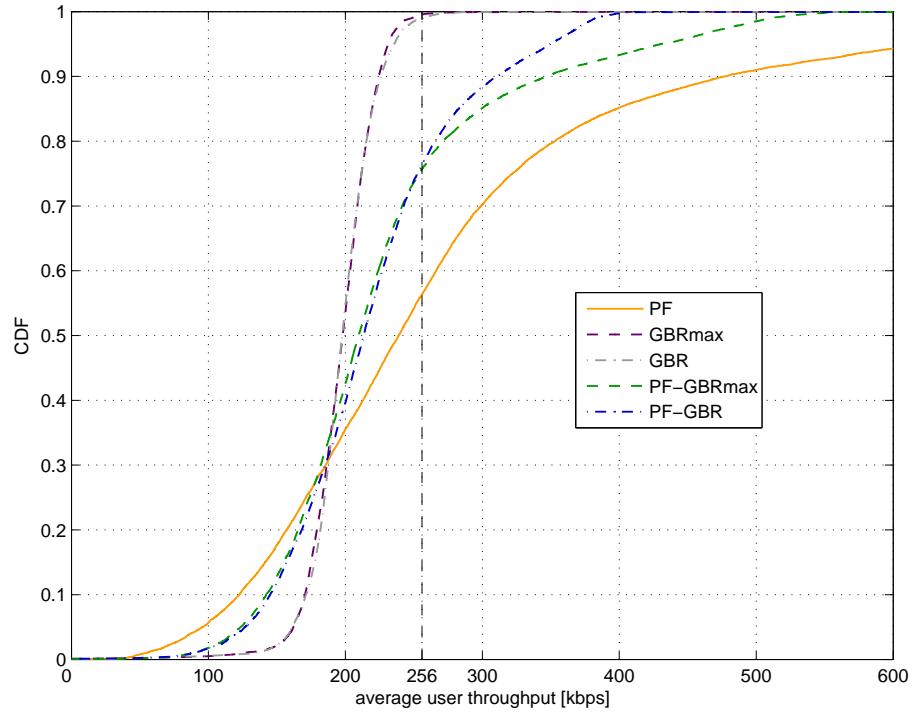


Figure 6.11: Average user throughput for different channel-aware and QoS-aware metrics in TD.

The average cell throughput and the percentage of users in outage (i.e., below GBR requirement) are shown in Figure 6.12.

In this scenario the impact that the different metrics have on the user throughput distribution is the same as in the case of 16 users per sector. On the other hand the average throughput per user is much lower than the GBR and, as result, the outage probability is considerably increased in all cases. Situations like this one, where the system is not able to provide the users with the required GBR call for the introduction of the AC which is introduced in next section.

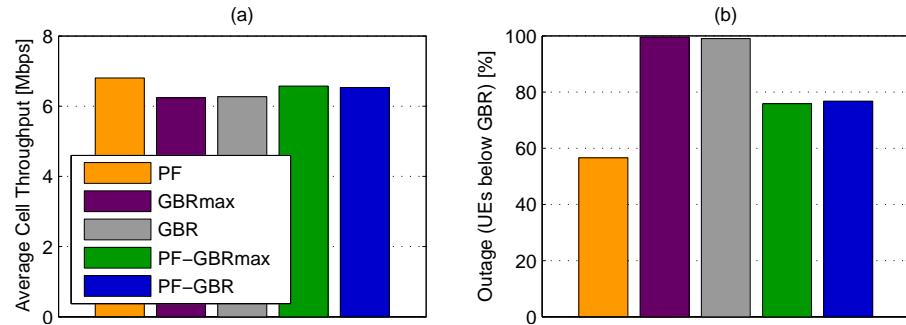


Figure 6.12: Average cell throughput and percentage of unsatisfied users for different QoS aware metrics in TD.

6.4 Poisson Arrival and Admission Control

The AC is a functionality which grants or denies the access of an incoming user based on the load of the local cell or of multiple cells. In this work the AC strategy is assumed to be single cell, that is, only the local cell load information is used. The aim of a QoS-based AC is to determine whether

a new user can be granted or denied access to the cell based on whether the QoS requirement of the incoming user can be fulfilled while guaranteeing the QoS of the existing users [25]. The introduction of the AC functionality requires the definition of another statistic, namely the blocking probability $P_{blocking}$, that is the ratio of the number of blocked users to the number of newly arriving users requesting admission. Further, a measure of unsatisfaction in the network, which includes the users blocked or in outage, is defined as:

$$P_{unsatisfaction} = 1 - (1 - P_{blocking})(1 - P_{outage}) \quad (6.11)$$

Table 6.9 summarizes the main simulation assumptions used for the rest of the chapter.

Table 6.9: Main simulations parameters under Poisson arrival, AC and EMG traffic (one and two GBRs).

Parameter	Setting
User arrival rate (Poisson distributed)	6, 8, 10 UE/s/cell
Max number of users handed to FD	10
TD scheduling	PF, PF-GBR, GBR
FD scheduling	PF-SINR, PF-SINR-GBR, PF-SINR-GBRsch
Propagation scenario	3GPP Macro case 1 (ISD of 500 m)
Traffic model	Finite buffer
GBR requirement	256 kbps or 128 kbps and 1000 kbps
Packet size call	1 Mbit

Figures 6.13 (a) and (b) show respectively the blocking probability and the outage probability for different arrival rates and different TD metrics.

In the case of a GBR of 256 kbps the metrics don't show a significant difference in performance for an arrival rate of 6 and 8 UE/s/cell. Larger differences can be seen at a user arrival rate of 10 UE/s/cell. In this case the GBR metric shows the lowest outage probability and, in turn, the largest blocking probability. At the other hand we have the PF metric which shows a larger outage probability and a lower blocking probability. In any case the outage probability is contained below 5% while the blocking probability ranges between 27% and 30%. Lower outage resulting from the adoption of a QoS metric leads to higher blocking as the users tend to stay longer in the system resulting in a higher blocking rate for incoming users.

In the case of a GBR of 512 kbps the blocking and outage probabilities are the same independently of the TD metric deployed.

The explanation of this findings can be found in the first place in the number of active users in the network shown in Figure 6.14 (a) and (b) for the GBR requirements of 256 kbps and 512 kbps.

When the number of active calls in the network is below the average of 10 users per cell the TD metric doesn't make any difference on how the users are prioritized as all of them are handed over to the FD scheduler. Because the number of users varies from cell to cell (given the dynamic arrival), some small differences in performance may occur in the cells where the number of users is temporarily above 10.

It is also worth noting in Figure 6.14 (a) that with the PF metric the number of users in the network is lower than with the PF-GBR or the GBR metrics. This happens because the PF metric doesn't have any QoS dependence and is therefore able to serve the user faster resulting in a lower blocking probability and a larger cell throughput.

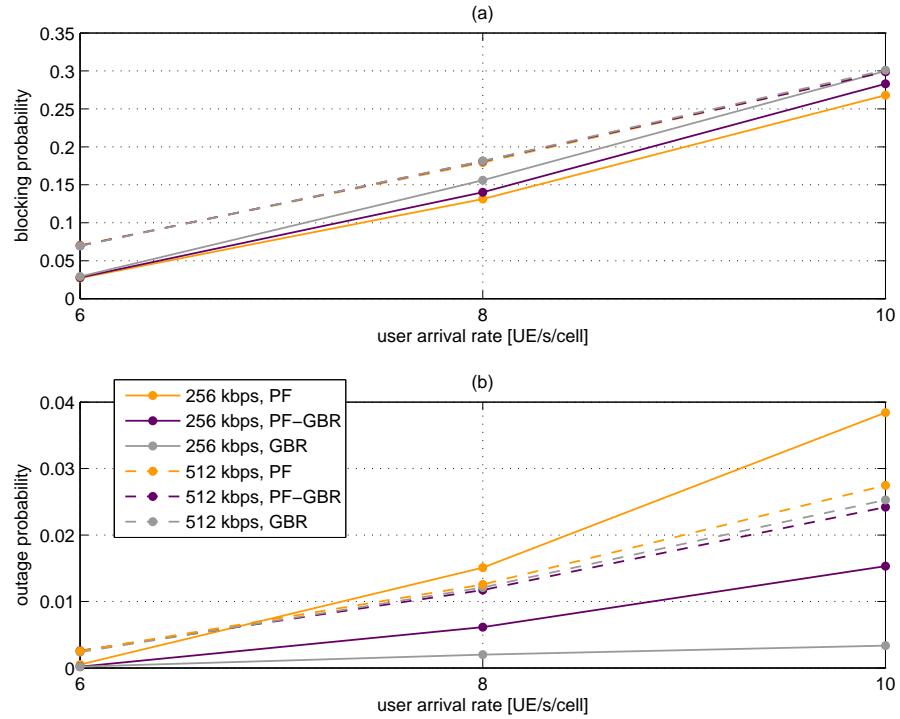


Figure 6.13: Blocking and outage probabilities vs arrival rate for different TD metrics and different GBR requirements (256 kbps and 512 kbps). Macro 1 case.

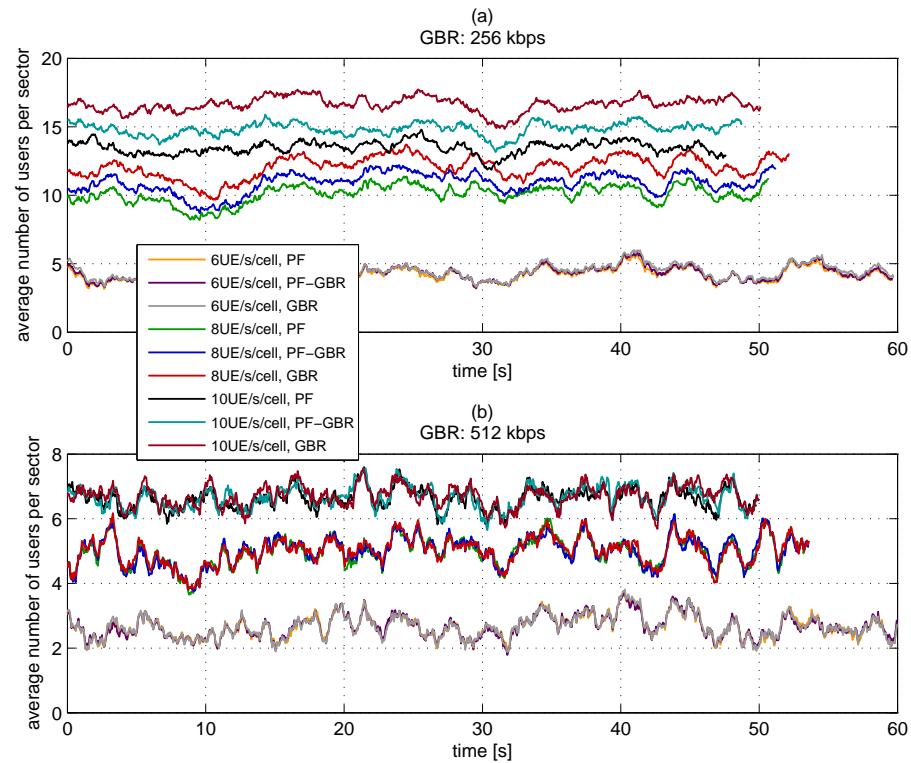


Figure 6.14: Average number of users per sector vs time for different TD metrics and user arrival rates. Figure (a) shows the trend for a GBR of 256 kbps and (b) for a GBR of 512 kbps. Macro 1 case.

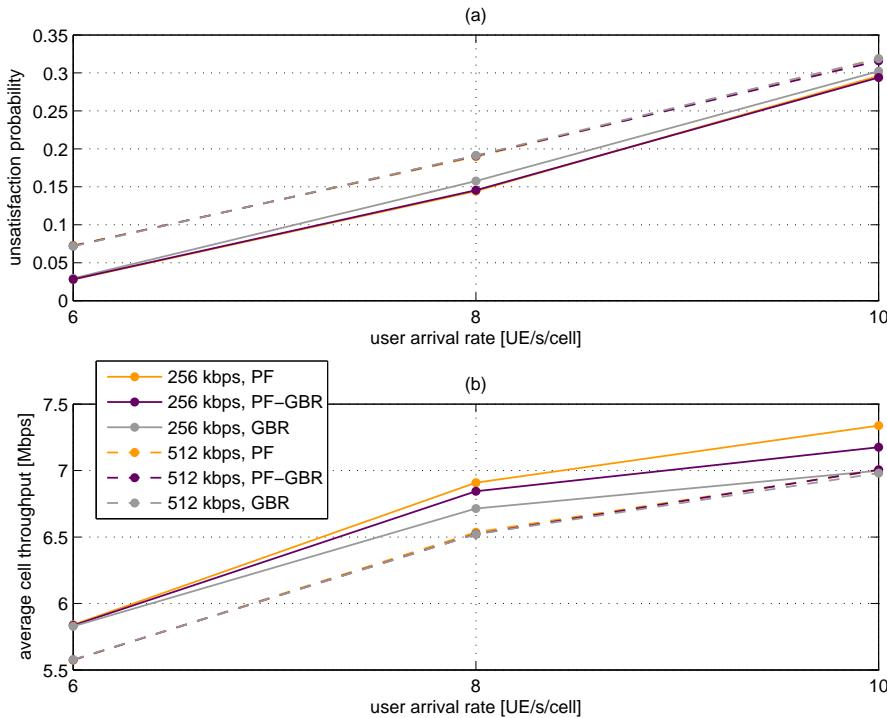


Figure 6.15: Unsatisfaction probability and average cell throughput vs arrival rate for different TD metrics and different GBR requirements (256 kbps and 512 kbps). Macro 1 case.

The unsatisfaction probability (which takes into account both the blocked users and the users in outage) and the average cell throughput are shown respectively in Figures 6.15 (a) and (b). In the case of 512 kbps GBR the different TD metrics don't exhibit any difference for the aforementioned reasons. In the case of a 256 kbps GBR the percentage of unsatisfied users is very similar while the average cell throughput shows a gain of up to 5% for the PF metric compared to the GBR metric.

Among the considered metrics the PF metric is the one that provides the largest average cell throughput and the lowest overall unsatisfaction in the network. This is mostly a consequence of the AC which filters the users not able to achieve their GBR, with the result that the admitted users are most probably not going to be in outage even though the PF metric is not QoS aware. For this reason the PF metric can more than make up for the loss in outage by increasing the average cell throughput and consequently reducing the blocking probability.

An important consideration pertains the location of the blocked users. Figure 6.16 shows that the AC algorithm is not only selective with respect to the GBR requirements but also with respect to the path gain of such users. Particularly, the users which experience the largest blocking probability are the ones with the lowest path gain, that is, the AC algorithm rejects the users at the cell edge thus leading to a cell shrinking effect which may not be desirable and should be taken into account in the network planning phase. Nonetheless, this effect should be taken into account when analyzing the results.

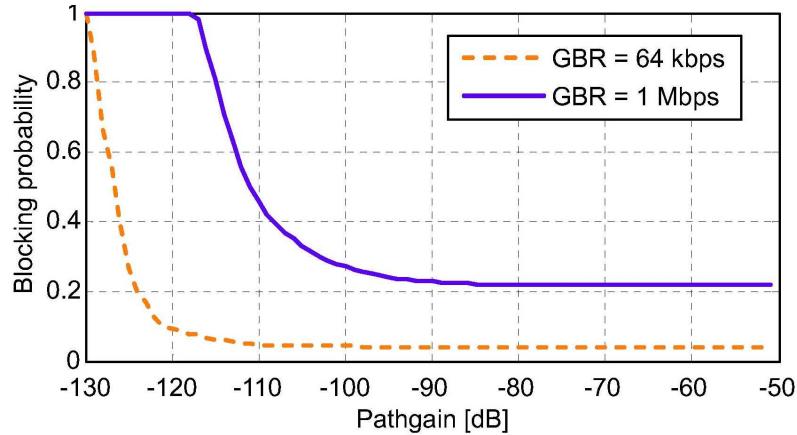


Figure 6.16: Blocking and outage probabilities vs pathgain for different GBR requirements (64 kbps and 1000 kbps). Macro 1 case.

As the case of single GBR combined with QoS-aware AC does not require QoS-aware scheduling in TD (PF gives, indeed, the best performance), it is worth investigating the case of users belonging to different GBR categories.

6.4.1 Elastic Traffic with Different Minimum Throughput Guarantees

In this section we are going to investigate the performance of different TD and FD metrics in a scenario where the incoming users are assigned one of two possible GBRs with equal probability. In order, for this scenario, to be significantly different from the previous one (where all the users had the same GBR requirement) we assume the two GBR requirements to be far apart from each other, specifically 128 kbps and 1000 kbps.

As in the previous case, the analysis is limited to the Macro 1 scenario, finite buffer with packet call size of 1 Mbit, dynamic user arrival and AC (see Table 6.9).

In Figure 6.17 the CDF of the average user throughput is represented for the two different GBR of 128 kbps and 1000 kbps. None of the three different TD metrics previously introduced, that is, PF, PF-GBR and GBR, is able to differentiate the two categories of users. Independently on the TD metric deployed, the outage probability of users with a GBR of 1000 kbps remains around 0.5.

This happens because whenever the users are passed to the FD scheduler, no differentiation is performed between the low GBR and the high GBR users. The consequence is that the low GBR users receive much more resources than the requirement and leave the system very fast while a large part (50%) of the high GBR users receive less resources than the requirement and stay in the system for longer.

This effect suggests that a weight dependent on the QoS requirement be applied also in FD. For this reason the following FD metrics, are introduced:

$$M_{PF-SINR-GBR}^{FD}(i, r, t) = \frac{SINR_{CSI}(i, r, t)}{SINR_{CSI,w}(i, t)} \cdot \frac{GBR(i)}{\bar{T}(i, t)} \quad (6.12)$$

$$M_{PF-SINR-GBRsch}^{FD}(i, r, t) = \frac{SINR_{CSI}(i, r, t)}{SINR_{CSI,w}(i, t)} \cdot \frac{GBR(i)}{\bar{T}_{sch}(i, t)} \quad (6.13)$$

The only difference between the two metrics is that in equation (6.12) the FD metric is weighted with the ratio of the GBR requirement with respect to the past averaged throughput while in equation (6.13) the FD metric is weighted with the ratio of the GBR requirement with respect to the past averaged throughput updated only in the TTIs in which the user is scheduled by the FD scheduler.

An important aspect of equation (6.13) is that the average metric value in FD for a user does not change if such user is not scheduled for one or more TTIs.

GBR/\bar{T}_{sch} is a measure representing the number of TTIs the user needs to be scheduled in order to achieve the GBR. When used in frequency domain and assuming the user is scheduled every TTI, it represents a factor that “corrects” the amount of resources the user would get if such weight was not there. The weight should represent exactly the fraction of resources the user should get in order to achieve the GBR assuming it is scheduled every TTI.

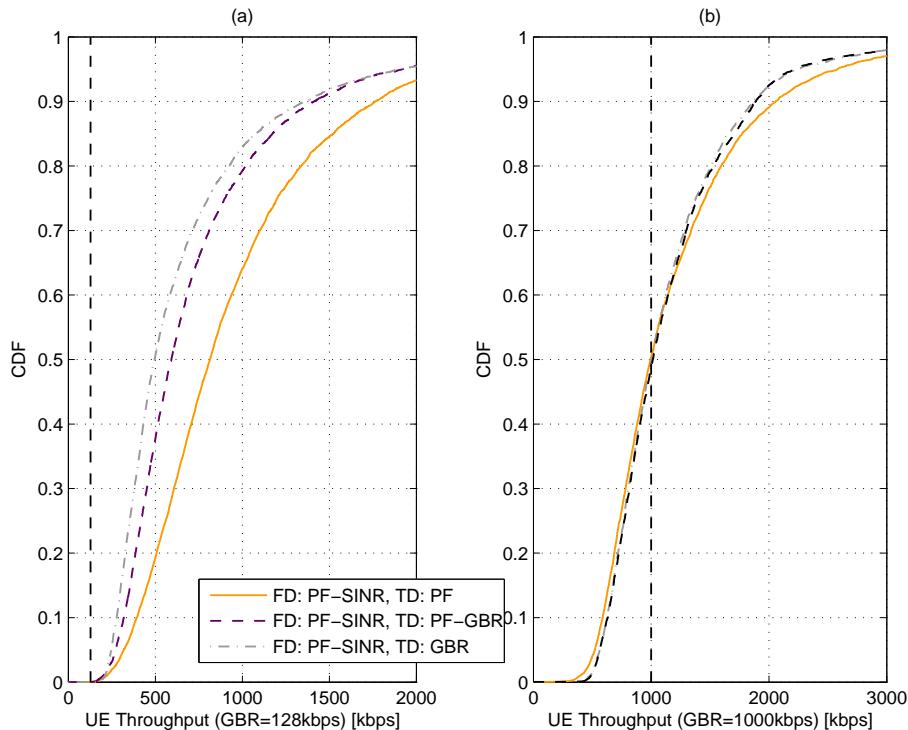


Figure 6.17: CDF of user throughput for the two different GBRs considered of 128 kbps and 1000 kbps. Only the TD metric is varied while the FD metric is kept to PF-SINR. Macro 1 case.

Figure 6.18 shows the performance in terms of throughput distribution for three different metric combinations which all assume the GBR metric in TD. In FD the PF-SINR and the QoS aware metrics defined in equations (6.12) and (6.13) are considered. Introducing a QoS component also in FD, for example using the PF-SINR-GBR metric, significantly improves the outage of the 1000 kbps users. However is only with the PF-SINR-GBRs_{sch} metric that the percentage of users in outage becomes almost 0. More specifically, only the combination of GBR/\bar{T} in the TD and GBR/\bar{T}_{sch} in the FD guarantees that the users are scheduled often enough by the TD scheduler and with the right amount of resources by the FD scheduler as to achieve the GBR requirement. Other combinations of metrics, which also use the PF-SINR-GBRs_{sch} metric in FD, are not as effective in providing the required GBR.

Figure 6.19 shows again the combination of GBR metric in TD and PF-SINR-GBRs_{sch} in FD as reference as well as two more metric combinations: one which is purely channel aware in TD and another one which is channel and QoS aware in TD: None of such additional combination is able to properly prioritize the users in the TD as to guarantee the required GBR.

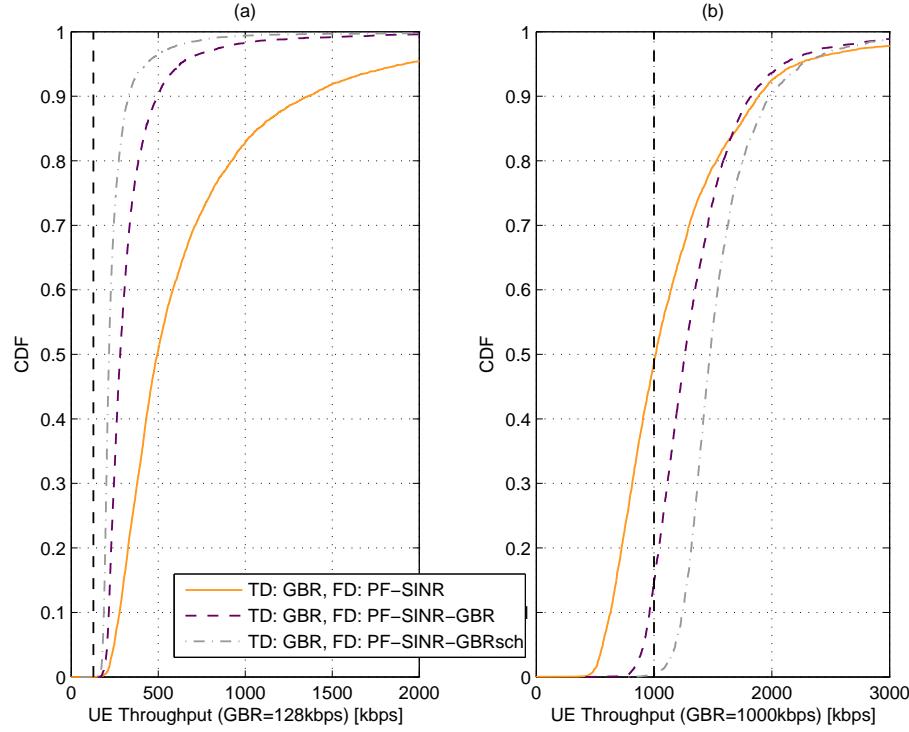


Figure 6.18: CDF of user throughput for the two different GBRs considered of 128 kbps and 1000 kbps. Only the FD metric is varied while the TD metric is kept to GBR. Macro 1 case. 10 UE/s/cell.

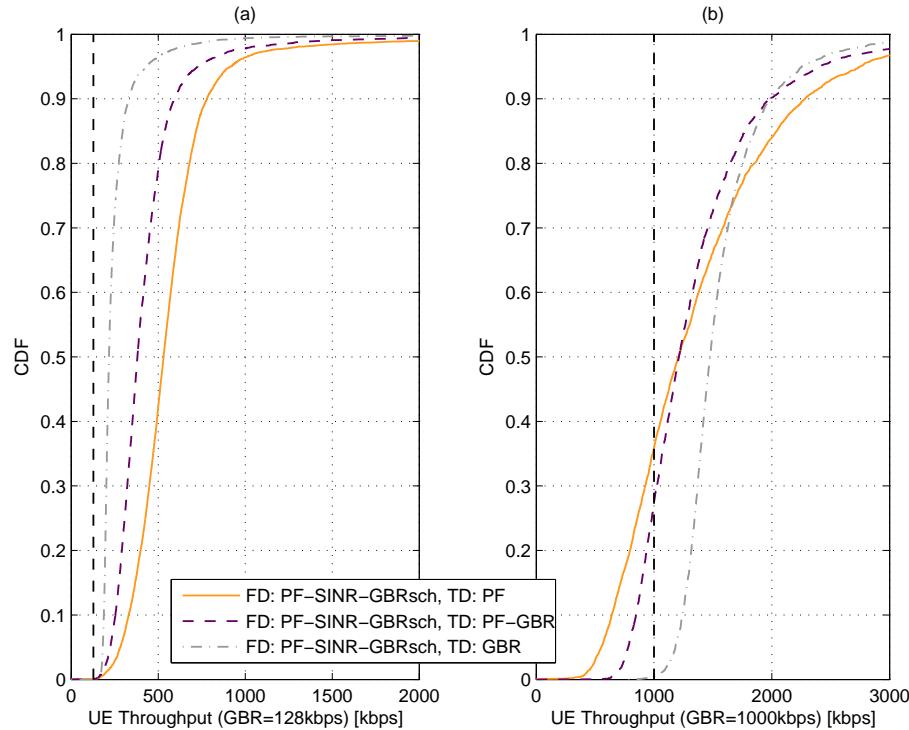


Figure 6.19: CDF of user throughput for the two different GBRs considered of 128 kbps and 1000 kbps. Only the TD metric is varied while the FD metric is kept to PF-SINR-GBRsch. Macro 1 case. 10 UE/s/cell.

The effect on the distribution of time-frequency resources vs APG for the GBR metric in TD

and the three considered metrics in FD is shown in Figure 6.20.

The results show that every metric combination chosen, being QoS-aware at least in TD, performs a differentiation in the allocation of resources between the two categories of GBR users. Moreover such differentiation is sensitive also to the APG so that the allocation of time-frequency resources decreases as the APG increases. The introduction of a QoS-aware factor also in FD (as in 6.20(b)) enlarges the gap between the two categories of users in terms of resource allocation. The metric adopted in Figure 6.20(c) further increases the gap between the two categories of users and is the only one able to provide the users with the required GBR. Thus the amount of resources depends both on the APG and the GBR requirement and such dependence becomes more pronounced as the weight of the QoS component becomes bigger. As a side note it is interesting to see how the users belonging to the higher GBR category are accepted by the AC at an higher APG value compared to the users of the lower GBR category.

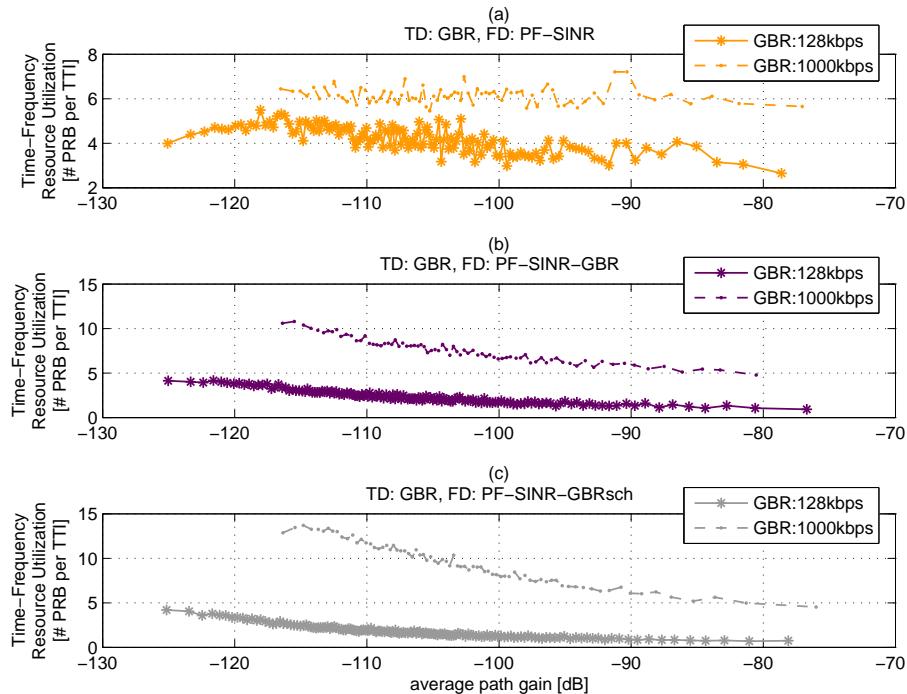


Figure 6.20: Time-Frequency resource utilization for the two different GBRs considered of 128 kbps and 1000 kbps. In TD only the metric indicated as GBR is used while the FD metric is varied. Macro 1 case. 10 UE/s/cell.

Another interesting effect is shown in Figure 6.21 where the number of active calls in the network grows as the metric becomes more QoS aware. A larger number of users is served at the same time when QoS-aware metric are deployed as a consequence of the low GBR users being allocated less resources and therefore staying in the system for a longer time.

Figure 6.22, finally, compares the performance of the three metric combinations in terms of blocking, outage, unsatisfaction and average cell throughput. The combination of GBR and PF-SINR-GBRsch shows a reduction of outage and unsatisfaction (for the 1000 kbps GBR) while keeping the average cell throughput almost unchanged compared to the other metrics and is therefore recommended over the other metrics combinations.

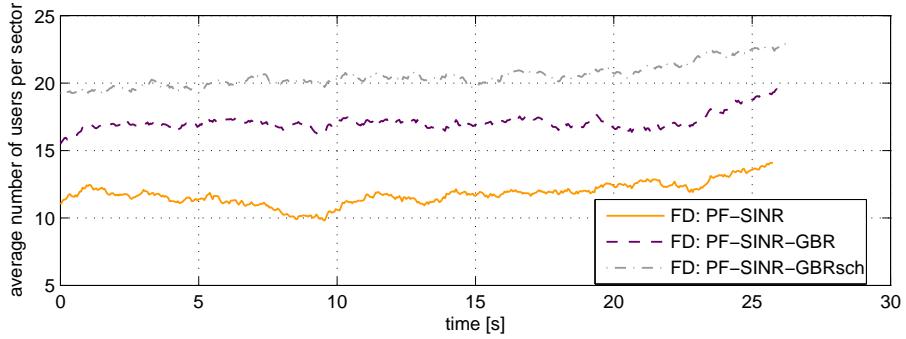


Figure 6.21: Active calls in the network vs time for different FD metrics. GBR metric in TD. Macro 1 case.

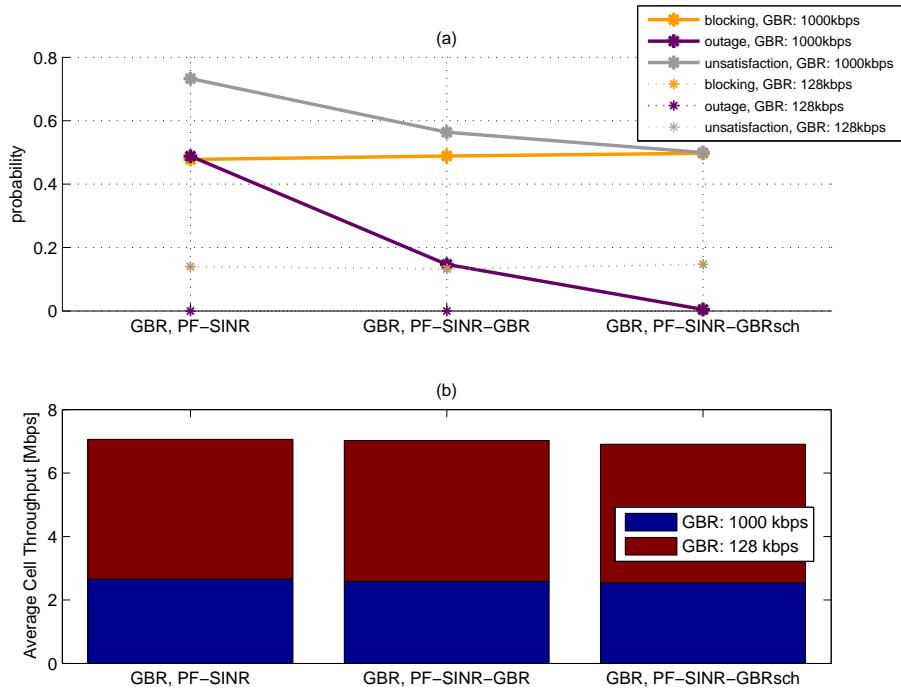


Figure 6.22: Blocking, outage and unsatisfaction probabilities for three different combination of TD and FD metrics which are, respectively, not QoS aware, QoS aware only in TD and QoS aware in TD and FD. Only the 1 Mbps GBR users are considered. Macro 1 case.

6.5 Conclusions

This first part of this chapter addresses the performance of different PF-like metrics for time and frequency domain scheduling. The performance analysis is carried out based on the way the resources are distributed in time and frequency among different users in different channel conditions. The results show that the PF-TTW metric, due to upper and lower limits of the wideband throughput, does not distribute resources fairly nor in an opportunistic way. The PF and PF-SINR metrics, instead, show a similar performance and exhibit different properties in terms of channel awareness and user scheduling activity. The combination of a wideband PF metric in TD and a PF-SINR metric in FD proves to be quite effective in scenarios where the experienced BLER deviates from the BLER target (due, for example, to power limitations) thus compromising the channel-awareness of the PF metric. The results show that in Macro 3 case such metric combination is able to achieve

a gain of approximately 21% in average cell throughput and 37.5% in outage user throughput with respect to a combination of PF both in TD and FD.

The second part of this chapter introduces the concept of QoS provisioning - limited to Guaranteed Bit Rate (GBR) - and focuses on the case of balanced load scenario where the users carry Elastic with Minimum Guarantee (EMG) traffic and are characterized by the same GBR requirement. If the aggregate minimum capacity requirement is lower than the average sector capacity then a GBR based metric is able to achieve the lowest outage. If the minimum capacity requirement is above the system capacity than the outage is very high, independently from the metric deployed. This result motivates the introduction of the Admission Control (AC) which is treated in the third part.

The third part introduces a Poisson user arrival process and the AC functionality. In this case, given the QoS-aware filtering performed by the AC, there is no need for a QoS-aware scheduling metric and the best performance in terms of low unsatisfaction and high cell throughput is achieved with a PF metric. For this reason a new scenario is introduced where the users, still carrying EMG traffic, have one of two possible GBR requirements with equal probabilities. The performance evaluation mostly revolves around the percentage of users in outage. The results show that a QoS-aware metric needs to be applied both in TD and FD in order to satisfy the requirements of both categories of users. The best combination is given by a GBR metric in TD and a PF-SINR-GBRsch metric in FD as it leads to the lowest percentage of unsatisfaction in the network without decreasing the average cell throughput.

Chapter 7

Scheduling for Service Differentiation

7.1 Introduction

This chapter continues the investigation on the QoS topic adding the CBR traffic to the scenario so far considered. More precisely there are two classes of users in the system. One is CBR which generates data at a fixed rate (e.g. streaming). The performance for CBR users severely degrades if the minimum throughput is not guaranteed. At the same time the allocation of more resources than the ones required would result in a waste because such resources will simply be filled with padding (see Appendix A.3.2 for more details). The other class is BE which does not require minimum throughput guarantee but can be allocated as few resources as possible (when not available) or as much resources as possible (when available). CBR users have priority over the BE ones which, ideally, should only be allocated the resources which are left after the QoS requirements of CBR users are met. As in the previous chapters AC is used to limit the number of connections for QoS-sensitive users so that the requirement of ongoing QoS connections can be maintained. For the BE users, which are not characterized by QoS requirements, no AC is performed.

Different studies have been carried out on the topic of scheduling and AC for service differentiation. In [52] an analytical framework is presented in the context of broadband packet-switched wireless networks. A system utility function as well as utility functions for BE and QoS-sensitive traffics are defined and used as mean for optimizing the parameters used in the proposed weighted fair queuing algorithm and threshold based AC. The framework is evaluated numerically using a queuing analytical framework.

In [53] is considered the scheduling problem of a cellular system where two classes of users exists, namely CBR users, which require exact minimum throughput average guarantee, and EMG users, which require minimum throughput average guarantee and more if possible¹. A combined scheduling and admission control algorithm is proposed to maximize a new generalized utility function, based only on minimum throughput guarantee, whose property of strict concavity guarantees the asymptotic convergence of proportional fairness. However the analysis is limited to a single-cell downlink scenario with only one user being served at a time, a persistent data traffic and a throughput calculated using the Shannon bound. Such approach, though interesting, is quite different from the line of thought followed throughout this thesis and no attempt is made to extend and adapt such work to the problem at hand.

In [54] an adaptive cross-layer scheduling algorithm is proposed and it is shown to outperform weighted-fair-queuing schedulers with respect to average normalized packet delay, average effective user throughput, user blocking, and user dropping for data services in downlink. The name is derived from the fact that the proposed algorithm adapts to the packet delay deadlines on link

¹The BE class considered in this chapter is a special case of the EMG class with a zero minimum throughput guarantee.

layer and channel qualities on the physical layer. The performance of the scheduling algorithm is evaluated under a simple AC algorithm which limits the number of active queues to grow beyond a specified threshold.

In [55] the PF scheduling metric is modified with QoS-dependent scaling factors which are based on the principle of Required Activity Detection (RAD) introduced in [56]. The proposed scheduler is shown to outperform the set of barrier-function-based schedulers [57] by a significant gain in cell throughput. Moreover it is shown to be as flexible as to support a general mix of real-time and non-real-time traffics. The scenario investigated is the HSDPA but the same principle could be extended to the TD of LTE uplink. Additionally, the proposed approach well fits within the designed scheduling framework as it just requires an adaptation of the metrics deployed to the LTE system.

The chapter is organized as follows: Section 7.2 presents the design of the metric used in TD to prioritize the CBR over the BE users. Section 7.3 introduces the modeling assumptions and the parameter settings while Section 7.4 presents the performance evaluation. Such section is divided in two subsection: the first about the impact on performance of the scheduling metric, the second about the impact on performance of the buffer knowledge. Finally Section 7.5 draws some conclusions.

7.2 TD Metric Design

The strategy followed, as in the previous scenarios, is to integrate in the metric additional aspects which enable the prioritization of the CBR users over the BE ones and ideally provide the first with the right amount of resources and the latter with the resources left. Unlike the previous scenarios, though, the problem is complicated from the fact that BE users do not have a data rate requirement and therefore cannot easily be assigned a scheduling metric value (or more specifically a weight) which depends on it.

For this reason the design follows a different approach where the user scheduling activity needed to meet the GBR is estimated for CBR users and then subtracted from the overall available resources to be allocated to BE users. The design is going to affect only the TD metric while possible modifications to the FD metric, which could improve the proposed solution, are left for future studies. More explicitly this means that only TD resources are controlled while the frequency resources are in average shared equally among the users handed over to the FD scheduler. This choice is also motivated by the fact that it is not straightforward to derive some weight in FD for the BE users as they are not characterized by a GBR requirement. At the same time a TD-only solution is appealing for implementation simplicity.

The estimation of the TD resources for all the CBR users is calculated as:

$$R_{CBR}(n) = \sum_{i=1}^{N_{CBR}} \frac{GBR(i)}{\bar{T}_{sch}(i, n)}$$

where N_{CBR} indicates the number of CBR users and $\bar{T}_{sch}(i, n)$ represents, for a user i at scheduling instant n , the past averaged acknowledged throughput updated only in the scheduling instants in which the user is allocated some resources. The quantity $GBR(i)/\bar{T}_{sch}(i)$ indicates the number of TTIs the user i should be scheduled in order to meet the GBR requirement. This means that if a CBR user happens to experience the condition

$$\frac{GBR(i)}{\bar{T}_{sch}(i)} > 1$$

then it won't be able to achieve the GBR unless some proper differentiation is applied also in frequency providing it with more resources than it would otherwise be allocated.

The resources available in time domain for the BE users at scheduling instant n can be derived as:

$$R_{BE}(n) = R - R_{CBR}(n)$$

where R indicates the total time-domain resources. As estimate of the resources available every TTI we take the number of scheduled users per TTI which, for further simplicity, is assumed to be equal to the number of users passed to the FD scheduler (N_{FD}) corrected with the BLER target ($BLER_T$) as specified in equation (7.1):

$$R = N_{FD} \cdot (1 - BLER_T) \quad (7.1)$$

Therefore the estimation of the time-domain resources for BE users can be written as:

$$R_{BE}(n) = N_{FD} \cdot (1 - BLER_T) - \sum_{i=1}^{N_{CBR}} \frac{GBR}{\bar{T}_{sch}(i, n)}$$

At this point we can define the time-domain metric, to which we are going to refer as RAD-metric, as:

$$M_{RAD}^{TD}(i, n) = RR(i, n) \cdot \left(\underbrace{\frac{GBR(i)}{\bar{T}_{sch}(i, n)}}_{w_{CBR}} + \underbrace{\frac{R_{BE}(n)}{N_{BE}} \cdot c(i)}_{w_{BE}} \right) \text{ with } c(i) = \begin{cases} 0 & \text{for } GBR(i) > 0 \\ 1 & \text{for } GBR(i) = 0 \end{cases} \quad (7.2)$$

where $RR(i, n)$ indicates the RR metric for user i at scheduling instant n , while N_{BE} indicates the number of BE users. The metric, apparently complex, is simply a modification of the RR metric via two weights, indicated as w_{CBR} and w_{BE} :

- w_{CBR} is used to establish priorities among the different CBR users and for the CBR users in comparison to the BE users.
- w_{BE} is used to establish priorities only in comparison to the CBR users while the priorities among different BE users are only managed via the RR metric.

It is interesting to note that, for $c = 0$, equation (7.2) becomes:

$$M_{RAD}^{TD}(i, n) = RR(i, n) \cdot \frac{GBR(i)}{\bar{T}_{sch}(i, n)} \approx \frac{GBR(i)}{\bar{T}(i, n)}$$

that is, the metric proposed here to handle the CBR users is similar to the one suggested in the previous chapters to handle the prioritization of GBR users as long as the number of TTIs represented by the RR metric does not increase significantly beyond 1. The advantage of the current metric over the previous one is that it facilitates the estimation of the extra resources which we would like to assign to BE users. Such metric is based on the principle proposed in [56] and it extends to LTE the work carried out in [55].

7.3 Modeling Assumptions

The main simulation parameters are listed in Table 7.1 while Table 7.2 indicates the user arrival for each of the two traffic types. The idea behind the chosen arrival values for BE and CBR users is to fill the available resources with BE traffic and then introduce a gradually increasing amount of CBR traffic to analyze whether the RAD metric is able to properly prioritize the CBR users over

the BE users. As mean of comparison also a case with only CBR traffic is considered. Each of the traffic proportion considered have also been given a letter (from (A) to (F)) to facilitate referring to it. The same QoS-aware AC algorithm deployed in Chapter 6 is used.

Table 7.1: Main simulations parameters

Parameter	Setting
Offered throughput and user arrival (Poisson distributed)	see Table 7.2
# UEs handed to FD	10
# PRBs per UE	1 to 48
TD scheduling	RR, RAD
FD scheduling	PF-SINR
BLER target at 1st transmission	30%
Propagation scenario	3GPP Macro case 1 (ISD of 500 m)
Call size (for BE users)	1 Mb
Knowledge of the buffer at eNode-B	Ideal
GBR (for CBR users)	512 kbps
Packet size (for CBR users)	6400 bytes
Packets inter-arrival (for CBR users)	100 ms
Call duration (for CBR users)	5 s

Table 7.2: Offered throughput and related user arrivals.

Case	(A)	(B)	(C)	(D)	(E)	(F)
Offered throughput (right) vs user arrival (below)	BE: 7 Mbps; CBR: 0 Mbps	BE: 7 Mbps; CBR: 1 Mbps	BE: 7 Mbps; CBR: 3 Mbps	BE: 7 Mbps; CBR: 5 Mbps	BE: 7 Mbps; CBR: 7 Mbps	BE: 0 Mbps; CBR: 7 Mbps
BE user arrival [UE/s/cell]	7	7	7	7	7	0
CBR user arrival [UE/s/cell]	0	~0.4	~1.2	~1.9	~2.7	~2.7

7.4 Performance Evaluation

7.4.1 Impact of scheduling metric on system performance

The performance of the proposed metric is evaluated in terms of user throughput distribution, average cell throughput figures and unsatisfaction probabilities. Additionally, in order to explain some results, also an analysis of the time-frequency resources utilization is provided.

Figure 7.1 (a) and (b) show the CDF of the user throughput for respectively the BE users and the CBR users. The metrics compared, both operating in time-domain, are the RR and RAD. In case of only BE traffic or only CBR traffic the metrics RR and RAD are expected to give the same results. The difference in performance between RR and RAD becomes appreciable in case of mix of BE and CBR traffics. Specifically, the impact of the RAD metric becomes more and more noticeable as the amount of offered CBR traffic increases. As expected the resources are subtracted to BE users and allocated to CBR users and, as result, the outage is decreased from about 100% to 15%. Even though considerably lower, such outage figure is still fairly high, therefore part of the

following results will be devoted to explaining the reasons behind such outage and possible ways for reducing it.

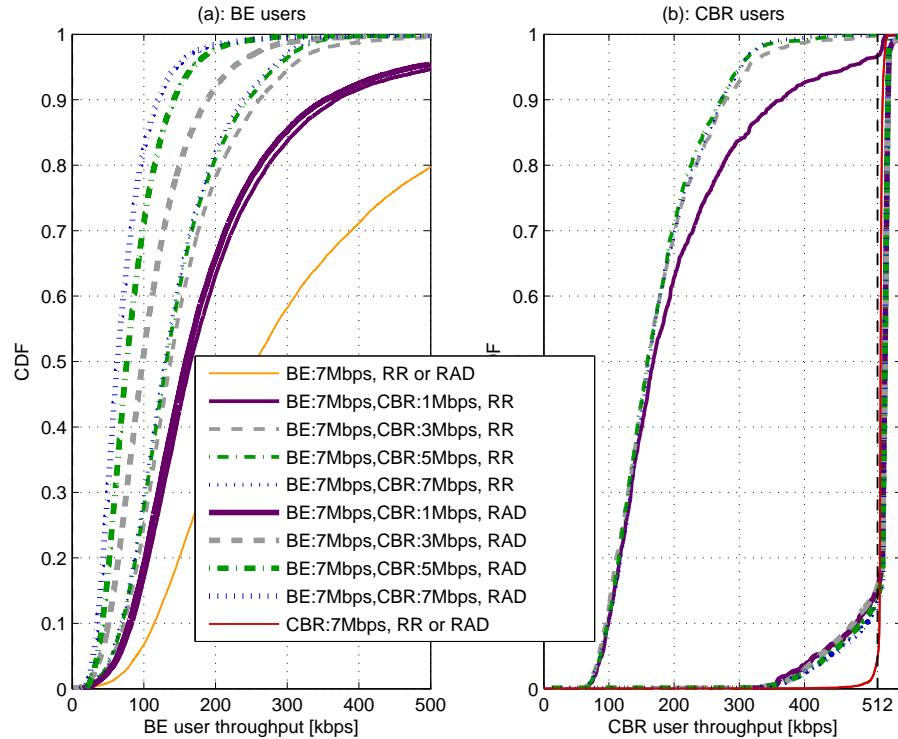


Figure 7.1: (a) CDF of the BE user throughput. (b) CDF of the CBR user throughput

Figure 7.2 compares the RR and RAD metrics in terms of carried cell throughput of CBR and BE users at different offered throughput (or equivalently user arrivals). The trends highlighted in Figure 7.1 are here quantified. Specifically it is shown that the RAD metric is much better capable of prioritizing the CBR users over the BE users. On the other hand we can also notice that not all of the offered CBR throughput can be carried out of the system. This becomes evident if we compare the cases indicated as (E) and (F) in Table 7.2. In case (E) the average cell throughput of CBR users is about 15% below the average cell throughput of CBR users in case (F).

Such lower average cell throughput is the result of mainly two factors:

- Higher blocking probability, as shown in Figure 7.3(a) and detailed in Table 7.3.
- Higher outage probability, as shown in Figure 7.3(b) and detailed in Table 7.3.

The higher blocking probability, in turn, is the consequence of two factors, that is:

- Higher bandwidth utilization.
- Higher outage probability.

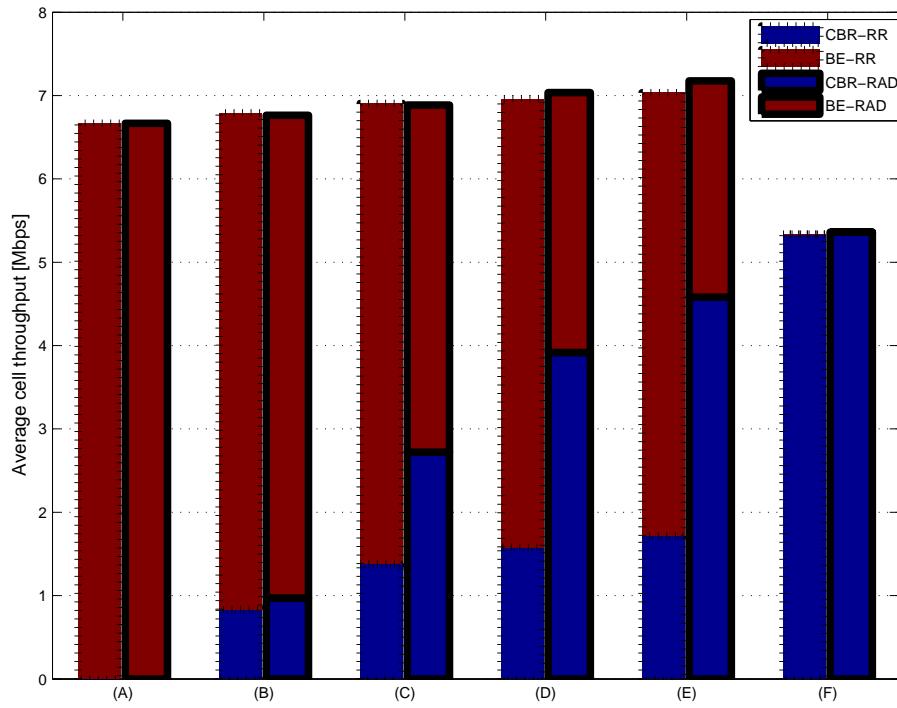


Figure 7.2: Average cell throughput for BE and CBR users for RR and RAD metrics.

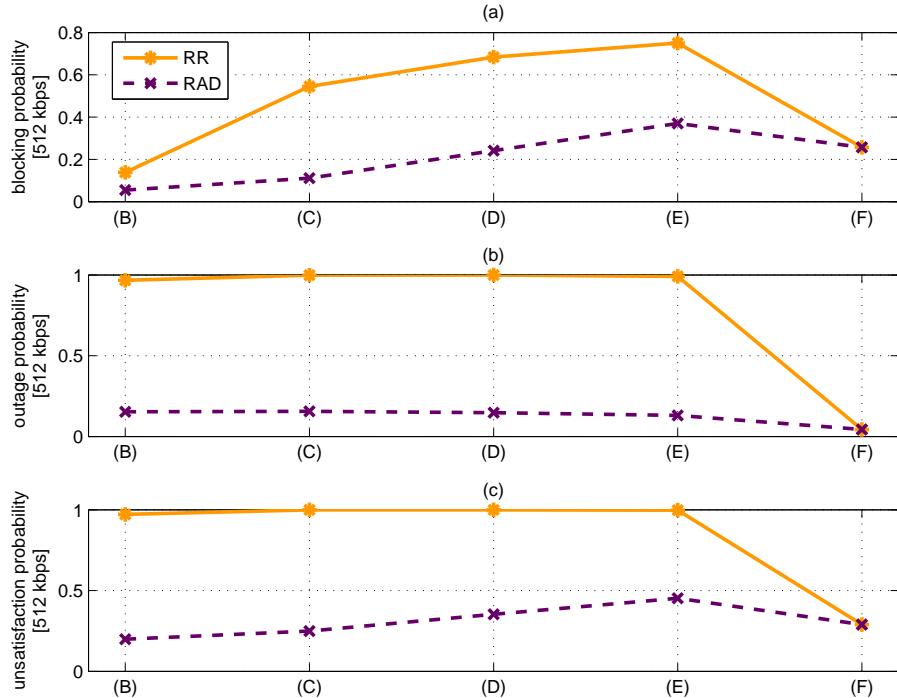


Figure 7.3: Blocking, outage and unsatisfaction for the CBR users at different arrival rates (that is, different portions of BE and CBR users).

The higher bandwidth utilization is caused by the presence of BE users and it leads to larger Noise Rise (NR) as shown in Table 7.3. This effect produces a larger rejection of incoming users as the available capacity estimated by the AC includes also the interference level. The higher

outage probability, instead, is investigated in the coming paragraphs but we can say it increases the blocking probability as a consequence of some users staying longer in the system. The negative impact of the outage on the user satisfaction is therefore twofold: not only it causes a percentage of users to be unable to achieve the required GBR, but it also increases the blocking probability. In other words a user in outage is unsatisfied (because unable to achieve the GBR) and causes unsatisfaction (by staying longer in the system and blocking other incoming users). It is therefore worth to try to investigate the reasons behind such outage value.

Table 7.3: Key Performance Indicator (KPI)s for case (E) (that is, mix of CBR and BE traffics) and case (F) (that is, CBR only).

	Bandwidth utilization	Noise rise median	Blocking probability	Outage probability	Aggregate CBR throughput
BE: 7 Mbps CBR: 7 Mbps (E)	99.6%	12.9 dB	37%	13%	4.58 Mbps
CBR: 7 Mbps (F)	68.7%	11 dB	25.7%	4.3%	5.36 Mbps

Figure 7.4 shows the time-frequency resource utilization for the RR and RAD metrics, for the BE users and the cases (C) and (E). Figure 7.5 shows the same statistics for the CBR users and the cases (C), (E) and (F). Figure 7.4(b) confirms that no differentiation is performed in frequency domain and the considered curves overlap with each other. Figure 7.4(a) shows how the RAD metric is able to significantly reduce the time resources allocated to the BE users compared to the RR metric according to the proportion of CBR traffic. Figure 7.5 adds the case (F) to the set of curves. The RR metric does not perform any differentiation between the cases (C) and (E). The case (F) under RAD metric shows in Figure 7.5(b) how the power limitations experienced in the lower path-gain range create the need for a much higher time scheduling activity (see Figure 7.5(a)) and, as result, a much higher time-frequency resource utilization (Figure 7.5(c)). In the cases (C) and (E) under the RAD metric, the presence of BE users considerably limits the resource allocated to CBR users in frequency domain. The only measure which can be taken by the RAD metric is to allocate more time resources to the users in disadvantaged channel conditions until the GBR is achieved. This is possible up to a path-gain value of approximately -115 dB. At lower path-gain values the CBR users cannot be scheduled more than 100% of the time and this results in an amount of time-frequency resources lower than the ones needed to achieve the GBR. This is clearly visible when comparing the cases (E) and (F) under RAD metric in the path-gain range -115 dB down to -120 dB which shows an increasing gap in the amount of allocated resources. Additionally in the (F) case, given the lower blocking probability, the path-gain range extends down to a value approximately 2 dB lower.

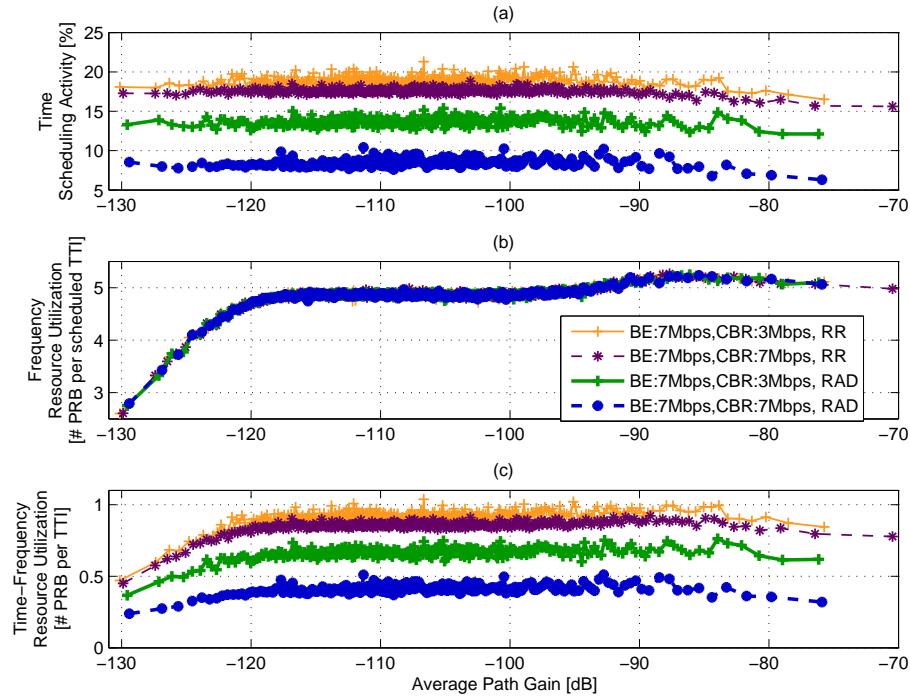


Figure 7.4: Time (a), Frequency (b) and Time-Frequency (c) resource utilization vs APG for BE users only and for RR and RAD metrics. Every point represents the mean over 50 consecutive users (with respect to APG).

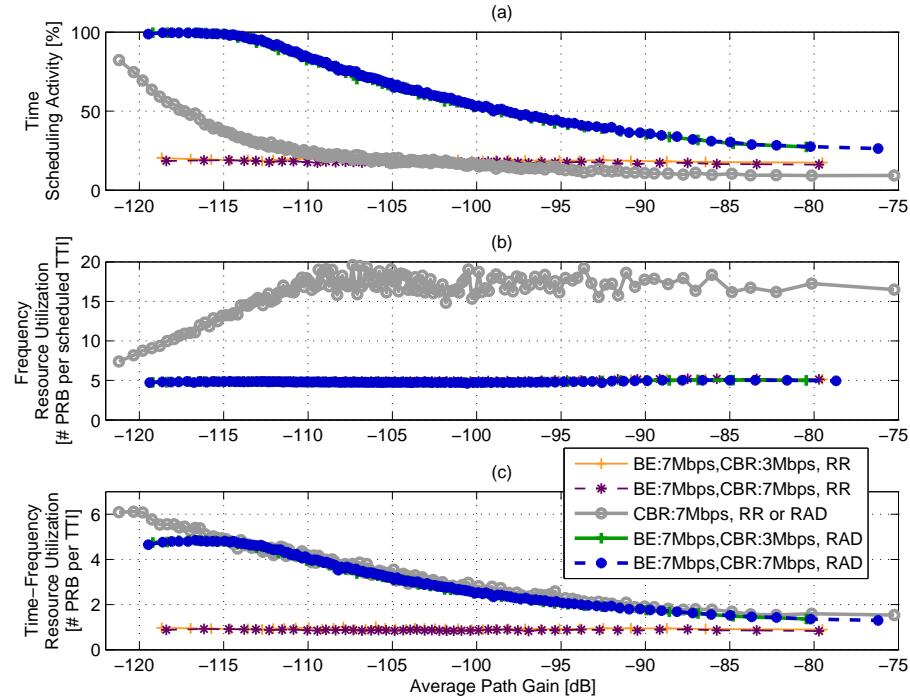


Figure 7.5: Time (a), Frequency (b) and Time-Frequency (c) resource utilization vs APG for CBR users only and for RR and RAD metrics. Every point represents the mean over 50 consecutive users (with respect to APG).

Figure 7.6 shows the user throughput vs the APG. Compared to the RR metric the RAD met-

ric shows how the user throughput is increased over the whole path-gain range until the GBR is reached but this is not possible in the lowest range because there are simply not enough resources in time to enable also the users in the lowest path-gain range to achieve the GBR. Different solutions can be implemented to overcome the problem. The simplest would consist in making a more restrictive estimation of the available capacity for BE users by lowering the overall capacity available in the first place. Another solution would consist in applying a proper weight also in frequency domain to differentiate among GBR users and BE users. Ideally, also the AC, by properly estimating the available resources, should be able to take care of this problem.

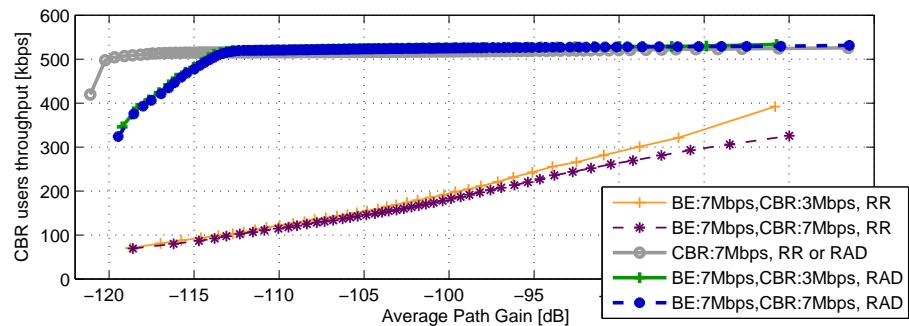


Figure 7.6: CBR users throughput vs APG. Every point represents the average over 50 consecutive users (with respect to APG).

7.4.2 Impact of buffer knowledge on system performance

The results presented in Section 7.4.1 are produced assuming a perfect buffer knowledge of the UE at the eNode-B. This means that the allocation of the resources, specifically the process of bandwidth expansion presented in Section 5.2, can stop as soon as there are no data bits left in the buffer. This knowledge could be of critical importance in a streaming service where the buffer can be emptied quite fast when the incoming packets are of small size compared to the Transport Block Size (TBS).

To evaluate the impact of such knowledge in this section we consider both the cases of ideal buffer knowledge (the eNode-B has perfect knowledge of the UE buffer) and the case of limited knowledge (the eNode-B knows only whether the user has data to transmit or not). Additionally, in order to relate such knowledge with the characteristics of the CBR traffic (that is, packet size and interarrival time), two types of CBR traffic (referred to as “CBR large packets” and “CBR small packets”) are considered as specified in Table 7.4. For certain types of traffic (like streaming) the eNode-B could predict the knowledge of the buffer without need for buffer status reports. On the other hand the aim here is simply to evaluate the impact of buffer knowledge and not to design algorithms to solve the problem of its lack at the eNode-B.

Table 7.4: CBR traffic specifications.

	Packet size	Inter-arrival time
CBR large packets	6400 bytes	100 ms
CBR small packets	768 bytes	12 ms

Figure 7.7 shows the CDF of the throughput for CBR users considering the cases (E) and (F) of traffic mixes and the RAD metric. The same statistic for the case of BE users is not presented as the buffer knowledge has in this case limited or no impact. As expected the knowledge of the

buffer has an impact on the allocation of the resources and it causes parts of the resources to be filled with padding. This aspect is more relevant for the case of CBR only compared to the case of CBR and BE mix, as shown by Figure 7.7 (a) and (b), and even more relevant for the “small packets” case compared to “large packets” case.

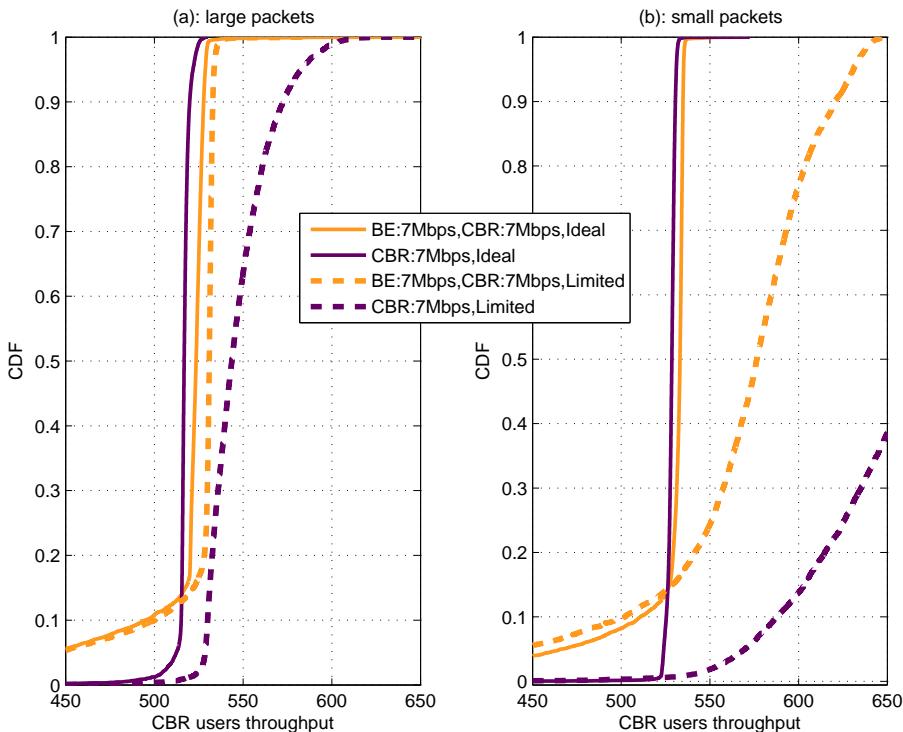


Figure 7.7: (a) CDF of the throughput under ideal and limited buffer knowledge for CBR users whose traffic is characterized by large packet size (6400 bytes) and large interarrival time (100 ms). (b) CDF of the throughput under ideal and limited buffer knowledge for CBR users whose traffic is characterized by small packet size (768 bytes) and small interarrival time (12 ms). Only the cases (E) and (F) and the RAD metric are considered.

In order to quantify the loss on data throughput due to the allocation of resources for padding, Figure 7.8 shows the average cell throughput (divided in BE and CBR) after the padding has been removed. The impact is limited in case of traffic mix or large packet size while it accounts for a loss of approximately 8% in case of small packet size and CBR only traffic. The reason is the same as the one given in Section 7.4.1 to justify the lower CBR throughput in case (E) compared to CBR throughput in case (F): A larger bandwidth utilization, in this case caused by padding, produces a higher NR and consequently higher blocking probability which results in the lower carried throughput. The related figures are shown in Table 7.5. For completeness a CDF of the NR is shown in Figure 7.9.

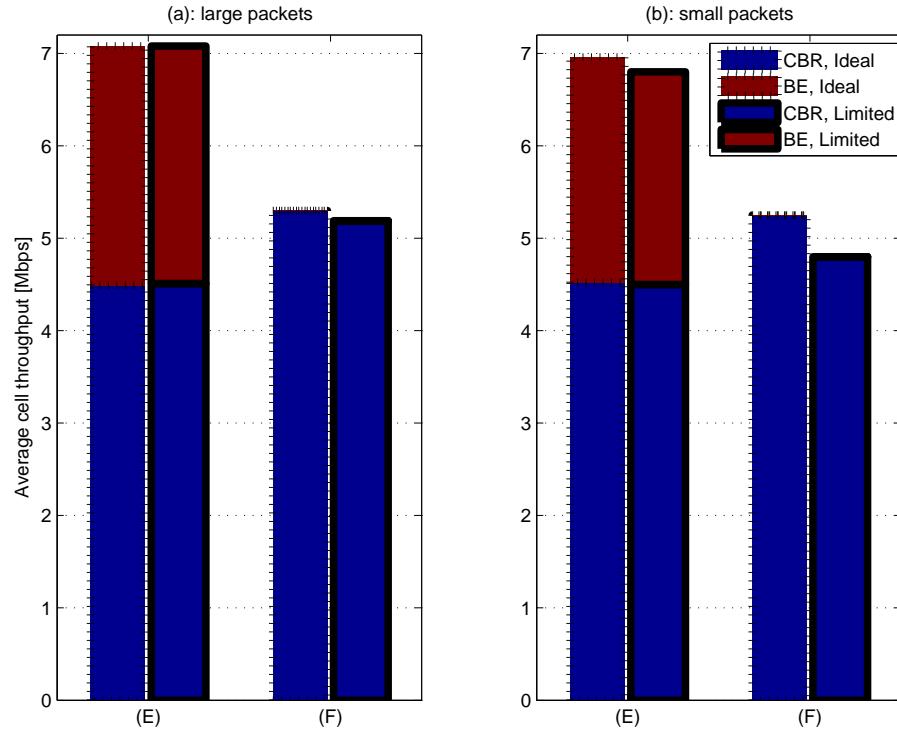


Figure 7.8: (a) Average cell throughput under ideal and limited buffer knowledge for BE and CBR users whose traffic is characterized by large packet size (6400 bytes) and large interarrival time (100 ms). (b) Average cell throughput under ideal and limited buffer knowledge for BE and CBR users whose traffic is characterized by small packet size (768 bytes) and small interarrival time (12 ms). Cases (E) and (F). RAD metric.

Table 7.5: KPIs for ideal and limited buffer knowledge (CBR only traffic and small packet size).

	Bandwidth utilization	Noise rise median	Blocking probability	Average cell throughput
CBR only, small packets, ideal knowledge	70%	11.1 dB	26.5%	5.25 Mbps
CBR only, small packets, limited knowledge	88%	12.3 dB	32.7%	4.8 Mbps

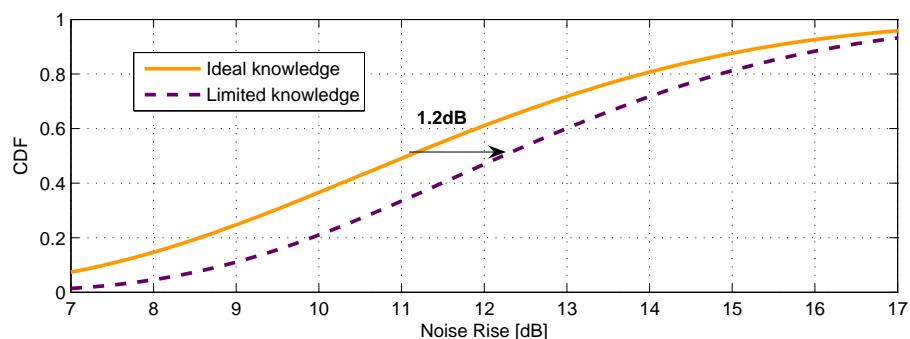


Figure 7.9: CDF of the NR for ideal and limited buffer knowledge (CBR only and small packet size).

7.5 Conclusions

In this chapter we have introduced a new scenario characterized by a mix of CBR and BE traffic and designed a TD metric to prioritize the allocation of CBR users over the BE users. The metric designed, named RAD, is able to prioritize the time resources allocation of CBR users over BE users. The result is a drop of the outage user throughput for CBR users from approximately 100% (under RR metric) to less than 15%. Interestingly enough, the presence of BE users affects the performance of CBR users by increasing bandwidth utilization, NR and consequently blocking probability, as well as by using frequency domain resources. In such situation assigning priority only in TD to CBR users over BE users is not enough to provide the most disadvantaged users with the required resources and therefore GBR. Possible solutions to this problem include the adoption of a proper prioritization also in FD as well as a more accurate estimate of the extra resources available for BE users.

Another topic investigated in this chapter pertains the impact that the knowledge of the buffer has on the system performance. Such knowledge is important only in the case of CBR only traffic and relatively small size of the packets. In the specific case considered (packet size of 768 bytes and inter-arrival time of 12 ms), the loss compared to a perfect knowledge is approximately 8%. The reason for the loss is again to be found in an increased bandwidth utilization, this time due to padding, which increases the NR and consequently the blocking leading to a lower average cell throughput.

Chapter 8

Overall Conclusions and Recommendations

The overall objective of this thesis has been to study the Radio Resource Management (RRM) functionalities in the context of LTE Uplink. The attention has been focused on the Power Control (PC), Adaptive Transmission Bandwidth (ATB) and Packet Scheduler (PS) entity but other functionalities, including fast Adaptive Modulation and Coding (AMC), Outer Loop Link Adaptation (OLLA), Admission Control (AC) and Hybrid Automatic Repeat reQuest (HARQ), have also been taken into account in the analysis of results due to their interaction with the considered entities.

The main tool used for the generation of the results has been the system level simulator described in Appendix A.1. Such simulator has been developed as part of the PhD project together with other colleagues at Nokia Siemens Networks. The development of the simulator has required a significant time investment in terms of modeling, software design, implementation and testing of the mentioned functionalities. Part of the PhD project has also been spent on the development of the dynamic system level simulator ELIISE whose results have not been reported here as they followed an entirely different approach than the one here presented.

The thesis is divided in three parts where the first addresses the topic of PC and evaluates the system performance after integrating interference measurements in the formula for setting the power. The second and third part encompass the topic of scheduling which represents the area where the main contribution of this project lies. Specifically, the second part, assuming an ideal traffic model, develops a scheduling framework (consisting of allocation algorithm and scheduling metric) on top of which the service differentiation, treated in the third part, is built.

8.1 Interference-based Power Control

The main idea behind Interference based Power Control (IPC) has been the introduction of the dependence from the user generated interference (expressed via the path-loss to other eNode-Bs) when setting the transmit power. Following the same approach of fractional PC, the dependence from the path-loss to the serving eNode-B and from the sum of path-loss to other eNode-Bs have each been weighted with a factor indicated as α and β . A relevant finding has been that the best set of performance points (in terms of cell coverage and cell capacity) is obtained under the condition $\alpha + \beta = 1$. Additionally we have shown that, under such condition, the formula for setting the power is equivalent to the formula proposed in [35] where the SINR target for a user is set as function of the path-loss difference between serving eNode-B and second strongest eNode-B.

The suggested approach provides a significant gain over the traditional Fractional Power Control (FPC). If we consider as reference the FPC case obtained for $\alpha = 0.6$, the IPC approach can pro-

vide a gain above 50% in outage user throughput by setting $\alpha = 0.6$ and $\beta = 0.4$ or a gain of approximately 15% in average cell throughput by setting $\alpha = 0.1$ and $\beta = 0.9$.

Such gains are preserved only in a full buffer traffic scenario while they are considerably reduced in a finite buffer traffic scenario where the path-loss distribution and therefore the interference patterns are significantly different.

Being the finite buffer traffic a closer model to a real scenario than the full buffer model and being the gain, according to the current evaluation, considerably reduced, the simpler FPC approach is recommended. On the other hand the development, in the future, of algorithms able to fine tune the P_0 (or, equivalently I_0) parameter may improve the performance of IPC also in a finite buffer scenario and lead to a gain sufficiently high as to being considered for implementation in a real system.

8.2 Channel-Aware Scheduling

The topic of channel aware scheduling is treated in its various aspects in Chapters 4, 5 and part of Chapter 6. Due to the limitation imposed by the SC-FDMA transmission scheme, which requires the Physical Resource Block (PRB)s assigned to the same user to be adjacent with each other, the allocation algorithm is severely complicated. For this reason the initial approach is to assume a Fixed Transmission Bandwidth (FTB) allocation as done in Chapter 4. The algorithm used initially is of a greedy type and is referred to as *matrix-search* as it consist of arranging the metrics in a matrix of users (the rows) and RCs (the columns) and searching the highest metric in the matrix.

The first objective is to evaluate the gain provided by a channel aware metric - like Proportional Fair (PF) - compared to a channel blind one - like Random (RAN). Depending on bandwidth and number of users, the average cell throughput gains using a greedy allocation algorithm range from 4% to 35% for the Macro 1 case with a gain of approximately 20% for a scenario with 10 users and 6 PRBs per user.

The FTB allocation offers a more clear separation of frequency selectivity and user diversity. We take advantage of that possibility by shifting our focus on the gains deriving from such mechanisms. As expected a higher frequency selectivity, corresponding to a narrower bandwidth allocation, leads to a higher cell throughput. The same can be stated for an increasing multi-user diversity. It is interesting to notice that the higher the frequency selectivity the larger, comparatively, is the number of users needed to reach user diversity gain saturation. Moreover there exist several cases in which a larger bandwidth combined with a higher number of users provides higher gain than a narrower bandwidth combined with a lower number of users, that is, in some cases the user diversity gain surpasses the frequency selectivity gain.

In order to evaluate the gain achievable from considering a larger search space a more general allocation algorithm, still based on FTB and that includes the matrix-search algorithm as special case, is proposed. Such algorithm, referred to as *tree-search*, is able to achieve only a marginal gain of a few percentage points in the considered scenarios. For this reason, and given the considerably lower computation complexity, the matrix-search algorithm is preferred to the tree-search algorithm and therefore recommended. Moreover it is used as default assumption in the FTB simulations for the rest of the study.

The evaluation provided using FTB, besides giving insights into the gains and the diversity mechanisms, offers a useful benchmark against which other algorithms can be evaluated. In Chapter 5 a new allocation algorithm is proposed which integrates the ATB functionality in the allocation algorithm. The motivation behind such algorithm is to achieve a more flexible bandwidth allocation in order, for example, to cope with different cell loads and traffic types or user power limitations.

The proposed algorithm shows, as expected, adaptation to user power limitations. The most interesting property, though, is the inbuilt adaptation to cell load. The curve of the average cell throughput vs the number of users obtained with the ATB follows the envelope of the different curves obtained with the FTB for different bandwidth sizes. This means that for a given number of users, the ATB results in an average bandwidth allocation which provides a throughput larger or comparable with the one of the FTB whose bandwidth has been fine tuned. Moreover, in case of unbalanced load the ATB is able to achieve a larger bandwidth utilization than the FTB. Thus, even though the gain achieved by the ATB is limited, it is still to be preferred to the FTB given the considerable flexibility offered.

8.3 Scheduling for Service Differentiation

Chapters 6 and 7 deal with the topic of service differentiation. As first step, in this sense, it is suggested to add a Time Domain (TD) unit to the scheduling framework. This is motivated not only by the need of assigning different priorities to users with different Quality of Service (QoS) requirements, but also by the need of taking into account limitations due to control channels and computational complexity.

The introduction of the TD unit requires also the adoption of a TD metric. For this reason a short study is conducted in the first part of Chapter 6, on the most appropriate combination of purely channel aware TD and Frequency Domain (FD) metrics without yet taking into account QoS requirements. The results indicate that the combination of a throughput-based PF metric in TD and an SINR-based PF metric in FD provide a significant gain (about 20% in average cell throughput and 35% in outage user throughput) over a combination of throughput-based PF both in TD and FD. Thus a combination of a throughput-based PF metric in TD, which embeds the memory of past scheduling decisions and uses such information to select the users to schedule, and an SINR-based PF metric in FD, which, being memory-less, simply aims at exploiting the channel gain, is to be preferred.

Afterwards the topic of QoS provisioning is introduced but limited to the Guaranteed Bit Rate (GBR) requirement only. Such requirement is first introduced in the context of Elastic with Minimum Guarantee (EMG) traffic. Initially the investigation is limited to the case in which all the users in the network are given the same GBR. In this case the proposed approach is to preserve the channel exploitation in the FD by deploying the SINR-based PF metric, and, instead, make possible the fulfillment of the GBR requirement via a modification of the scheduling activity performed by the TD scheduler. For this reason the TD metric is modified with a GBR-dependent weight. The results show that in such case the modification of scheduling activity (and therefore distribution of available resources) via TD metric only is sufficient to guarantee proper user prioritization and satisfaction as long as the system has enough capacity to serve all the users in the system.

The case of a number of users larger than the system can support (because the minimum aggregate throughput requirement is larger than the average cell throughput) motivates the introduction of the AC which guarantees system stability. Under such assumptions is considered the more complex case of two classes of EMG users each characterized by a different GBR. In this case it is shown how a GBR-dependent metric in TD only is not sufficient to give proper user prioritization and the introduction of a GBR-dependent metric in FD is also needed. With the adoption of GBR-dependent metrics in TD and FD the user outage is eliminated and the users achieve the largest satisfaction. For this reason a GBR-dependent metric should be deployed also in FD when the modification of the scheduling activity only is not sufficient to provide the users with the proper amount of resources.

Finally, Chapter 7 introduces the Constant Bit Rate (CBR) traffic type and deals with the

problem of scheduling with the right priority CBR users and Best Effort (BE) users (the latter being a specific case of EMG users with GBR set to 0). In this case the problem is solved via a modification of the TD metric only which is referred to as Required Activity Detection (RAD) metric. The outage of CBR users is reduced from 100% with the Round Robin (RR) metric to 13% with the RAD metric. The residual outage cannot be eliminated unless a more sophisticated estimation of resources available for BE users is done or proper prioritization is introduced also in FD.

8.4 Topics for Future Research

The scheduling framework developed in the course of this research project aims at providing system capacity enhancements and proper service differentiation. Further developments would therefore naturally address one or the other topic.

On the topic of capacity enhancement further advances could be made by improving the ATB-based allocation algorithm developed and used throughout this project. A possibility in this sense is offered by the allocation of retransmissions which in the version presented and used in this project consist in placing them sequentially and blindly at the beginning of the bandwidth. A better mechanism, which would still preserve unfragmented the bandwidth for first transmission users while at the same time exploit the channel, would consist in allocating the retransmissions from both edges of the bandwidth choosing every time the retransmission user and the bandwidth edge characterized by the highest scheduling metric. This would result beneficial especially in case of high experienced BLER due to adverse channel conditions or simply to the type of service in use.

On the topic of service differentiation different directions are possible including a refinement of the proposed framework to better cope with the considered traffic types, the consideration of other QoS parameters (like delay) as well as the consideration of additional traffic types like VoIP.

The above proposal for future research are in some measure related as the handling of VoIP traffic inherently calls for the consideration of delay. On the other hand each of them can be seen as a complete topic for future research.

The refinement of the framework could consist in developing a more general strategy as well as scheduling metrics which could handle any mix of CBR and EMG traffics (including BE as special case of EMG). This would involve prioritizing the allocation of time-frequency resources in order to fulfill the QoS requirements of the admitted users in order to eliminate the outage, as well as dropping the users from lowest to highest guarantees in case of resource limitations.

As for the consideration of other QoS parameters and the consideration of other traffic types, it is important to highlight that this study has been done for delay tolerable traffic while it is important to develop scheduling metrics and AC strategies able to cope with delay sensitive traffic like VoIP which is one of the main interests to operators. This would involve modifying the metrics as to include the delay budget as an additional criterion to prioritize time and frequency resource allocation. Still within the topic of VoIP, it would be relevant to study semi-persistent scheduling techniques which are necessary when a very high number of VoIP users would require more control capacity than is available by using dynamic scheduling.

Appendix A

Semi-Static System Level Simulator Description

This appendix describes the models used in the semi-static simulator used for the performance assessment of the algorithms and framework built in the course of the Chapters 3, 4, 5, 6 and 7. Section A.1 contains the general description of the simulator including the network layout. Section A.2 gives a short description of how the link-to-system performance mapping is realized. Section A.3 contains a description of the different traffic models utilized while Section A.4 describes the Poisson process used to model the user arrival. Finally, A.5 defines the KPIs mostly used in this research project.

A.1 Semi-Static System Simulator

The semi-static system simulator consists of a detailed multi-cell deployment, based on the LTE guidelines provided in [20]. The framework consists of a hexagonal regular grid cellular setup, where the center three cells are surrounded by the two tiers of cells, as shown in Figure A.1. There are nineteen cell sites in the simulation area, each consisting of three sectors per site, giving a total of fifty seven sectors. The orientation of the main lobes of directional sector antenna elements is indicated by arrows in Figure A.1. In order to avoid the drawback of limited network layout the wrap around techniques is used. A 3-sector network topology with 70 degrees half power beam width eNode-B is assumed for the Macro cell deployment. The propagation modeling consists of the path-loss, shadowing and fast fading. The path loss model for the Macro cell case includes a 20 dB outdoor-to-indoor penetration loss. Fast fading is simulated according to the Typical Urban power delay profile for user speed of 3 kmph [58], which is a tapped delay line implementation with uncorrelated Rayleigh fading paths [59]. The user creation/arrival process is different depending on the traffic model deployed.

- In case of full buffer traffic model the users are created at the beginning of each simulation run and remain in the system until the end of the run.
- In case of finite buffer traffic model and no AC, the users are created at the beginning of the simulation. A user is killed and replaced by a new user once the buffer is empty.
- In case of finite buffer traffic model combined with AC, the users are created in the system according to a Poisson call arrival process (described in Section A.4). If the proposed AC decision criterion is fulfilled, the user is admitted otherwise the user is rejected or blocked.

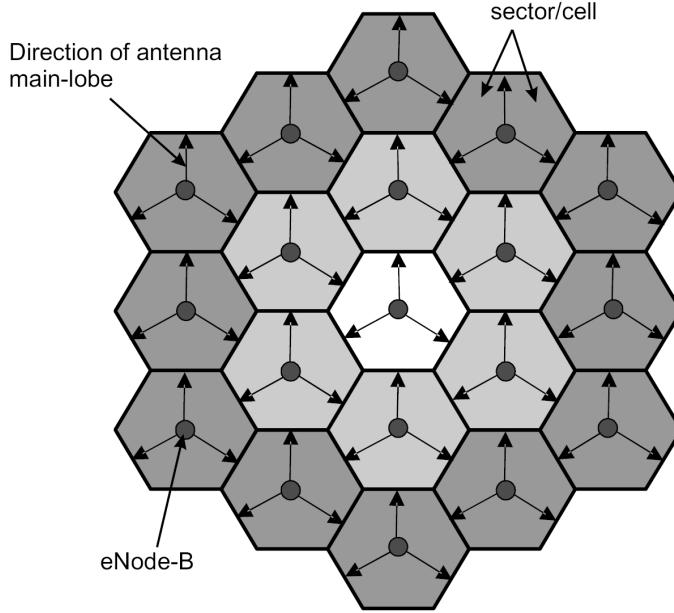


Figure A.1: The hexagonal regular grid cellular setup used in the system simulator according to the LTE guidelines [20].

Additionally the load is said to be balanced if the users are created uniformly within each cell and each cell contains the same number of users while it is unbalanced if the users are uniformly distributed within the whole network so that different cells may contain a different number of users.

During a packet call the path-loss and shadowing components are assumed to be constant for each user, while the fast fading is time-varying. Shadowing is fully correlated between cells of the same site, while the correlation is 0.5 between sites. The serving cell for a new user is selected according to the lowest total path-loss including distance dependent path-loss, shadowing, and effective antenna gains. The RRM functionalities such as LA, PS, and HARQ are accurately modeled for all the sectors. The LA is modeled as fast AMC based on CSI. To maintain the BLER target in the first transmission OLLA offset is used to bias the CSI before using it for MCS selection. The PS is modeled as decoupled TD and FD scheduler as shown in Figure 2.5. The total number of PRBs used for the data transmission is 48, while 2 PRBs are reserved for control signaling transmission. The default simulation assumptions and parameters are listed in Table A.1.

A.2 Link-to-System Performance Mapping

In order to estimate the performance at the system-level with reasonable accuracy, an evaluation based on extensive simulations under a variety of scenarios is crucial. A single simulator approach would be preferable, but the complexity of such a simulator including everything from link-level processing to multi-cell network is too high for the required simulation resolution [60]. Therefore, separate link-level and system-level simulators are needed. The link and system levels are connected through a link-to-system performance mapping function, which is used to predict the instantaneous BLER at system-level without performing detailed link-level processing steps. This function is estimated using link-level simulations, and it takes into account factors such as MCS format, receiver type, and channel state [30]. The desired characteristics of the link-to-system mapping function are that it should be general enough to cover different multiple access strategies and transceiver types, including different antenna techniques. Further, it should be possible to de-

Table A.1: Main simulations parameters

Parameter	Setting
Cellular layout	Hexagonal grid, 19 cell sites, 3 cells/sectors per site
Inter-site distance	500 m (Macro case 1) 1732 m (Macro case 3)
System bandwidth	10 MHz
Number of sub-carriers per PRB	12
Number of PRBs	50 (180 kHz per PRB). 2 of the 50 PRBs are used for control channel
Sub-frame/TTI duration	1 ms
Maximum User transmit power	24 dBm (250 mW)
Distance dependent path loss	$128.1 + 37.6 \log_{10}(d [km])$
Penetration loss	20 dB
Log-normal shadowing	standard deviation = 8 dB correlation distance = 50 m correlation between sectors = 1.0 correlation between sites = 0.5
Minimum distance between UE and cell	35 m
Power delay profile	TU3, 20 taps[58]
CSI log-normal error standard deviation	1 dB
CSI resolution	2 dB
OLLA offset	0.5 dB
OLLA offset range	[-4.0, 4.0] dB
Control channel overhead	3/14 symbols
Link adaptation	Fast AMC
Modulation/code rate settings	QPSK [R = 1/10, 1/6, 1/4, 1/3, 1/2, 2/3, 3/4], 16-QAM [R=1/2, 2/3, 3/4, 5/6]
HARQ model	Synchronous and adaptive with ideal chase combining
Max. No. of HARQ transmission attempts	4
Ack/Nack delay	2 ms
Channel estimation	Ideal
Carrier frequency	2 GHz
eNode-B antenna gain	14 dBi
UE antenna gain	0 dBi
UE noise figure	9 dB
UE speed	3 kmph
UE receiver	2-Rx Maximal Ratio Combining (MRC)
Frequency re-use factor	1

rive the parameters of the model from a limited number of link-level evaluations. In this study, the AVI method is used for the link-to-system mapping [61, 11]. The AVI tables constructed from an extensive link-level simulations are used to map the average received SINR to the corresponding BLEP. The target BLER is given in input to the OLLA algorithm to determine the offset to be applied to the input SINR.

A.3 Traffic Models

A.3.1 Best Effort

In this study the BE traffic models services for which no priority or guaranteed QoS is given. As a consequence best effort users will obtain variable bit rate depending on the current traffic load. If coupled with a full (infinite) buffer model then users experiencing different channel conditions and consequently different data rates will upload a different amount of data in the same simulation time. If coupled with a finite buffer model then each user will upload the same amount of data and the session is terminated once the upload is completed. In this case users at the cell edge will stay longer in the system due to lower data rates compared to users at the cell center.

A.3.2 Constant Bit Rate

The CBR streaming traffic model is used to represent a realistic GBR service. The CBR traffic is modeled as one ON period of an ON/OFF traffic source. An ON/OFF traffic source, e.g. Voice over Internet Protocol (VoIP), is characterized by the ON periods where the data packets are generated at a certain inter-packet arrival time followed by the OFF periods at a certain interval as illustrated in Figure A.2. The source activity factor is defined as sum of ON periods over the call duration, that is, the sum of ON and OFF periods.

For each CBR user a fixed amount of data packets are generated at the user with a constant packet size and constant inter-arrival time. In case the system is able to fulfill the GBR requirement of CBR services, the session time of each CBR user will be same as the streaming duration irrespective of the user location. This means that if the scheduling and AC are working effectively the CBR users will fulfill their GBR requirements.

A.4 Poisson process

A Poisson process can be described by the counter process $N(t)$. The counter tells the number of arrivals in the interval $(0, t)$ or, more generally, in the interval (t_1, t_2) .

$$\begin{cases} N(t) = \text{number of arrivals in the interval } (0, t) \\ N(t_1, t_2) = N(t_2) - N(t_1) = \text{number of arrivals in the interval } (t_1, t_2) \end{cases}$$

The number of arrivals $N(t)$ in a finite interval of length t obeys the Poisson distribution:

$$P\{N(t) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

Moreover, the number of arrivals $N(t_1, t_2)$ and $N(t_3, t_4)$ in non-overlapping intervals ($t_1 \leq t_2 \leq t_3 \leq t_4$) are independent.

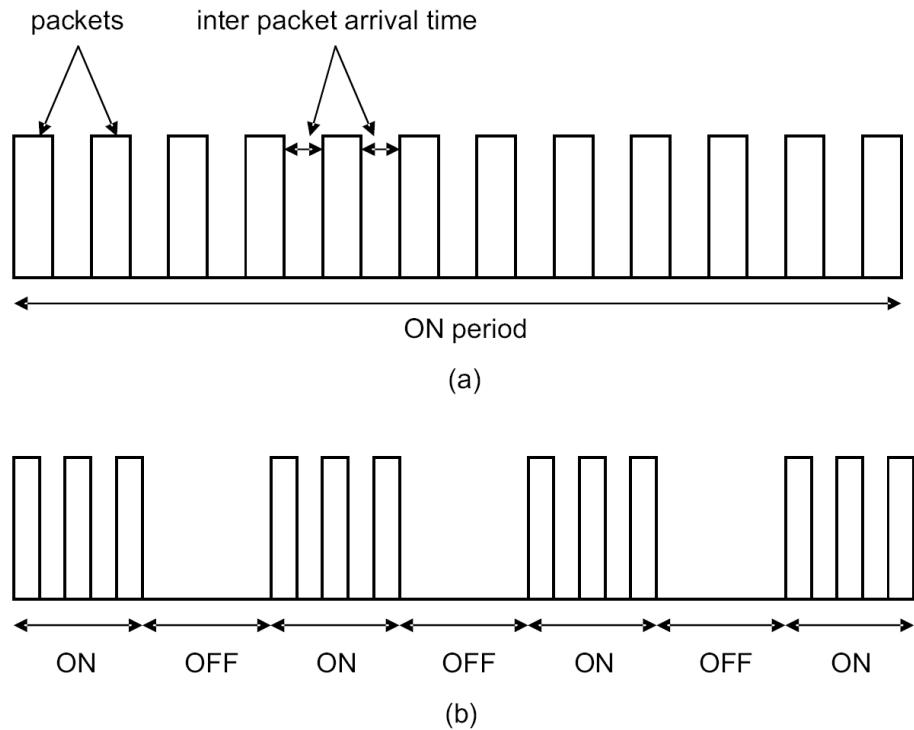


Figure A.2: (a) CBR streaming traffic with only ON spurt, (b) ON/OFF traffic model.

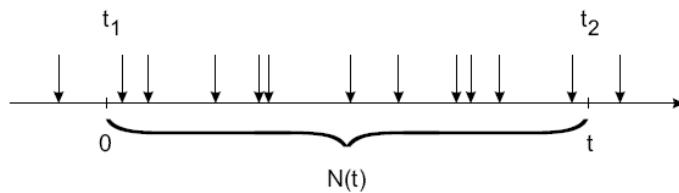


Figure A.3: Poisson arrival process.

A.5 Key Performance Indicators

The main performance indicators used in this system-level study to evaluate the performance of PC, scheduling and AC are:

- The average cell throughput is defined as:

$$\bar{T}_{cell} = \frac{\sum_i^{N_{cells}} b_i}{N_{cells} \cdot t_{sim}}$$

where N_{cells} is the number of cells in the system, b_i is the number of correctly received bits during the simulation time (t_{sim}).

- The user throughput at 5% outage (more precisely indicated as the 5th percentile value of the user throughput distribution) is calculated via linear interpolation as:

$$v = v_k + \frac{N}{100}(p - p_k)(v_{k+1} - v_k)$$

where N represents the number of user throughput values, p the 5th percentile, k an integer such that $p_k \leq p \leq p_{k+1}$ and v the value corresponding to the percentile p .

- The blocking probability P_b is simply defined as the ratio of the number of blocked users to the number of new users requesting admission.
- The outage probability P_o is defined as the ratio of the number of users not fulfilling their GBR requirements to the total number of users admitted.
- The unsatisfied user probability P_u is defined as:

$$P_u = 1 - (1 - P_b)(1 - P_o)$$

A.6 Acknowledgment

The system simulator was developed in collaboration with other colleagues in research group as well as in partnership with Nokia Siemens Networks.

Appendix B

SC-FDMA link level performance and assumed propagation conditions

B.1 Introduction

This chapter discusses the performance of the SC-FDMA at link level and the propagation conditions assumed for the generation of the results. They are addressed respectively in Section B.2 and B.3.

B.2 SC-FDMA Link Level Performance

The AVIs used in the system level simulator are produced according to the latest LTE physical layer specifications for Frequency Division Duplex (FDD). The antennas have a Single Input Single Output (SISO) 1x1 configuration. Real channel estimation (via pilots) is assumed. The receiver type used is Linear Minimum Mean Square Error (LMMSE).

Figure B.1 represents the AVIs used in the system level simulator for a case of 2 PRBs. On the x-axis the SINR averaged over fast fading is used (as the fast fading depends on the radio channel and the user speed, different AVIs for different channels are generated). As expected the BLER decreases as the SINR increases for all the considered MCSs and a higher MCS requires a larger SINR in order to keep the BLER constant.

Figure B.2 represents the required SINR to keep a 30% BLER target as function of the bandwidth. The slight dependency on the bandwidth is due to the following three factors: the code block size (resulting in higher coding gain as the bandwidth increases), the frequency diversity (which improves as the bandwidth increases) and the accuracy of channel estimation (which worsens as the bandwidth increases compensating the other effect).

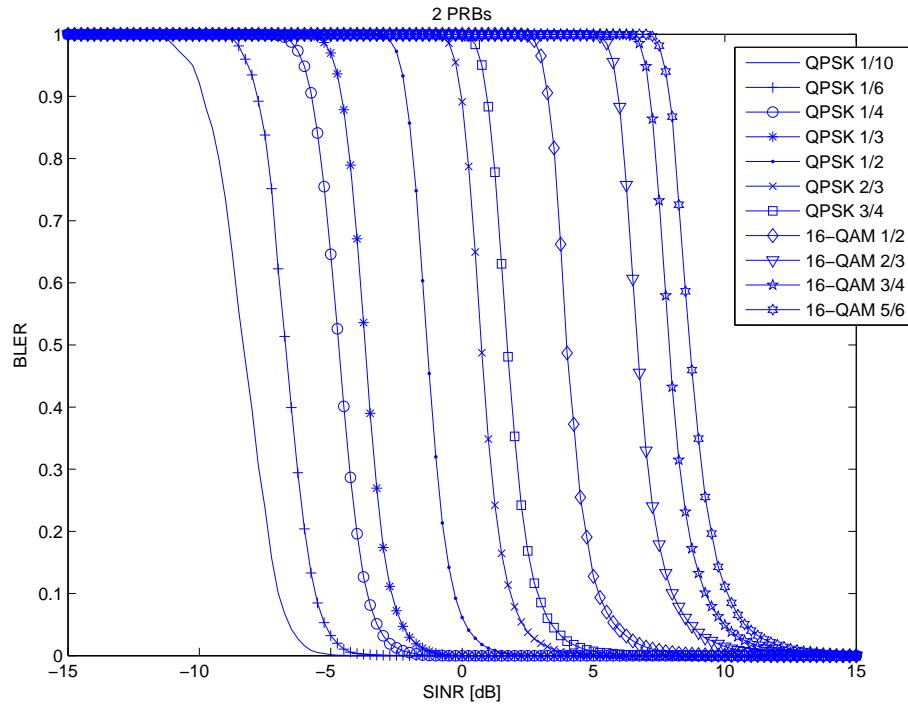


Figure B.1: AVI for 2 PRBs.

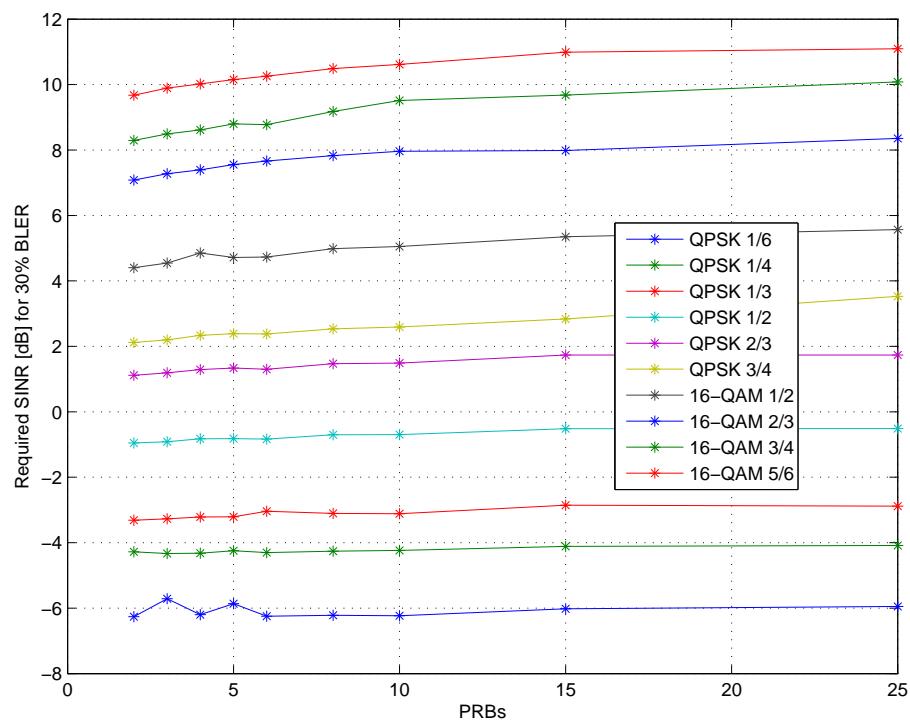


Figure B.2: Required SINR to have 30% BLER depending on number of PRBs for different MCSs.

B.3 Propagation Conditions

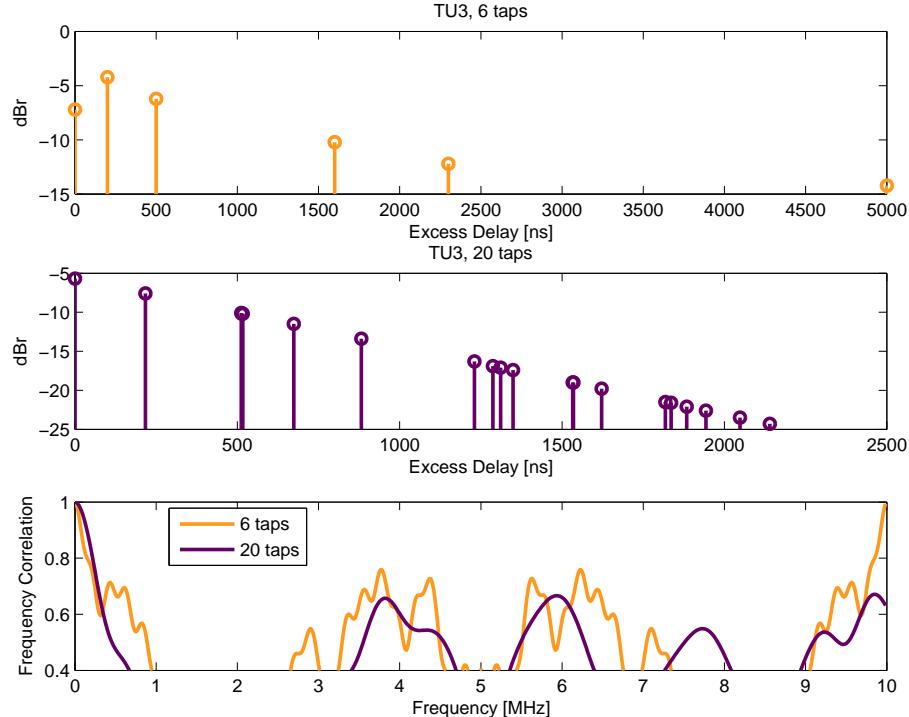


Figure B.3: Power Delay Profile and Frequency Correlation for TU 3 kmph, 6 taps and 20 taps.

The Power Delay Profile (PDP) depicts the relative power and the time delay of the different reflections. The channel profile used throughout this study is the TU 3 kmph with 6 taps. The TU 3 kmph with 20 taps is also recommended in the specifications to obtain an accurate frequency domain correlation function (and such profile was used, for example, to generate the fast fading file of ELIISE, the first simulator developed). The time domain responses of the channels are represented by the PDPs given in Figure B.3. For the TU3 with 6 taps the values are also listed in Table B.1.

Table B.1: Typical Urban Channel Power Delay Profile, 6 taps [37].

Tap number	Delay (μs)	Power (dBm)
1	0	-7.22
2	0.195	-4.22
3	0.488	-6.22
4	1.595	-10.22
5	2.311	-12.22
6	5.013	-14.22

The time dispersive properties of the multipath channel are characterized by maximum delay spread (τ_{max}), mean excess delay (τ_{mean}) and the r.m.s. delay spread (τ_{rms}).

τ_{max} represents the maximum time interval during which reflections of significant energy are received.

τ_{mean} is defined as:

$$\tau_{mean} = \sum_{i=1}^N \tau_i \cdot |h(t, \tau_i)|^2$$

where N is the number of taps, while τ_{rms} is defined as:

$$\tau_{rms} = \sqrt{\sum_{i=1}^N (\tau_i - \tau_{mean})^2 |h(t, \tau_i)|^2} = \sqrt{\tau_{mean}^2 - (\tau_{mean})^2}$$

where τ_{mean}^2 is the second moment and $(\tau_{mean})^2$ is the mean squared.

The values of such parameters are listed in Table B.2.

Table B.2: Main parameters for TU3, 6 taps.

Parameter	Value
τ_{max}	$5 \mu s$
τ_{mean}	$0.6704 \mu s$
τ_{rms}	$1.8715 \mu s$
$B_c \approx \frac{1}{\tau_{max}}$ [62]	$200 kHz$
B_c at a correlation level of 0.6 (from inspection of frequency correlation plot)	$32 kHz$

The value of τ_{max} of $5 ms$ is comparable with the duration of the cyclic prefix (which is a guard time consisting of the last part of the useful OFDM symbol being copied at the beginning of the same symbol in order to protect the transmission from ISI). The frequency selectivity of the CSI in this study has been set to two PRBs ($360 kHz$) which is also comparable with the coherence bandwidth of the channel. The TU3 profiles with 6 and 20 taps show similar correlation properties up to a bandwidth of approximately $320 kHz$ (corresponding to a correlation of 0.6), after which they diverge. Such correlation properties are of particular importance for the evaluation of wideband system concepts with frequency dependent characteristics like frequency domain link adaptation and packet scheduling as their performance is closely related to it[59].

Appendix C

Statistical Significance Assessment and Convergence

C.1 Introduction

This chapter assesses the statistical significance of the KPIs obtained from the semi-static simulator described in Appendix A.1 by using standard statistical methods. The chapter is organized as follows: The modeling assumptions including the list of selected simulation scenarios is outlined in Section C.2. The results and discussion are presented in Section C.3.

C.2 Modeling Assumptions

The statistical significance analysis is performed by running a certain number of simulations (50 simulations) with identical parameter setup but different seed for the random number generator. The variation in the KPIs is investigated by means of the box and whiskers diagram (or box plot) [63]. The following KPIs, which have been used throughout the thesis, are considered:

- Average cell throughput
- User throughput at 5% outage

The investigation of statistical significance is carried out for the case of full buffer as this is the scenario under which the number of calls is directly related to the number of runs. This is useful considering that the statistics investigated are call statistics and therefore their accuracy depends on the number of completed calls. The other main parameters are indicated in Table C.1.

C.3 Results

Figure C.1 shows the box and whiskers diagrams for the average cell throughput and outage user throughput for the FTB strategy and Macro case 1. The KPIs are normalized to the sample mean i.e. the mean value obtained from all the simulations. In the box and whiskers diagram the box is drawn from the lower hinge defined as the 25% percentile, to the upper hinge corresponding to the 75% percentile. The median value is shown as a line across the box. The length of the box gives the inter-quartile range, while the whiskers on each side of the box is extended to the most extreme data value within 1.5 times of the inter-quartile range. Data values lying beyond the ends of the whiskers are marked as outliers.

Table C.1: Main simulations parameters

Parameter	Setting
# UEs in TD	8
# UEs in FD	8
Allocation strategy	FTB, ATB
TD scheduling	not relevant
FD scheduling	PF and PF-SINR
Propagation scenario	Macro case 1 (ISD of 500 m); Macro case 3 (ISD of 1732 m)
Traffic model	Full buffer

In Figure C.1 the deviation in average cell throughput from its corresponding sample mean decreases with the number of calls and it is contained within $\pm 0.7\%$ in case of 20 runs corresponding to 9120 calls. The user throughput at 5% outage is necessarily less stable and shows a larger deviation from the mean which is anyway contained in the range between -2% and $+3\%$ for the case of 20 runs. Such values are sufficiently accurate for the gains exhibited throughout, moreover most of the results in the thesis have been obtained with at least 30 runs and 10 users per cell (which results in a significantly larger number of calls) and, as expected, a higher stability.

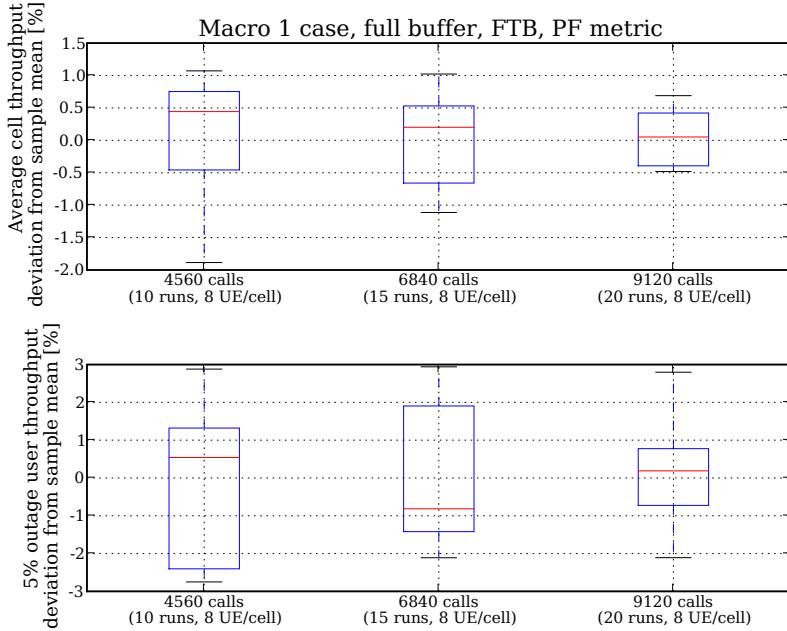


Figure C.1: Box plots of the average cell throughput and outage user throughput for the FTB strategy, Macro 1 case.

Figure C.2 shows the same performance indicators and their variation for the same setup but considering the ATB allocation strategy. For a lower number of calls the ATB shows a larger instability which is anyway contained within a range similar to the FTB in case of 20 runs. With 30 runs only a small improvement is noticeable.

Figure C.3 shows the same results for the Macro case 3 but in this case also the cases of 30 and 40 runs, the latter being the most common setup in the presented results, is considered. For the 30 runs case, corresponding to 13680 calls, the outage shows a box plot whose values are contained

between +4% and -3%. The ATB allocation strategy shows again a larger instability, as shown in Figure C.4, but the adoption of the PF-SINR metric, which is the default assumption for the largest part of simulations using ATB, increases the stability. In the 40 runs case the outage is contained between +4% and -4%.

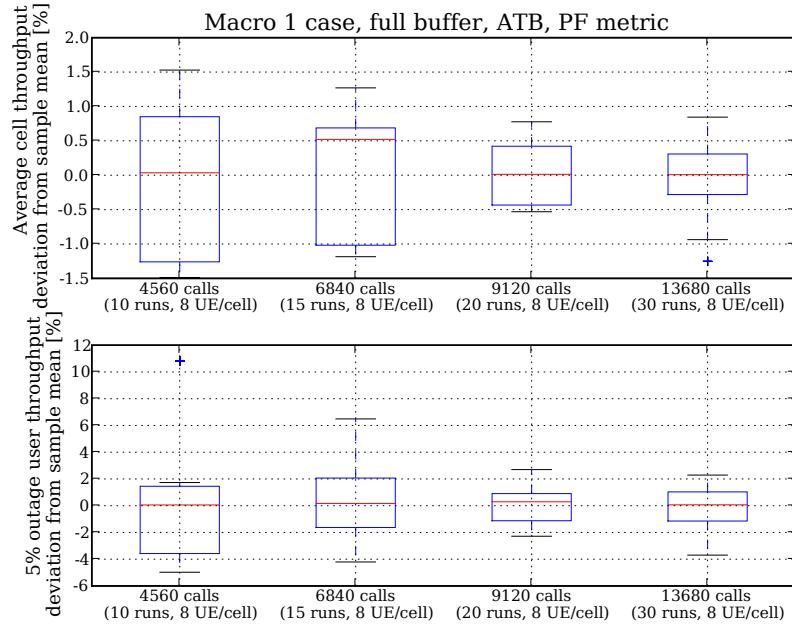


Figure C.2: Box plots of the average cell throughput and outage user throughput for the ATB strategy, Macro 1 case.

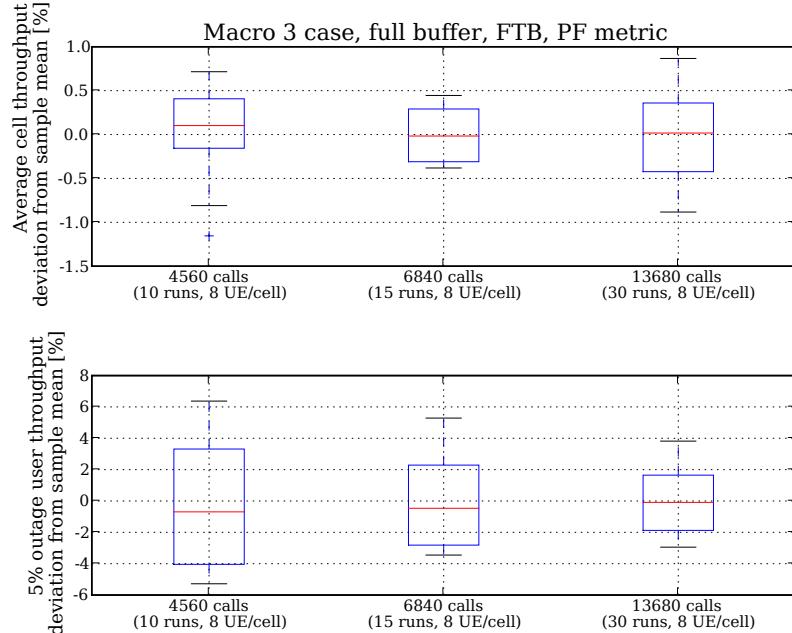
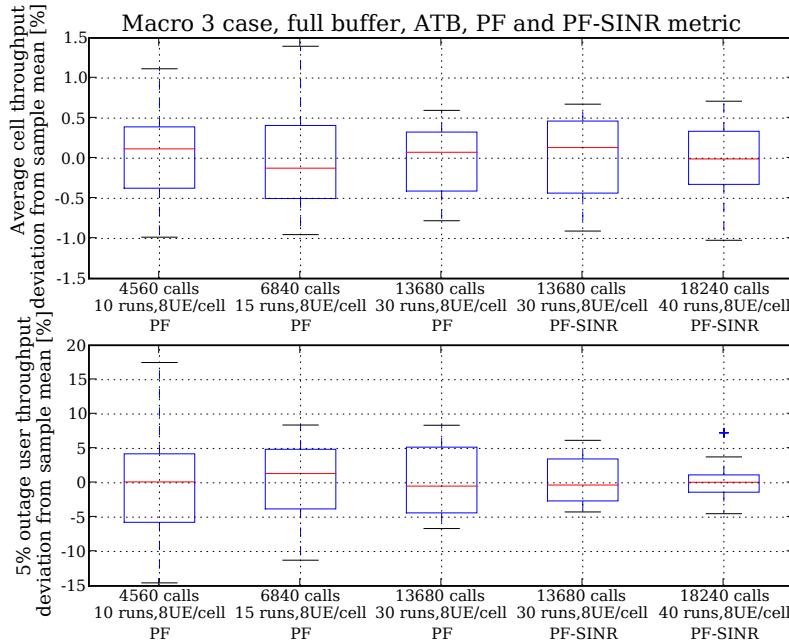


Figure C.3: Box plots of the average cell throughput and outage user throughput for the FTB strategy, Macro 3 case.

Table C.2: Main simulations parameters

Parameter	Case I	Case II	Case III
User arrival rate	8 UE/s/cell (Poisson arrival)	8 UE/s/cell (Poisson arrival)	8 UE/cell (constant)
Network load distribution	Unbalanced	Unbalanced	Balanced
AC strategy	Channel-aware AC	No AC	No AC
Traffic model	Finite buffer (1 Mbps)	Finite buffer (1 Mbps)	Finite buffer (1 Mbps)

**Figure C.4:** Box plots of the average cell throughput and outage user throughput for the ATB strategy, Macro 3 case.

C.4 Convergence to Steady-State: Assumptions and Results

In the following Section we are going to show some results regarding the convergence of the path-gain of users in the network to a stable state. Such results are relevant as they have been used as criterion to decide the warm-up time of the simulator. For this reason we are going to use three different scenarios (in terms of user distribution and user arrival) which have been used throughout the thesis and are indicated in Table C.2. The other parameters are as indicated in Table C.1.

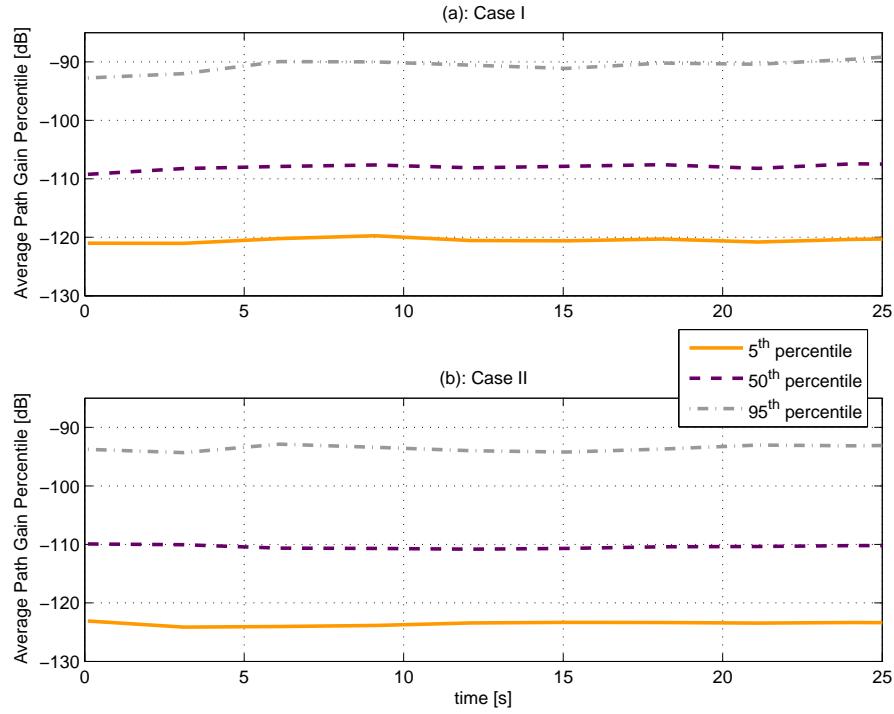


Figure C.5: Time-evolution of the 5th, 50th and 95th percentiles of the path-gain of the users in the network. (a): Case I, (b): Case II. Macro 1 case.

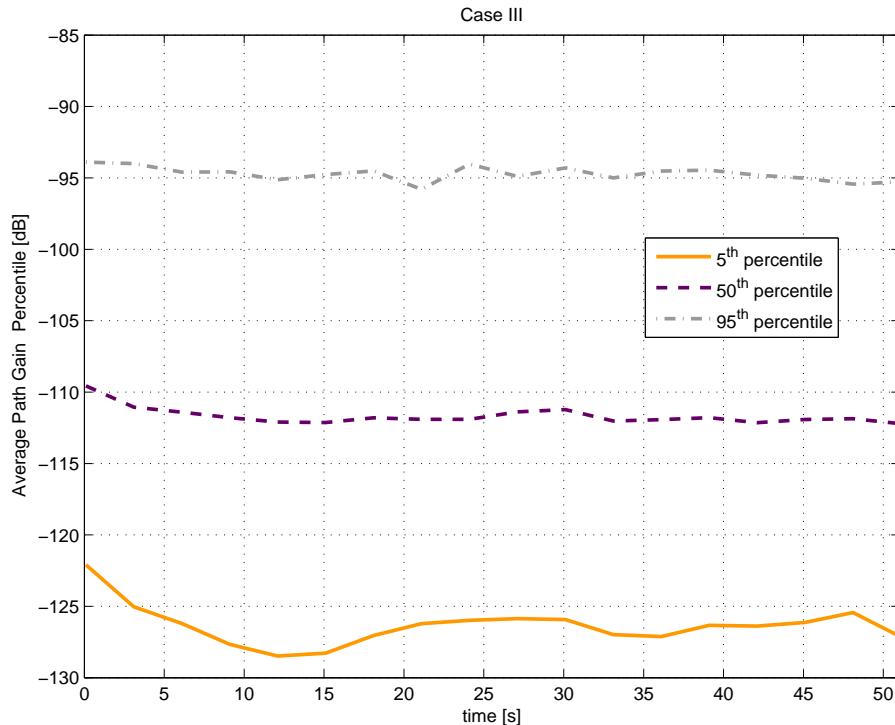


Figure C.6: Time-evolution of the 5th, 50th and 95th percentiles of the path-gain of the users in the network for the Case III. Macro 1 case.

Figure C.5 shows the evolution of the chosen percentiles vs time for the cases I and II according

to Table C.2 and assuming a null warm-up. As expected the adoption of AC in Case I leads to higher path-gain values than in Case II. In both cases the percentiles of the path-gain is very stable since the beginning. The situation is considerably different in Case III where the users do no enter the network following a Poisson user arrival process but are created at the beginning of the simulation and then migrate towards the cell-edge. This happens because their time to empty the buffer is different depending on the location within the cell (and therefore the channel quality). It is more difficult in this case to identify a precise point in time after which the 5th percentile can be said to be stable. A warmup time of 10 s has been chosen because considered long enough to stabilize the path-gain in all the specified cases.

Bibliography

- [1] GSMA, “GSM Association (GSMA) - www.gsmworld.com,” April 16th 2008.
- [2] H. Holma and A. Toskala, Eds., *WCDMA for UMTS – HSPA Evolution and LTE*. John Wiley & Sons Ltd, 2007.
- [3] 3GPP Technical Report 25.913, version 7.3.0, *Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN)*, March 2006.
- [4] B. Classon, P. Sartori, V. Nangia, X. Zhuang, and K. Baum, “Multi-dimensional Adaptation and Multi-user Scheduling Techniques for Wireless OFDM systems,” in *Proceedings of the IEEE International Conference on Communications (ICC)*, vol. 3, Orlando, Florida, USA, May 2003, pp. 2251–2255.
- [5] 3GPP Technical Specifications 36.101, version 8.2.0, *Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Radio Transmission and Reception (Release 8)*, May 2008.
- [6] 3GPP Technical Specifications 36.104, version 8.2.0, *Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) Radio Transmission and Reception (Release 8)*, May 2008.
- [7] A. Toskala, H. Holma, K. Pajukoski, and E. Tiirola, “UTRAN Long Term Evolution in 3GPP,” in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Helsinki, Finland, 2006.
- [8] B. Classon, K. Baum, V. Nangia, R. Love, Y. Sun, R. Nory, K. Stewart, A. Ghosh, R. Ratasuk, W. Xiao, and J. Tan, “Overview of UMTS Air-Interface Evolution,” in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Montreal, Canada, October 2006.
- [9] H. Yang, “A Road to Future Broadband Wireless Access: MIMO-OFDM-Based Air Interface,” *IEEE Communications Magazine*, vol. 43, no. 1, pp. 53–60, January 2005.
- [10] M. Rumney, “Demystifying Single Carrier FDMA,” White Paper, Agilent Technologies, April 2008.
- [11] C. Rosa, “Enhanced Uplink Packet Access in WCDMA,” Ph.D. dissertation, Aalborg University – Department of Communication Technology, Aalborg, Denmark, December 2004.
- [12] A. J. Goldsmith and S.-G. Chua, “Adaptive Coded Modulation for Fading Channels,” *IEEE Transactions on Communications*, vol. 46, no. 5, pp. 595–602, May 1998.
- [13] X. Qiu and K. Chawla, “On the Performance of Adaptive Modulation in Cellular Systems,” *IEEE Transactions on Wireless Communications*, vol. 47, no. 6, pp. 884–895, June 1999.
- [14] K. L. Baum, T. A. Kostas, P. J. Sartori, and B. K. Classon, “Performance Characteristics of Cellular Systems With Different Link Adaptation Strategies,” *IEEE Transactions on Vehicular Technology*, vol. 52, no. 6, pp. 1497–1507, November 2003.
- [15] C. Rosa, D. L. Villa, C. U. Castellanos, F. D. Calabrese, P. H. Michaelsen, K. I. Pedersen, and P. Skov, “Performance of Fast AMC in E-UTRAN Uplink,” in *Proceedings of the IEEE International Conference on Communications (ICC)*, Beijing, China, May 2008, pp. 4973–4977.
- [16] H. Holma and A. Toskala, Eds., *WCDMA for UMTS – Radio Access for Third Generation Mobile Communications*, 3rd ed. John Wiley & Sons Inc, 2004.
- [17] K. B. Letaief and Y. J. Zhang, “Dynamic Multiuser Resource Allocation and Adaptation for Wireless Systems,” *IEEE Wireless Communications*, vol. 13, no. 4, pp. 38–47, August 2006.
- [18] A. Jalali, R. Padovani, and R. Pankaj, “Data Throughput of CDMA-HDR High Efficiency-High Data Rate Personal Communication Wireless System,” in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, vol. 3, Tokyo, Japan, May 2000, pp. 1854–1858.
- [19] P. Ameigeiras, “Packet Scheduling and Quality of Service in HSDPA,” Ph.D. dissertation, Aalborg University – Department of Communication Technology, Aalborg, Denmark, December 2003.

- [20] 3GPP Technical Report 25.814, version 7.1.0, *Physical Layer Aspects for Evolved UTRA*, October 2006.
- [21] H. Holma and A. Toskala, Eds., *HSDPA/HSUPA for UMTS – High Speed Radio Access for Mobile Communications*. John Wiley & Sons Ltd, 2006.
- [22] Orange, China Mobile, KPN, NTT DoCoMo, Sprint, T-Mobile, Vodafone, Telecom Italia, “LTE physical layer framework for performance verification,” *3GPP TSG-RAN1#48, RI-070674*, February 2007.
- [23] 3GPP Technical Specification 36.213, version 8.4.0, *Physical layer procedures*, September 2008.
- [24] M. Anas, “Uplink Radio Resource Management for QoS Provisioning in Long Term Evolution With Emphasis on Admission Control and Handover,” Ph.D. dissertation, Aalborg University – Department of Electronic Systems, Aalborg, Denmark, January 2009.
- [25] 3GPP Technical Specification 36.3, version 8.3.0, *Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2*, January 2008.
- [26] 3GPP Technical Specification 23.401, version 8.4.0, *General Packet Radio Service (GPRS) Enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access*, December.
- [27] 3GPP Technical Specification 36.211, version 8.4.0, *Physical Channels and Modulation*, September 2008.
- [28] 3GPP Technical Report 25.813, version 7.0.0, *E-UTRA and E-UTRAN; Radio Interface Protocol Aspects*, June 2006.
- [29] A. Sampath, P. S. Kumar, and J. M. Holtzman, “On Setting Reverse Link Target SIR in a CDMA System,” in *Proceedings of IEEE Vehicular Technology Conference (VTC)*, Phoenix, AZ, USA, May 1997.
- [30] K. Brueninghaus, D. Astley, T. Salzer, S. Visuri, A. Alexiou, S. Karger, and G.-A. Seraji, “Link Performance Models for System Level Simulations of Broadband Radio Access Systems,” in *Proceedings of the IEEE Personal Indoor and Mobile Radio Communications Conference (PIMRC)*, vol. 4, Berlin, Germany, September 2005, pp. 2306–2311.
- [31] 3GPP Technical Specification 36.321, version 8.4.0, *E-UTRA Medium Access Control (MAC) protocol specification*, December 2008.
- [32] M. Boussif, N. Quintero, F. D. Calabrese, C. Rosa, and J. Wigard, “Interference Based Power Control Performance in LTE Uplink,” in *Proceedings of the IEEE International Symposium on Wireless Communication Systems (ISWCS)*, Reykjavik, Iceland, October 2008.
- [33] C. U. Castellanos, D. L. Villa, C. Rosa, K. I. Pedersen, F. D. Calabrese, P. H. Michaelsen, and J. Michel, “Performance of Uplink Fractional Power Control in UTRAN LTE,” in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Singapore, May 2008, pp. 2517–2521.
- [34] W. Xiao, R. Ratasuk, A. Ghosh, R. Love, Y. Sun, and R. Nory, “Uplink Power Control, Interference Coordination and Resource Allocation for 3GPP E-UTRA,” in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Montreal, Quebec, September 2006.
- [35] A. Rao, “Reverse Link Power Control for Managing Inter-cell Interference in Orthogonal Multiple Access Systems,” in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Baltimore, Maryland, September 2007.
- [36] W. Zhengdao, M. Xiaoli, and G. Giannakis, “OFDM or single-carrier block transmissions,” *IEEE Transactions Communications*, vol. 52, no. 3, pp. 380–394, March 2004.
- [37] B. E. Priyanto, “Air Interfaces of Beyond 3G systems with User Equipment Hardware Imperfections: Performance and Requirements Aspects,” Ph.D. dissertation, Aalborg University – Department of Electronic Systems, Aalborg, Denmark, February 2008.
- [38] F. D. Calabrese, P. H. Michaelsen, C. Rosa, M. Anas, C. U. Castellanos, D. L. Villa, K. I. Pedersen, and P. E. Mogensen, “Search-Tree Based Uplink Channel Aware Packet Scheduling for UTRAN LTE,” in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Singapore, May 2008, pp. 1949–1953.
- [39] C. Wengerter, J. Ohlhorst, and A. G. E. v. Elbwart, “Fairness and Throughput Analysis for Generalized Proportional Fair Frequency Scheduling in OFDMA,” in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, vol. 3, Stockholm, Sweden, May 2005, pp. 1903–1907.
- [40] A. Pokhariyal, T. E. Kolding, and P. E. Mogensen, “Performance of Downlink Frequency Domain Packet Scheduling for the UTRAN Long Term Evolution,” in *Proceedings of IEEE Personal Indoor and Mobile Radio Communications Conference (PIMRC)*, Helsinki, Finland, September 2006, pp. 1–5.
- [41] J. Lim, H. Myung, K. Oh, and D. Goodman, “Proportional Fair Scheduling of Uplink Single-Carrier FDMA Systems,” in *Personal, Indoor and Mobile Radio Communications (PIMRC)*, September 2006.

- [42] M. Al-Rawi, R. Jäntti, J. Torsner, and M. Sågfors, "Opportunistic Uplink Scheduling for 3G LTE Systems," in *4th International Conference on Innovations in Information Technology*, November 2007.
- [43] F. D. Calabrese, C. Rosa, M. Anas, P. H. Michaelsen, K. I. Pedersen, and P. E. Mogensen, "Adaptive Transmission Bandwidth Based Packet Scheduling for LTE Uplink," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Calgary, Canada, September 2008.
- [44] G. Monghal, K. Pedersen, I. Kovács, and P. Mogensen, "QoS Oriented Time and Frequency Domain Packet Schedulers for the UTRAN Long Term Evolution," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Singapore, May 2008.
- [45] F. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 33–37, January 1997.
- [46] J. M. Holtzman, "CDMA Forward Link Water Filling Power Control," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, vol. 3, Tokyo, Japan, May 2000, pp. 1663–1667.
- [47] ———, "Asymptotic Analysis of Proportional Fair Algorithm," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 2, San Diego, California, USA, September 2001, pp. F33–F37.
- [48] A. Pokhariyal, K. I. Pedersen, G. Monghal, I. Z. Kovacs, C. Rosa, T. E. Kolding, and P. E. Mogensen, "HARQ Aware Frequency Domain Packet Scheduler with Different Degrees of Fairness for the UTRAN Long Term Evolution," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Dublin, Ireland, April 2007, pp. 2761–2765.
- [49] 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," in *submitted to IEEE Vehicular Technology Conference (VTC)*, Anchorage, Alaska, September 2009.
- [50] M. Anas, C. Rosa, F. D. Calabrese, P. H. Michaelsen, K. I. Pedersen, and P. E. Mogensen, "QoS-Aware Single Cell Admission Control for UTRAN LTE Uplink," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Singapore, Singapore, May 2008.
- [51] M. Anas, C. Rosa, F. D. Calabrese, K. I. Pedersen, and P. E. Mogensen, "Combined Admission Control and Scheduling for QoS Differentiation in LTE Uplink," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Calgary, Canada, September 2008.
- [52] D. Niyato and E. Hossain, "Service Differentiation in Broadband Wireless Access Networks with Scheduling and Connection Admission Control: A Unified Analysis," *IEEE Transactions on Wireless Communications*, vol. 6, no. 1, pp. 293–301, January 2007.
- [53] H.-W. Lee and S. Chong, "Combined Packet Scheduling and Call Admission Control with Minimum Throughput Guarantee in Wireless Networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 8, pp. 3080–3089, August 2007.
- [54] K. B. Johnsson and D. C. Cox, "An Adaptive Cross-Layer Scheduler for Improved QoS Support of Multiclass Data Services of Wireless Systems," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 334–343, February 2005.
- [55] D. Laselva, J. Steiner, F. Khokhar, T. E. Kolding, and J. Wigard, "Optimization of QoS-Aware Packet Schedulers in Multi-Service Scenarios over HSDPA," in *IEEE International Symposium of Wireless Communication Systems*, Trondheim, Norway, October 2007.
- [56] T. E. Kolding, "QoS-Aware Proportional Fair Packet Scheduling with Required Activity Detection," in *Proceedings of IEEE Vehicular Technology Conference (VTC)*, Montreal, Canada, September 2006.
- [57] M. Lundevall, B. Olin, J. Olsson, N. Wiberg, S. Wänstedt, J. Eriksson, and F. Eng, "Streaming Applications over HSDPA in Mixed Service Scenarios," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, vol. 2, Los Angeles, California, USA, September 2004, pp. 841–845.
- [58] 3GPP Technical Report 25.943, version 5.1.0, *Deployment Aspects (Release 5)*, June 2002.
- [59] T. B. Sørensen, P. E. Mogensen, and F. Frederiksen, "Extension of the ITU Channel Models for Wideband (OFDM) Systems," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, vol. 1, Dallas, Texas, USA, September 2005, pp. 392–396.
- [60] H. Holma and A. Toskala, Eds., *WCDMA for UMTS – Radio Access for Third Generation Mobile Communications*, 2nd ed. John Wiley & Sons, 2002.
- [61] S. Hämäläinen, P. Slanina, M. Hartman, A. Lappetäinen, H. Holma, and O. Salonaho, "A Novel Interface between Link and System Level Simulations," in *Proceedings of ACTS Summit*, Aalborg, Denmark, October 1997, pp. 599–604.
- [62] B. Sklar, "Rayleigh fading channels in mobile digital communication systems," *IEEE Communications Magazine*, pp. 90–109, July 1997.
- [63] R. L. Mason, R. F. Gunst, and J. L. Hess, *Statistical Design and Analysis of Experiments - With Applications to Engineering and Science*. John Wiley & Sons, 2003.

