Dynamic resource allocation for cellular networks with interference

Tania Villa Trapala

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2013-ENST-0054

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Wireless Communications »

présentée et soutenue publiquement par

Tania VILLA TRAPALA

le 26 Septembre 2013

Gestion dynamique de ressources appliquée aux réseaux cellulaires avec interférence

Directeur de thèse : **Raymond Knopp**

Jury

M. Angel LOZANO, Prof., Universitat Pompeu Fabra	Rapporteur
M. Tommy SVENSSON, Prof., Chalmers University of Technology	Rapporteur
M. David GESBERT, Prof., EURECOM	Examinateur
M. Ruben MERZ, Dr., Swisscom	Examinateur
M. Antonio CIPRIANO, Dr., Thales Communications & Security	Examinateur

TELECOM ParisTech
école de l'Institut Télécom - membre de ParisTech

THÈSE

DISSERTATION

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
from TELECOM ParisTech

Specialization: Wireless Communications

Tania Villa Trapala

Dynamic resource allocation for cellular networks with interference

Defense scheduled on the 26th of September 2013 before a committee composed of:

Reviewers	Prof. Angel Lozano, Universitat Pompeu Fabra
	Prof. Tommy Svensson, Chalmers University of Technology
Examiner	Prof. David Gesbert, EURECOM
	Dr. Ruben Merz, Swisscom
	Dr. Antonio Cipriano, Thales Communications & Security
Thesis Supervisor	Prof. Raymond Knopp, EURECOM

THESE

présentée pour obtenir le grade de

Docteur de TELECOM ParisTech

Spécialité: Wireless Communications

Tania Villa Trapala

Gestion dynamique de ressources appliquée aux réseaux cellulaires avec interférence

Soutenance de thèse prévue le 26 Septembre 2013 devant le jury composé de :

Rapporteurs	Prof. Angel Lozano, Universitat Pompeu Fabra
	Prof. Tommy Svensson, Chalmers University of Technology
Examinateur	Prof. David Gesbert, EURECOM
	Dr. Ruben Merz, Swisscom
	Dr. Antonio Cipriano, Thales Communications & Security
Directeur de thèse	Prof. Raymond Knopp, EURECOM

Abstract

Mobile networks have experienced dramatic growth during the past decades. Modern communication systems require high data rates and better quality of service control for services such as voice telephony, online gaming, web browsing, etc. The main obstacle found in wireless communication networks is the time-varying nature of the physical channel. Taking this into account, the goal of a system designer is to simplify the overall network design and optimize the performance.

The aim of this thesis is to design, implement and evaluate practical cross-layer algorithms to handle interference and allocate the radio resources in an efficient way for LTE and post-LTE uncoordinated networks. We develop mathematical and computational interference models that allow us to understand the behavior of such networks and we apply an information-theoretic approach to different interference scenarios and traffic characteristics. We have tried to remain as close as possible to practical systems to be able to test the feasibility of the proposed techniques.

As a part of the evolution to 4G systems, the introduction of small-cells that overlay the existing cellular network has been envisioned to fill in the coverage white spots or serve mobile and outdoor users where the cellular network is not deployed.

This thesis deals with performance evaluation of interference scenarios in 4G networks, in particular those arising from small-cell deployments. The work in this thesis also deals with analysis of resource-allocation and incremental-redundancy based hybrid automatic repeat request (HARQ) with bursty interference (or more general time-varying channels) which allows for only partial channel state information at the transmitter. The work is then applied to practical scheduler design for LTE base stations and includes performance analysis for real LTE modems.

We showed that, in general, by adapting the number of physical dimensions across rounds, we can exploit the interference mitigation effects of HARQ using it not only to recover from errors but for interference cancellation. We proposed efficient resource allocation algorithms to increase the throughput, which can potentially come very close to optimal performance.

Table of Contents

List of Figures

Acronyms

Here are the main acronyms used in this document. The meaning of an acronym is usually indicated once, when it first appears in the text.

3G	Third Generation
4G	Fourth Generation
3GPP	Third Generation Partnership Project
ABSF	Almost Blank Subframe
AMC	Adaptive Modulation and Coding
AWGN	Additive White Gaussian Noise
BLER	Block Error Rate
CDF	Cumulative Distribution Function
CP	Cyclic Prefix
CQI	Channel Quality Indicator
CSI	Channel State Information
DCI	Downlink Control Information
DL	Downlink
DRX	Discontinuous Reception
DSL	Digital Subscriber Line
EPA	Extended Pedestrian
eNB	eNodeB (LTE base station)
HeNB	Home-eNodeB (or femtocell)
HetNet	Heterogeneous Network
HARQ	Hybrid Automatic Repeat Request
iid	Independent and Identically Distributed
IoT	Internet of Things
IR	Incremental Redundancy
LTE	Long Term Evolution
M2M	Machine-to-Machine
MAC	Multiple Access Channel
MCS	Modulation and Coding Scheme
MIESM	Mutual Information Effective SINR Mapping
MIMO	Multiple Input Multiple Output
NDI	New Data Indicator

NLOS	Non-Line of Sight
OAI	OpenAirInterface
OFDMA	Orthogonal Frequency Division Multiple Access
PDU	Packet Data Unit
PDCCH	Physical Downlink Control Channel
PDSCH	Physical Downlink Shared Channel
PRB	Physical Resource Block
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Shared Channel
QoS	Quality of Service
QPSK	Quadrature Phase-Shift Keying
RB	Resource Block
RBG	Resource Block Group
RE	Resource Element
RLC	Radio Link Control
SC-FDMA	Single Carrier Frequency Division Multiple Access
SDR	Software-Defined Radio
SINR	Signal to Interference and Noise Ratio
SISO	Single Input Single Output
SNR	Signal to Noise Ratio
SRS	Sound Reference Signal
TBS	Transport Block Size
TTI	Transmission Time Interval
UE	User Equipment
UL	Uplink

Notations

Here is a list of the main operators and symbols used in this document. We have tried to keep the notation consistent throughout the thesis, but rarely symbols have different definitions in different chapters and in that case they are defined very explicitly to avoid any confusion.

x or X	Input signal
$h, \mathbf{H}$	Channel
$z, \mathbf{Z}$	Noise
σ^2	Variance of the noise
$y, \mathbf{Y}$	Received signal
$\|x\|$	Absolute value of scalar
$\mathbb{E}$	Palm Expectation
$E[.]$	Mathematical expectation
$\mathbb{M}_{max}$	Maximum number of HARQ rounds
$\mathcal{CN}$	Complex Normal Distribution
$\mathcal{H}(.)$	$= -E \log p(.)$ Entropy of the argument
$F_\chi(\chi)$	Cumulative distribution function of χ
$Q_M(a, b)$	Marcum Q-function with M degrees of freedom and paramters a and b
$\log_2$	All logarithms are to the base 2
$I(.;.)$	Mutual information between the two arguments
μ	Activity factor
$\bar{R}$	Spectral effiency or throughput
ρ	Ratio of physical dimensions used across HARQ rounds
$\Im\{\}$	Imaginary part
$\Phi_I(\omega)$	Characteristic function of the mutual information I
$\beta_{1,2}$	Calibration factors
S_1	$\|x_1\|$ i.e. cardinality of the constellation of x_1
S_2	$\|x_2\|$ i.e. cardinality of the constellation of x_2

Chapter 1

Introduction

Modern communication systems, like the Long Term Evolution (LTE) standard [6], require high data rates and better Quality of Service (QoS) control for services such as voice telephony, online gaming, web browsing, etc. In order to cope with the requirements of these new types of services while simultaneously offering high data rates, the standards have to evolve and adapt. Both the challenges from the Physical layer (PHY) and the QoS demands from the applications have to be taken into account. Future wireless communications networks require optimization of parameters in all layers. In a cross-layer design, rate, power and coding at the PHY can be adapted to meet the requirements of the applications given the current channel and network conditions (See figure 1.1).

By designing policies that combine interference mitigation at the PHY layer with scheduling algorithms and rate adaptation at the Medium Access Control layer (MAC), and further radio resource management at the Radio Link Control layer (RLC), we can obtain higher throughput, more bandwidth and therefore, more efficient networks.

As a part of the evolution to Fourth Generation (4G) systems, the introduction of small-cells that overlay the existing cellular network has been envisioned to fill in the coverage white spots or serve mobile and outdoor users where the cellular network is not deployed. However, since these small-cells may not be connected to the operator backhaul network, coordination between them for resource management is hardly feasible (see figure 1.2).

The aim of this thesis is to design, implement and evaluate practical cross-layer algorithms to handle interference and allocate the radio resources in an efficient way for LTE and post-LTE uncoordinated networks. We develop mathematical and computational interference models that allow us to

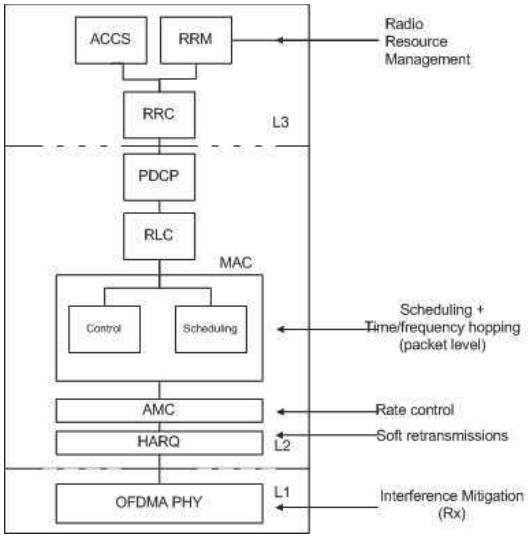


Figure 1.1: Cross-Layer Techniques.

understand the behavior of such networks. We first do a theoretical study by applying an information-theoretic approach to different interference scenarios and traffic characteristics. We have tried to remain as close as possible to practical systems to be able to test the feasibility of the proposed techniques. Secondly, we perform a full simulation study which ends in the implementation and evaluation of the proposed techniques in the OpenAirInterface platform (OAI) [3].

1.1 Contributions and Thesis Outline

The main obstacle found in wireless communication networks is the time-varying nature of the physical channel. Taking this into account, the goal of a system designer is to make the PHY/MAC smarter to simplify the overall network design and optimize the performance. In our strategies, we have tried to take into account scenarios and parameters that make them applicable to practical systems.

In order to better understand the relevance of our study, we begin in

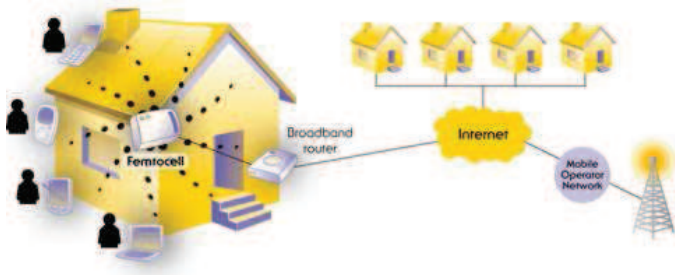


Figure 1.2: Femtocells are an example of small-cell deployments [4].

Chapter 2 by describing the evolution of wireless communication networks, together with the new interference scenarios resulting from this evolution. We then explain the fundamentals of LTE and interference networks and we also describe the basics of scheduling and link adaptation in LTE.

Chapter 3 deals with performance evaluation of interference scenarios in 4G networks, in particular those arising from femtocell deployments. In this chapter, we analyze the throughput of the network with the help of an inhomogeneous discrete-time Markov chain and basic information theory quantities. We consider one or two dominant interferers and we study a decentralized interference mitigation scheme that combines Hybrid Automatic Repeat and Request (HARQ) and Incremental Redundancy (IR) with an interference cancellation decoder. For comparison purposes, we also study an ARQ scheme where the information is not accumulated across transmission rounds. Our performance evaluation based on analytical modeling and Monte Carlo evaluation of throughput shows that our scheme is effective at combating interference without requiring any coordination. The results were published in

- Villa, Tania; Merz, Ruben; Knopp, Raymond, "**Interference management in femtocell networks with Hybrid-ARQ and interference cancellation**", in the proceedings of IEEE Asilomar Conference on Signals, Systems, and Computers, November 2011, Pacific Grove, CA, USA.

The fourth chapter starts by describing some of the key emerging applications where our information-theoretic analysis can be applied. This chapter is crucial for understanding the importance of dynamic resource allocation schemes in LTE. Unlike any previous work, we consider the allocation of physical resources not fixed across HARQ transmissions. The latter is a real possibility in schedulers for LTE base stations (eNodeB or eNB) and, to the best of our knowledge, no well-known methodology exists for adapting physical resources across HARQ rounds when subject to time-varying channels either due to fading or time-varying interference or a combination of both. By doing this, we exploit the interference mitigation effects of HARQ using it not only to recover from errors but for interference cancellation and we propose efficient resource allocation algorithms to increase the throughput, which can potentially come very close to optimal performance.

We first present an analysis of interference-free networks with time-varying channels. Rather than performing extensive simulations, we take an information theoretic approach to derive analytical expressions that represent the long-term throughput of a point-to-point link and consider practical cases where there is a constraint on the outage probability representing the latency of the protocol. We consider Gaussian input signals and we consider

cases where Channel Quality Indicator (CQI) feedback is either unavailable, or outdated.

The fourth chapter then examines the case of networks with interference. We motivate the use of inter-round resource allocation through a simple but illustrative analysis with Gaussian signals and interference. We include the use of activity factors which model sporadic interference patterns characteristic of future heterogeneous networking deployments, in particular the interference seen from small-cell base stations with bursty traffic in the receiver of a macrocell user.

Finally, we look at practical interference scenarios to illustrate the applications of our analytical framework. We model a Manhattan-like topology which represents a block of apartments with femtocells creating interference. We then model a macro-femto scenario where a macrocell is overlaid by a femtocell and we look at the downlink (DL) channel of the macrocell user when the interference is coming from the interfering femtocell. With the use of an activity factor we model the fact that the femtocell is not active all the time. Finally, we explain the procedure that has to be followed to perform PHY abstraction with the use of our analytical framework, given the importance to accurately model the link performance in order to speed-up simulations. The results were published in

- Villa, Tania; Merz, Ruben; Knopp, Raymond, "**Adaptive modulation and coding with Hybrid-ARQ for latency-constrained networks**", in the proceedings of IEEE European Wireless Conference (EW2012), April 2012, Poznan, Poland.

- Villa, Tania; Merz, Ruben; Knopp, Raymond, "**Adaptive transmission and mutiple-access for sparse-traffic sources**", in the proceedings of IEEE European Signal Processing Conference (EUSIPCO), August 2012, Bucharest Romania.

- Villa, Tania; Knopp, Raymond; Merz, Ruben, "**Dynamic resource allocation in heterogeneous networks**", in the proceedings of IEEE Global Communications Conference (GLOBECOM), December 2013, Atlanta, USA.

and has been accepted for publication in

- Villa, Tania; Knopp, Raymond; Merz, Ruben, "**Dynamic resource allocation for time-varying channels in next generation cellular networks, Part I: a mathematical framework**", submitted to IEEE Transactions on wireless communications

and will be submitted as one part of

- Villa, Tania; Knopp, Raymond; Merz, Ruben, "**Dynamic resource allocation for time-varying channels in next generation cellular networks, Part II: applications in LTE**", under preparation.

Chapter 5 deals with practical scheduler design for LTE base stations. LTE offers a lot of flexibility in terms of resource allocation and, in particular, resource allocation algorithms can be tailored for a particular class of traffic with specific requirements. Nevertheless, work is still needed to exploit this flexibility efficiently for key emerging applications. In this chapter, we study the performance of our dynamic resource allocation polices for the scheduling of IR-HARQ transmissions under the constraints of LTE coded-modulation. In this capther, we show the results of the implementation of the scheduler in the OAI software-defined radio (SDR) platform [3] in order to test the performance and compliance of our resource allocation strategies in LTE. We show that the implementation of such policies in a real system is feasible withouth requiring any coordination or complex and time-consuming optimization procedure. We show that our scheduling techniques work for different environments and most importantly, we show that the results are in agreement with the theoretical results presented in Chapter 4. The results will be submitted as one part of

- Villa, Tania; Knopp, Raymond; Merz, Ruben, "**Dynamic resource allocation for time-varying channels in next generation cellular networks, Part II: applications in LTE**", under preparation.

Chapter 2

Background

Mobile networks have experienced dramatic growth during the past decades. Third Generation systems (3G) are expanding the possibilities of information transfer and communication. Besides high data rate, 3G systems also envisioned providing better QoS control for a variety of applications, from voice telephony and gaming, to web browsing, e-mail, and streaming multimedia applications. With the roll-out of 3G technology in many countries, researchers and standardization organizations have increased their efforts to offer even higher data rates and more services to the users than 3G. The capability of high data rate transmission determines the kind of service that can be provided to users and the QoS they receive. Users now expect to have the same on-demand access to multimedia content from anywhere and while "on the move".

2.1 Evolution of Wireless Communication Systems

The evolution of wireless communication systems is mainly driven by the introduction of new services and the availability of more advanced technologies. Over the last couple of decades, cellular networks have grown exponentially and the demand for new and improved services has become an important issue for operators. There is a need for new technologies to alleviate the capacity limitations of the network and to maintain the QoS demanded by the users. This has motivated the development of new standards like the Third Generation Partnership Project's (3GPP) LTE standard [6] in order to provide higher data rates and an improved QoS in wireless networks.

The 3GPP LTE standard has chosen Orthogonal Frequency Division Multiple Access (OFDMA) as the underlying modulation technology [7]. It is

a spectrally efficient version of multi-carrier modulation, where the subcarriers are selected in such a way that they are all orthogonal to one another over the symbol duration. Physical layer DL transmissions are implemented using OFDMA while Uplink (UL) transmission uses Single Carrier Frequency Division Multiple Access (SC-FDMA). The main difference between both schemes is that in OFDMA data detection is performed in the frequency domain while in SC-FDMA it is done in the time domain averaging the noise over the entire bandwidth [52].

Despite being in the early days of roll-out, LTE has become the fastest developing system due to the fast speeds and high quality user experience that it offers. It has now been launched on all continents, by 156 operators in 67 countries. By 2013, there are 6.4 billion of mobile subscriptions globally [1]. Current commercial LTE deployments are based on 3GPP Release 8 and 9 [5].

Data traffic will continue to grow, along with mobile data subscriptions and an increase in the average data volume per subscription. In fact, overall mobile data traffic is expected to continue the trend of doubling each year [1]. This growth in traffic and services will bring new technical challenges to the operators, interference being one of the most performance-limiting.

The demand for new mobile services and for higher peak bit rates and system capacity is tackled with the evolution of the technology to 4G. The 4G systems are based on 3GPP LTE and are progressing on a large scale, with 55 million users as of 2012 and from 1.6 to 2 billion users anticipated in 2018 [1].

2.2 Interference Scenarios in 4G Networks

During the past 20 years, there has been a massive growth in traffic volume, number of devices connected, and an increased demand for video data. Future cellular networks should be able to cope with this increased demand and handle all the traffic in an efficient way.

There are new technical challenges, and potential interference scenarios that vary with the type of deployments, requirements, high data rate and QoS levels. These new interference scenarios have been considered during the LTE Release 10 standardization:

- Macro-picocell interference (see figure 2.1).
- Macro-Home-eNodeB (HeNB) interference (see figure 2.2).

Enhancements in the operation of the base stations, advanced terminal receivers and a new carrier type with reduced transmission of “always-on” signals are a requirement for high network energy efficiency [28]. With the transmission of control signaling, a non-negligible amount of energy is used by the power amplifier. Minimizing the transmission of “always-on” signals

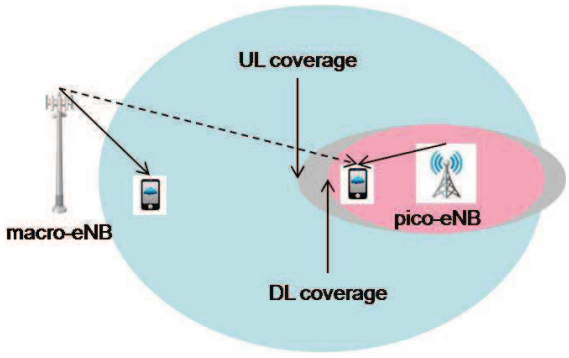


Figure 2.1: Macro-eNB vs pico-eNB.

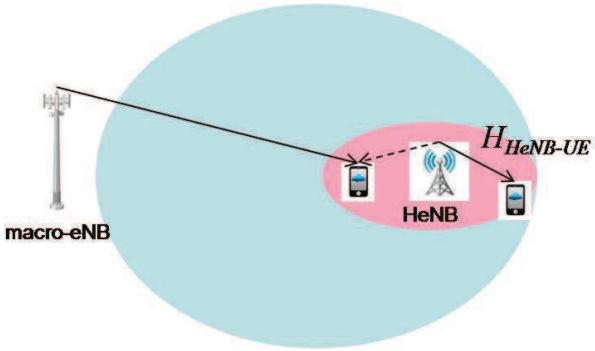


Figure 2.2: Macro-eNB vs HeNB.

allows the base station to turn off circuitry if it has no data to transmit. By eliminating unnecessary transmissions, the interference is reduced, leading to improved data rates in the network. Another aspect to consider is that with the use of small-cells, sporadic traffic, and a reduction in the transmission of control signals, interference becomes time-varying. When designing new techniques, these new characteristics and requirements have to be taken into account.

2.2.1 Heterogeneous Networks and Interference

As the traffic demand grows and the RF environment changes, the network relies on cell splitting or additional carriers to overcome capacity and link budget limitations and maintain uniform user experience. Moreover, site acquisition for macro base stations with towers becomes more difficult in dense urban areas. A more flexible deployment model is needed for operators to improve broadband user experience in a ubiquitous and cost-effective way.

The concept of Heterogeneous Networks (HetNets) has been introduced in LTE to address the capacity and coverage challenges resulting from the enormous and continuous growth of data services. The traditional macro network is deployed to provide umbrella coverage and smaller nodes are added as an underlay network to alleviate coverage holes and traffic hot zones. These

low-power nodes provide very high traffic capacity and user throughput locally, for ex. indoor and outdoor hotspot positions. The macro layer ensures service availability and QoS over the entire coverage area. HetNets include microcells, picocells, femtocells, and distributed antenna systems (remote radio heads), which are distinguished by their transmit powers, coverage areas, physical size, backhaul and propagation characteristics [33].

Energy-efficient load-balancing can be achieved by turning off the low-power nodes when there is no ongoing data transmission. A macro base-station can use Almost Blank Subframes (ABSF) to reserve some subframes for small-cells [42].

2.2.2 Femtocells

A femtocell, or HeNB, is a low power, low cost wireless access point installed by the end user to improve voice and data performance. They are deployed on top of the macro network and its primary purpose is to provide enhanced capacity for busy outdoor areas and improved coverage for indoor areas (see figure 2.3). HeNBs connect mobile devices to the operator network through the broadband connection of the user. Since a femtocell can operate using the licensed spectrum of the operator, it can allow operators to guarantee an acceptable level of QoS to the user, both in terms of coverage and capacity, avoiding the need for additional network interfaces. Since HeNBs require only low transmission power, battery life of mobile phones increases and the interference decreases.

The installation for such devices is planned to be plug and play, so the user does not have to worry about cumbersome installation, and configuration. Updates and maintenance to the software and configuration will be done transparently and controlled by operators.

2.2.3 Machine-to-Machine Communications

Over the last years, there has been a dramatic growth in communication between devices. Their traffic can be characterized as small, delay-tolerant data packets which are sent infrequently [61]. Applications of Machine-to-Machine communications (M2M) include those related to public safety, surveillance cameras, sensoring, monitoring, etc. M2M communications imply some technical and non-technical challenges to the network such as:

- Allowing for very low-cost of device types
- Achieving low device energy consumption to ensure long battery life for relevant applications
- Providing extended coverage options in challenging locations
- Handling a very large number of devices per cell

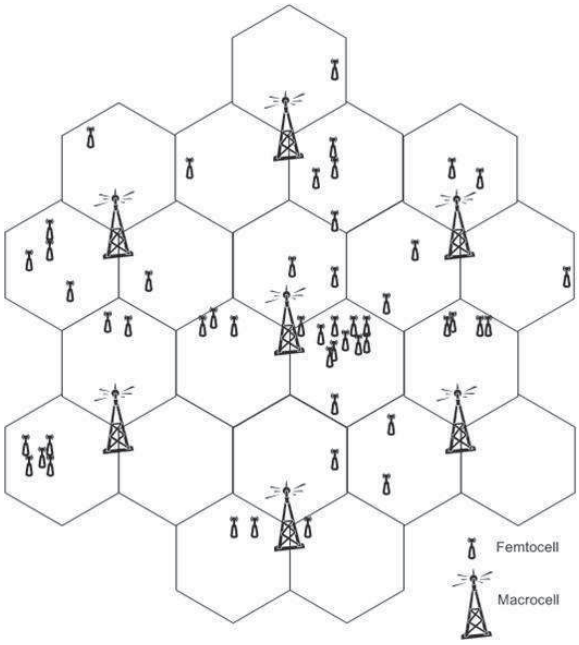


Figure 2.3: Macrocell network overlaid by HeNBs.

With the introduction of a large amount of communicating devices throughout the network, interference is increased. As a mean to control inter-cell interference, scheduling algorithms can be used together with link adaptation to adapt to a changing-channel and also to increase the throughput.

2.3 Scheduling and Link Adaptation in LTE

In LTE, the scheduling algorithms and rate adaptation at the MAC layer can be combined with radio resource management at the RLC layer to obtain higher throughput, more bandwidth and therefore, more efficient networks. Implementing efficient algorithms for radio resources management, packet scheduling, admission control or power and interference control are important to optimize the capacity and performance.

Scheduling consists on allocating the transmission resources, Physical Resource Blocks (PRBs) in LTE, to the users, every transmission opportunity. Given the variations experienced in the quality of a wireless channel, the choice of other parameters such as Modulation and Coding Scheme (MCS) can be adapted with the goal of maximizing the capacity in the cell, while satisfying the QoS requirements of every user. In this way, the randomness of the radio link can be taken into account and exploited to use the resources in the most efficient way. The scheduler interacts closely with the HARQ manager who is responsible for the scheduling of the retransmissions in case

of an incorrect reception. The LTE standard supports dynamic, channel-dependent scheduling to enhance overall system capacity.

In LTE, the scheduler resides at the eNB. Capacity is shared among multiple users on an on-demand basis. The purpose of the scheduler is to decide which terminal or base station transmit and on which set of resources.

Similar to OFDMA schedulers used on the DL, SC-FDMA schedulers for the UL can be both time and frequency-opportunistic. An important difference between DL and UL scheduling is that CQI reporting is not needed since the scheduler is located at the eNB which can measure UL channel quality through Sound Reference Signals (SRS) [61].

2.3.1 Resource Allocation in LTE

In LTE, the available bandwidth is divided into N subcarriers. From the N subcarriers, 12 or 24 adjacent subcarriers are grouped together forming what is called a Resource Block (RB), which represents the minimal scheduling resource for both UL and DL transmissions and it corresponds to 180 KHz of spectrum (see figure 2.4). LTE frames are divided into two slots of duration $T_{slot} = 0.5$ms. A slot is formed by N_{RB} RBs in the frequency domain for the duration of 6 or 7 OFDMA symbols in the time domain, depending on the length of the Cyclic Prefix (CP) used. The CP is used for synchronization purposes and is attached to each slot. A specific subcarrier inside the RB is called a Resource Element (RE). Since the subcarriers in OFDMA are orthogonal there is no interference from within the cell, but interference is experienced from the neighboring cells.

The number of RBs available depends on the bandwidth of the channel (see figure 2.5), and depending on the length of the CP, a different number of OFDMA symbols is accommodated in a slot. Table 2.1 gives the different number of RBs available for each of the channel bandwidths specified in the 3GPP standard with the corresponding number of RBs.

Table 2.1: N_{RB} vs Downlink System Bandwidth

Bandwidth	1.4 MHz	3 MHz	5 MHz	10 MHz	15 MHz	20MHz
N_{RB}	6	15	25	50	75	100

Radio resource management aims at scheduling the available resources in the best way to allow users to achieve a specific QoS. An intelligent mechanism has to consider the interference created with already assigned physical resources.

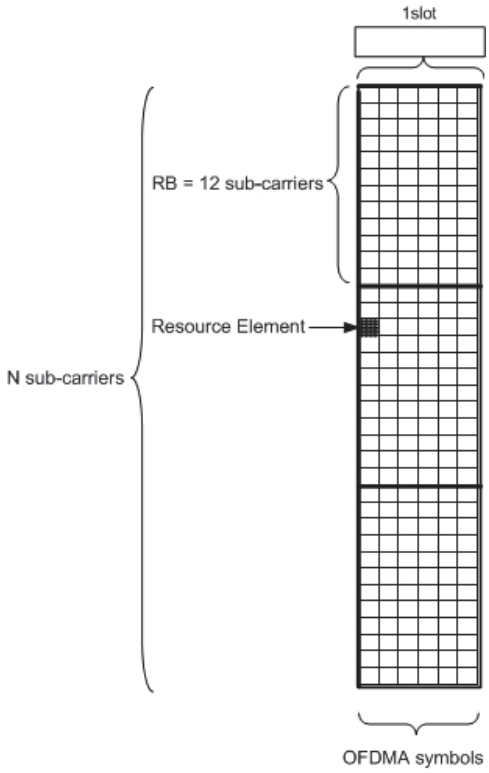


Figure 2.4: LTE resource grid [9].

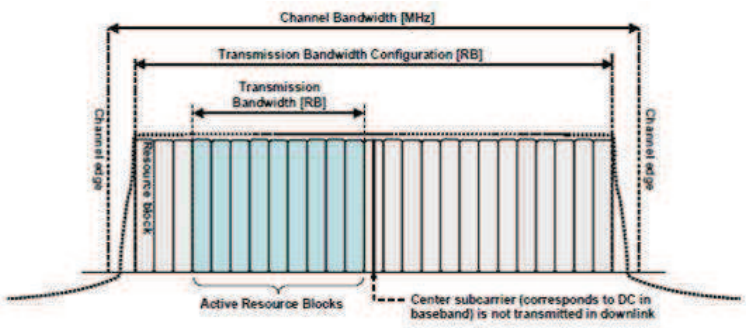


Figure 2.5: LTE bandwidth and resource blocks [28].

2.3.2 Discontinuous Reception (DRX)

Packet-data traffic is often highly bursty, with sporadic transmissions followed by inactivity periods. However, the control signals have to be monitored in order to receive UL grants or DL data transmissions and to adapt to the traffic variations. The latter consumes a non negligible amount of battery power. To reduce the power consumption, LTE introduces mechanisms for DRX by configuring a DRX cycle in the terminal which allows the terminal to monitor the control signaling only for a predetermined duration of time and turning off the transmission circuitry otherwise, allowing for savings in power consumption. HARQ retransmissions take place regardless of the DRX cycle [28]. By implementing DRX procedures, the network is able to trade-off between scheduling flexibility and power performance.

Chapter 3

Performance Evaluation of Small-cell Deployments

The current and growing expansion need of cellular networks represents a challenge as laying out and deploying new infrastructure is extremely expensive. Small-cells are believed to be a cost effective solution to expand the coverage and capacity of LTE and post-LTE networks [4].

Small-cells are low-power wireless access points that operate in licensed spectrum, used outdoor to enhance coverage, or indoor for enterprise or in-home usage. The concept of small-cells include femtocells, picocells and microcells.

In the case of in-home usage, femtocells provide high quality, high speed cellular access. They are deployed by end-users and connected to the operator network by a digital subscriber line (DSL), cable modem or optical fiber connection [24]. Because of the unplanned nature of femtocell deployments, they can suffer from severe inter-cell interference with neighboring femtocells in dense deployments [26, 54, 68]. In addition, coordination is hardly feasible due to delays induced by the backhaul infrastructure of these home femtocell networks.

In this chapter, we study a decentralized interference mitigation scheme that combines IR-HARQ with an interference cancellation decoder. Our performance evaluation based on analytical modeling and Monte Carlo experiments shows that our scheme is effective at combating interference without requiring any coordination.

3.1 Interference in Femtocell Deployments

We concentrate on LTE and LTE-advanced technologies (so-called 4G) with OFDMA physical layers and explore alternative strategies to mitigate interference. OFDMA ensures orthogonality of the subcarriers and therefore, there is no intra-cell interference. However, interference can be experienced from users in adjacent cells.

In LTE, HARQ is used to reduce errors in transmissions by retransmitting and combining the information when a frame is received with errors [27]. With IR, additional and new redundancy information is transmitted and combined with information already transmitted offering a coding gain. A frame is retransmitted until it is discarded or a maximum number of retransmissions is reached. From an information-theoretic perspective, mutual information is accumulated over retransmissions, increasing the probability to decode [20].

The throughput of HARQ has been investigated for Gaussian input signals [20] over a Gaussian channel with fading and in [70], rate adaptation is performed to maintain a fixed outage probability. In [29] the diversity-multiplexing-delay tradeoff has been studied for the Multiple-Input Multiple-Output (MIMO) ARQ channel.

Unlike the previous work, we take advantage of the non-Gaussian nature of interference in home femtocell deployments where there are typically only one or two strong dominant interferers [60]. We consider signals coming from discrete alphabets in order to be able to benefit from the structure of the interference. Gaussian signals achieve the maximum spectral efficiency. However, practical systems make use of small, finite-size input alphabets.

Under these assumptions, we propose a decentralized strategy that combines interference cancellation decoding [32] with an incremental redundancy HARQ policy [20].

To evaluate the performance of this strategy, we develop an analytical model of the throughput achieved by an HARQ protocol with an interference cancellation decoder. In particular, our model builds on an information-theoretic characterization of the achievable rate with interference cancellation.

3.1.1 System Model and Assumptions

We focus on a downlink scenario. Without loss of generality, we currently consider single antenna transmission. Furthermore, transmissions are slotted and perfectly synchronized. We have N_u transmitters, where node 0 is the transmitter of interest and the remaining $N_u - 1$ transmitters are interferers. We let d_k be the distance between the node k and the receiver. 4G systems are based on OFDMA physical layer [61]. We let $y[m]$ be the received signal in a particular RB at time m. In LTE, every RB is defined as a group of K

subcarriers. Within a given cell, we assume that RBs are orthogonal to each other. Hence, we can write the received signal $y[m]$ at time m as

$$y[m]=\sum_{j=0}^{K-1}\sum_{k=0}^{N_u-1}\sqrt{P_k d_k^{-\alpha}}h_{k,j}[m]\mu_{k,j}[m]x_{k,j}[m]+z_j[m]. \tag{3.1}$$

where $x_{k,j}[m]$ is the transmitted signal from node k in the jth subcarrier of a particular RB, $\mu_{k,j}[m]$ is a so-called *activity factor*, $z_j[m]$ is thermal noise, P_k is the transmission power, α is the path loss exponent and $h_{k,j}[m]$ is the channel coefficient. We model $z_j[m]$ as an independently and identically distributed (iid) zero-mean Additive White Gaussian Noise (AWGN) process with variance σ^2. If we concentrate on a particular subcarrier, the received signal in the jth subcarrier at time m is

$$y_j[m] = \sum_{k=0}^{N_u-1}\sqrt{P_k d_k^{-\alpha}}h_{k,j}[m]\mu_{k,j}x_{k,j}[m]+z_j[m] \tag{3.2}$$

The random variable $h_{k,j}[m]$ is iid for each slot with a Rayleigh distribution. Hence, the channel coefficient remains constant during the duration of a slot. The activity factor models the traffic load and/or Discontinuous Transmission (DTX) features of LTE systems [61]. We model $\mu_{k,j}$ with an iid Bernoulli distribution with parameter p. These features are taken into account as they have a direct effect on the interference distribution.

The retransmission protocol is an HARQ scheme using IR [20]. For comparison purpose, we also consider a simple ARQ scheme that retransmits the same data block in case of unsuccessful transmission. The parameter $\mathbb{M}_{max}$ is the maximum number of ARQ rounds. Therefore, a given frame can be retransmitted at most $\mathbb{M}_{max}$ times and is discarded if $\mathbb{M}_{max}$ is reached. We assume perfect Channel State Information (CSI) of the desired and interference signals at the receiver and we let R define the transmission rate given by a particular MCS.

In the next section, we present the throughput analysis for our HARQ protocol with an interference cancellation receiver.

3.1.2 Average Throughput Analysis of HARQ with Interference Cancellation

Without loss of generality, our analysis can consider a single subcarrier and we can drop the index j from (3.2). We consider a slotted (i.e. discrete-time) system where each slot corresponds to a frame transmission. We define the following symbols:

- $\bar{R}$ is the average throughput expressed in frames per second.
- $p[m]$ is the probability that a frame is successfully decoded at slot m.

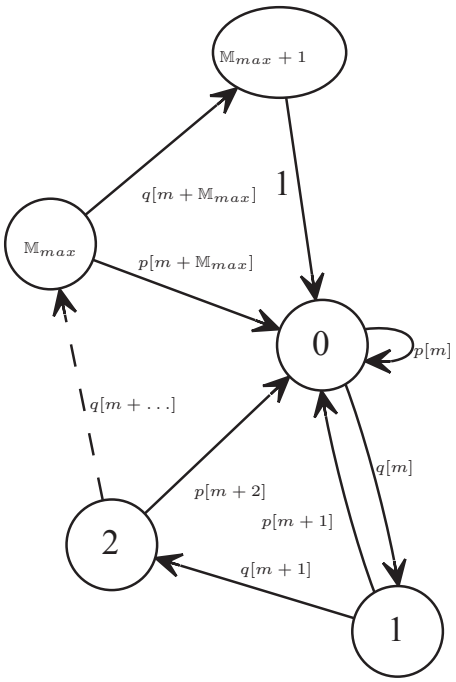


Figure 3.1: Retransmission Markov chain $\mathbb{X}_n$: a frame transmission attempt always initiates and finishes in state 0. A frame retransmission corresponds to a transition from state i to $i+1$. A successful frame transmission corresponds to a transition from any state $i = 0, \ldots, \mathbb{M}_{max}$ to the state 0. Finally, the frame is dropped if state $\mathbb{M}_{max}+1$ is reached. Note that $q[m] = 1 - p[m]$, $\forall m$.

The behavior of our system can be modeled by a discrete-time Markov chain [18]. However, because (1) a new channel coefficient $h_{k,j}[m]$ is possibly drawn at every slot, and (2) the activity factor can change the number of active sources at every slot, the Markov chain is inhomogeneous i.e. the state transition probabilities can change over time.

Remember from Section 3.1.1 that a frame can be retransmitted at most $\mathbb{M}_{max}$ times before being discarded. Hence, let $\mathbb{X}_n$ be the retransmission state of the source (see figure 3.1). The Markov chain $\mathbb{X}_n$ has $\mathbb{M}_{max}+2$ states (numbered from 0 to $\mathbb{M}_{max}+1$): a frame transmission attempt always initiates and finishes in state 0. A frame retransmission corresponds to a transition from state i to $i+1$. A successful frame transmission corresponds to a transition from any state $i = 0, \ldots, \mathbb{M}_{max}$ to the state 0. Finally, the frame is dropped if state $\mathbb{M}_{max}+1$ is reached. The transition probabilities

are the following:

$$\begin{array}{ll} p_{\mathbb{X}}(i,i+1) = 1 - p[m+i] = q[m+i], & i = 0,\ldots,\mathbb{M}_{max} \\ p_{\mathbb{X}}(i,0) = p[m+i], & i = 0,\ldots,\mathbb{M}_{max} \\ p_{\mathbb{X}}(\mathbb{M}_{max}+1,0) = 1 & \end{array} \tag{3.3}$$

where $p_{\mathbb{X}}(i,j) = \Pr(\mathbb{X}_{n+1} = j | \mathbb{X}_n = i)$. Each new frame transmission attempt[1] corresponds to a trip on the chain $\mathbb{X}_n$ starting in state 0 and returning back to the state 0. The average throughput $\bar{R}$ can be computed by dividing the average number of successful frame transmissions per trip by the average duration of a trip. For a trip from state 0 back to state 0, we can define two random variables:

- N_{s} is the number of successful frame transmissions per trip. Observe that N_{s} is equal to 0 or 1.
- T is the duration of a trip.

Now, following the approach in [51], we can write

$$\bar{R} = \frac{\mathbb{E}^0(N_{\mathrm{s}})}{\mathbb{E}^0(T)} \tag{3.4}$$

where $\mathbb{E}^0$ is a Palm expectation [12] (see [51, eq. 9] for a definition specific to the context of (3.4)). We do not have a closed-form expression for $\mathbb{E}^0(N_{\mathrm{s}})$ and $\mathbb{E}^0(T)$, but they can be evaluated by simulation. However, we need to be able to compute the transition probability $p[m+i]$ from state i to state 0 at time $m+i$.

We take an information-theoretic approach. Namely, without any interference cancellation, the transition probability is computed as a function of the instantaneous mutual information conditioned on the channel knowledge. For interference cancellation, we follow the mechanism and modeling of [32] and modify the mutual information expression appropriately. Let I_m denote the mutual information between the received and the desired signal at time m and let R denote the target operating rate. In the following, the conditioning on the channel realizations and number of active users is implicitly assumed. For IR-HARQ, it is shown in [20] that the mutual information is accumulated over retransmissions when IR is used. Following [70], we have

$$p[m+i] = \Pr\left(\sum_{k=m}^{m+i} I_k > R \,\middle|\, \sum_{k=m}^{m+i-1} I_k \right) \tag{3.5}$$

for $0 < i \leq \mathbb{M}_{max}$ and $p[m] = \Pr\left(I_m > R\right)$, for $i = 0$. For ARQ, we have

$$p[m+i] = \Pr(I_{m+i} > R) \tag{3.6}$$

[1] As opposed to a retransmission, which is one transition on the chain.

for $0 \leq i \leq \mathbb{M}_{max}$. To evaluate equations (3.5) and (3.6) we need to compute the mutual information between node 0 and the receiver. To model the effect of interference cancellation, we follow [32]. For a receiver that cancels two interferers, the mutual information is (we ignore the conditioning on the channel for simplicity)

$$I(Y; X_0|X_1, X_2) = \log \mathcal{S}_0 - \frac{1}{\mathcal{S}_0\mathcal{S}_1\mathcal{S}_2} \sum_{x_0} \sum_{x_1} \sum_{x_2} \int_y p(y|x_0) \times \log \frac{\sum_{x_2'} \sum_{x_1'} \sum_{x_0'} p(y|x_0', x_1', x_2')}{\sum_{x_2'} \sum_{x_1'} p(y|x_0, x_1', x_2')} dy.$$

where Y is the received signal, X_0 is the desired signal, X_1 and X_2 are the interference signals, and $\mathcal{S}_i$, $i = \{0, 1, 2\}$ is the size of the constellation for X_i. If the structure of the interference is known, i.e. the constellation, then the receiver can use it to cancel the interferers. Note the discrete input distribution used to more accurately model practical systems. Also, to take the remaining transmitters with index 3 to $N_u - 1$ into account, we perform a Gaussian approximation. With a receiver that does not cancel interference, we simply compute $I(Y; X_0)$.

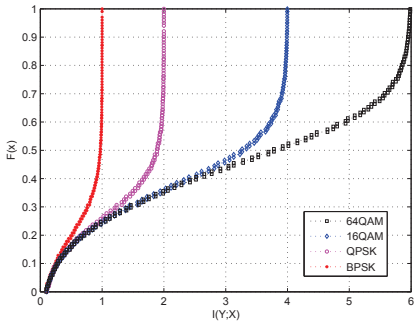
To visualize the mutual information with discrete signals, we plot in figure 3.2 the Cumulative Distribution Function (CDF) of the mutual information for the case when there is no interference and we compare it with the case of canceling two strong interferers. When interference is not taken into account, it is clear that higher modulation orders will translate into higher mutual information, but in an scenario with interference, there is not always a clear gain from using higher modulation orders, it will depend on the interference strength.

3.1.3 Performance of the HARQ Protocol in Femtocell Deployments with Interference

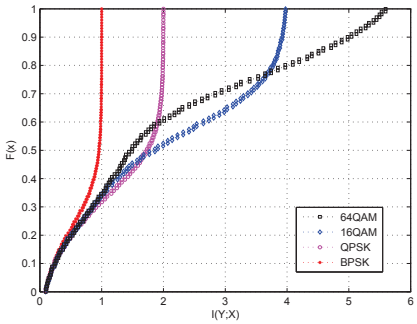
In this section, first we give results for a comparison between ARQ and HARQ. For a different number of retransmissions, we obtain the throughput with HARQ for two cases: without interference and with two strong interferers. Then, we study a Manhattan-like topology when the two strongest interferers can be canceled. Finally, we evaluate a scenario where interference is randomized by using an activity factor. This can be understood as the re-use factor of the network.

All results are obtained evaluating the throughput expressions in Section 3.1.2 by Monte Carlo experiments averaged over fading and noise distributions. The target rate R remains fixed for each simulation.

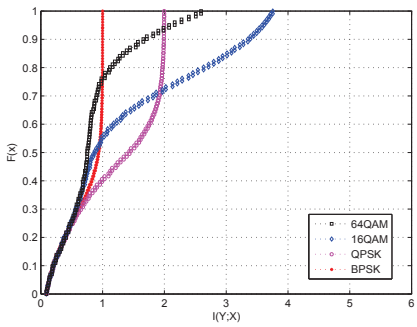
figure 3.3 compares the average throughput of the two ARQ protocols. We study a scenario with different modulations (BPSK, QPSK, 16QAM



(a) Mutual information CDF.



(b) Mutual information CDF, one interferer.



(c) Mutual information CDF, two interferers.

Figure 3.2: In (a) we have the CDF of the mutual information under Rayleigh fading and without any interference, (b) shows the CDF of the mutual information with one interferer, and (c) for two interferers. The interferers have the same power as the user of interest.

and 64QAM), without interference, and we set $\mathbb{M}_{max}$ to 3. The Signal-to-Noise Ratio (SNR) goes from -10 to 25 dB and the target rate is $R = 2$ bits/second. We clearly observe an improvement in throughput with HARQ in comparison with the simple ARQ protocol without combining at the receiver.

We can also see that it is possible to identify SNR intervals for each modulation order and that generally, at high SNR, 64QAM gives the highest throughput. In figure 3.4, we plot the average throughput for different values of maximum retransmissions $\mathbb{M}_{max}$ (in our terminology, retransmission is equivalent to round). We focus on QPSK modulation, HARQ and $\mathbb{M}_{max} = 1, 2, 3, 5, 10$. We consider both no interference and two strong interferers cancellation.

When there is no interference (figure 3.4(a)), at most two rounds instead of one gives a gain of around four times at a 0 dB SNR. However, going from two rounds to three rounds gives a gain of only 0.1 in throughput.

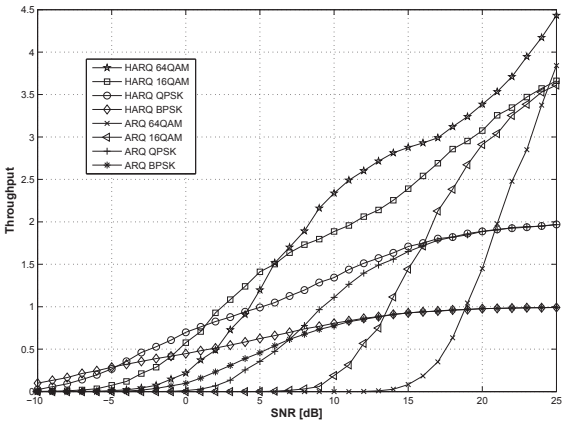
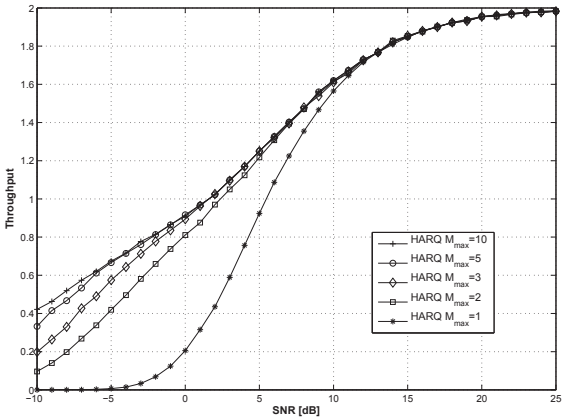


Figure 3.3: We consider different modulations, for HARQ and ARQ, with $\mathbb{M}_{max} = 3$ retransmissions. We show the average throughput for both retransmission protocols and we see the improvement in throughput of HARQ over ARQ for all modulations.

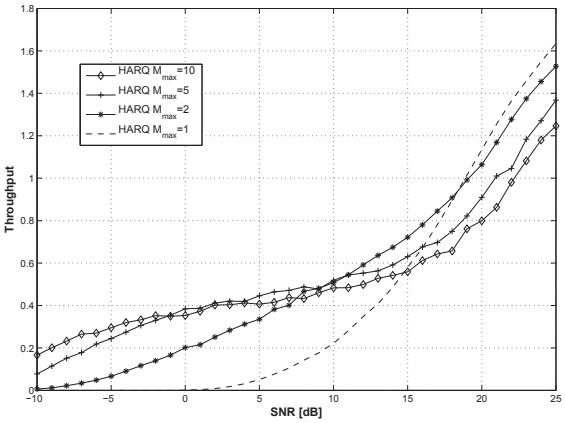
The later means that just increasing the maximum number of rounds is not sufficient to get a significant gain in throughput. Now, on figure 3.4(b) is a scenario with interference. We can observe that in the low SNR region, more retransmission rounds gives the highest throughput, however, at higher SNR, having only one retransmission gives the highest throughput. This is in contradiction of what is expected from having more opportunities to get the information decoded correctly. The reason for this result is that we fix the target rate for each case, which happens when there is no rate adaptation. This behavior suggests that the rate must be adapted depending on the instantaneous SNR.

Next we present our results when combining IR-HARQ with interference cancellation of the two strongest interferers. The topology is the Manhattan-like scenario in figure 3.5 and is a three by three grid topology, of size 30 by 30 meters, with node 0 in the middle. The receiver is located 5 meters away from the transmitter. This is a typical residential scenario.

In figure 3.6, we show the results for this topology. We consider QPSK modulation, a target rate of $R = 2$ bits/second, maximum number of retransmissions $\mathbb{M}_{max} = 3$ for both HARQ and ARQ. We show the average throughput with and without interference cancellation for an SNR from -10 to 25 dB. For interference cancellation, two interferers can be canceled. We perform a Gaussian approximation on interferers that are not canceled. Hence, the SNR without interference cancellation is around -7 dB and the SNR if the two strongest interferers are decoded is around 6 dB. The respective average throughputs are of 0.15 and 1. We observe a gain of around ten times in throughput from canceling only the two strongest interferers.



(a) Different number of ARQ rounds without interference



(b) Different number of ARQ rounds with two strong interferers.

Figure 3.4: We consider QPSK modulation, for HARQ, and we show the average throughput for different numbers of retransmissions $\mathbb{M}_{max} = 1, 2, 3, 5, 10$. In (a) we plot the case without any interference and in (b) we see a different trend for the case with two strong interferers, (b) shows that adapting the rate is an optimal way to increase the throughput.

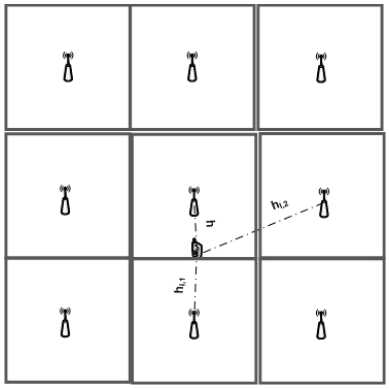


Figure 3.5: Manhattan-like topology with the user of interest at the edge of the apartment and the rest of the interfering femtocells placed at the middle of the surrounding apartments

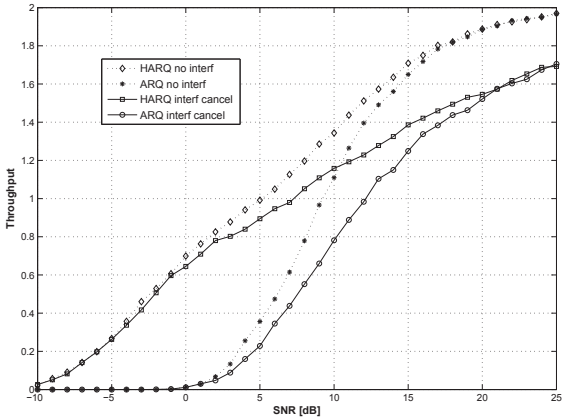


Figure 3.6: We look at the scenario in figure 3.5 and we consider QPSK modulation, $R = 2$, $\mathbb{M}_{max} = 3$ for HARQ and ARQ. We show the average throughput with no interference and with interference canceling two interferers for an SNR from -10 to 25 dB. We do a Gaussian approximation for the interferers that are not canceled. The SNR without interference canceling is -7 dB and 6 dB for interference cancellation of the two strongest interferes. The corresponding throughputs are 0.15 and 1, therefore there is a gain of ten times in throughput.

Finally, we investigate a bursty interference scenario where the interference is not constant. It is important to investigate how the throughput is affected if we randomize the interference. In femtocell networks, being deployed by end-users, network and frequency planning becomes complicated, since the placement of the femtocells will be unknown for the operator. Moreover, the users may be able to turn them on and off. In our model, if we let the nodes to be either active or inactive, this becomes equivalent to having a re-use factor in the network which creates an interference process that is no longer ergodic. We investigated the case of interferers being present 50% and 75% of the time, this is equivalent to setting the activity factor $\mu = \{0.5, 0.75\}$ and we compare it against an activity factor $\mu = 1$. There are two interferers and we assume that all nodes transmit with unit power for simplicity. The SNR is defined then by $\text{SNR} = \frac{1}{N_0+2}$

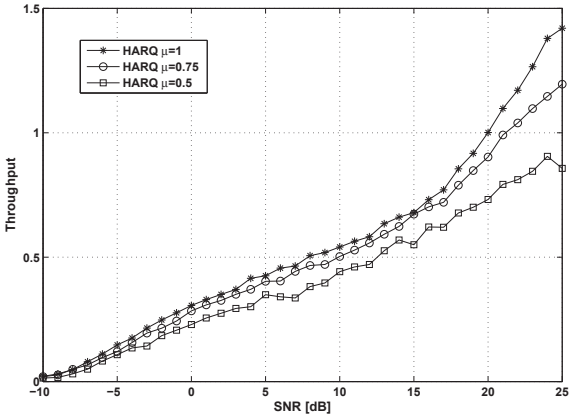


Figure 3.7: We consider QPSK modulation, we show the average throughput when the interference is randomized with an activity factor $\mu = 0.5, 0.75$, which means that the interferers will be active either half or 75% of the time and we compare with the case of the interference present all the time. We see that activity factor of 0.5 has the lowest throughput and that activity factor of 75% is closer to the corresponding curve for $\mu = 1$.

In figure 3.7 we show the curves for activity factors $0.5, 0.75$ and 1 for the case of canceling the two strongest interferers. We let all the users to have an activity factor of $\mu = \{0.5, 0.75, 1\}$. It means that both the user of interest and the interferers are not active all the time. This is equivalent to having a re-use factor of 2 and 3/4 in time and frequency. If we look at figure 3.7, the throughput when we use these values for the re-use factor is less than the corresponding throughput with activity factor of 1. This result tells us that

under this scenario, it is better to use the whole bandwidth all the time. It is important to highlight the fact that in this scenario, we consider a receiver that can cancel the strong interferers. If we have such a receiver, then it is optimal to use all resources or have a re-use factor of one. Having a re-use factor of one also means that the spectral efficiency is higher.

Chapter 4

Mutual Information Analysis of Interference Networks

LTE performance in terms of spectral efficiency and available data rates is, relatively speaking, more limited by interference from adjacent cells as compared to previous communication standards [27]. Means to reduce or control the inter-cell interference can potentially provide substantial benefits to LTE performance, especially in terms of the QoS provided to every user.

Efficient resource allocation algorithms are one of the key features in providing high spectrum efficiency in LTE [34]. An intelligent scheduling algorithm can, at the same time, help to reduce the impact of interference.

The performance of any particular scheduling algorithm can be evaluated with the help of exhaustive simulations, but it is time-consuming and has a high computational cost. By using information theory, we can analyze the achievable throughput, or spectral efficiency, under different system model assumptions.

Rather than performing extensive simulations, we take an information theoretic approach to derive analytical expressions that represent the long-term throughput of the network and consider practical cases where there is a constraint on the outage probability representing the latency of the protocol. We consider signals from discrete alphabets to model practical systems and we consider cases where CQI feedback is either unavailable, or outdated.

4.1 Key Challenging Applications

In this section, we talk about some of the key emerging applications to which our information theoretic analysis can be applied.

4.1.1 Heterogeneous Networks

Cellular wireless networks can be divided into homogeneous and HetNets. Heterogeneous implies that there are different types of cells in the network, i.e. low-power base stations are distributed throughout a macro cell network. These low power base stations can be microcells, picocells, relays, femtocells or distributed antenna systems [33]. On the one hand, microcells, picocells and relays are deployed by the operator to increase the capacity and coverage in public places, enterprises buildings, etc. On the other hand, femtocells are user-deployed at home to improve capacity. We generally denote low-power base stations by small-cells.

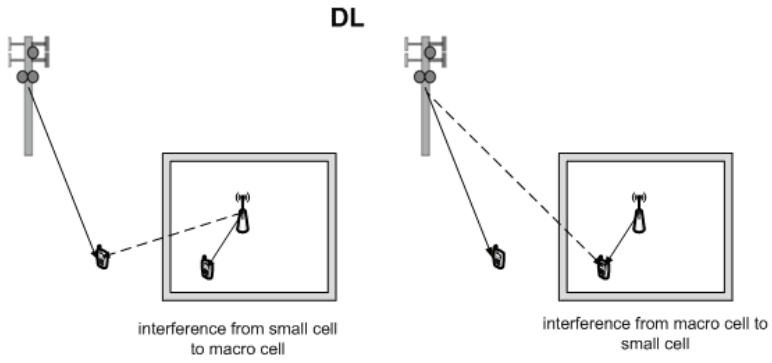
Cellular HetNets typically operate on licensed spectrum owned by the network operator. The most severe interference is experienced when the small-cells are deployed on the same frequency carrier as the macrocells [47]. More challenging interference scenarios are identified since interference can come across layers (macro–small-cell, small–macrocell), for example, a macrocell user far from the base station is transmitting at a very high power hurting the femtocells in the vicinity. Interference can also be experienced between small cells in both the UL and DL channels (see figure 4.1). In the case of inter-layer interference, the macrocell scheduler has to take into account the bursty interference from the femtocells since they will be serving only a couple of users.

Given the fact that in HetNets there is no controller managing the allocation process [11], operators will not be able to handle interference between small-cells in a centralized manner (centralized frequency planning). The design of distributed algorithms and techniques allowing for an efficient utilization of infrastructure is one of the key challenges in HetNets [44]. In this type of approach, the small-cell adapts its performance independently from other cells avoiding the need for any *a priori* centralized frequency planning and without having to exchange information or sending it to a central controller. They rely only on feedback, avoiding uncontrolled delays.

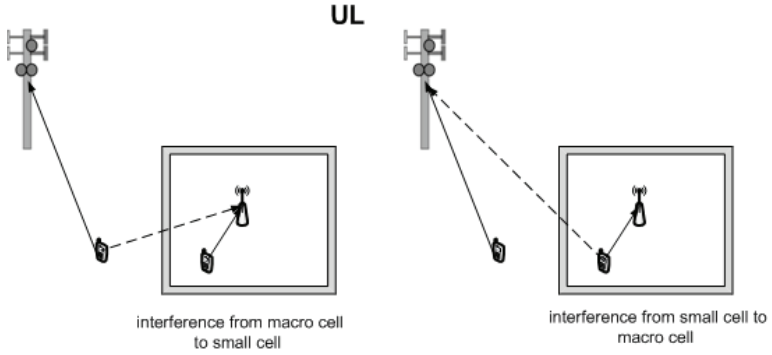
Link adaptation is one of the key features of HetNets. For instance, for an OFDM system, the allocation of subcarriers can be tailored according to the interference conditions. However, mutual interference should be carefully considered when designing heterogeneous systems.

4.1.2 M2M and Sparse Latency-Constrained Traffic

High-performance online gaming, M2M and sensor data communications are emerging massive applications for cellular networks. A typical example of M2M applications in mobile environments is sensors connected to public transport vehicles, to trains or to equipment in factories. It is predicted that these applications, in addition to voice and Internet traffic, will be an integral part of the traffic transported by LTE [61] and LTE-Advanced [57]



(a) Interference scenarios DL.



(b) Interference scenarios UL.

Figure 4.1: Figure (a) shows the interference scenarios for HetNets in the DL, figure (b) shows the interference scenarios for HetNets in the UL

networks. M2M is expected to account for a considerable amount of the traffic of such networks [25].

M2M communications, part of the Internet of Things (IoT) revolution, is expected to create an increasing number of connected devices, which will exceed human-to-human communications over the following years (50 billions machines against seven billion people for 2011) [25, 53]. A large class of the traffic generated by these emerging applications can require low-latency [41,71]. For example, for online gaming applications, the low-latency is critical to offer the best game experience as possible [50]. Large portions of M2M applications are expected to produce sparse traffic with low-delay constraints. The two main reasons being power reduction through DRX and transport of small sporadic packets to M2M devices. Concretely, a User-Equipment (UE) terminal emerging from an idle state to deliver a small packet to the network should reconnect for the smallest amount of time possible to conserve power. Moreover, the paradigm of many sparsely connected UEs transmitting sporadic traffic poses interesting problems related to re-

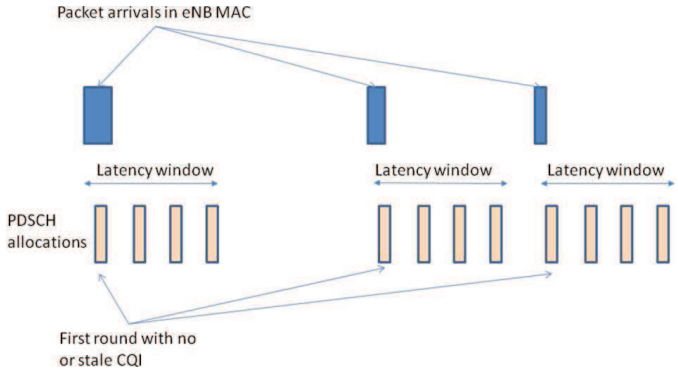


Figure 4.2: Sparse traffic in a delay-constrained scenario. Traffic arrivals in the eNB MAC layer are sparse as depicted in blue (there are three of them). The latency constraint is four slots, i.e. there are up to four possible PDSCH channel allocations. Because of the sparse traffic, CQI is outdated or unavailable on the first slot.

source allocation policies in a scheduled-access MAC protocol such as that of 3GPP LTE.

In sparse and latency-constrained traffic scenarios, packet arrivals are sporadic and must be scheduled under a latency constraint (see figure 4.2). In this context, CQI is typically outdated or unavailable. Note that outdated CQI also occurs because of moderate to high mobility, of insufficient uplink CQI periodicity or of non-stationary inter-cell interference. The latter will become more and more important with LTE Release 10 networks and their inherent heterogeneity. Hence, the scheduler must operate blindly for AMC and can only benefit from feedback after the first HARQ transmission round in the form of ACK/NACK signaling.

The extremely large number of devices connected to the network, the expected reliability of the service regardless of the operation environment, and low-latency requirements from applications such as emergency messages, video surveillance or health-care will require some enhancements to the network including link adaptation protocols, modulation and coding, and HARQ schemes [48, 69]. All these network optimizations will be included as a part of the LTE-Advanced standard since M2M communication is one of the main focuses in LTE-Advanced [25].

4.2 Related Work

Extensive research has explored variable-rate adaptation techniques. However, very little attention has been paid to the more performance-limited case of interference. Early work in [35] suggests a gain from adaptive policies. By deriving the Shannon capacity regions of variable rate and power, it is shown

that the maximum capacity is achieved when the rate is varied based on the channel variations. More recently, [55] explores rate adaptation with successive interference cancellation receivers for MIMO systems with outdated channel state information and Gaussian signals. When no outage constraint is considered, [13] presents a Signal-to-Interference-plus-Noise-Ratio (SINR) threshold-based adaptation without the use of HARQ and in [66], the rate is adapted to the fading conditions under the concept of maintaining a certain "fairness" between users.

The throughput of HARQ has been investigated for Gaussian input signals [20] over a Gaussian channel with fading and in the limit of infinite block length. In [67], the long-term throughput analysis of a HARQ protocol under slow-fading channels is presented for fixed-rate, variable-power transmissions under the framework of the renewal-reward theory of [20]. Rate adaptation for HARQ protocols under delay constraints is studied in [70], and for time-correlated channels in [36] and [43]. In [16], rate and transmit power are adapted under perfect CSI. Power adaptation is also presented in [19] to minimize the outage probability and in [17], both power and rate control are derived through dynamic programming without outage constraints. Combined power and rate adaptation is also presented in [22], and in [23], the optimization of either the packet drop probability or the average transmit power is shown for the case of IR-HARQ with a maximum number of retransmissions. In [64], the information-theoretic approach of [20] is adapted to variable rate transmissions in the case of HARQ with IR. In [39], without considering HARQ, a mathematical framework based on a sum-rate analysis for heterogeneous networks with partial feedback is developed.

The idea of changing the MCS for retransmissions is presented in [31], for IP video surveillance camera traffic by assigning additional redundancy to the retransmissions and reducing the estimated CQI.

Recent so-called *rateless or fountain* coding techniques with IR for additive-noise channels are also reported in [30], [58], [14], [37]. When combined with a HARQ link-layer protocol, these coding schemes allow for transmission over unknown channels without the need for sophisticated rate adaptation policies, and whose instantaneous rate (or spectral-efficiency) depends on the time the decoder is able to decode the message. The basic principle for this type of transmission was introduced for content distribution over the internet and broadcast networks by Luby [49] using so-called *LT-codes* for erasure channels. These were improved by Shokrollahi with his invention of *Raptor codes* [63]. The latter were then adapted for AWGN channels in [56].

All of these coding strategies are structured, and, in particular Perry *et al*'s *Spinal Codes*, can approach Shannon's AWGN channel capacity with varying degrees of encoding and decoding complexity provided the number of transmissions is allowed to grow without bound. Although not shown in [30] [14] it may very well be true for any ergodic time-varying additive-

noise channel. An extension of the promising superposition coding technique designed for successive decoding at the receiver considered by Erez *et al* was also described for time-varying channels without an *a priori* stochastic model [30]. This considered the performance of their rateless coding construction for a small number of transmission rounds.

In this work, we consider similar rateless strategies for time-varying channels for a finite and small number of transmission rounds, potentially allowing for a residual outage probability after the maximum number of rounds. Imposing a quasi-finite duration for transmission is often required to minimize latency in data transmission networks. For instance, the HARQ protocol of LTE reference channels [8] is tuned to offer an approximate 1% outage rate after two transmission rounds which allows for a one-way latency of 10ms for 99% of transmissions. This can, of course, be tuned to offer different latency-throughput tradeoffs. Since the maximum number of transmission rounds is fixed in such protocols, it seems natural that the number of dimensions used in each round should be optimized in order to maximize throughput, by progressively decreasing code rate across rounds. We should note that the latter is not a requirement in rateless coding with an unbounded number of transmission rounds.

4.3 Initial Analysis for Interference-free Networks

We start our analysis by looking into interference-free networks, i.e. we do not consider interference created by neighboring transmitters and we focus on single antenna systems, Single-Input Single-Output (SISO), although our model can be extended to MIMO.

4.3.1 Signal Model and Assumptions

In the following, we present the signal model and assumptions for this section. Without loss of generality, we consider OFDM signaling. The UL of an LTE system uses SC-FDMA modulation. Our joint HARQ and rate adaptation policy applies equally, but the signaling details differ. Therefore, for a particular subcarrier j, let x denote the complex-valued transmitted symbol, z denote the additive white Gaussian noise (AWGN), and h denote the channel gain. Both z and h are modeled with a zero-mean and unit variance complex Gaussian random variable. With a total of K subcarriers, the received signal (y_j), in a particular subcarrier is

$$y_j = h_j x_j + z_j, \qquad j = 1, 2, \ldots, K. \tag{4.1}$$

We consider a block-stationary Rayleigh fading channel model. Fading remains static for the duration of a HARQ round but varies between retransmissions. The HARQ feedback channel is assumed to be error-free. CQI

can be received after each round. However, prior to the first round, CQI may or may not be available. As explained earlier, this can occur because of sparse traffic. Furthermore, because of fast-fading, (low) mobility or non-stationary inter-cell interference, CQI can at any round be simply unusable. Consequently, we analyze cases where, at the first transmission round, CQI is either outdated or simply unavailable. On the further rounds, we keep on assuming that CQI is not available. But, we take advantage of the one bit of ACK/NACK information given to the transmitter after each HARQ round.

For the outdated CQI case, it is assumed that the fading statistics are available to the transmitter. This assumption is reasonable because the eNB scheduler can maintain a database of channel measurements in its cell, allowing it to derive the fading statistics over time.

4.3.2 Modeling and Optimization of a Resource Scheduling Policy

We consider a one-shot transmission model where one transport-block of size N_{TB} arrives in sub-frame n and must be served at maximum spectral-efficiency under a latency constraint. We denote by $\mathbb{M}_{max}$ the maximum number of transmission rounds. To characterize code performance and the effect of the channel, we use the instantaneous mutual information in each transmission round. This theoretic measure, although asymptotic, provides a very accurate indication of potential performance in LTE, whose coded-modulation subsystem performs close to asymptotic limits. Let H_r denote the vector of channel realizations in the fth transmission round. Then $I(H_r)$ denotes the corresponding instantaneous mutual information. Accordingly,

$$I(H_1, \ldots, H_{\mathbb{M}_{max}})$$

defines the mutual information accumulated over $\mathbb{M}_{max}$ transmission rounds. In order to compute the mutual information, we assume Gaussian input signals (upper-bound on QAM modulation). For example, let us consider one subcarrier of a SISO system without interference and let P denote the received power, h_r is the channel response at round r and N_0 is the noise power, then

$$I(H_1, \ldots, H_{\mathbb{M}_{max}}) = \sum_{r=1}^{\mathbb{M}_{max}} \log_2 \left(1 + \frac{P|h_r|^2}{N_0}\right). \tag{4.2}$$

Generalizing the notation from [20], the probability of decoding a transport-

block in round r with N_j as the number of dimensions used in round j is

$$\Pr\Big(I(H_1,\cdots,H_r) > R_r \sum_{j=1}^{r} N_j, \\ I(H_1,\cdots,H_n) < R_n \sum_{j=1}^{n} N_j, \forall n < r\Big). \tag{4.3}$$

Let $P_{out,n}$ denote the target transport-block error probability after n transmission rounds. The latency constraint is expressed by ensuring that the probability that the transport-block is not served after $\mathbb{M}_{max}$ transmission rounds is below $P_{out,\mathbb{M}_{max}}$. Under this framework, rate adaptation is the optimization of the rate sequences R_r such that (1) the packet error probability remains below $P_{out,\mathbb{M}_{max}}$ after $\mathbb{M}_{max}$ transmission rounds and (2) the spectral-efficiency is maximized. The optimization is carried out as a function of the distribution of $I\left(H_1,\ldots,H_{\mathbb{M}_{max}}\right)$.

For simplicity, we consider at most two retransmission rounds (ARQ rounds), but our policy can also be applied for more than two. We consider three scenarios

1. Minimal-latency: a trivial case of serving the packet in one round which corresponds to the minimal-latency rate adaptation policy.

2. Latency-constrained with no prior CQI: we consider two transmission rounds and no information about the channel.

3. Latency-constrained with outdated CQI: we consider again two transmission rounds, but unlike the previous case, we assume that we have outdated information about the channel with some correlation with the actual channel.

For simplicity and in the interest of obtaining semi-analytical results, we concentrate on one subcarrier, i.e. that H_r is a scalar. In Chapter 5, using a fully-compliant LTE modem implementation, this is reconsidered using a multipath 3GPP Extended Pedestrian (EPA) channel model.

Scenario Analysis: Minimal-latency

We consider first the trivial case of serving the transport-block in one round. This is the minimal-latency rate adaptation policy. The rate allocation law for R_1 is given by the solution to

$$\Pr\big(I(H_1) < R_1\big) = P_{out,1}. \tag{4.4}$$

Without any a priori information regarding the channel statistics, this essentially says that the best that can be done is to transmit with the lowest

spectral-efficiency coding scheme (i.e. lowest MCS) to minimize latency. With a priori information, the largest MCS such that the probability of channel realizations requiring a smaller MCS is still below the threshold is chosen.

Let H_{out} denote the channel corresponding to outdated CQI. If stale CQI is available prior to transmission of the transport-block, then the rate should be chosen such that

$$\Pr\left(I(H_1) < R_1 | H_{out}\right) = P_{out,1} \tag{4.5}$$

in order to take into account the outdated CQI.

Scenario Analysis: Latency-constrained with no Prior CQI

We now consider the case with two transmission rounds. Let B define the number of information bits to be transmitted. Let N_T denote the total number of dimensions available and let N_1 denote the number of dimensions used in the first round. Hence, the rate in the first round is $R_1 = \frac{\log_2 B}{N_1}$, and the rate in the second round $R_2 = \frac{\log_2 B}{N_T}$. We define

$$\rho = \frac{N_1}{N_T}. \tag{4.6}$$

and we can relate R_1 to R_2 with $R_2 = \rho R_1$. Let $\bar{R}$ denote the overall spectral efficiency. With $P_{out,1}$ as the outage probability after the first round, we have

$$\begin{aligned} \bar{R} &= R_1\left(1 - P_{out,1}\right) + P_{out,1} R_2 \\ &= R_1\left(1 - P_{out,1}\right) + P_{out,1}\rho R_1. \end{aligned} \tag{4.7}$$

We want to maximize $\bar{R}$ such that the probability of outage after the second round is below the given constraint $P_{out,2}$ For the first round, there is no feedback information. The outage probability $P_{out,1}$ is unknown but it depends on H_1 and the SNR. We can relate R_1 to $P_{out,1}$ as follows. From equation (4.4), we have

$$\Pr\left(I(H_1) < R_1\right) = \Pr\left(\log_2\left(1 + \mathrm{SNR}|h_1|^2\right) < R_1\right) = P_{out,1}. \tag{4.8}$$

Consequently, we obtain

$$R_1 = \log_2\Big(1 - \mathrm{SNR}\ln(1 - P_{out,1})\Big). \tag{4.9}$$

In the second round, feedback about the previous round is available. The outage probability is now given by

$$\Pr\Big(I(H_1, H_2) < R_2 | I(H_1) < R_1\Big) = P_{out,2} \tag{4.10}$$

We can rewrite equation (4.10) as follows

$$\begin{aligned}\Pr\Big(I(H_1,H_2)<R_2|I(H_1)<R_1\Big) &= \frac{\Pr(I(H_1,H_2)<R_2, I(H_1)<R_1)}{\Pr(I(H_1)<R_1)}\\ &= \frac{\int_0^{\frac{2^{R_1}-1}{\mathrm{SNR}}} e^{-|h_1|^2} d|h_1|^2}{P_{out,1}} - \frac{\int_0^{\frac{2^{R_1}-1}{\mathrm{SNR}}} e^{-a-|h_1|^2} d|h_1|^2}{P_{out,1}}\\ &= P_{out,2} \end{aligned} \tag{4.11}$$

where

$$a = \left(\left(\frac{2^{R_1}}{1+\mathrm{SNR}|h_1|^2}\right)^{\frac{\rho}{1-\rho}} \frac{1}{\mathrm{SNR}}\right) - \frac{1}{\mathrm{SNR}} \tag{4.12}$$

and the limits stem from the fact that if $I(H_1) < R_1$ then $|h_1|^2 < \frac{2^{R_1}-1}{\mathrm{SNR}}$. The integrals in equation (4.11) are evaluated numerically.

To find the optimal value of R_1 in the first round, we perform an extensive exploration on $P_{out,1}$, given that we want to maximize equation (4.7) and subject to the constraint $P_{out,2}$ in equation (4.11).

Scenario Analysis: Latency-constrained with Outdated CQI

Because of the sparse traffic characteristic, of moderate to high mobility, of insufficient uplink CQI periodicity or of inter-cell interference, we investigate cases where the UL CQI is outdated or unavailable. In such cases, the scheduler can only benefit from binary feedback after the first HARQ transmission round (in the form of ACK/NACK signaling [61]).

We make the additional assumption that the channel remains constant over the two transmission rounds and let $h = h_1 = h_2$. Furthermore, we denote by h_{out} the channel value that corresponds to the outdated CQI. In order to model a possible correlation between h_{out} and h, we use the following model. Let λ be the correlation parameter, then

$$h = \sqrt{\lambda} h_{out} + \sqrt{1-\lambda} h^{'}$$

where h_{out} and $h^{'}$ are i.i.d. Gaussian-distributed random variables. Note that in this case,

$$\lambda = \mathbb{E}\left[h_{out} h^{*}\right].$$

In addition, $|h|^2$ is a non-central Chi-square random variable with two degrees of freedom. We follow the same general procedure to obtain the throughput and probability of outage than in the previous cases. However, the spectral efficiency is a function of the outdated CQI and we have to average over the distribution of $|h_{out}|^2$.

First, let $\gamma_1 = \sqrt{\frac{2^{R_1}-1}{\mathrm{SNR}}}$ be the outage threshold in the first round and $\gamma_2 = \sqrt{\frac{2^{R_2}-1}{\mathrm{SNR}}}$ be the outage threshold in the second round. Then, $\Pr\left(h > \gamma_1\right)$

represents the probability of having a successful transmission in the first round, $\Pr(h < \gamma_1, h > \gamma_2)$ is the probability of being unsuccessful in the first round but successful in the second round, and $\Pr(h < \gamma_2)$ gives the probability of being in outage. All these probabilities are a function of the CDF of $|h|^2$. The non-centrality parameter of $|h|^2$ is $s^2 = \lambda|h_{out}|^2$. Let $F_\chi(\chi)$ denote the CDF of $|h|^2$. It can be expressed in terms of the Marcum Q-function [59], i.e.

$$F_\chi(\chi) = 1 - Q_1(s, \chi) \tag{4.13}$$

where $Q_M(a, b)$ is the Marcum Q-function with M degrees of freedom and paramters a and b.

The overall spectral efficiency $\bar{R}$ over the two ARQ rounds can be written as

$$\bar{R} = \Pr(h > \gamma_1) R_1 + \Pr(h > \gamma_2, h < \gamma_1) R_2. \tag{4.14}$$

To find the optimal rates, we first obtain R_2 from the outage constraint $P_{out,2}$. Since we know that $h < \gamma_2$ implies an outage, R_2 is given by solving equation (4.15) for R_2. Therefore

$$P_{out,2} = \Pr\left(|h|^2 < \frac{2^{R_2} - 1}{\text{SNR}}\right) = F_\chi\left(\gamma_2^2\right) \tag{4.15}$$

Next, to find the value of R_1 that will maximize the overall spectral efficiency, we first write (4.14) in terms of the Marcum Q-function. We have

$$\bar{R} = Q_1(a_1, b_1)R_1 + (Q_1(a_2, b_2) - Q_1(a_1, b_1))R_2 \tag{4.16}$$

where $a_1 = a_2 = s$, $b_1 = \gamma_1$, and $b_2 = \gamma_2$. We now take the derivative of equation (4.16) with respect to R_1, and we solve for R_1 when the derivative is zero. We obtain

$$\begin{aligned} \frac{\partial \bar{R}}{\partial R_1} &= \frac{\partial Q_1(a_1, b_1)}{\partial R_1} R_1 + Q_1(a_1, b_1) - \frac{\partial Q_1(a_1, b_1)}{\partial R_1} R_2 \\ &= \frac{\partial Q_1(a_1, b_1)}{\partial R_1}(R_2 - R_1) + Q_1(a_1, b_1) = 0. \end{aligned} \tag{4.17}$$

To find the derivative of the Marcum Q-function in (4.17), we used [10].

4.3.3 Numerical Results

In this section, we present numerical results in terms of (1) the probability of outage and (2) the achieved spectral efficiency. Remember that we assume Gaussian signaling to compute the mutual information. Throughout this section, we fix the spectral efficiency to 2 bits per channel use. The maximum number of retransmissions is one round (at most two rounds).

Figure 4.3 presents the minimum SNR necessary to achieve a given outage probability $P_{out,2}$. For a given value of $P_{out,2}$, we calculate the corresponding

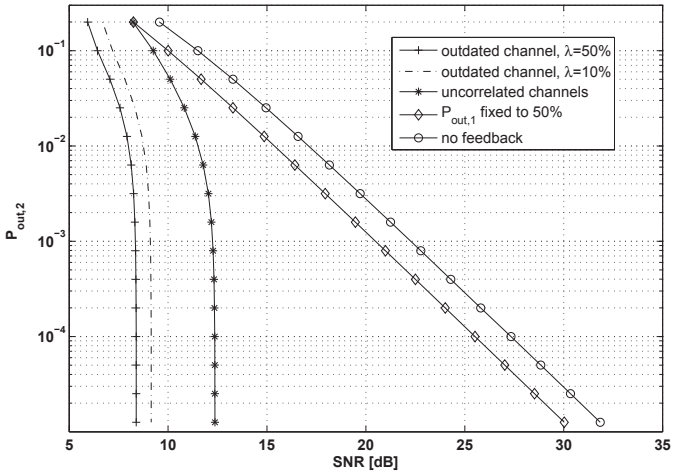


Figure 4.3: For different values of the probability of outage after the second round $P_{out,2}$, we calculate the corresponding SNR for the different scenarios. The symbol λ is the correlation coefficient between the actual channel and the channel corresponding to outdated CQI information. We compare a correlation coefficient value of $50\%, 10\%$ and uncorrelated case. For comparison purposes, we also plot two more cases. First when no ACK/NACK feedback is available from the HARQ process. Second, when $P_{out,1}$ is fixed to 50% with $\rho = 0.5$ to make sure that 50% of the dimensions are used in each round.

SNR for our rate adaptation policy. For the scenario (3) from Section 4.1.2, the outdated CQI has a correlation coefficient with the actual channel of $\lambda = 10\%$ or $\lambda = 50\%$. For comparison purposes, we consider two more cases in addition to the scenario (2) and (3) from Section 4.1.2. First we evaluate a case where we force the probability of outage after the first round to 50%, fixing $\rho = 0.5$ to make sure that 50% of the dimensions are used in each round. Typically, while conventional systems try to ensure a 10% outage probability per slot, we observe from our results that a higher value gives, in fact, a higher overall spectral efficiency. Second, we evaluate a case where no feedback at all is available i.e. when we can not even receive ACK/NACK from the HARQ process. This highlights the significant gain from adapting the rate across rounds with only one bit of feedback, even in the case without any CQI information. The gain is even higher when only outdated CQI information is available. Our rate adaptation policy gives a zero probability of outage without the need of having a high SNR. From the results in figure 4.3, we can observe that it does not make a difference to increase the SNR above 12.5 dB for the case without CQI. We show that adjusting the dimensions used in each round results in almost causal feedback performance. In our scenarios, the two rates are simply controlled by $\rho = \frac{R_2}{R_1}$, which depends on the SNR and target outage probability $P_{out,2}$. We only need one bit of feedback, which we get causally from HARQ. In fact, it is the state of the channel that chooses the code rate. By choosing the rate in the first round as high as possible, we can guarantee a probability of outage after the second round while maximizing the spectral efficiency.

Figure 4.4 presents the overall spectral efficiency obtained for a given SNR. We set $P_{out,2}$ to 1%. For the outdated CQI case, we consider $\lambda = 50\%$ and $\lambda = 10\%$. For reference purposes, we also plot the ergodic capacity (Rayleigh channel capacity), i.e. perfect rate adaptation. Finally, we consider a scenario where the rate in the first round is chosen as the one that corresponds to a probability of outage after the first round of 50%. This value is chosen because it gives the highest spectral efficiency. Fixing the probability of outage after the first round to more or less than 50% gives, in fact, a lower overall spectral efficiency.

From our results, we see a significant improvement in spectral efficiency even in the case without CQI. When we can benefit from outdated CQI, we achieve a performance close to the ergodic capacity. If we look at the results for the ratio of dimensions for the second round (right axis and dashed lines), we can see that as the SNR gets higher, more dimensions are used in the second round.

4.4 Interference Networks Analysis

We now consider a network with interference from adjacent cells.

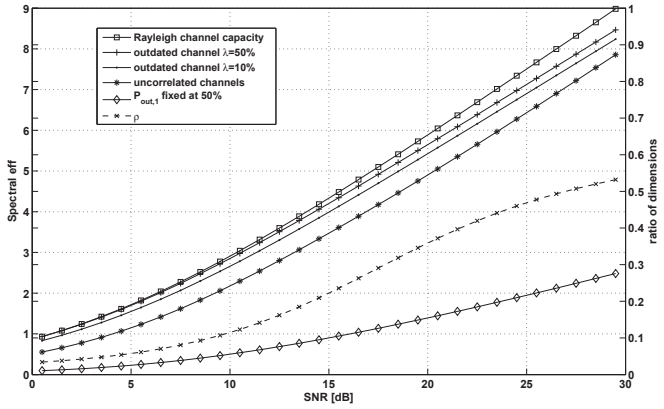


Figure 4.4: The axis on the left (solid lines) shows the spectral efficiency versus SNR for the different scenarios. We set $P_{out,2}$ to 1%. The symbol λ is the correlation coefficient between the actual channel and the outdated/stale CQI information. We compare correlation coefficients of $50\%, 10\%$ and uncorrelated case. For comparison purposes we plot the curve for the ergodic capacity and $P_{out,1}$ fixed to 50% with $\rho = 0.5$ to make sure that 50% of the dimensions are used in each round. The axis on the right (dashed lines) shows the ratio of dimensions used in the two rounds.

4.4.1 Modeling and Assumptions

We consider a slotted transmission scheme and we take an information-theoretic approach to analyze the throughput performance. When there is more than one user, we assume that all transmissions in every slot are synchronized and we randomize the interference process with the use of activity factors. The latter models sporadic interference patterns characteristic of future heterogeneous networking deployments, in particular the interference seen from small-cell base stations with bursty traffic in the receiver of a macrocell user. They can also model dual-carrier networks with cross-carrier scheduling. In this type of network, we can talk about clean and dirty carriers. On the one hand, clean carriers are used by the macrocell to carry their data plus signaling for small-cells because of their controlled interference property. On the other hand, dirty carriers are interfering carriers where the "cleaning" is done with the use of HARQ.

We consider a maximum of $\mathbb{M}_{max}$ HARQ transmission rounds and the channel is iid or constant over all the transmission rounds of the protocol. After each transmission we receive an error-free acknowledgment (ACK or NACK) indicating a successful or unsuccessful transmission. We define the probability of outage as being unsuccessful to correctly receive the information at the end of the HARQ protocol. This probability translates to the

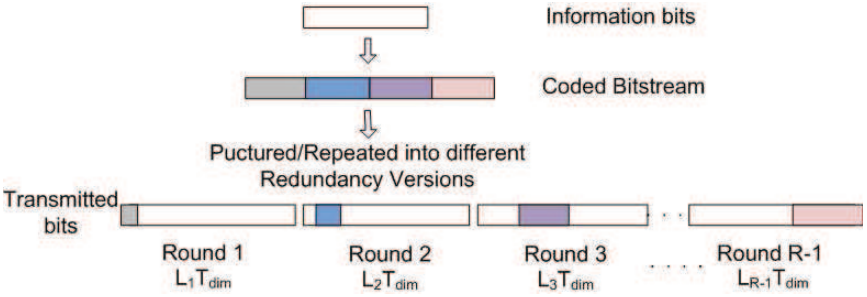


Figure 4.5: Coding Model

latency of the protocol and QoS in our system.

In general, we define R_r as the code rate at the rth round. For a particular user, we define the number of dimensions in time as T_{dim} and the number of dimensions in frequency as L_r. Let L'_r be the number of dimensions in frequency up to round r. Then, at each transmission round, the total number of dimensions is $L'_r T_{dim}$. Assuming the channel does not vary during T_{dim} time dimensions and for a packet length of B information bits, the rate R_r at the rth round, in bits/dim is given by:

$$R_r = \frac{\log_2 B}{L'_r T_{dim}} \quad \text{bits/dim.} \tag{4.18}$$

In IR-HARQ, the retransmission consists of the same set of information bits as the original, however, the set of coded bits are chosen differently and they may contain additional parity bits. In each of the transmission rounds there are $L_r T_{dim}$ dimensions, however, this number is not necessarily the same across rounds according to the LTE standard [7] (see figure 4.5).

In the context of LTE, the number of physical dimensions $L_r T_{dim}$ refers to the number of resource blocks allocated to one user in one subframe of 1 ms duration, i.e. one Transmission Time Interval (TTI). There are at most two transport blocks delivered to the physical layer in the case of spatial multiplexing [28]. In a single-user LTE system, there is only one transport block in one TTI, representing only one codeword "in the air" at the same time. Each transport block is carried by an HARQ process, and each process is assigned to a subframe (number of processes is fixed). In our model, if the number of dimensions for a user is less than the maximum number of available resources N_T, ($L_r T_{dim} < N_T$), then the rest will not be utilized. Although not possible in the current LTE standard, one could propose to assign the unused resources to transmit multiple codewords in parallel (at the same time), to increase the throughput. In a multiuser system, the remaining dimensions would be allocated to other users and thus the efficiency of the protocol should be chosen to maximize the aggregate spectral efficiency of the cell.

Let P_{succ_1} be the probability of having a successful transmission in the first round, and $P_{succ_r,fail_{r-1}}$ the probability of not having a successful transmission in the $(r-1)$th round, but being successful in the rth round. Finally, let P_{out} represent the probability of outage at the end of the protocol. The overall throughput can thus be expressed as:

$$\bar{R} = P_{succ_1} R_1 + \sum_{r=2}^{\mathbb{M}_{max}} P_{succ_r,fail_{r-1}} \left(\frac{R_r}{r} \right) \quad \text{bits/dim.} \tag{4.19}$$

where the outage probability is given by $P_{out} = \Pr(r = \mathbb{M}_{max} + 1) = 1 - \sum_{r=1}^{\mathbb{M}_{max}} P_{succ_r}$.

In the case of an upper layer ARQ, the throughput expression becomes [20]:

$$\bar{R} = [1 - P_{out}] \left[P_{succ_1} R_1 + \sum_{r=2}^{\mathbb{M}_{max}} P_{succ_r,fail_{r-1}} \left(\frac{R_r}{r} \right) \right] \quad \text{bits/dim.} \tag{4.20}$$

We consider N_u transmitters, where user 0 is the transmitter of interest, and the remaining $N_u - 1$ transmitters are interferers. We model an OFDMA physical layer with K subcarriers. We let $\mu_{j,k}$ be the activity factor, P_j the transmission power, and $x_{j,k}$ the input signal of the jth user on the kth subcarrier. We assume discrete signals with equal probabilities and size of the constellation $|x_{j,k}| = \mathcal{S}$, z_k is the zero mean complex Gaussian noise with variance σ^2. Since we assume Rayleigh fading, $h_{j,k}$ is a circularly symmetric complex Gaussian random variable with unit mean. The received signal y is given by:

$$y = \sum_{j=0}^{Nu-1} \sqrt{P_j} \sum_{k=0}^{K-1} h_{j,k} \mu_{j,k} x_{j,k} + z_k \tag{4.21}$$

Variations in the channel are caused at the receiver because of the activity factor plus the frequency shifting from the resource allocation process. For the interfering users, the channel variation depends on whether we consider the UL or DL. In the UL, it is caused by the interference coming from different user terminals. In the DL, the activity factors will introduce variations originated from the fact that the interfering cells are not active the whole time.

To model heterogeneous networks where the interference is not constant, we let the user of interest to be active all the time (macrocell user) and we let the small interfering cells to transmit with a certain probability (see figure 4.6).

In the next section, we provide a motivating example on the importance of adapting the rate and physical dimensions across the rounds of the HARQ protocol.

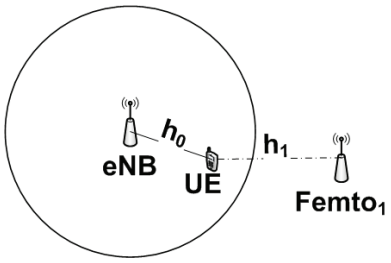


Figure 4.6: Downlink of a macrocell with a femtocell interfering.

4.4.2 Simple Interference Analysis in Zero-outage

For the sake of analytical tractability, we start by looking at the special case of (4.21) where the $h_{j,k}$ are fixed (AWGN channel) and we assume Gaussian signals. We consider one interferer and we model it with an activity factor, which means that the interferer could be active or inactive. The activity factor is Bernoulli distributed with probability p.

The rate with Gaussian codebooks that can be achieved by the protocol depends on the interference state (interference active or inactive). Let R_H be the capacity that can be achieved without interference, and R_L the corresponding capacity with interference, which are given by:

$$R_H = \log_2(1 + \mathrm{SNR}_1) \tag{4.22}$$

$$R_L = \log_2\left(1 + \frac{\mathrm{SNR}_1}{1 + \mathrm{SNR}_2}\right) \tag{4.23}$$

where SNR_1 is the SNR for the user of interest, SNR_2 is the corresponding SNR for the interferer and we assume unitary noise variance.

We consider a HARQ protocol with two rounds and we define $\rho = \frac{N_1}{N_T}$, where N_1 is the number of dimensions used in the first round and N_T has been previously defined as the total number of dimensions. Then for a packet of length B bits, the rate in the first round is $R_1 = \frac{1}{\rho N_T}\log_2 B = R_H$ and in the second round $R_2 = \frac{1}{N_T}\log_2 B = R_L$. Therefore, $R_L = \rho R_H$, and $\rho = \frac{R_L}{R_H}$.

In the remainder of this section, we derive the zero-outage throughput with and without feedback and we consider also the case of a residual outage at the end of the protocol with feedback.

Zero-outage Throughput Without Feedback and no Delay Constraint

We now look at the case of no feedback. Let the activity factor μ define the state of the interference. We consider ON/OFF interference, therefore, if $\mu = 0$ there is no interference and $\mu = 1$ means that the interference is

active and it happens with probability p. Then the throughput $\bar{R}$ with zero outage (without delay) is given by:

$$\bar{R} = \mathbb{E}_{\mu} I(X;Y|\mu) \tag{4.24}$$

$$= (1-p)R_H + pR_L \tag{4.25}$$

It is interesting to note that (4.25) is the ergodic capacity (average over all possible states). In the next section we explore the case when feedback becomes available and we look at the case of more than two transmission rounds.

Zero-outage Throughput with Feedback

In this case, we assume that we have feedback from the HARQ protocol and we vary the tolerable latency by fixing the maximum number of transmission rounds $\mathbb{M}_{max}$, but still assume zero-outage probability. Then, given that we want zero-outage at round $\mathbb{M}_{max}$, we choose the rate that guarantees successful decoding (i.e. R_L). We choose the rate in the first round to be as high as possible, and the intermediate rates are at the optimal value between R_L and R_H. Therefore, the rate at the rth round is given by:

$$R_1 = \frac{\log_2 B}{\rho_1 N_T} = R_H \qquad r = 1 \tag{4.26}$$

$$R_r = \frac{\log_2 B}{(\sum_{j=1}^{r} \rho_j) N_T} = \left(\frac{\rho_1}{\sum_{j=1}^{r} \rho_j} \right) R_H \qquad 2 \leq r < \mathbb{M}_{max} \tag{4.27}$$

$$R_{\mathbb{M}_{max}} = \frac{\log_2 B}{N_T} = R_L = \rho_1 R_H \;\Rightarrow \rho_1 = \frac{R_L}{R_H} \qquad r = \mathbb{M}_{max} \tag{4.28}$$

In this case, the throughput expression for $\mathbb{M}_{max}$ rounds is given by:

$$\bar{R} = (1-p)R_H + \sum_{r=2}^{\mathbb{M}_{max}-1} p^{r-1}(1-p) \left(\frac{\rho_1}{\sum_{j=1}^{r} \rho_j} \right) R_H + p^{(\mathbb{M}_{max}-1)} R_L \tag{4.29}$$

For the rates to be achievable, we observe that there is a restriction on the ratio of dimensions after the second round ρ_r, $r > 1$. This restriction comes from the fact that the rate after round r is $\left(\sum_{j=1}^{r-1} \rho_j \right) N_T R_L + \rho_r N_T R_H$ which means that after round r we decode if:

$$\left(\sum_{j=1}^{r} \rho_j \right) N_T R_r < \left(\sum_{j=1}^{r-1} \rho_j \right) N_T R_L + \rho_r N_T R_H \tag{4.30}$$

$$R_r = \left(\frac{\rho_1}{\sum_{j=1}^{r} \rho_j} \right) R_H < \frac{\left(\sum_{j=1}^{r-1} \rho_j \right) R_L + \rho_r R_H}{\sum_{j=1}^{r} \rho_j}$$

$$\rho_r > \rho_1 \left(1 - \sum_{j=1}^{r} \rho_j \right) \tag{4.31}$$

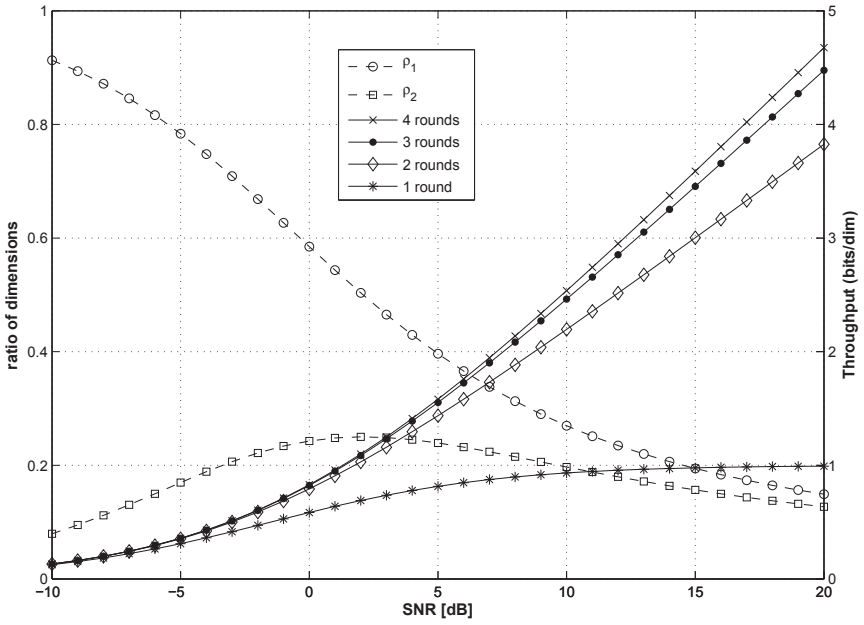


Figure 4.7: The axis on the right (solid lines) shows the zero-outage throughput for the HARQ protocol with different number of rounds, while the axis on the left (dashed lines) shows the ratio of dimensions per round for the three rounds, zero-outage HARQ protocol. In both cases the channel is AWGN with Gaussian signals and there is one interferer with probability $p = 0.5$. The interference strength is the same as the user of interest ($\mathrm{SNR}_1 = \mathrm{SNR}_2$).

If we look at figure 4.7, the solid lines and the right axis show the zero-outage throughput for the HARQ protocol with a maximum number of rounds $\mathbb{M}_{max} = \{1, 2, 3, 4\}$. We can see that there is a high gain when going from one to two rounds and after three rounds there is only a marginal gain. The dashed lines and left axis show how the dimensions are being distributed across the rounds of the protocol. We illustrate the case of three rounds (i.e. $\mathbb{M}_{max} = 3$) and we look at the proportion of physical dimensions used in each round (ρ_r). In both cases, the interference strength is the same as the user of interest ($\mathrm{SNR}_1 = \mathrm{SNR}_2$), the channel is AWGN and we assume Gaussian signals with one interferer active with probability 50%. If we look at 10 dB SNR, we observe that 20% of the dimensions are used in the first round, 15% in the second round and the remaining 65% are left for the third round. In general, we observe that at high SNR, the number of dimensions used in the first round decreases, leaving progressively more dimensions to

the last round. If we think about the almost blank subframes (ABS) feature of LTE, which restricts the transmission in the cell if there is interference, we would have a lower spectral efficiency than the average rate and therefore, by adapting the dimensions one can achieve a higher spectral efficiency even in the presence of interference.

Throughput With Outage and Feedback

In this case, we allow the protocol to have a residual outage probability which is overcome by an upper layer ARQ process on top of the IR-HARQ [46], [20], and we assume that we have feedback. For two rounds, from (4.20) the throughput is given by:

$$\bar{R} = (1 - P_{out}(\rho, R_2)) \left[(1 - P_{out,1}(\rho, R_2)) \frac{R_2}{\rho} + P_{out,1} \left(1 - P_{out,2}(\rho, R_2|out_1)\right) R_2 \right] \tag{4.32}$$

where $P_{out,2}(\rho, R_2|out_1)$ is the outage probability at the second round, given that there was an outage in the first round, and $P_{out,r}$ is the probability of outage at round r.

$I(\mu_r)$ is the mutual information as a function of the state of the interference at round r, and it is defined by μ_r:

$$I(\mu_r) = \begin{cases} \log_2(1 + \mathrm{SNR}_1) & \mu_r = 0 \\ \log_2\left(1 + \frac{\mathrm{SNR}_1}{1+\mathrm{SNR}_2}\right) & \mu_r = 1 \end{cases} \tag{4.33}$$

Now, we can define the probabilities in (4.32) as follows: $P_{out,1}(\rho, R_2)$, the outage probability at the first round, is given by:

$$\begin{aligned} P_{out,1}(\rho, R_2) &= \Pr(R_2 > \rho I(\mu_1)) \\ &= \begin{cases} 1 & \text{if } R_2 > \rho I(0) \\ 0 & \text{if } R_2 < \rho I(1) \\ \Pr(\mu_1 = 1) = p & \text{if } \rho I(1) < R_2 < \rho I(0) \end{cases} \end{aligned} \tag{4.34}$$

$P_{out,2}(\rho, R_2|out_1)$ is given by:

$$\begin{aligned} P_{out,2}(\rho, R_2|out_1) &= \Pr(R_2 > \rho I(\mu_1) + (1-\rho) I(\mu_2) \ldots \\ &\qquad \ldots | R_2 > \rho I(\mu_1)) \\ &= \begin{cases} 1 & R_2 < \rho I(1) \\ \Pr((1-\rho) I(\mu_2) < R_2 - \rho I(1)) & \rho I(1) < R_2 < \rho I(0) \end{cases} \end{aligned}$$

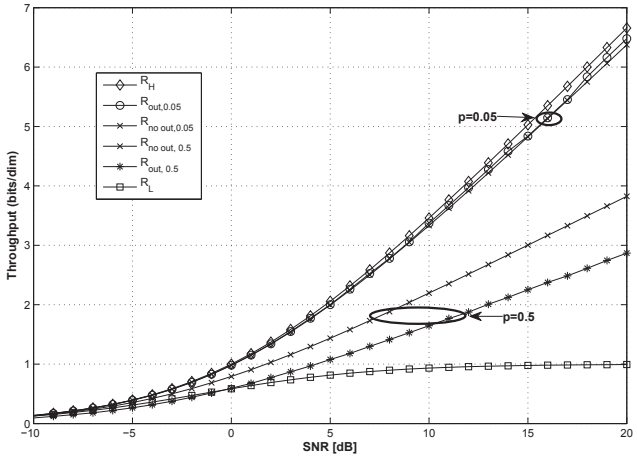


Figure 4.8: Throughput of the two rounds HARQ protocol in an AWGN channel with Gaussian signals. There is one interferer with probability $p = 0.05, 0.5$.

$$\text{where } \Pr((1-\rho)I(\mu_2) + \rho I(1) < R_2) = \begin{cases} p & \text{if } I(1) < R_2 < \rho I(1) + (1-\rho)I(0) \\ 0 & \text{if } I(1) > R_2 \\ 1 & R_2 > \rho I(1) + (1-\rho)I(0) \end{cases} \tag{4.35}$$

Finally, $P_{out}(\rho, R_2)$ is the probability of outage after the second round, independently of the interference state at the first round and it is given by:

$$P_{out}(\rho, R_2) = \begin{cases} p(1-p) & R_2 > \rho I(1) + (1-\rho)I(0) \\ p^2 & R_2 < \rho I(1) \end{cases} \tag{4.36}$$

Figure 4.8 shows the throughput of the HARQ protocol with two transmission rounds. There is one interferer with probability $p = \{0.05, 0.5\}$. We compare the zero-outage throughput against the throughput that allows an outage at the end of the protocol. We also plot the maximum capacity achieved with one round and no interference R_H and the corresponding capacity for interference R_L. If we look at the case of 50% probability of interference, we can see that the zero-outage throughput is higher for all SNR values, however, if we look at a case with a lower probability of having interference ($p = 0.05$, or 5%), we have almost the same throughput, except at high SNR, where the throughput with an outage is slightly higher. In this case, we also see that the capacity that can be achieved by adapting the rate and dimensions gets close to the capacity achieved without interference.

Discussion

If we consider the case with the ergodic capacity and no feedback, we transmit N_T dimensions per channel realization. Therefore, we have the average capacity:

$$\mathbb{E}_\mu = \begin{cases} \log_2(1+\mathrm{SNR}_1) & \mu = 0 \\ \log_2\left(1+\frac{\mathrm{SNR}_1}{1+\mathrm{SNR}_2}\right) & \mu = 1 \end{cases} \tag{4.37}$$

where μ is the state of the interference. Now, if we consider a channel with feedback of the state of the interference (non-causal feedback). Then at round r, the transmit signal is a function of the message W and the interference state μ:

$$\begin{cases} x_r = f(W,\mu) & r > 1 \\ x_1 = f(W) \end{cases} \tag{4.38}$$

To get an insight into how a rate-adaptive scheme performs when changing the number of dimensions across rounds, we focus on the case of the HARQ protocol with two transmission rounds. At round r, if $\mu = 1$, then there is no transmission, and it happens with probability $\Pr(\mu_r = 1) = p$. However, if there is no interference, $\mu_r = 0$, it transmits with $\frac{N_T}{1-p}$ dimensions, and in this case we get a throughput$= (1-p)\left(\frac{\log_2(1+\mathrm{SNR}_1)\frac{N_T}{(1-p)}}{N_T}\right) = \log_2(1+\mathrm{SNR}_1)$ which is the maximum achievable spectral efficiency. When feedback becomes available, it allows the scheme to perform better than the ergodic capacity. The latter is in contrast to the work in [20] where in the infinite delay case, the authors conclude that the maximum that can be achieved is the ergodic capacity. The difference comes from the fact that in [20] there is always a fixed bandwidth allocation for each user, regardless of the state of the channel. In our case, we dynamically adapt the bandwidth for each user depending on the interference conditions of past transmissions for the same codeword. From the perspective of the scheduler, the bandwidth is better distributed.

From our initial analysis, we can conclude that the highest spectral efficiency that can be achieved happens in the case of the zero-outage protocol where increasing the delay becomes beneficial to a certain point and brings only a marginal gain after this point. In the next section, we look at practical interference scenarios where having zero-outage throughput is not possible. However, a constraint on the outage probability can be imposed.

4.5 Practical Interference Networks Analysis

To model practical systems, we derive expressions for the mutual information assuming discrete constellations. We target LTE Release 10 networks with an OFDMA physical layer, and we study both the single-user and one dominant

interferer cases. For the sake of simplicity, we show the derivations focusing on one subcarrier, with unitary power and distance, and we drop the indexes.

Let H_r denote the vector of channel realizations in the rth round, then $I_r(H) = I_r(Y;X|H)$ denotes the corresponding instantaneous mutual information at round r. For IR-HARQ, mutual information is accumulated over retransmissions. In the case of bursty interference, this permits some averaging of the fading and interference affecting the signal [65].

For a particular user, we define the mutual information at round r, in bits, as:

$$I_r(H) = T_{dim} \sum_{j=1}^{r} \sum_{k=1}^{L_j} I_{j,k}(H_j) \tag{4.39}$$

where $I_{j,k}(H_j)$ is the mutual information for the user at round j and subcarrier k, and it is given by

$$I_k(H) = I_k(Y;X|H) = \frac{1}{\mathcal{S}_1\mathcal{S}_2} \sum_{x_1} \sum_{x_2} \int_y f(y|x_1,x_2,H) \times \log_2 \left[\frac{\sum_{x_2'} f(y|x_1,x_2',H)}{\frac{1}{\mathcal{S}_1} \sum_{x_1'} \sum_{x_2'} f(y|x_1',x_2',H)} \right] dy \tag{4.40}$$

In the rest of this section, we refer to the mutual information in bits/dim. For this purpose, we define $I'_r(H)$ as the mutual information in bits/dim as

$$I'_r(H_r) = \frac{1}{L'_r T_{dim}} I_r(H_r) \tag{4.41}$$

where L'_r is the number of dimensions up to round r, $\left(\sum_{j=1}^{r} L_j = L'_r\right)$. We can relate the generic throughput expression to the mutual information by defining the probabilities in (4.19) as:

$$P_{succ_r} = \Pr(I'_r(H) > R_r) \tag{4.42}$$

$$P_{succ_r, fail_{r-1}} = \Pr(I'_r(H) > R_r, I'_{r-1}(H) < R_{r-1}) \tag{4.43}$$

For a given channel realization h_r and a particular value of SNR, the maximum rate of reliable communication supported by the channel at round r is $I'_r(h_r)$ bits/s/Hz, which is a function of the random channel gain h_r and is therefore random. If the transmitter encodes data at a rate R_r bits/s/Hz, then at round r, if the channel realization h_r is such that $I'_r(h_r) < R_r$, the transmission is called unsuccessful and this happens with probability $\Pr(I'_r(h_r) < R_r)$.

In this case, zero-outage is impossible since power control and channel state feedback are not assumed [19]. However, we assume an outage constraint at the end of the HARQ protocol. To model this constraint, we

consider an IR-HARQ protocol with a maximum of $\mathbb{M}_{max}$ rounds, and we say that the constraint is met whenever the packet error probability after $\mathbb{M}_{max}$ rounds is smaller than a predefined threshold P_{out}.

We look at the case of two transmissions rounds where for a given R_1, we have successful transmission in the first round if $\Pr(I_1'(H_1) > R_1)$, after the second round, outage corresponds to $\Pr(I_2'(H_2) < R_2)$. Let $R_1 = \frac{\log_2 B}{L_1 T_{dim}}$ be the rate at the first round, and $R_2 = \frac{\log_2 B}{N_T T_{dim}}$ the rate at the second round. Then, the overall throughput expression is:

$$\bar{R} = \left(\Pr(I_1'(H_1)) > R_1)\right) R_1 + \left(\Pr(I_1'(H_1) < R_1, I_2'(H_2) > R_2)\right) R_2 \quad (4.44)$$

where the outage probability is $P_{out} = \Pr(I_1'(H_1) < R_1, I_2'(H_2) < R_2)$.

For the sake of obtaining long-term average throughput, we isolate the target channel h and average (4.44) over the channel distribution

$$\Pr\left(I_1'(H_1) > R_1\right) = \mathbb{E}_H \Pr\left(I_1'(H_1) > R_1 | h\right) \quad (4.45)$$

To find the operating rates of the protocol, we start by choosing the rate in the second round as the rate that will satisfy the outage constraint, i.e. R_2 that satisfies $\Pr(I_2'(H) < R_2)$. The rate in the first round (R_1), is chosen as the one that maximizes the throughput expression in (4.44) while satisfying the given constraint. In this case, we also optimize the number of dimensions used in each of the retransmission rounds.

Since there is no closed-form expression for the probability of outage of discrete signals, we notice that $\Pr(I(H_2) < R_2)$ represents the CDF of the mutual information evaluated at R_2, i.e. $F_I(R_2)$. With the help of the inversion formula in [62], we use the characteristic function of the mutual information $\Phi_I(\omega)$ to find the CDF as:

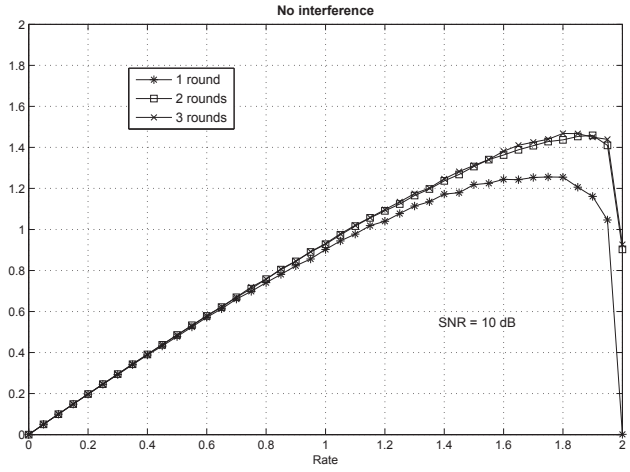
$$F_I(R_2) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\Im\left\{\exp(-j\omega R_2)\Phi_I(\omega)\right\}}{\omega} d\omega \quad (4.46)$$

The characteristic function is defined as $\Phi_I(\omega) = \mathbb{E}\left[\exp(j\omega I)\right]$, where $\mathbb{E}$ denotes expectation. Since we assume Rayleigh fading, the expectation is over the channel squared magnitude probability density function (PDF), which is exponentially distributed:

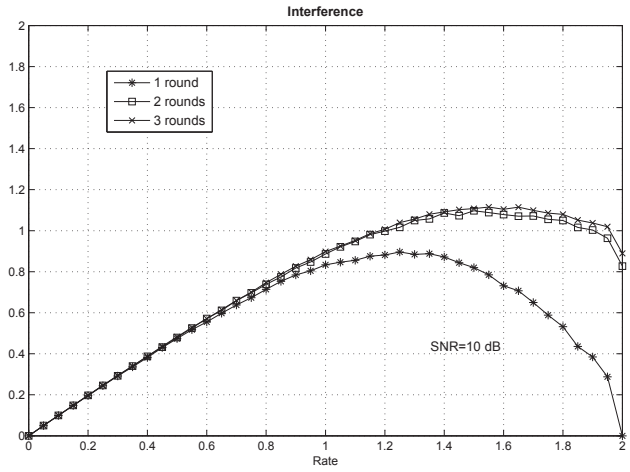
$$\Phi_I(\omega) = \int_0^\infty \exp(j\omega I(h)) \exp(-h) dh \quad (4.47)$$

Finally, we can use (4.46) and the outage constraint to solve the outage probability expression for R_2.

We start with the case of fixed rates across rounds when $R_1 = R_2$, and we want to show the existence of one optimal rate for the protocols. For this purpose, we compute the throughput for the different cases by setting different rates R_r at round r.



(a) Reliable throughput, no interference



(b) Reliable throughput, 1 interferer.

Figure 4.9: In (a) we have the reliable throughput for the HARQ protocol with different number of rounds under Rayleigh fading, constant over rounds, and without interference, (b) shows the reliable throughput different number of rounds and one dominant interferer. The rates across rounds are fixed and the operating SNR is 10 dB.

In figure 4.9, we plot the reliable throughput for the HARQ protocol without interference and one strong interferer. We can identify that there exists a value of the rate that gives the highest throughput or spectral efficiency. The latter means that there is an optimal value for the rate. If we look at the curves for one transmission round, we can observe an optimal rate of around 1.8 without interference and 1.3 with one interferer. We can also see that with more than one transmission round, the optimal rate is higher but after two rounds, it does not change. From these results, we can see that if we let the maximum number of transmission rounds to increase, one could have the option to fix the coding rate to a value close to the maximum rate and let the HARQ protocol work across the rounds until successful decoding. However, the latency in this case would be high. If we perform rate adaptation, we can choose the optimal rate in each of the HARQ rounds and decode the information sooner, therefore minimizing the latency.

4.5.1 Rate Optimization (fixed across rounds)

We can now proceed to optimize the rate for different operating SNR points. This means that at every SNR point, we choose the rate that gives the maximum throughput. In this case, we do not consider an outage constraint.

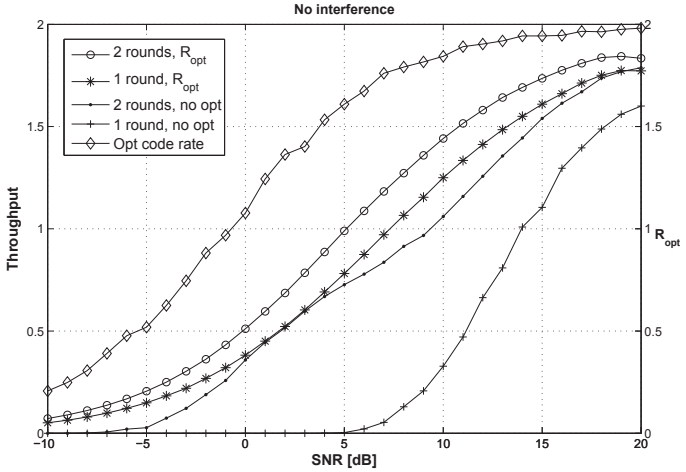
In figure 4.10, we show the spectral efficiency and the probability of outage when we optimize the rate per transmission. The rate is the same across rounds. We compare rate optimization with a fixed rate operation for a maximum number of HARQ rounds $\mathbb{M}_{max} = 1, 2$ and we can see that, for example, optimizing the rate with one HARQ round gives more or less the same gain as having an additional transmission round, but minimizing the delay. If we allow two HARQ rounds with rate optimization, then the gain in throughput is even higher. In this case, the channel is constant and assumed unknown to the transmitter.

Figure 4.11 shows the spectral efficiency and the probability of outage with one or two transmission rounds in a Rayleigh fading channel with QPSK modulation. Across the transmission rounds, the rates are fixed. We consider that there is one interferer present all the time, so the activity factor $\mu = 1$. If we allow the activity factor to take other values in time than one, then we can talk about the DL channel.

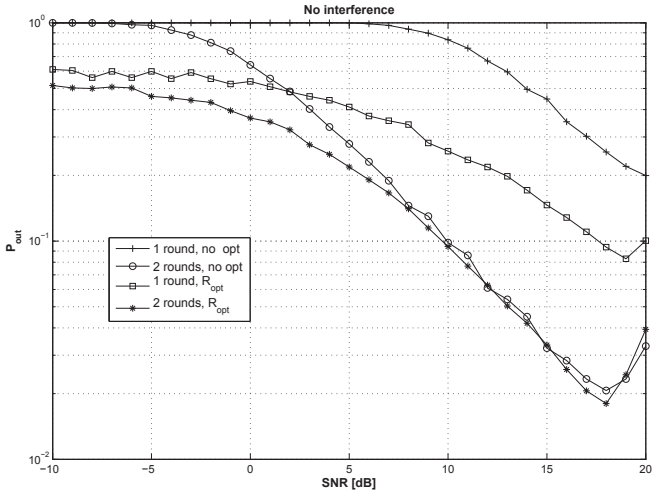
As a next step, we proceed to optimize the number of dimensions used in each of the HARQ rounds.

4.5.2 Rate Optimization with an Outage Constraint

We first give an example on how to obtain the rates for the case of iid channels across the HARQ rounds. According to the rate definition in (4.18), we define the rates R_1 and R_2 as $R_1 = \frac{\log_2 B}{T_{\dim} L_1}$ and $R_2 = \frac{\log_2 B}{T_{\dim}(L_1+L_2)}$ where

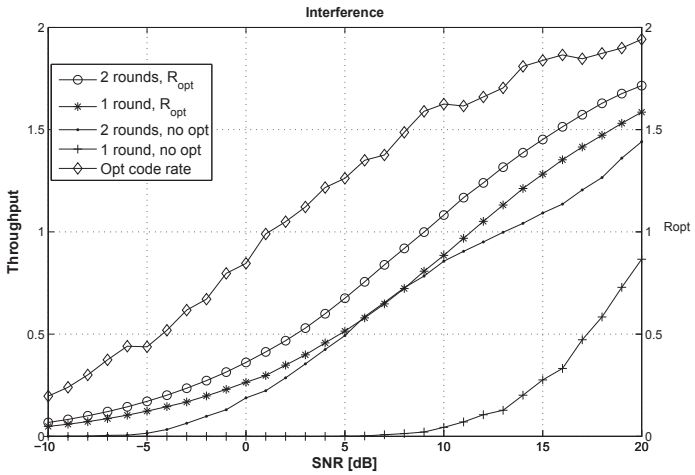


(a) Rate optimization, no interference.

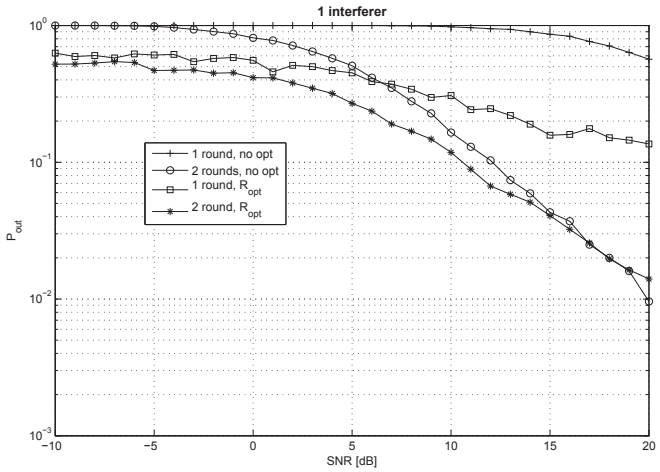


(b) Probability of outage, no interference

Figure 4.10: In (a) we show the rate optimization of the HARQ protocol for a diferent number of rounds $\mathbb{M}_{max} = 1, 2$ in a Rayleigh fading channel with QPSK modulation. The rates are fixed across rounds $R_1 = R_2 = R$. Figure (b) shows the corresponding probability of outage.



(a) Rate optimization, one interferer.



(b) Probability of outage, one interferer

Figure 4.11: In (a) we show the rate optimization of the HARQ protocol for a diferent number of rounds $\mathbb{M}_{max} = 1, 2$ in a Rayleigh fading channel with QPSK modulation. The rates are fixed across rounds $R_1 = R_2 = R$. There is one interferer all the time. Figure (b) shows the corresponding probability of outage.

L_r is the number of dimensions used in round r. If we define $\rho = \frac{L_1}{L_1+L_2}$ (i.e. $R_2 = \rho R_1$), then, the mutual information in bits/dim per round is given by:

$$I_1'(H_1) = \frac{T_{\text{dim}} L_1 I_1(H_1)}{T_{\text{dim}} L_1} = I_1(H_1) \tag{4.48}$$

$$I_2'(H_2) = \frac{T_{\text{dim}} \times (L_1 I_1(H_1) + L_2 I_2(H_2))}{T_{\text{dim}}(L_1 + L_2)} = I_1(H_1) + I_2(H_2)\left(\frac{1-\rho}{\rho}\right) \tag{4.49}$$

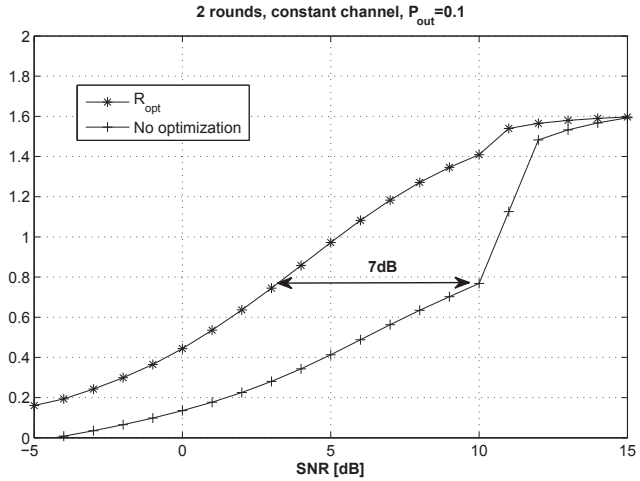
In this case, the characteristic function of the mutual information at the second round is given by:

$$\begin{aligned}
\phi_{I_2}(\omega) &= E\left[\exp(j\omega I_2'(H_2))\right] && (4.50)\\
&= E\left[\exp\left(j\omega\left(I_1(H_1) + I_2(H_2)\left(\frac{1-\rho}{\rho}\right)\right)\right)\right] \\
&= E\left[\exp\left(j\omega I_1(H_1)\right)\exp\left(j\omega I_2(H_2)\left(\frac{1-\rho}{\rho}\right)\right)\right] && (4.51)\\
&= E\left[\exp\left(j\omega I_1(H_1)\right)\right] E\left[\exp\left(j\omega I_2(H_2)\right)^{\left(\frac{1-\rho}{\rho}\right)}\right] && (4.52)\\
&= \phi_{I_1}(\omega)\phi_{I_2}\left(\omega\left(\frac{1-\rho}{\rho}\right)\right)
\end{aligned}$$

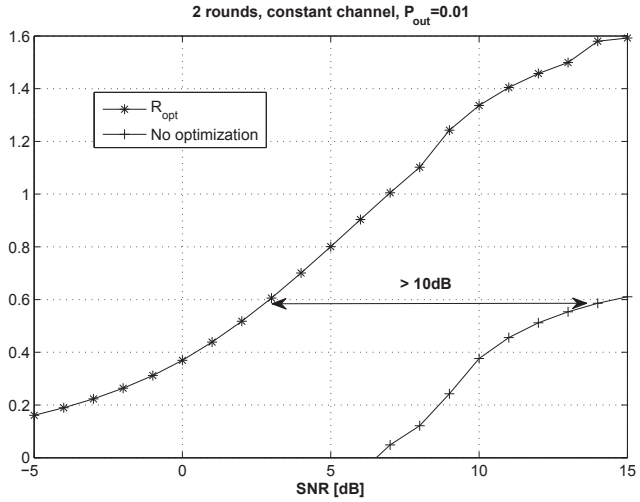
where going from (4.51) to (4.52) is possible because the channels are independent across rounds. Finally, we can use (4.46) and the outage constraint to find R_2.

In figure 4.12, we show the results for the rate optimization with an outage constraint of 10% and 1% when the channel remains constant across the HARQ rounds. There is a clear advantage on optimizing the rate across the HARQ rounds with a maximum gain of more than 10dB for the constant channel case. The gain is higher when we have a more strict outage constraint of 1%. If we compare these results to those in section 4.5.1, without outage constraint, we can see the significant gain introduced by optimizing the rate for latency-constrained scenarios. It can also be noticed that the throughput with rate optimization and an outage constraint of 1% is only lower by a small quantity as compared to the outage constraint of 10%. This last observation tells us that optimizing the rate can, indirectly, minimize the outage constraint (achieving almost the same throughput for both outage constraints).

In figure 4.13, we show the case for iid channels across transmission rounds and the two outage constraints as in the previous case. To obtain the maximum throughput, we also investigate the case of removing the constraint on the modulation. To do this, we choose the rate in the first round according to the mutual information expression for Gaussian inputs, which is not bounded by a particular modulation order, and we choose the modulation



(a) Rate optimization with an Outage constraint of 10%.



(b) Rate optimization with an Outage constraint of 1%.

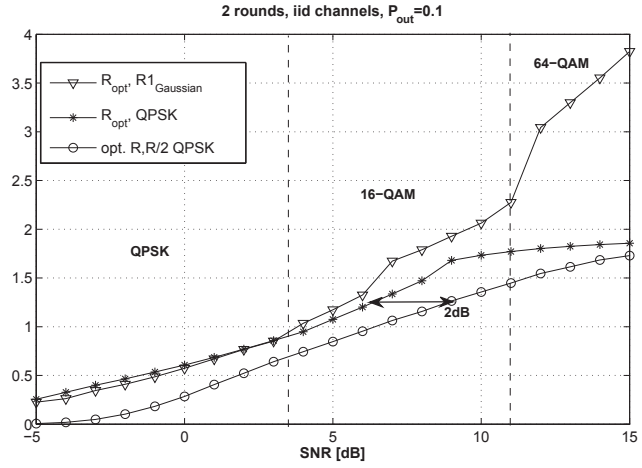
Figure 4.12: In (a) we show the rate optimization with an outage constraint of 10%. The channel is constant across the HARQ rounds and there is no interference. This is equivalent to a Non-Line-Of-Sight (NLOS) with slow fading channel. Figure (b) shows the corresponding curves for an outage constraint of 1%

that allows us to achieve this rate. We define threshold values for changing modulation between QPSK, 16-QAM and 64-QAM according to the maximum rate achieved with each modulation for a particular SNR value. We confirm once more that the gain is higher for the 1% case (around 7dB for the case of a more strict outage constraint of 1% while a gain of around 2dB is observed fot the 10% case). When we change the modulation with respect to the SNR, we observe a higher throughput in the high SNR region. This is caused by allowing the protocol to use higher modulation orders. In the following section, we show results from bringing interference into the scenario.

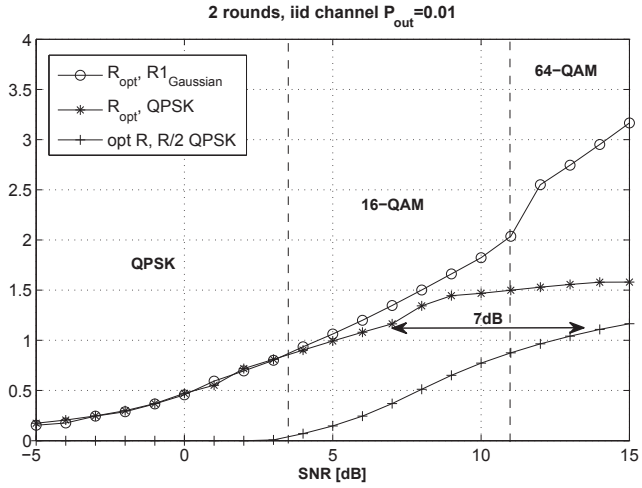
Interference case

We consider one dominant interferer and we assume a constant channel on the desired user. We analyze the case for the UL and DL differently. We first consider the DL of a femtocell with one interferer, and we consider an activity factor for the interferer which means that it is active only a portion of the time. We consider the channel of the femtocell user as AWGN since it is likely to experience a good channel. For the interference, we consider the channel as Rayleigh. Figure 4.14 shows the results for an activity factor of 50%, i.e. interference is present only half of the time. We fix the outage constraint at 1% and we plot the throughput for the user of interest. In figure 4.15, we show the DL case for an outage constraint of 10%. We observe the same behavior as the case without interference, with a higher gain for a more strict outage constraint (2dB for the 10% case against 10dB for the 1% case). For the UL, since the interference is coming from other users, it changes in time. Therefore, we consider the interference constant and iid across the HARQ rounds. Figure 4.16 shows the results for rate optimization in the 10% outage constraint case.

In this chapter, we have demonstrated the benefits of adapting the rate and physical dimensions across transmission rounds of HARQ protocols. We obtained a throughput higher than the ergodic capacity in the case of zero-outage throughput and we have showed that having an upper layer ARQ in case of a residual outage probability results in a lower throughput. In practical scenarios without power control and channel state information, it is not possible to get zero-outage throughput. However, we benefit from the dynamic resource allocation and by imposing a constraint on the outage probability, we can improve the throughput by varying the latency of the protocol. By motivating the use of inter-round resource allocation, we derived distributed resource allocation policies that are applicable for the UL and DL channels. We considered interference that can be bursty due to the characteristics of HetNets deployments. Rather than performing extensive simulations, we derived analytical expressions, based on mutual information



(a) Rate optimization with an Outage constraint of 10%.



(b) Rate optimization with an Outage constraint of 1%.

Figure 4.13: In (a) we show the rate optimization with an outage constraint of 10%. The channel is iid across the HARQ rounds and there is no interference. In this case, it is equivalent to a NLOS channel with fast fading or frequency hopping. Figure (b) shows the corresponding curves for an outage constraint of 1%

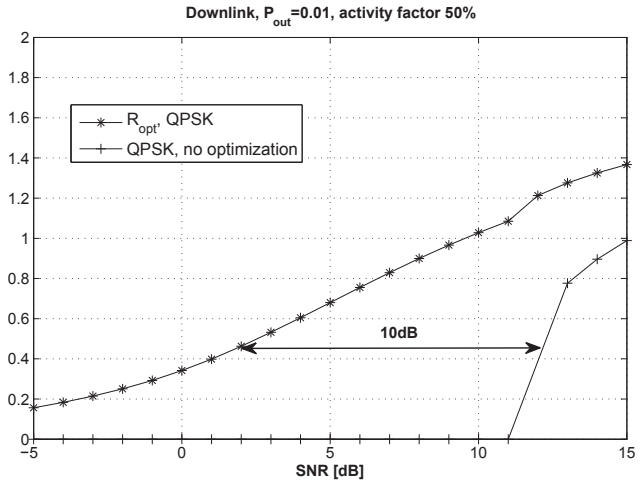


Figure 4.14: Rate optimization for Rayleigh fading on the downlink channel. There is one dominant interferer with an activity factor of 50%. The outage constraint is 1%

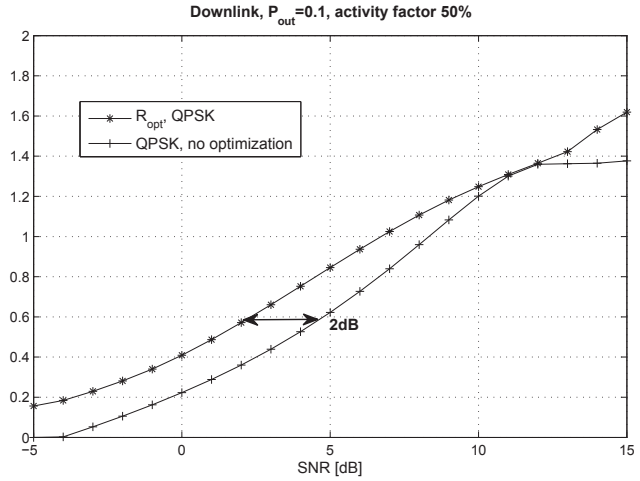


Figure 4.15: Rate optimization for Rayleigh fading on the downlink channel. There is one dominant interferer with an activity factor of 50%. The outage constraint is 10%

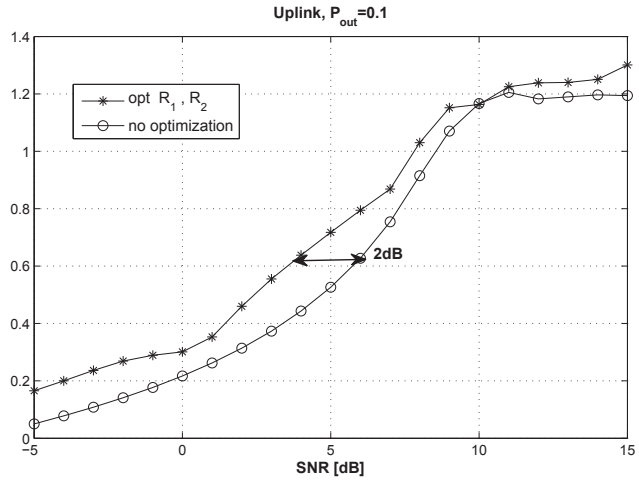


Figure 4.16: Rate optimization for Rayleigh fading on the uplink channel. There is one dominant iid interferer. The outage constraint is 10%

modeling, that capture the throughput performance of the network with or without a latency constraint.

4.6 Methodology for Resource Allocation in Practical Interference Scenarios

To illustrate some of the applications of the analytical framework presented in Chapter 4, we provide in the following sections several examples in more practical scenarios. First, we present a Manhattan-like topology for the case of small-cells inside a block of apartments. Then, we present a cross-layer interference case, with a macrocell overlaid by a single femtocell and we take a look at the DL channel of the macrocell user. Finally, we explain the procedure to perform PHY abstraction for interference scenarios. It should be noted that all the results in this section are using the rate optimization procedures described in section 4.4.2.

4.6.1 Manhattan-topology for Small-cells

We can model small-cell networks with non-constant interference by using the activity factors. A block of apartments is modeled as a rectangular grid with small-cells positioned at the center of each apartment. If we assume that the small-cells are active with a certain probability, then the interference process is no longer ergodic and interference is randomized in time. We study the performance of our rate adaptation policy in this apartment block scenario. For this purpose, we look at the DL of the small-cells Manhattan-

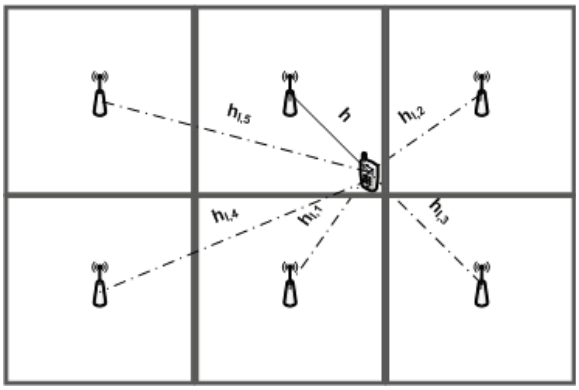


Figure 4.17: Manhattan-like topology with the user of interest at the edge of the apartment and the rest of the interfering small-cells placed at the middle of the surrounding apartments

like topology in figure 4.17 with the user of interest located inside the upper central apartment. To model indoor scenarios, we consider the path loss exponent in (4.21) to be $\alpha = 3$ [41]. The size of the apartments is $10 \times 10m^2$ and we measure, at every point, the distance to the neighboring interferers and scale the noise accordingly. We assume that all nodes transmit at equal power and the interference is known at every slot, i.e. we assume that the constellation and the channel can be estimated. This can be achieved, for instance, through a loose-cooperation policy where the base stations agree on the bandwidth and constellation they use.

Figure 4.18 shows the throughput of the small-cell user under Rayleigh fading, located along a 45° line from the center towards the corner of the apartment. The neighboring small-cells are located at the center of every apartment. We take into account one dominant interferer. In the case there are three interferers located at the same distance, we consider the maximum of three Rayleigh random variables and we apply a Gaussian approximation to the rest of the interferers by scaling the noise with the corresponding power. We consider QPSK 16-QAM, and 64-QAM modulations and an activity factor of 50%. We assume that, at the corner, the SNR for the user is equal to the strength of the interferer and is 10dB and we considered an attenuation loss (penetration of walls/floors) of 10dB. As expected, the throughput decreases proportionally to the distance from the small-cell. In distances close to the small-cell, the throughput with 64-QAM is substantially higher than when using 16-QAM or QPSK. After 5.5 m, they all achieve the same throughput.

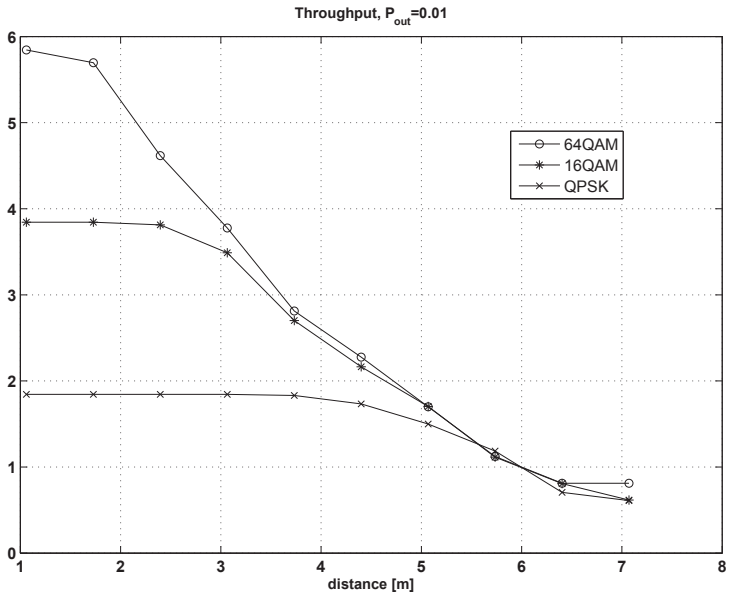


Figure 4.18: Downlink of a small-cell user. Interference is coming from the neighboring small-cells with an activity factor $\mu = 50\%$. The modulation is QPSK, 16-QAM, or 64-QAM and the SNR is equal to the interference strength and is 10dB. We consider the strongest interferer and we make a Gaussian approximation for the rest. The rate is adapted across rounds for all cases.

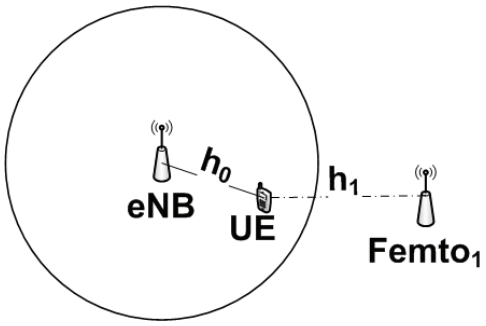


Figure 4.19: Downlink of a macrocell with a femtocell interfering.

4.6.2 Macro/Small-cell Scenario

In this case, we look at a scenario where the cellular network is overlaid by a femtocell network. We model a scenario where a single macrocell is overlaid by one femtocell and we look at the DL of the macrocell user when the interference is coming from the interfering femtocell (see figure 4.19).

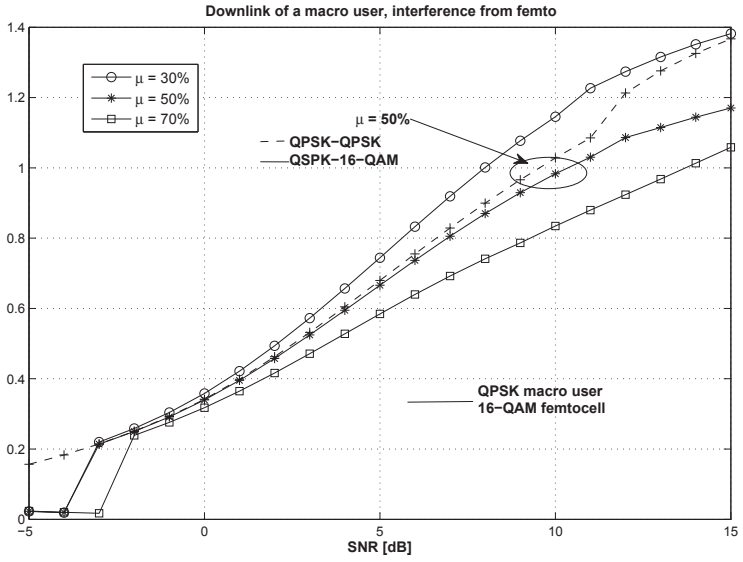


Figure 4.20: Downlink of a QPSK macro cell user under Rayleigh fading. Interference is coming from a 16-QAM femtocell active only part of the time. Curves are shown for different activity factors 30%, 50% and 70%.

We model a femtocell that is active only with a certain probability specified by the activity factor. We consider an activity factor of 30%, 50% and 70%. The macro cell user is considered to be active all the time. Figure 4.20 shows the results when the modulation for the macro cell user is QPSK and for the interfering femtocell, we consider 16-QAM, since a femtocell user will

experience a better channel. For an activity factor of 50%, we make the comparison of a femtocell using either QPSK or 16-QAM and we observe that for low to moderate SNR, the achieved throughput is very close. The higher constellation of the interference results in a lower throughput only after 10 dB. We have to remember that we assumed the strength of the interference equal to the user of interest.

4.6.3 Physical layer Abstraction Models

Simulating an entire physical layer can be extremely time consuming and in some cases computationally unfeasible. By using analytical expressions that model the performance of interference networks, we can abstract the physical layer. PHY abstraction methods allow us to accurately predict the coded Block Error Rate (BLER) for any given channel realization.

The BLER can be obtained through mapping mechanisms from instantaneous channel state measures such as SINR. When the mapping mechanism is a function of the mutual information capacity, it is called Mutual Information Effective SINR Mapping (MIESM) and it makes use of stored look-up tables. Performing PHY abstraction is out of the scope of this work, however, we explain the procedure that has to be followed in order to do the abstraction.

The first step to perform a PHY abstraction is to compute the SINR given a set of parameters such as power, interference, fading, etc. This SINR value is then used to compute the corresponding mutual information. The mutual information will be different at every round of the HARQ protocol since we use a different number of dimensions per round and the interference conditions may vary (due to the burstiness of the traffic). Finally, we have to map the mutual information to the corresponding BLER and store it in a look-up table. The effective SINR mapping for MIESM is given by:

$$\mathrm{SINR}_{eff} = \beta_1 f^{-1}\left(\frac{1}{P}\sum_{p=1}^{P} f\left(\frac{\mathrm{SINR}_p}{\beta_2}\right)\right) \tag{4.53}$$

where the f function is a mapping function to an information metric, i.e. the mutual information expression for discrete constellations, $\beta_{1,2}$ are calibration factors that have to be adjusted according to the channel model and modulation schemes used. The look-up table is then constructed mapping the mutual information values to the corresponding BLER (or outage probability). We have derived the necessary expressions to compute the BLER for a range of mutual information values.

In the next chapter, we investigate the performance of our resource allocation strategies on practical LTE MODEMs in an interference environment under the constraints of LTE coded-modulation.

Chapter 5

Practical Scheduler Design for LTE Base Stations

A key feature of LTE is the adoption of advanced radio resource management procedures in order to increase the system performance. The efficient use of radio resources is essential to meet the system performance targets [21]. In this chapter, we study the performance of our resource allocation policies for the scheduling of HARQ transmissions under the constraints of LTE coded-modulation.

The HARQ and rate adaptation algorithm is implemented and managed by the scheduler at the MAC layer. The scheduling mechanism plays a fundamental role because it is responsible for choosing, with fine time and frequency granularity, how to distribute the radio resources among different users, taking into account the current channel conditions and QoS requirements. The scheduling algorithm is a base station-implementation issue and it is not specified in the standard. The standard only gives channel-quality measurements/reports and specific procedures for dynamic resource allocation [28] needed for the terminal's implementation (and consequently the base station as well). This allows to have vendor-specific algorithms which can be optimized for specific scenarios.

The design of effective resource allocation mechanisms becomes crucial for an efficient network operation. Low complexity and scalability are fundamental requirements for finding the best allocation decision. Complex and non-linear optimization problems or exhaustive research over all the possible combinations would be too expensive in terms of computational cost and time [45]. An algorithm should be easily implemented and, above all, should require very low computational cost. Many theoretical solutions can

be found in the literature, but when studied closely, they cannot be deployed in real systems given the difficulty to be implemented in real devices and the high computational cost required. Robust strategies should guarantee the ability to work in very different scenarios. It should not require strong parameter settings, or it should at least dynamically adapt such parameters to environmental changes [21].

5.1 Handling Interference in LTE Networks with HARQ and AMC

Traditionally, HARQ has been used as a way to recover from errors occurring during the transmission of information, decreasing the probability of unsuccessful decoding. HARQ can be combined with Adaptive Modulation and Coding (AMC). When implementing AMC with HARQ, the MCS is adapted between each retransmission. The coding rate can be adapted across transmissions to further improve the performance. The code rate can be fine-tuned by puncturing, generating different redundancy versions to match the number of coded bits to the channel. The code rate, rate matching, and the number of resources allocated for one transmission determine the transport-block size (TBS) [61]. In essence, rate adaptation adapts the MCS to the current channel conditions which translates to the data rate or error probability of each link. The MCS along with signaling overhead (reference signals, control channel resource elements), in terms of spectral efficiency, represents the number of information bits per modulation symbol. The use of rate adaptation provides manufacturers an incentive to implement more advanced receivers since it will result in higher end-user data rates than standard receivers [27].

Based on the available CQI, it is the scheduler at the eNB that can address the different quality and service requirements of all the associated UEs in a spectrally efficient way.

LTE also supports the adaptation of the so-called TBS and the amount of PRBs used per transmission. A transport-block is the name given to a block of data at the MAC layer [9] and corresponds to a codeword, or more precisely a set of codewords, at the physical layer. For LTE, HARQ is supported for the Physical Uplink Shared Channels (PUSCH) in both UL and DL (PDSCH), and separate control channels with configurable levels of error probabilities on the uplink are used to send the associated acknowledgment feedback. HARQ can be classified as either synchronous or asynchronous and the retransmissions can be adaptive or non-adaptive [61]. In a synchronous system, retransmissions occur at a predefined time. In an asynchronous system, the retransmissions can occur at any time (and must be signaled). With adaptive HARQ, the MCS and other transmission attributes can be changed after each round. In a non-adaptive context, transmission attributes

are fixed or pre-defined. In LTE, HARQ is asynchronous and adaptive in the DL and synchronous in the UL. Retransmissions can be adaptive or not in the UL [61]. Feedback for HARQ in LTE comprises a simple ACK/NACK signal.

An important factor for the eNB scheduling algorithm is the accuracy of the available CQI for the active UEs in the cell. In the DL, CQI is reported back by the UE. For the UL, the eNB can use Sounding Reference Signals (SRS) or other signals transmitted by the UEs to estimate CQI [61]. On the UL, two channels are used to send CQI reports. Namely the Physical Uplink Control Channel (PUCCH) and the PUSCH. Reporting can be periodic or aperiodic. Periodic reports are normally transmitted on the PUCCH. However, the eNB can request the UE to send aperiodic CQI reports on the PUSCH because it is more appropriate to transmit large and detailed reports [38]. The key issue with respect to resolving channel quality is the ability to obtain accurate information in the presence of a large number of connected UEs (even idle) and the presence of sporadic interference (primarily on the uplink). The eNB can adjust the periodicity and granularity of the CQI feedback (down to 2 ms periodicity for both CQI and SRS in Rel-10 LTE [7]) allowing it to trade-off between the amount of overhead and the accuracy of the channel information. Of course, when a long delay occurs with respect to the scheduling time, the performance can be significantly affected. Short feedback periodicity is difficult to achieve in heavily loaded cells.

In summary, LTE offers a lot of flexibility in terms of resource allocation and, in particular, resource allocation algorithms can be tailored for a particular class of traffic with specific requirements. Nevertheless, work is still needed to exploit this flexibility efficiently for key emerging applications. To this end, we consider the dynamic resource allocation schemes for IR-HARQ presented in Chapter 4 under the constraints of LTE coded-modulation. Our goal is to analyze the feasibility of such techniques in real LTE MODEMs and to measure the performance and possible limitations with respect to the implementation in real systems. Furthermore, we want to measure how close we can get to theoretical predictions of the performance and distribution of the physical dimensions across HARQ rounds.

5.2 OpenAirInterface Implementation

We first present the OAI SDR platform used to test the performance and compliance of our resource allocation strategies in LTE. OAI is an open-source hardware/software development platform and open-forum for innovation in the area of digital radio communications [3]. It was created by the Mobile Communications Department at EURECOM based on its experience in publicly-funded R&D carried out in the context of collaborative research

projects [15]. One of the most significant benefits from a full SDR platform is that the same code can be used for emulation as in a real implementation, providing a smooth transition from the simulations to the real tests.

In the following sections, we give the detailed procedure for the resource allocation algorithm implementation on the downlink channel.

5.2.1 Physical Layer and Resource Allocation

In LTE, data is encapsulated in the so-called Packet Data Units (PDUs) at the MAC layer and forwarded to the PHY for transmission over the air. The physical layer provides services to the MAC layer in the form of transport channels and is responsible for the coding, modulation, and TBS determination and assignment of the physical resources for transmission of information. LTE defines a number of downlink physical channels to carry information blocks received from the MAC and higher layers. These channels are categorized as transport or control channels. In the DL, the PDSCH is the main physical channel allocated to users on a dynamic and opportunistic basis. It carries all user and control-plane information and it is scheduled using the Downlink Control Information (DCI) on the Physical Downlink Control Channel (PDCCH). The UE decodes the PDCCH every sub-frame to know which PDSCH to decode and where to find them in time or frequency. There are different formats for the DCI, but we use DCI format 1 defined for the scheduling of one PDSCH codeword (Transmission modes 1, 2 or 7). The DCI includes information about the resource allocation type, RBs assignment, MCS, HARQ information and redundancy version, and the power control command for the PUCCH. The PDSCH carries data in TBs. They are passed from the MAC layer once per TTI which in LTE is 1 ms. The physical resources are assigned on a basis of two resource blocks for one TTI. This is referred to by "pair of resource blocks" which is the minimum of resources that can be allocated.

In general, a compromise between flexibility and signaling overhead has to be met when sending the resource allocation assignments in terms of a set of RBs. The simplest solution would be to send a bitmap for each UE in which each bit indicates a particular RB, but this would be costly for large bandwidths (> 5 MHz). The resource allocation in DCI is described by two parts: the header (type of allocation : RA type 0,1,2), and the resource block assignment itself. We give the details of the resource allocation type that we adopted which is part of the LTE standard Rel-8.

Resource Allocation Type 0

As mentioned before, the bandwidth is divided into groups of RBs and a bitmap indicates the actual resource block assignment, i.e. the Resource Block Groups (RBGs) that are allocated to the scheduled UE, where RBG

is a set of consecutive PRBs. The RBG size (P) is a function of the system bandwidth (see Figure 5.1 [7]). The total number of RBGs (N_{RBG}) for a downlink system bandwidth of $N_{\mathrm{RB}}^{\mathrm{DL}}$ PRBs is given by $N_{\mathrm{RBG}} = \left\lceil \frac{N_{\mathrm{RB}}^{\mathrm{DL}}}{P} \right\rceil$ where $\left\lfloor \frac{N_{\mathrm{RB}}^{\mathrm{DL}}}{P} \right\rfloor$ of the RBGs are of size P and if $N_{\mathrm{RBG}} \mod \mathrm{P} > 0$ then one of the RBGs is of size $N_{\mathrm{RBG}} - P * \left\lfloor \frac{N_{\mathrm{RB}}^{\mathrm{DL}}}{P} \right\rfloor$.

Table 5.1: RBG Size vs Downlink System Bandwidth

System Bandwidth ($N_{\mathrm{RB}}^{\mathrm{DL}}$)	RBG size (P)
≤ 10	1
11-26	2
27-63	3
64-110	4

Modulation and TBS Determination

Modulation and coding formats are limited with respect to achieving all spectral efficiencies (which is possible with LTE rate matching), but the granularity is quite fine. There are 5 bits that provide a table entry (I_{MCS}) giving the modulation order (Q_m) : QPSK (2), 16-QAM (4), 64-QAM (6) and TBS indication (I_{TBS}). This information is combined with the number of allocated RBs to yield the transport block size (table lookup). The UE in the physical layer delivers I_{TBS} along with HARQ process ID (from PDCCH) and New Data Indicator (NDI) bit to the MAC layer in parallel to decoding PDSCH/DLSCH.

The selection of the TBS, MCS and antenna mapping, together with the logical-channel multiplexing for downlink transmissions are controlled by the eNB. In the case of a retransmission, the TBS is, by definition, unchanged, but the same TBS should ideally appear for several different resource-block allocations as this allows the number of resource blocks to be changed between retransmission attempts, providing increased scheduling flexibility [7].

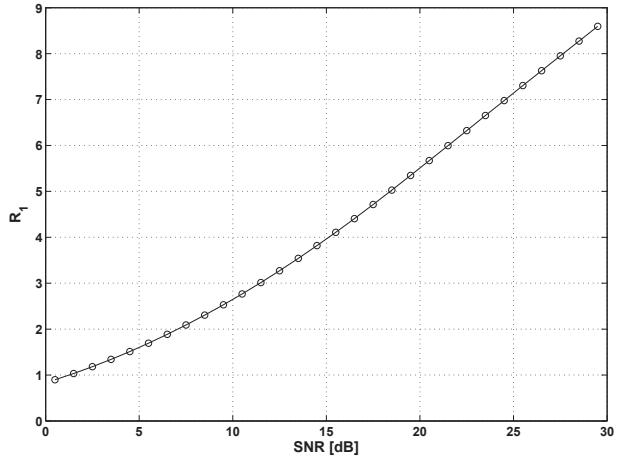
5.3 Application of the Scheduling Policies in LTE

The policies considered in the previous chapters require an incremental-redundancy coded-modulation system with the possibility of adjusting the code rate (physical resources) via puncturing and repetition (rate matching). This is possible on the LTE Rel-8/10 DL and UL where the allocation can be adapted across transmission rounds, the granularity of which depends on the transport block size. The only difference between the UL and DL allocation

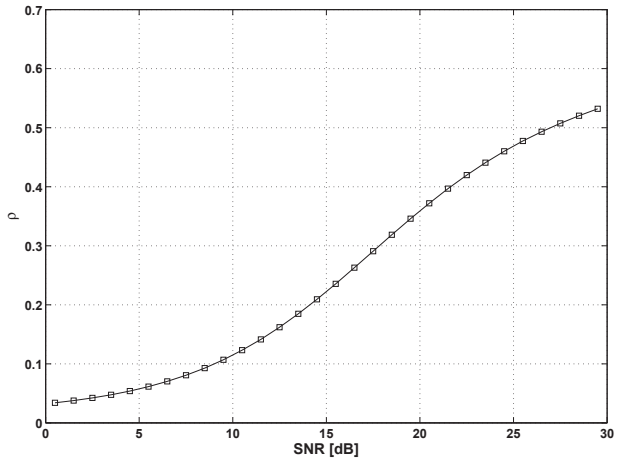
is that the modulation order must remain fixed on the UL for each round. It is likely that this restriction is insignificant with respect to performance. The slight penalty for inter-round rate adaptation is the requirement to send a new DCI packet with updated resources rather than an automatic retransmission. Let us now explain with an example the rate-adaptation policy for two-round transmission in an LTE context. For simplicity, we consider the case with Gaussian signals and without interference in Chapter 4.3.

For illustration purposes, consider a DL transmission in transmission mode 1 (SISO transmission). The policy must choose the initial spectral efficiency, MCS_1 corresponding to the first transmission round and N_1 and N_2 corresponding to the number of allocated physical resource blocks. An additional constraint (on the downlink) is that these must be multiples of an integer P which is the allocation granularity dictated by the transmission bandwidth (for a 10MHz carrier, $P = 3$ [7].) Figure 5.1 shows the result of an optimized two-round protocol, specifically the rate R_1 which must be translated to MCS_1 and ρ which provides the ratio of the physical dimensions used in the two rounds. This can now be transformed as a function of the target transport block size. Now consider an average SNR of 7 dB, this corresponds to R_1 of 2 bits/dimension and $\rho = 0.07$. Assume an allocation in the first round using 3 resource blocks (minimal allocation for a 10 MHz carrier). In a PDSCH-only normal prefix subframe with one PDCCH symbol, the total number of resource elements is $N_1 = 450$ (13 PDSCH symbols, 3 with 10 PDSCH resource elements per resource block and 10 with 12). The target transport block size would therefore be around 900 bits, so the closest transport block size for 3 resource blocks is 904 bits ([7, Table 7.1.7.2.1-1]) with $\mathrm{MCS}_1 = 16$ (16QAM). The second-round dimensions would ideally be $N_2 = N_1(\frac{1-\rho}{\rho}) = 5978$ yielding 39.9 (39 or 42) resource blocks.

The previous example exposes the difficulty of applying such latency-constrained policies for large transport block sizes, at least under the constraints of the current LTE specifications. For instance, if we were to use the same operating point (7 dB average SNR) with twice the number of dimensions in the first round, corresponding to a transport block size of 1800 bits, the number of required resource blocks in the second round jumps to 78 or 81 which is only realizable on a 20 MHz carrier. This can be easily overcome by allowing a third transmission round, i.e. by increasing the acceptable latency of the transmission. Another important consideration regarding application on the UL is that the nature of the resource allocation policy is fundamentally related to power control, since we are assuming a constant transmit energy per channel dimension. This is also the adopted policy in LTE (assuming power adjustments are not made during retransmission rounds). Basically, low power is used during the first transmission and significantly more power is used in the second transmission if required. In the numerical example



(a)



(b)

Figure 5.1: In (a), we consider the scenario without CQI (uncorrelated channels), and we plot the rate in the first round (R_1) for different values of SNR. We fix the probability of outage after the second round to 1%. In (b), for the scenario without CQI (uncorrelated channels), we plot the ρ parameter against different values of SNR. We fix the probability of outage after the second round to 1%. ρ determines the rate used in the second round according to equation [4.6]

above the power boost is $10\log_{10}\left(\frac{1-\rho}{\rho}\right) = 11.2\text{dB}$. This, of course, requires that the UE has signaled sufficient *power headroom* [7] for the eNB to allow this allocation. This clearly shows that, on the UL, latency can be controlled through a combination of rate adaptation, HARQ and power control. An instance of this appearing in the literature in the case of MIMO transmission with HARQ can be found in [29], although in that example the number of dimensions across tranmsission rounds remained fixed and the energy per channel dimension increased across rounds.

5.4 Performance Analysis of the Scheduler

A scheduling algorithm responds to a pre-formulated capacity-related metric which is optimized across all possible resource allocation solutions satisfying a set of predetermined requirements such as QoS, spectral efficiency, or latency. To compute the spectral efficiency, we have to consider the TBS, the modulation order and the outage probability:

$$\text{Spec}_{\text{eff}} = \frac{Q_{m,1}}{\text{G}_1}\text{TBS}\Big(1 - P_{\text{out},1}\Big) + \left(\frac{Q_{m,1}}{\text{G}_1} + \frac{Q_{m,2}}{\text{G}_2}\right)\text{TBS}\Big(P_{\text{out},1}(1 - P_{\text{out},2})\Big) \tag{5.1}$$

where G_i is the number of coded bits per codeword, and $Q_{m,i}$ the modulation order used at the ith round, previously defined in section 5.2.1. $P_{\text{out},1}$ is the outage probability after the first round, and $P_{\text{out},2}$ is the outage probability after the second round.

As an example of the optimization of (5.1), we fix the TBS for all possible allocations to around 1000 bits and from the table 7.1.7.2.1-1 in [7], we get the TBS index depending on the number of PRBs used in the first round and the TBS closest to 1000. Finally, table 7.1.7.1-1 (see [7]) gives us the corresponding MCS index and modulation order. Table 5.2 shows the different allocations used to generate the results in figures 5.2 and 5.3, where $\left(N_{\text{RB}}^{\text{DL}}\right)_1$ represents the number of PRBs allocated in the first round. We use this table to test different allocations and investigate the performance of the LTE codes. By using the DCI, we can signal the new information with respect to the number of PRBs in the consecutive rounds of the HARQ protocol. The MCS can also be adapted, but the TBS remains fixed.

Table 5.2: Resource Allocation vs Modulation and Transport Block Sizes

$\left(N_{\mathrm{RB}}^{\mathrm{DL}}\right)_1$	$\left(N_{\mathrm{RB}}^{\mathrm{DL}}\right)_2$	TBS	$\mathrm{I_{TBS}}$	$\mathrm{I_{MCS}}$	$\mathrm{Q_m}$
21	4	936	2	2	2
19	6	1096	3	3	2
17	8	968	3	3	2
15	10	1064	4	4	2
13	12	904	4	4	2
12	13	1032	5	5	2
10	15	1032	6	6	2
8	17	968	7	7	2
6	19	1032	10	11	4
4	21	1000	13	14	4
2	23	1000	21	23	6

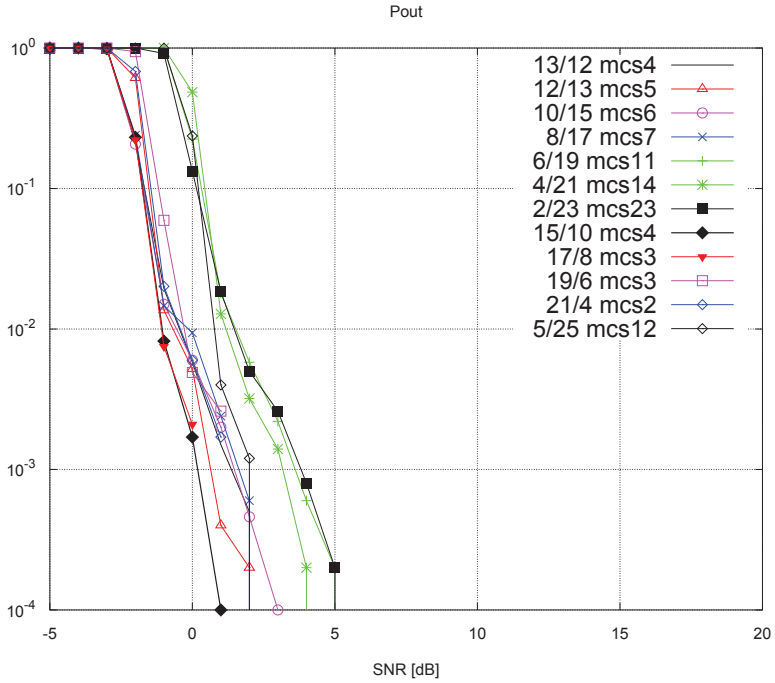


Figure 5.2: Probability of Outage after two HARQ rounds. The channel is AWGN and there is no interference.

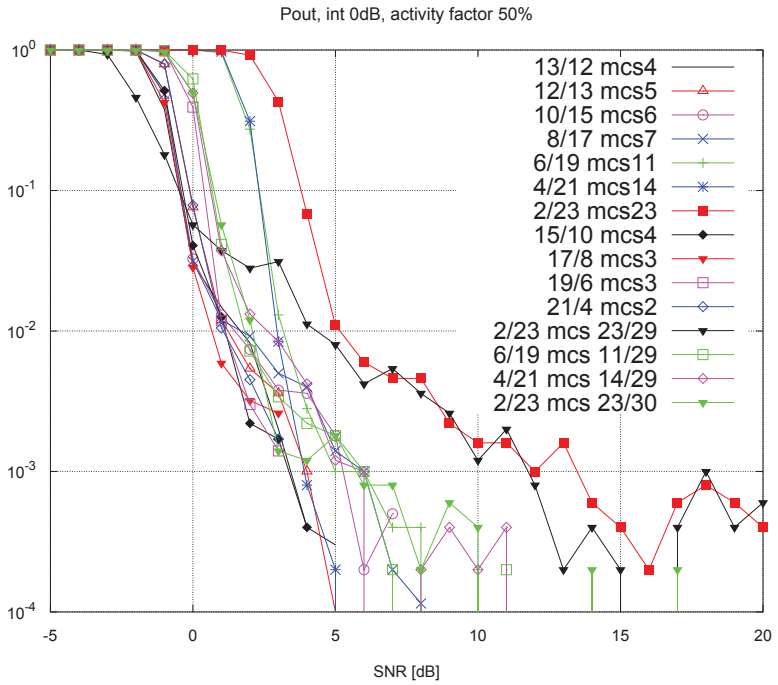


Figure 5.3: Probability of Outage after two HARQ rounds. There is one interferer of the same strength as the user of interest with a 50% probability of being active. The channel is AWGN for both users.

We can see in figures 5.2 and 5.3 the performance of the LTE constrained coded-modulation without interference and with one dominant interferer which is active 50% of the time. Theoretically, all the codes should behave in the same way, meaning that they should have the same outage probability. However, the results show that for the smallest allocations and therefore the highest modulations (MCS values higher than 11) there is a difference with respect to the lower MCS values. If we consider the modulation used by these MCS values, we can see that MCS $= 11, 12, 14$ use 16-QAM and MCS 23 uses 64-QAM modulation. In the case of interference shown in figure 5.3, we show also the performance of the codes when forcing the retransmissions of these MCSs to use QPSK (MCS 29) or 16-QAM (MCS 30). As we see, in general, the best result is achieved by using QPSK in the second round (except for MCS 23 where is better to use 16-QAM). In our results, to reduce the gap for different MCSs, we force the retransmissions to use QPSK by setting the MCS value to 29 as indicated in the 3GPP standard (MCS values $29, 30, 31$ are reserved in the LTE standard to adapt the modulation order across rounds). This has no impact on spectral efficiency but changes the behavior of the rate-matching algorithm by allowing the code to operate with less repetition which may be more efficient.

In the following sections, we show the results for the optimization of the number of resources in the second round (the total number of PRBs used in both rounds can be less than $N_{\mathrm{RB}}^{\mathrm{DL}}$. To optimize this number, we follow the next procedure:

1. fix the TBS (eg. 1000 bits)
2. run over the different MCS and SNR values
3. compute the outage in the first round
4. run over all combinations of PRBs in the second round until an outage of 1% is satisfied (we fix$\mathrm{Q}_m = 2$ for retransmissions)
5. choose the corresponding rate at the second round R_2 and compute the spectral efficiency over the rounds

In the following sections, we give the results of this procedure in the case without interference and one dominant interferer, where we also take into account the activity factors which model bursty interference.

The combination of the fixed-rate turbo-code with dynamic resource allocation and the rate-matching permutation amounts to doing rateless coding such as those proposed in [30] and [58] where they consider an unbounded number of transmission rounds. Our dynamic adaptation works over a small number of transmission rounds (2-4) and can actually be applied to such coding schemes. The latter would require appropriate link-layer HARQ protocols (e.g. [40]) adapted to such dynamic rateless coders.

It should be noted that all simulations in this chapter assume imperfect channel estimation. If we would like to truly compare to rateless codes, we should assume perfect channel estimation or channel state information at the receiver (CSIR).

5.4.1 Results without interference

In this section, we present the results for the resource allocation strategies with different channel models. We simulated 10000 frames for each of the SNR points and MCS values using the OAI unitary LTE PHY link simulator and we give the results in terms of the spectral efficiency. For reference purposes, we also give the probability of outage after the first round $\mathrm{P}_{\mathrm{out},1}$. The outage probability after the second round $\mathrm{P}_{\mathrm{out},2}$ is fixed to 1% for all simulations.

AWGN channel

First, we present the case with AWGN channel and no interference. We optimize the allocation of PRBs in the second HARQ round by following the procedure described in section 5.4. We save the number of erroneous transmissions after each round to compute the probability of outage and we use it together with the number of coded bits per codeword to compute the spectral efficiency as in (5.1).

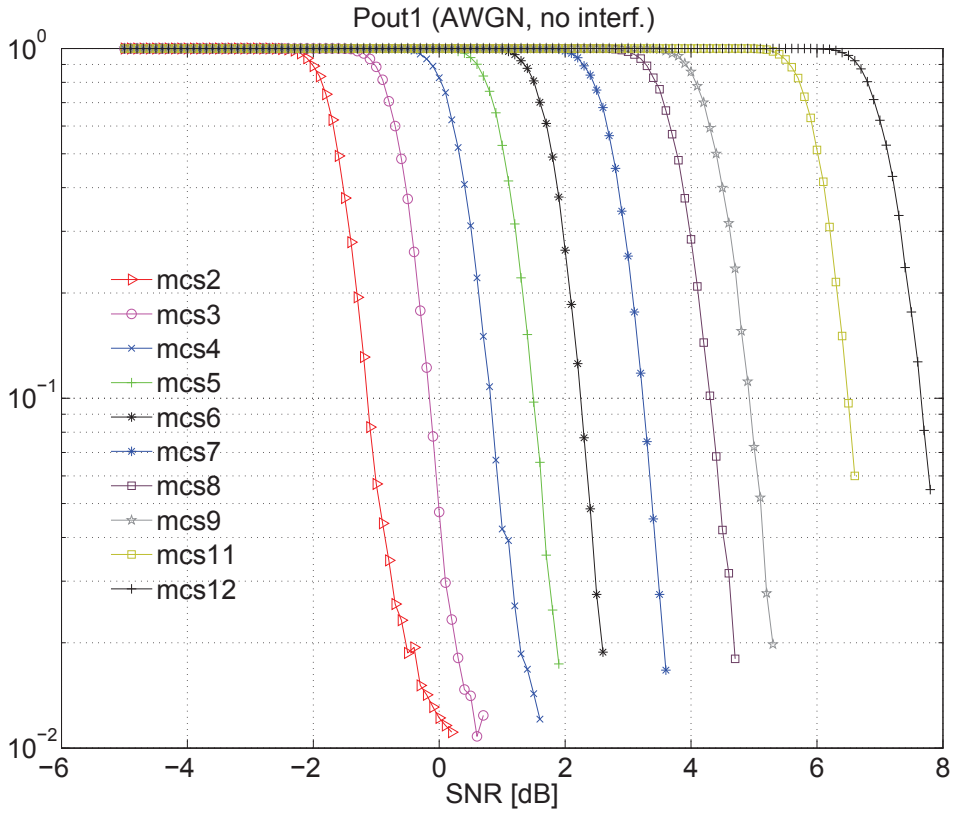


Figure 5.4: Probability of Outage after the first round of the HARQ protocol. The channel is AWGN and there is no interference.

Figure 5.4 shows the results for the probability of outage after the first HARQ round. In figure 5.5, the left axis and solid lines show the spectral efficiency depending on the resource allocation chosen. The right axis and dashed lines show the corresponding (optimized) number of PRBs used in the second round. We show the MCS used in each case (for the retransmissions the MCS is fixed to force them to use QPSK), the TBS remains fixed across the HARQ rounds. We also show the spectral efficiency values specified in the LTE standard for terminal feedback signaling (see Table 7.2.3-1 in [7]). If

we look at the lines showing these values, we can see the corresponding MCS that would be used in each case. For example, if we consider the spectral efficiency of 0.3770, we would be using MCS 3 which gives the highest achievable spectral efficiency. If we consider the next spectral efficiency 0.6016, we would use MCS 6. However, for the large number of intermediate values that the protocol can use, one could change the allocation and/or MCS depending on the SNR to maximize the spectral efficiency. This could be done by the base station scheduler based on estimates of the first and second round error probabilities. The latter could be achieved based on the statistics of received ACK/NACK signals. By doing so, one has more liberty to adapt parameters such as PRB allocation or MCS and modulation across rounds to those values that result in a higher spectral efficiency given the latency constraint.

The desired latency of the protocol translates as the probability of outage for the different operating SNR points. For our latency constraint of 10^{-2}, we can look at different SNR values and see what the best allocation is. If we consider an SNR= 0dB, the corresponding MCS is 3 and the allocation corresponds to 18 PRBs in the first round and 3 in the second round. This corresponds to a ratio of dimensions of $\rho = 0.8571$. If we now consider an SNR= 5dB, the MCS becomes 9 with 6 PRBs in the first round and 2 in the second round ($\rho = 0.75$). As the SNR gets higher, we use less resources in the second round. Interestingly, if we look at the outage probability after the first round, we see that the second case of MCS 9, has a higher probability of outage when compared to the case with MCS 3, which means that in this case, almost all the data transmission occurs at the second round of the HARQ protocol.

Rayleigh Channel

We investigate the effect of the distribution of dimensions across rounds (related to section 4.3.2, Chapter 4) and we look at the results when the channel is Rayleigh distributed giving a worse channel quality than the AWGN case. This is equivalent to the outdated CQI case with uncorrelated channels presented in section 4.3.2 of Chapter 4.

Figure 5.6 show the probability of outage after the first round of the HARQ protocol. Figure 5.7 shows the spectral efficiency for the different optimized allocations and those specified in the LTE standard. When we consider an SNR of 2dB, the best combination of MCS and PRBs allocation is MCS 7 with 8 PRBs in the first round and 25 in the second round, which gives a ratio of dimensions $\rho = 0.2424$. However, if we look at high SNR (15dB), the best MCS becomes 16 with 3 PRBs in the first round and 6 in the second round, with $\rho = 0.3333$.

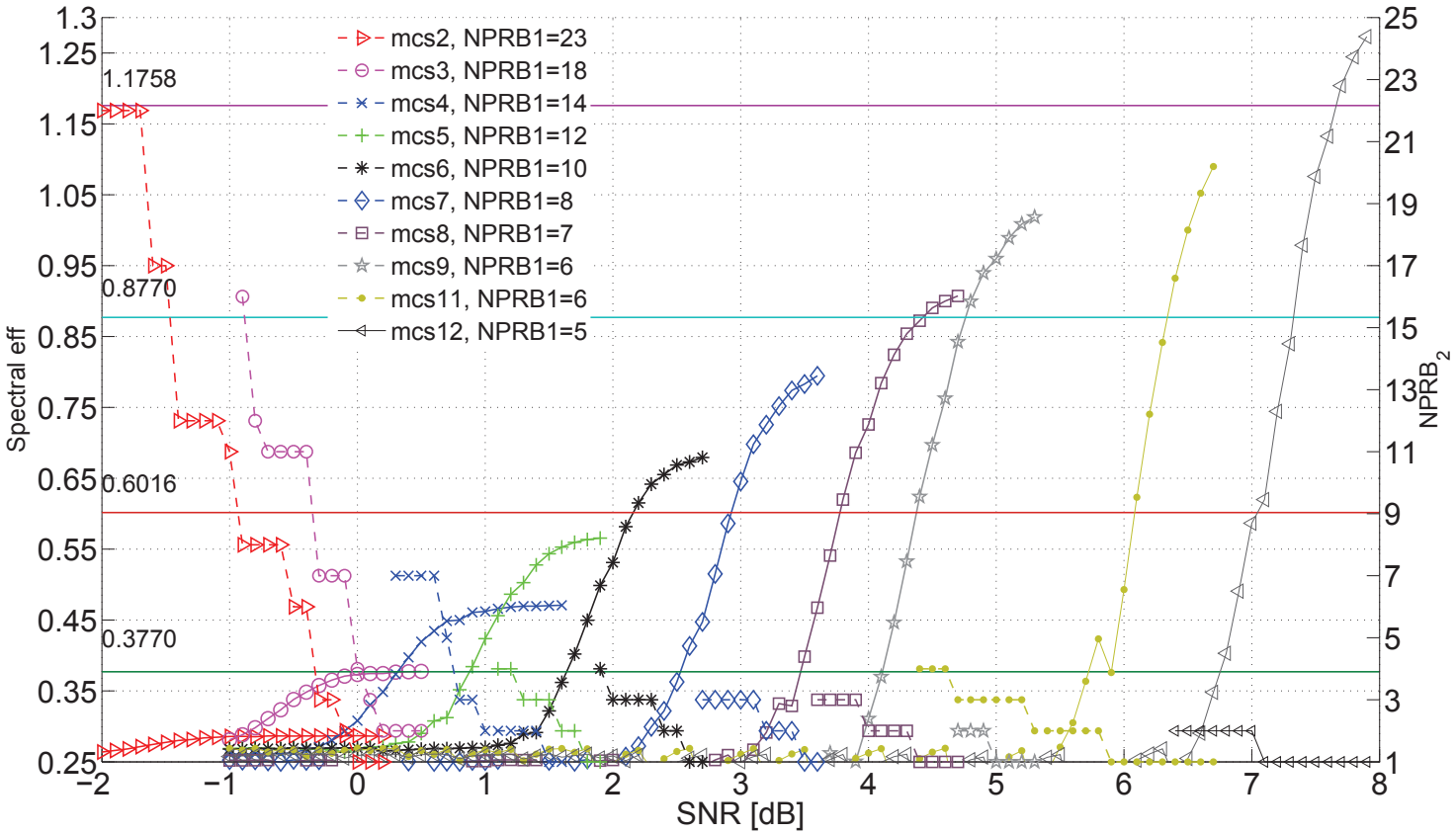


Figure 5.5: Spectral efficiency for MCS= $\{2, 3, 4, \ldots, 9\}$. There is no interference.

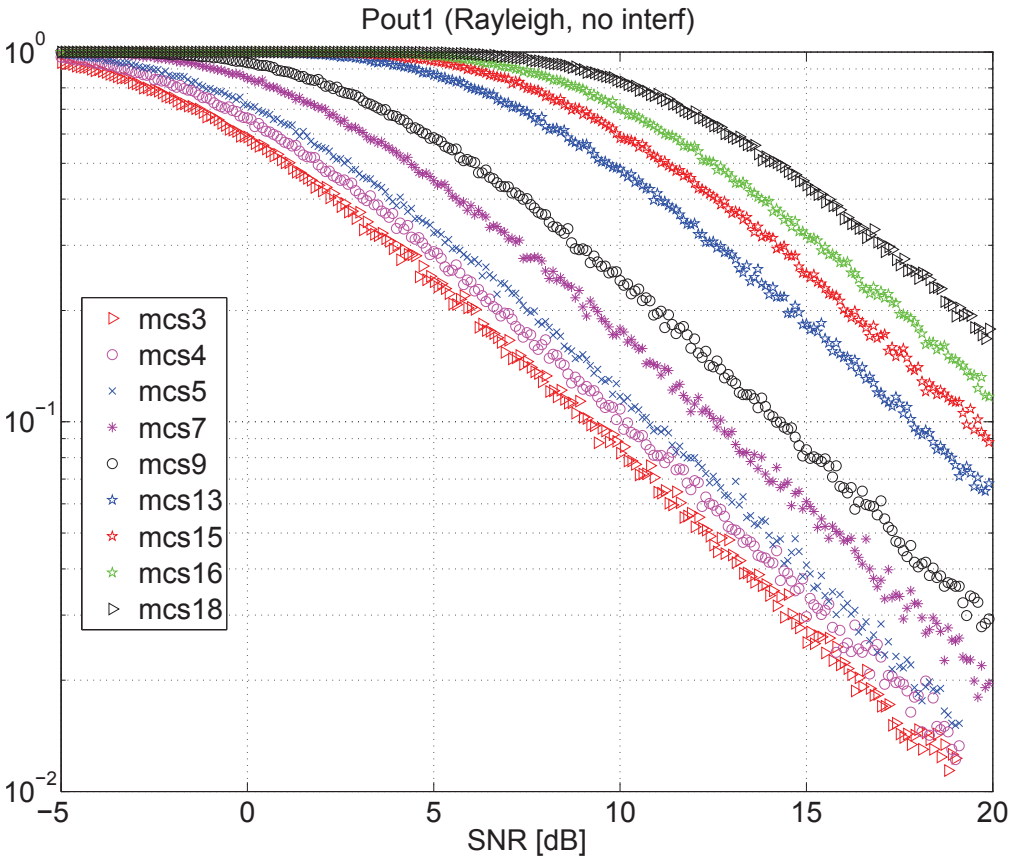


Figure 5.6: Probability of Outage after the first round of the HARQ protocol under Rayleigh fading. There is no interference.

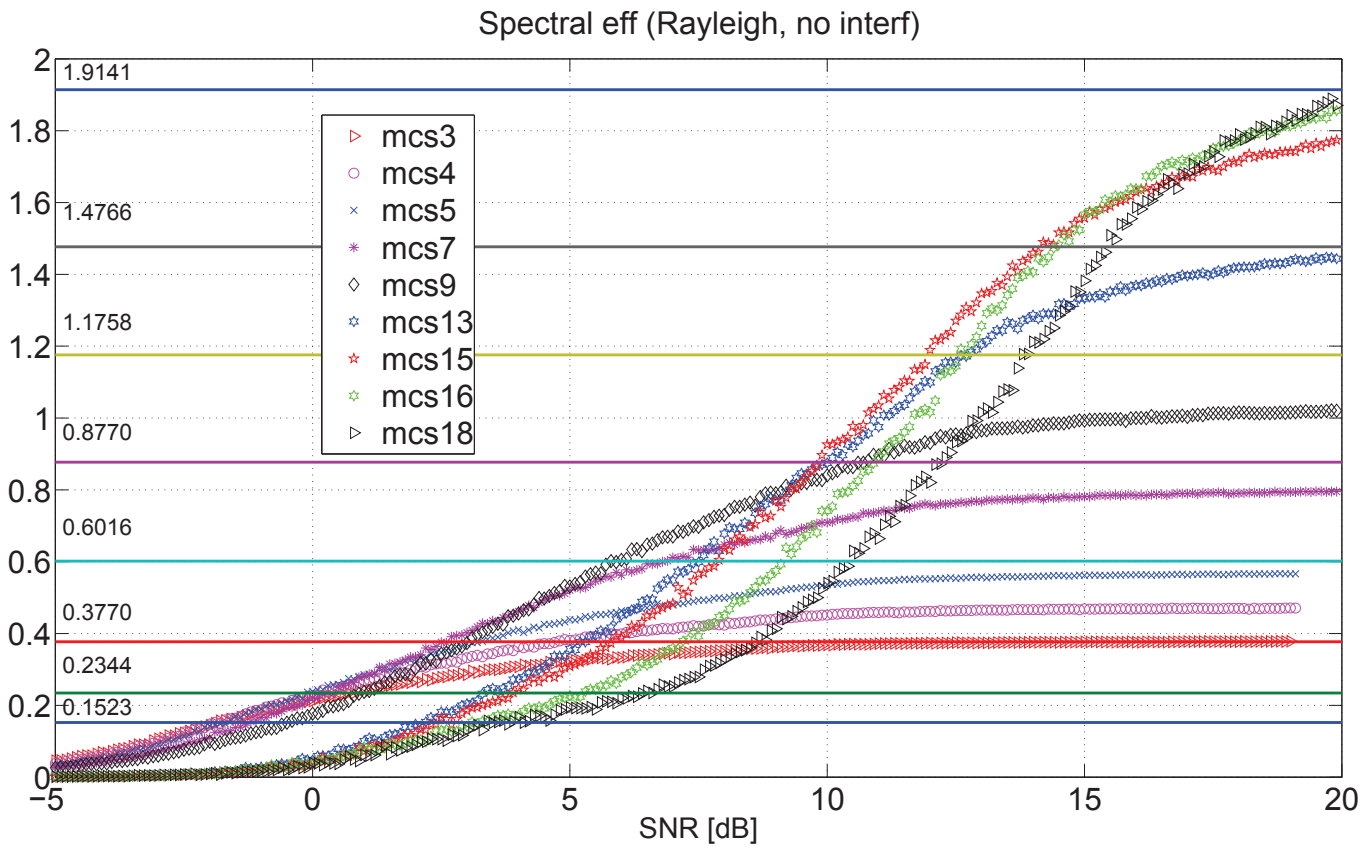


Figure 5.7: Spectral efficiency for different resource allocations under Rayleigh fading. There is no interference.

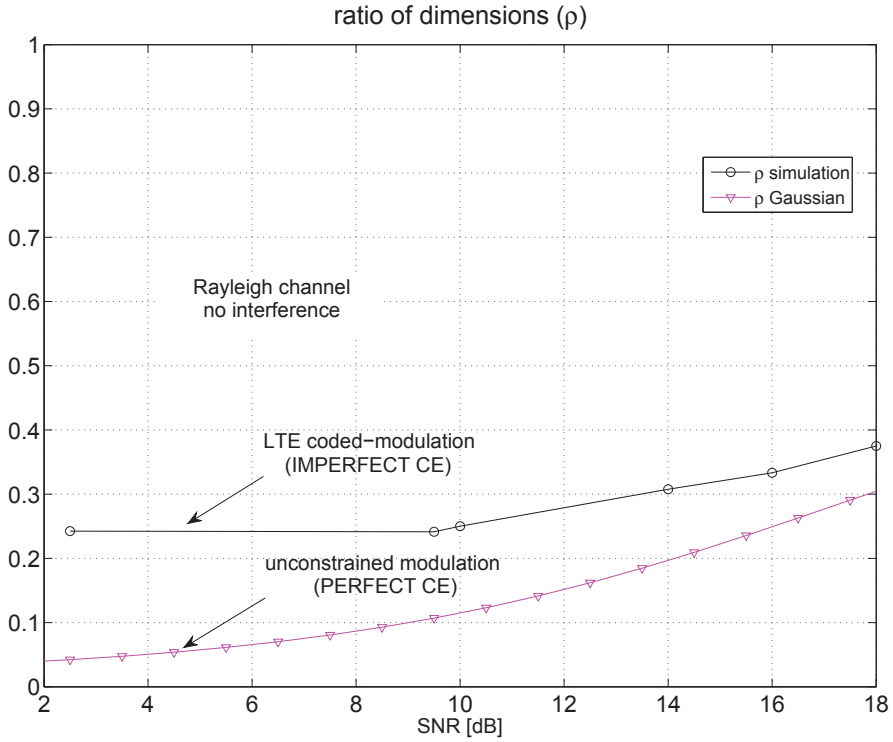


Figure 5.8: Ratio of dimensions.

We can compare these values with the theoretical example in figure 5.1(b). For this purpose, in figure 5.8, we plot the ratio of dimensions between first and second round and we also plot the results from figure 5.1(b). The difference in the curves comes from the fact that in the theoretical example, the results were obtained from the Gaussian expressions for mutual information which represent an unconstrained modulation and in the case of the simulation results, we study the LTE codes with imperfect channel estimation. However, we see that the results follow the same trend in both cases with ρ becoming higher in the high SNR region. Also, from figure 5.6, we observe that as the SNR becomes higher, the outage probability after the first round becomes higher. The latter is important if we consider that commercial systems are designed to operate in the order of 10^{-1}, but allowing a higher outage probability results in a higher spectral efficiency. We can relate our results to those from the rateless coding schemes with AWGN channels. In this case, more dimensions are used in the first round than in the second since it is possible to overcome errors due to finite block-length (inducing a gap from the Shannon capacity).

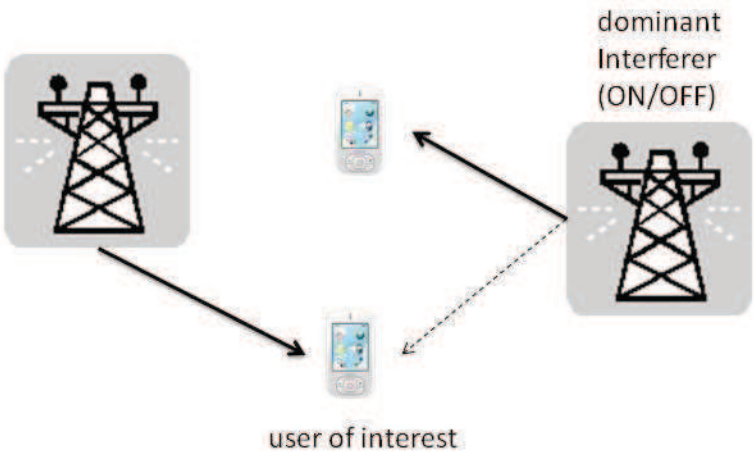


Figure 5.9: Scenario.

5.4.2 Results with one interferer

In this section, we present the results for the interference case in figure 5.9. We consider one dominant interferer with the same power strength as the user of interest and an activity factor of 50% which means that it is active only half of the time. We performed simulations for different channel models for both the user of interest and the interference. We followed the same procedure to optimize the resource allocation as for the case with no interference. We generate data signals for both the user and the interferer and we model explicitly the interference by adding both signals. At the receiver we do not perform any type of interference cancellation. However, we observe that using the best resource allocation decreases, to a certain extent, the impact of the interference.

AWGN / AWGN

For the simulations with interference, we considered 3 different channel models. First, we consider the case of having an ideal AWGN channel for both users, which also means that the interference experienced is the worst.

Figure 5.10 shows the outage after the first round of the HARQ protocol. The results show that the outage probability after the first round is around 50%, which corresponds to the residual outage caused by the interference (the activity factor for the interference is set to 50%). In figure 5.11, we show the spectral efficiency for the different MCS and PRB allocations. We can see the how, for different SNR values, there is a combination of MCS and

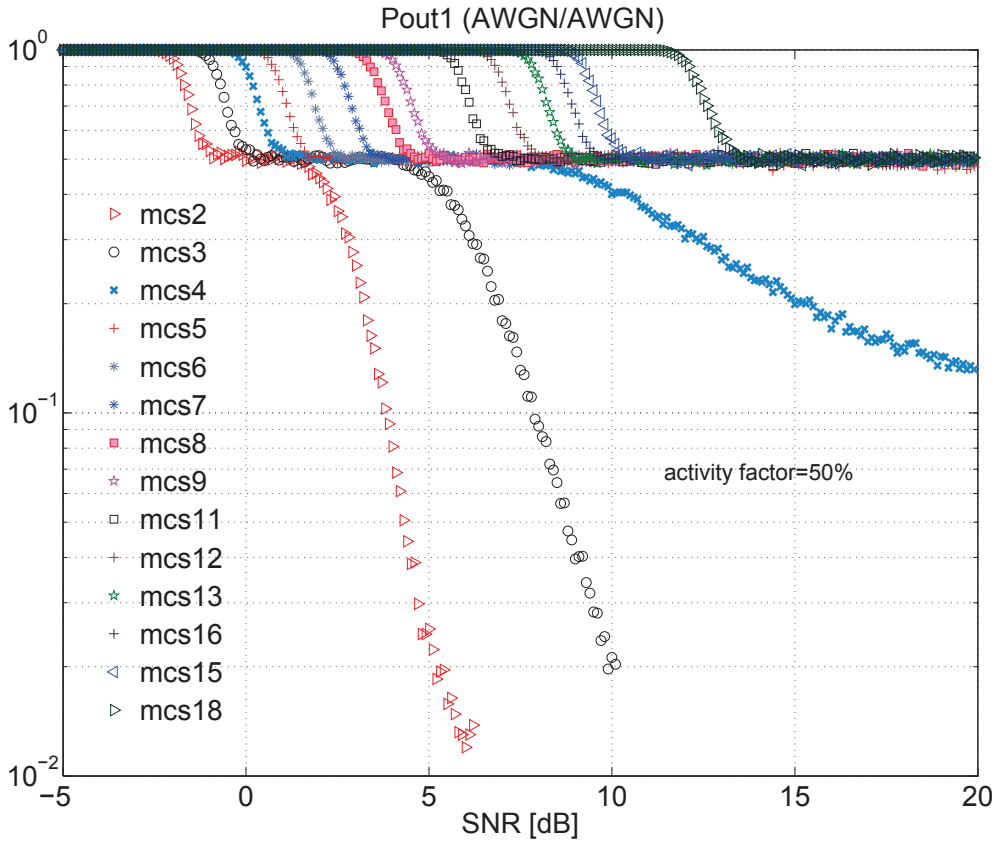


Figure 5.10: Probability of Outage after the first round of the HARQ protocol. There is one interferer of the same strength as the user of interest with a 50% probability of being active. The channel is AWGN for both users.

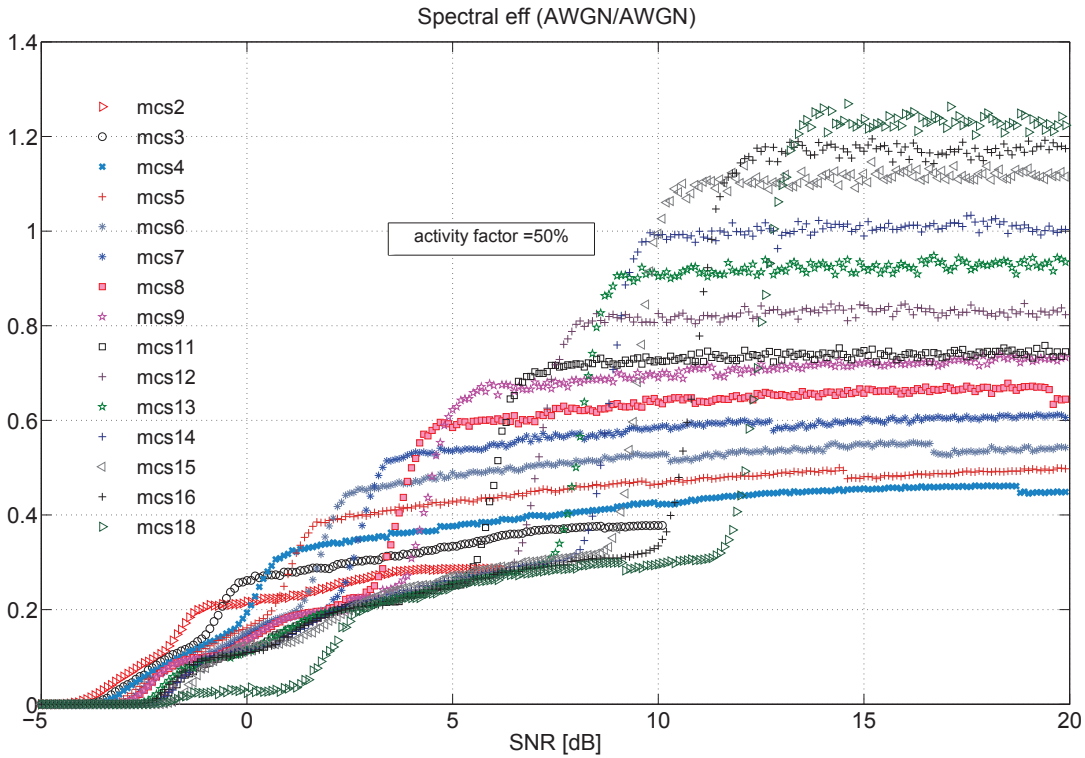


Figure 5.11: Spectral efficiency for different resource allocations. There is one interferer of the same strength as the user of interest with a 50% probability of being active. The channel is AWGN for both users.

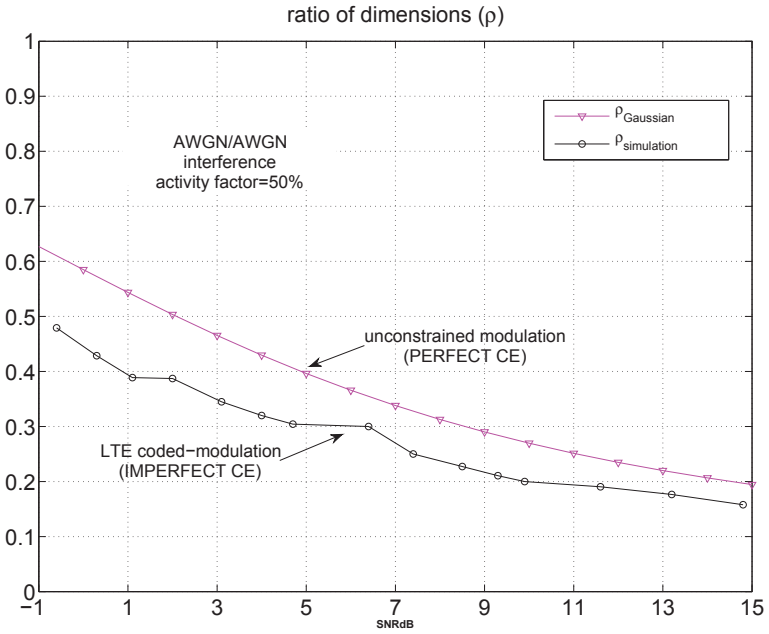


Figure 5.12: Ratio of dimensions.

PRBs distribution that gives the maximum spectral efficiency. It should be mentioned that some of the MCS values like MCS 17 are never used, since there is not an SNR region where it gives the highest spectral efficiency

We can compare this case to the one in the example with Gaussian signals from section 4.4.2 in Chapter 4. For an outage constraint of 10^{-2}, the ratio of dimensions is decreasing using every time, more dimensions in the second round than in the first round. In the LTE case, the results show the same trend, with a difference that can be explained by the fact that we do not assume perfect channel estimation and the constraints of the LTE coded-modulation against an unconstrained modulation represented by the mutual information expressions coming from the Gaussian expressions.

AWGN/Rayleigh

Next, we choose the channel of the user of interest as AWGN and a Rayleigh fading channel for the interference. Figures 5.13 to 5.14 show the results for this case. As we can see from figure 5.13, for the same latency constraint of 10^{-2}, higher MCSs have higher probability of outage at the first round. And from figure 5.14, we observe that depending on the SNR, there is a combination of MCS and PRBs allocation that gives the highest spectral efficiency and, in general, more dimensions are used in the second round. This in contrast with the AWGN channels case.

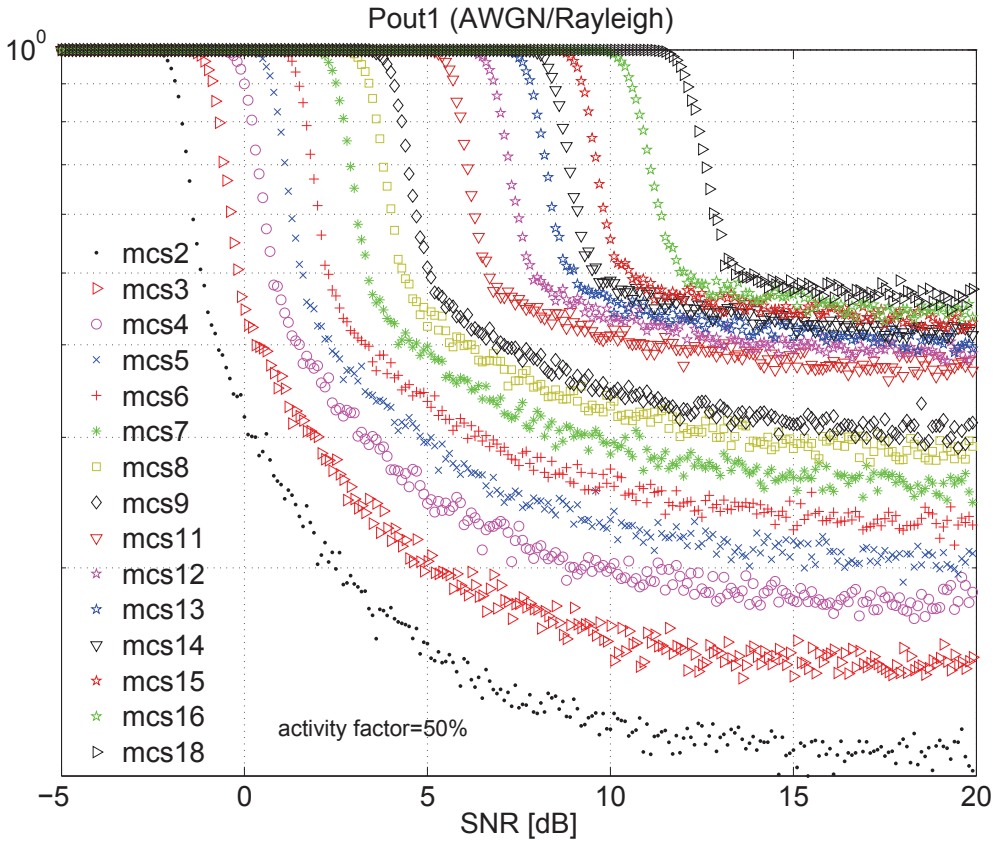


Figure 5.13: Probability of Outage after the first round of the HARQ protocol. There is one interferer of the same strength as the user of interest with a 50% probability of being active. The channel is AWGN for the user of interest and Rayleigh for the interferer.

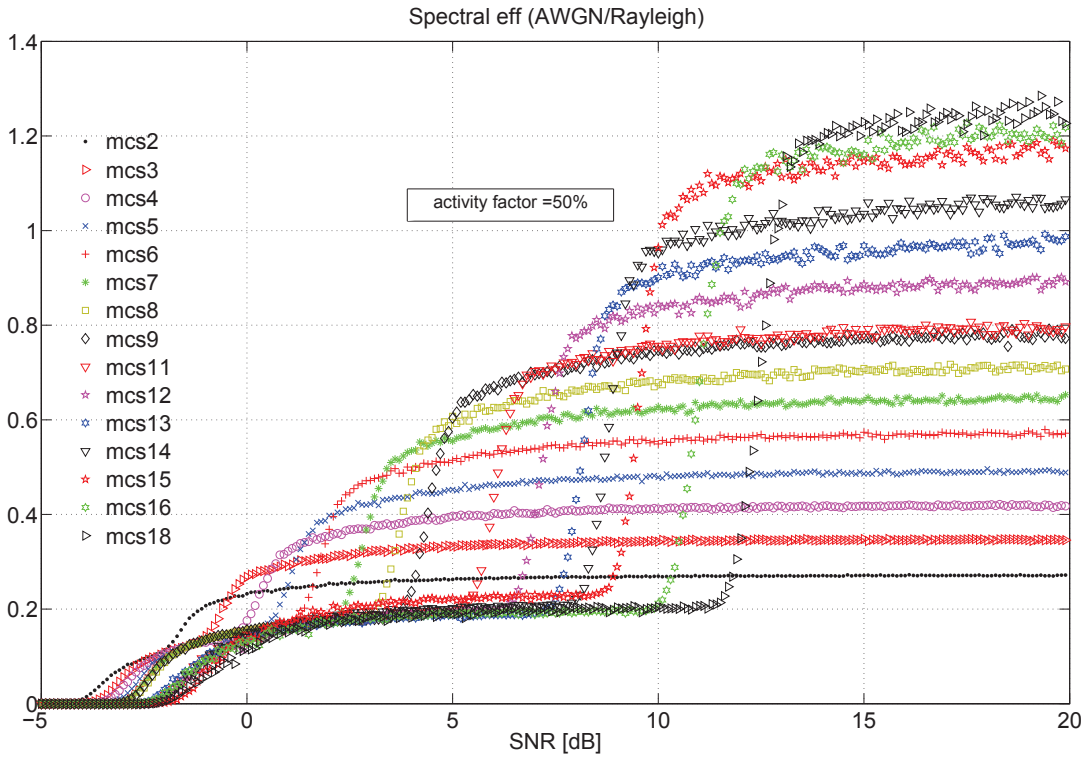


Figure 5.14: Spectral efficiency for different resource allocations. There is one interferer of the same strength as the user of interest with a 50% probability of being active. The channel is AWGN for the user of interest and Rayleigh for the interferer.

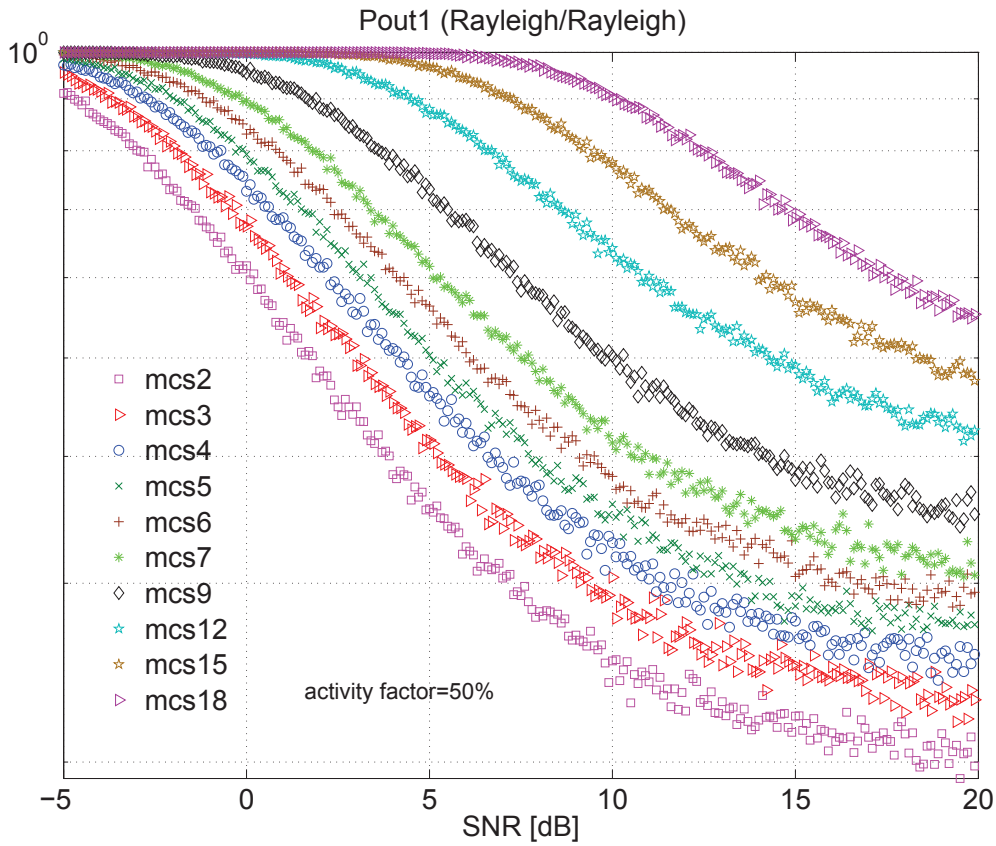


Figure 5.15: Probability of Outage after the first round of the HARQ protocol. There is one interferer of the same strength as the user of interest with a 50% probability of being active. Both the user of interest and the interferer experience a Rayleigh fading channel.

Rayleigh/Rayleigh

Now, we look at the case of Rayleigh fading channel for both users. In this case, the outage probability after the first HARQ round (figure 5.15) is higher because of the bad conditions of the channel. However, if we look at figure 5.16, we see that not all of the MCS values are used for this case with latency constraint of 1%. For example, MCS $4, 5, 12$, have always a lower spectral efficiency. This result may change for different latency constraints.

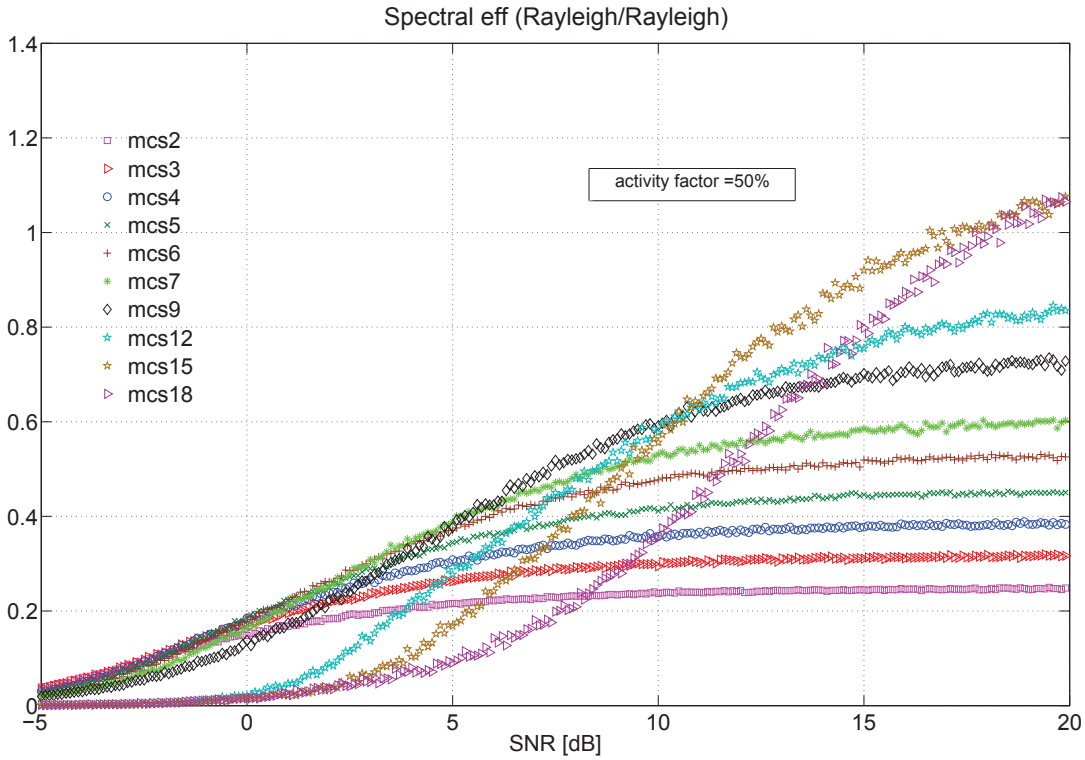


Figure 5.16: Spectral efficiency for different resource allocations. There is one interferer of the same strength as the user of interest with a 50% probability of being active. Both the user of interest and the interferer experience a Rayleigh fading channel.

EPA/AWGN

For a more realistic setup, we simulate a multipath EPA model, which is one of the LTE channel models [2], employed in an urban environment with small cell sizes (low delay spread). We model the channel of the interferer as AWGN since we assume that interference is coming from an indoors small-cell near the macro user. We show in figure 5.17 the outage probability after the first round of the HARQ protocol. In figure 5.18, we show the spectral efficiency. The performance is similar to the case with Rayleigh channel for the user of interest with a high probability of outage and the best allocation corresponding to the use of less resources in the first round for the high SNR region. For the sake of visibility, we do not show all of the MCS values, but only those that give the highest spectral efficiency, for different SNR regions.

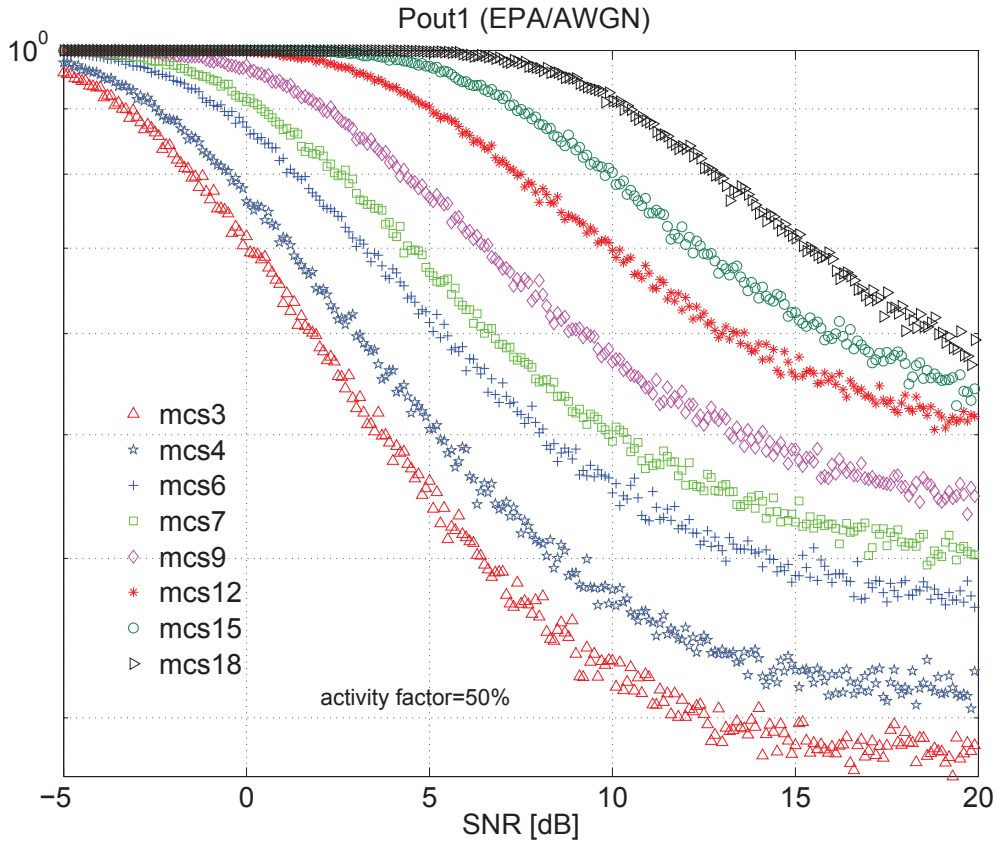


Figure 5.17: Probability of Outage after the first round of the HARQ protocol. There is one interferer of the same strength as the user of interest with a 50% probability of being active. The channel model is EPA for the user of interest and AWGN for the interferer.

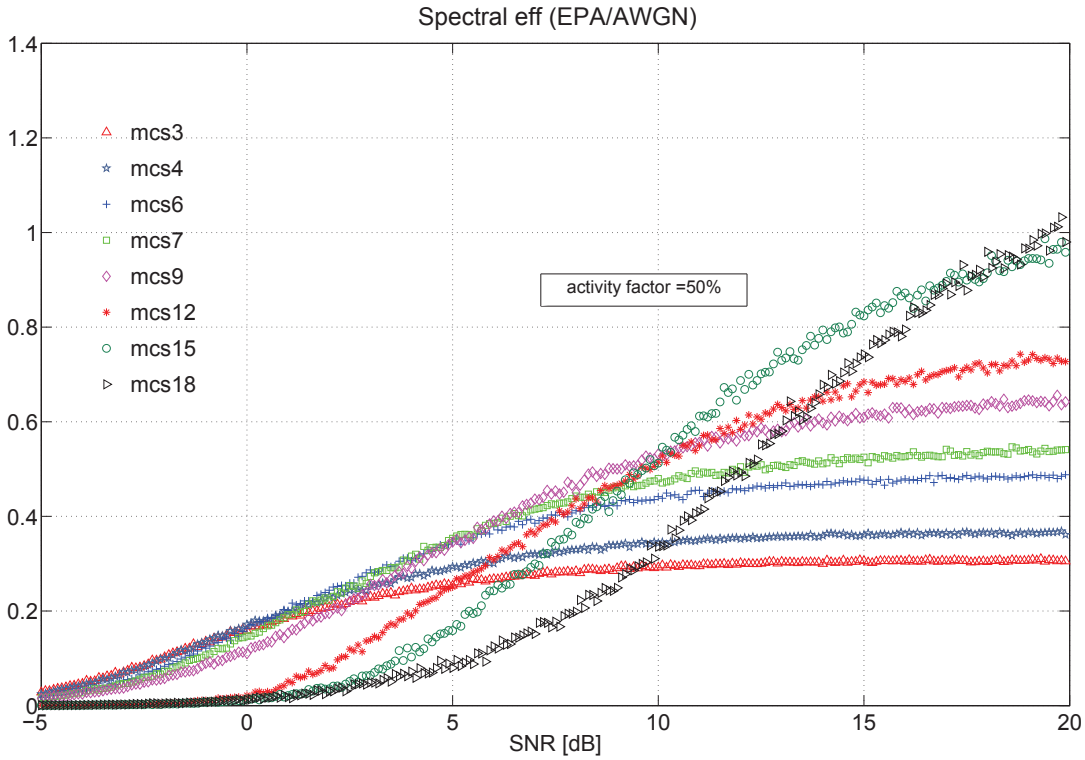


Figure 5.18: Spectral efficiency for different resource allocations. There is one interferer of the same strength as the user of interest with a 50% probability of being active. The channel model is EPA for the user of interest and Rayleigh for the interferer.

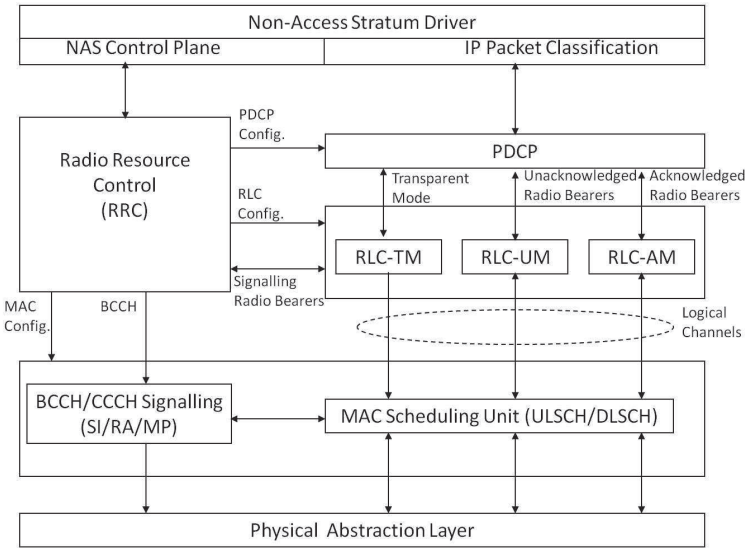


Figure 5.19: Openair Emulation Protocol Stack.

5.5 Scheduler under the Full LTE PHY/MAC Protocol Stack

The OAI platform provides a complete wireless protocol stack and radio hardware (see Figure 5.19) and comprises C-language implementations of the protocol stack and the PHY abstraction unit each corresponding to a particular node in the network. At the network layer, OAI implements radio resource management, routing, multi-casting, topology control, and proxy mobile IP.

To implement dynamic resource allocation policies such as we have described in Chapter 4 in real systems, changes have to be introduced in the MAC scheduling unit. These changes allows us to modify the MCS, and number of PRBs used across HARQ rounds. The allocation of resources has to be signaled on the PDCCH with the use of the DCI every retransmission (retransmissions are always scheduled through PDCCH), instead of an automatic retransmission. The redundancy version has to be explicitly signaled, i.e. the redundancy version index (RVI) corresponding to the redundancy version provided in the HARQ information. With respect to figure 5.19, this involved updating the downlink scheduling policy in MAC Scheduling Unit (DLSCH). In addition, the LTE mechanisms for adapting the modulation order was not available in the OAI implementation in neither the eNB or UE. This was added to both.

5.5.1 Feasibility Evaluation

To test the feasibility of our resource allocation strategies on real LTE modems, we evaluated the spectral efficiency of such shemes in a complete 3GPP PHY/L2 implementation and compared it to the results obtained from the OAI unitary PHY link simulation. As a validation test, we considered the case without interference and AWGN channel in section 5.4.1.

Table 5.3, shows the comparison of the spectral efficiency performance of the PHY and PHY/MAC variable resource allocation implementations. As we can see, the results match those from the AWGN case in section 5.4.1. To make the comparison, we chose some SNR points and then run both cases for 10000 frames.

Table 5.3: Results with PHY/MAC protocol stack

SNR (dB)	MCS	$NPRB_1$	$NPRB_2$	Spectral efficiency	
				PHY	PHY/MAC
0	3	18	4	0.3782	0.3732
3	7	8	3	0.6658	0.6455
5	9	6	1	1.0248	0.9598
7	11	6	1	1.1440	1.0897

The next step is to test our strategies under the presence of interference. For this purpose, in the next section we describe the scenario that we are currently investigating.

5.5.2 Interference Scenario Description

In this case, we consider the interference scenario of figure 5.20 where there is a single dominant interferer with sporadic traffic (eNB_2) creating interference on the downlink of a macrocell user ($UE_{1,1}$) who is also generating sporadic traffic. In this case, the interference can model a small cell (pico/femtocell) which is transmitting only a portion of the time.

To simulate the fully-loaded macrocell (eNB_1), we add a second user ($UE_{1,2}$) connected to the same eNB as the user of interest which is transmitting constant traffic. This user can be seen as a traffic generator used to "saturate" the traffic flow from eNB_1. In a scenario with two small cells, the second UE ($UE_{1,2}$) would have to be inactive.

In this scenario, the interference seen at the user of interest ($UE_{1,1}$) is strong while the interference coming from the macrocell on the small cell user is weak which can be due to the position of the small cell user (indoors) or the distance to the macrocell, etc. In our simulation, the interfering UEs

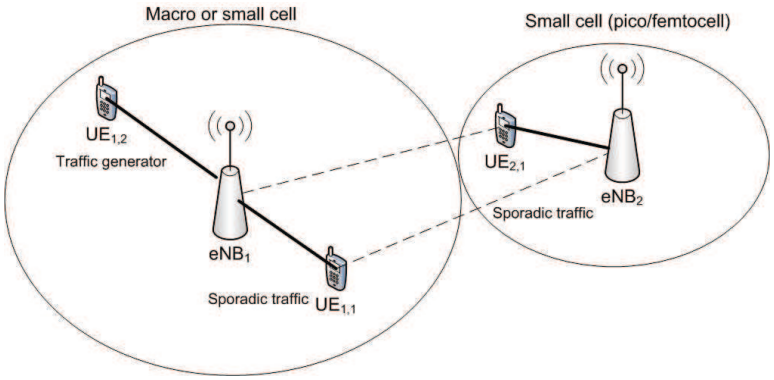


Figure 5.20: Interference Scenario.

occupy artificially (by design) the same subframe which allows us to control the activity factor.

One target is the comparison of the distributed techniques with the ABS feature (see section 2.2.1) introduced in LTE Release 10.

We choose the resource allocation based on statistics of the CQI that the eNB can keep over time and from this information we can infer, in the long time, if the user is experiencing interference or not. Another way to get information about the state of the interference, would be to use an improved feedback assuming two possible SNR values to be fed back. One level of SNR would correspond to the state with interference and the other to the no-interference state. Finally, we can use the expressions in section 4.4.2 to choose the rate and dimensions to use in each of the transmission rounds.

To simulate an interference scenario, an activity factor can not be explicitly simulated, however, we choose the load on the interference to be low compared to the user of interest. We are currently in the process of performing a full protocol simulation with the allocation strategies in section 5.4.2 with two UEs and two eNBs that corresponds to the downlink of the user of interest with one interferer. Table 5.4, shows the comparison of the spectral

Table 5.4: Interference results with PHY/MAC protocol stack

SNR (dB)	MCS	$NPRB_1$	$NPRB_2$	Spectral efficiency	
				PHY	PHY/MAC
5	9	6	16	0.3882	0.3694
8	12	5	17	0.6074	0.5710
10	15	4	18	0.7026	0.7011

efficiency performance of the PHY and PHY/MAC variable resource allocation implementations in the case of one interferer. As we can see, the results match those from the AWGN case in section 5.4.2. To make the comparison,

we chose some SNR points and then run both cases for 10000 frames. We let the interference to transmit on the same subframe as the user of interest. It transmits over all the PRBs to make sure there is interference. The results in table 5.4 are for the case when both users have the same strength.

Chapter 6

Conclusions and Areas for Further Research

6.1 Conclusions

We now summarize the main results of this thesis. In Chapter 3, we analyzed the performance of small-cell networks, in particular, femtocell networks with inter-cell interference. We proposed a decentralized HARQ retransmission protocol that employs incremental redundancy combined with a receiver that can cancel two strong interferers. With the use of Monte Carlo simulations to analyze the throughput, we showed that our scheme is effective at combating interference without requiring any coordination.

In Chapter 4, we first investigated the throughput of a point-to-point link for time-varying channels. By taking an information-theoretic approach, we derived the throughput of IR-HARQ with dynamic physical resource allocation mechanisms. We presented rate-adaptation policies in the case of sparse and latency-constrained traffic. These policies can be applied for both downlink (DL) and uplink (UL) data. We then treated the case when outdated CSI is available at the transmitter and we showed that even in the case of outdated information with a low correlation with the actual channel provides a gain in throughput.

We showed that, in general, by adapting the number of physical dimensions across rounds, we can exploit the interference mitigation effects of HARQ using it not only to recover from errors but for interference cancellation. We proposed efficient resource allocation algorithms to increase

the throughput, which can potentially come very close to optimal performance. Later in this chapter, we studied the case of interference networks, and we demonstrated the benefits of adapting the rate and physical dimensions across transmission rounds of HARQ protocols. We motivated the use of inter-round resource allocation with an example using Gaussian input signals and we obtained a throughput higher than the ergodic capacity in the case of zero-outage throughput and we showed that having an upper layer ARQ in case of a residual outage probability results in a lower throughput. We studied practical scenarios (with signals coming from discrete constellations) without power control and CSI. In this case it is not possible to get zero-outage throughput, however, we benefit from the dynamic resource allocation and by imposing a constraint on the outage probability, we showed that the throughput can be improved by varying the latency of the protocol. We also obtained results for the case of bursty interference in the UL channel when the interference is time-varying because is coming from other users. In this scenario, we use the activity factors to represent the probability of the interference being active.

In the last part of this chapter, we showed some examples of practical applications for the analytical framework. We showed that adapting the resources across HARQ rounds brings a benefit for scenarios like a Manhattan-like topology or a macrocell overlaid by a femtocell. We ended the chapter with a description of the procedure that has to be followed in performing PHY abstraction with the use of our information-theoretic quantities.

Finally, in Chapter 5, we presented the design of practical schedulers for LTE base stations. We showed results from the implementation of our resource allocation strategies in the OAI SDR platform. With the use of a fully-compliant LTE modem implementation, we tackled the case of adapting the resource allocation under the constraints of LTE coded-modulation. By using the DCI, we adjusted the number of PRBs across HARQ transmission rounds and we proved that by adapting the PRBs, we obtained a higher throughput with respect to a static allocation. Moreover, we showed that the results are in agreement with those obtained from the theoretical analysis of Chapter 4. We also found that the performance of LTE codes differs from the theory since not all code rates behave in the same way. As a result of the rate-matching algorithm, higher MCS have a worse performance with respect to lower MCS. The latter can be overcome, by lowering the MCS in the consecutive rounds with the use of the LTE reserved MCS indexes for HARQ operation. In the last part, we presented a full PHY/MAC protocol stack implementation of the scheduling strategies.

6.2 Areas for further research

We have concentrated on single antenna systems, however, our polices can be also applied to MIMO systems. The framework can be extended to consider systems with multiple antennas both at the transmitter and receiver.

The model can also be extended to account for the scheduling of multiple UEs. In this case, the resources have to be allocated taking into account the number of active users and their QoS requirements. Priority-based polices can be also considered within this approach.

Our analytical framework can be used to derive the metrics needed to perform PHY abstraction. In order to perform large scale network performance evaluations, it is useful to abstract the PHY, since the simulation time can grow exponentially and become computationally not feasible. PHY abstraction helps reducing the simulation time without a high computational cost.

Appendices

Appendix A

Summary of the thesis in French

A.1 Abstract en français

L'objectif de cette thèse est de concevoir, implémenter et évaluer les algorithmes cross-layer pratiques. Nous nous concentrons sur la technologie LTE et les réseaux non coordonnés post-LTE où l'interférence est un enjeu majeur compte tenu des nouvelles tendances du trafic. L'objectif est d'allouer les ressources radio d'une manière efficace. Nous développons des modèles d'interférence mathématiques et informatiques qui nous permettent comprendre le comportement de ces réseaux et nous appliquons une approche basée sur la théorie de l'information à différents scénarios d'interférence et caractéristiques du trafic. Nous avons essayé de s'approcher le plus possible de systèmes réels pour être en mesure de tester la faisabilité des techniques proposées.

La thèse porte sur l'évaluation de la performance des scénarios avec interférence dans les réseaux 4G, en particulier celles qui découlent du déploiements de cellules de petite taille ("small cell"). Le travail dans cette thèse s'adresse également à l'analyse de l'allocation des ressources et la requête de répétition automatique hybride (HARQ) à redondance incrémentale pour les interférences sporadiques (de façon plus générale les canaux variables dans le temps) qui permet uniquement des informations partielles de l'état du canal à l'émetteur. Ce travail est ensuite appliquée à la conception d'ordonnanceur pour les stations de base LTE et inclut une analyse de performance pour les modems LTE réels.

A.2 Introduction

Les systèmes de communication modernes, comme le Long Term Evolution (LTE) [6], exigent des taux élevés de données et une meilleure qualité de service (QoS) de contrôle pour les services tels que la téléphonie vocale, les jeux en ligne, la navigation Web, etc. A fin de faire face aux exigences de ces nouveaux types de services tout en offrant simultanément des débits élevés, les normes doivent évoluer et s'adapter. Tant les défis de la couche physique (PHY) et les exigences de qualité de service des applications doivent être pris en compte. Les futurs réseaux de communication sans fils requièrent une optimisation des paramètres dans toutes les couches. Dans une conception inter-couches, le taux, la puissance et le codage à la PHY peuvent être adaptés pour répondre aux exigences des applications, compte tenu des conditions de canal(voir la figure A.1)

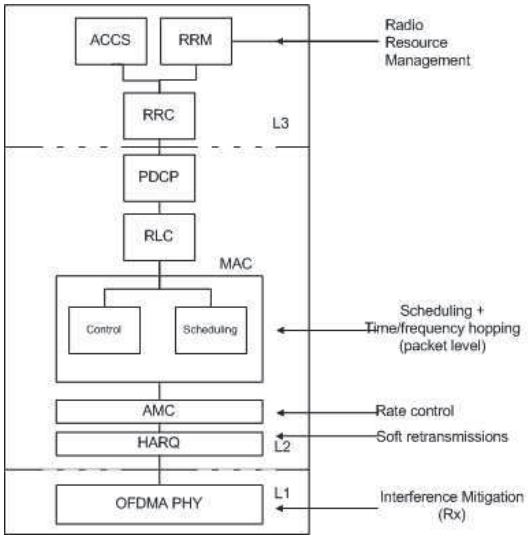


Figure A.1: Techniques Cross-Layer.

En élaborant des politiques qui combinent l'atténuation des interférences sur la couche PHY avec des algorithmes d'ordonnancement et de l'adaptation du taux de la couche Medium Access Control (MAC), et en outre la gestion des ressources radio à la couche de contrôle de liaison radio (RLC), nous pouvons obtenir un débit plus élevé, plus la bande passante et donc des réseaux plus efficaces. Dans le cadre de l'évolution de quatrième génération (4G) des systèmes, la mise en place de cellules de petite taille qui recouvrent le réseau cellulaire existant a été envisagée pour remplir les taches blanches de couverture ou servir les utilisateurs mobiles et de plein air où le réseau cellulaire n'est pas déployée. Cependant, comme ces petites cellules ne peuvent pas être connectés au réseau opérateur de raccordement, la coordination entre eux pour la gestion des ressources est difficilement réalisable (voir la

figure A.2).

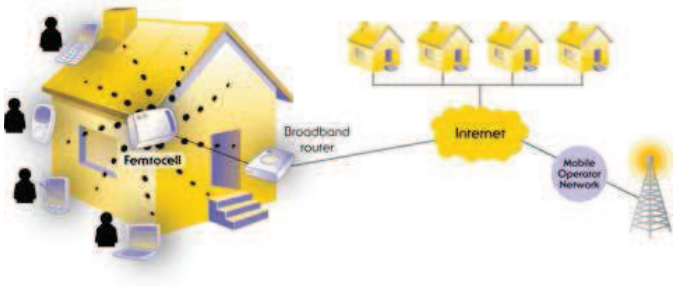


Figure A.2: Femtocells sont un exemple de déploiements avec cellules de petite taille [4].

L'objectif de cette thèse est de concevoir, implémenter et évaluer les algorithmes cross-layer pratiques. Nous nous concentrons sur la technologie LTE et les réseaux non coordonnés post-LTE où l'interférence est un enjeu majeur compte tenu des nouvelles tendances du trafic. L'objectif est d'allouer les ressources radio d'une manière efficace. Nous développons des modèles d'interférence mathématiques et informatiques qui nous permettent comprendre le comportement de ces réseaux et nous appliquons une approche basée sur la théorie de l'information à différents scénarios d'interférence et caractéristiques du trafic. Nous avons essayé de s'approcher le plus possible de systèmes réels pour être en mesure de tester la faisabilité des techniques proposées. Finalement, nous effectuons une étude de simulation complet qui termine par la implémentation et l'évaluation des techniques proposées dans la plate-forme de OpenAirInterface (OAI) [3].

A.2.1 Contributions et cadre de cette thèse

Le principal obstacle trouvée dans les réseaux de communication sans fils est le caractère variable dans le temps du canal physique. Compte tenu de cela, l'objectif d'un concepteur de système est de rendre le PHY/MAC intelligents pour simplifier la conception globale du réseau et optimiser les performances. Dans nos stratégies, nous avons essayé de prendre en compte les scénarios et les paramètres qui les rendent applicables aux systèmes pratiques.

Afin de mieux comprendre la pertinence de notre étude, nous commençons au chapitre 2 en décrivant l'évolution des réseaux de communication sans fils, ainsi que les nouveaux scénarios avec interférence résultant de cette évolution. Nous expliquons alors les fondements de réseaux LTE et interférence et nous décrivons également les bases de la programmation et de l'adaptation de liaison LTE.

Le Chapitre 3 traite de l'évaluation des performances des scénarios avec

interférence dans les réseaux 4G, en particulier celles qui découlent de déploiements de femtocells. Dans ce chapitre , nous analysons le débit du réseau à l'aide d'un temps discret chaîne de Markov non homogène et quantités de théorie de l'information de base. Nous considérons un ou deux brouilleurs dominantes et nous étudions un système d'atténuation d'interférence décentralisée qui combine Répétez hybride automatique et demande (HARQ) et redondance incrémental (IR) avec un décodeur d'annulation d'interférence. Pour fins de comparaison , nous étudions également un système ARQ où l'information n'est pas accumulée sur tours de transmission . Notre évaluation de la performance basée sur la modélisation analytique et l'évaluation de Monte Carlo de débit montre que notre système est efficace à la lutte contre l'interférence sans nécessiter de coordination . Les résultats ont été publiés dans

- Villa, Tania; Merz, Ruben; Knopp, Raymond, "**Interference management in femtocell networks with Hybrid-ARQ and interference cancellation**", in the proceedings of IEEE Asilomar Conference on Signals, Systems, and Computers, November 2011, Pacific Grove, CA, USA.

Le quatrième chapitre commence par décrire quelques applications émergentes où notre analyse de l'information théorie peut être appliquée. Ce chapitre est essentiel pour la compréhension de l'importance des systèmes de répartition des ressources dynamiques LTE. Contrairement à un travail précédent, nous considérons l'allocation des ressources physiques ne sont pas fixées dans les transmissions HARQ. Ce dernier est une possibilité réelle dans ordonnanceurs pour les stations de base LTE (eNodeB ou ENB) et, au meilleur de notre connaissance, existe aucune méthode connue pour l'adaptation des ressources physiques sur tours HARQ lorsqu'il est soumis à des canaux variant dans le temps, soit à cause de la décoloration ou variant dans le temps des interférences ou une combinaison des deux. En faisant cela, nous exploitons les atténuation des effets d'interférence de HARQ utilisant non seulement de récupérer des erreurs mais pour annulation d'interférence et nous proposons des algorithmes d'allocation des ressources efficaces pour augmenter le débit, qui pourrait venir très proche de la performance optimale.

Nous présentons d'abord une analyse des réseaux sans interférence avec les chaînes de variables dans le temps. Plutôt que d'effectuer des simulations, nous adoptons une approche théorique de l'information pour obtenir des expressions analytiques qui représentent le débit à long terme d'une liaison point-à-point et d'envisager des cas pratiques où il ya une contrainte sur la probabilité d'interruption représentant la latence du protocole. Nous considérons les signaux d'entrée de Gauss et nous considérons le cas où l'indicateur

de qualité de canal (CQI) rétroaction est indisponible, ou pas à jour.

Le quatrième chapitre examine ensuite le cas des réseaux d'interférences. Nous motivons l'utilisation de l'allocation des ressources inter-retour à travers une analyse simple mais illustratif avec des signaux de Gauss et les interférences. Nous incluons l'utilisation des facteurs d'activité qui modèle des figures d'interférence sporadiques caractéristique de futurs déploiements de réseaux hétérogènes, en particulier l'ingérence vu des stations de base à cellules de petite taille avec du trafic en rafales dans le récepteur d'un utilisateur macrocellulaire.

Enfin, nous examinons les scénarios d'interférence pratiques pour illustrer les applications de notre cadre d'analyse. Nous modélisons une topologie Manhattan comme ce qui représente un bloc d'appartements avec femtocells créer des interférences. Nous modélisons alors un scénario macro-femto où une macro-cellule est recouverte par une femtocell et nous regardons la liaison descendante (DL) canal de l'utilisateur macrocellulaire lorsque l'interférence provient de la femtocell. Avec l'utilisation d'un facteur d'activité, nous modélisons le fait que la femtocell n'est pas active en permanence. Enfin, nous expliquons la procédure à suivre pour effectuer PHY abstraction de l'utilisation de notre cadre d'analyse, étant donné l'importance de modéliser avec précision les performances de liaison afin d'accélérer les simulations. Les résultats ont été publiés dans

- Villa, Tania; Merz, Ruben; Knopp, Raymond, "**Adaptive modulation and coding with Hybrid-ARQ for latency-constrained networks**", in the proceedings of IEEE European Wireless Conference (EW2012), April 2012, Poznan, Poland.
- Villa, Tania; Merz, Ruben; Knopp, Raymond, "**Adaptive transmission and mutiple-access for sparse-traffic sources**", in the proceedings of IEEE European Signal Processing Conference (EUSIPCO), August 2012, Bucharest Romania.
- Villa, Tania; Knopp, Raymond; Merz, Ruben, "**Dynamic resource allocation in heterogeneous networks**", in the proceedings of IEEE Global Communications Conference (GLOBECOM), December 2013, Atlanta, USA.

et a été accepté pour publication dans

- Villa, Tania; Knopp, Raymond; Merz, Ruben, "**Dynamic resource allocation for time-varying channels in next generation cellular networks, Part I: a mathematical framework**", submitted to IEEE Transactions on wireless communications

et sera présenté comme une partie de

- Villa, Tania; Knopp, Raymond; Merz, Ruben, "**Dynamic resource allocation for time-varying channels in next generation cellular networks, Part II: applications in LTE**", under preparation.

Le chapitre 5 traite avec la conception de planificateur pratique pour les stations de base LTE . LTE offre une grande flexibilité en termes d'allocation des ressources et, en particulier, les algorithmes d'allocation de ressources peut être adapté pour une classe particulière de trafic à des exigences spécifiques. Néanmoins, les travaux sont encore nécessaires pour exploiter efficacement cette flexibilité pour les applications émergentes. Dans ce chapitre, nous étudions les performances de nos politiques d'allocation de ressources dynamique pour la programmation des transmissions IR-HARQ sous les contraintes de la LTE code-modulation. Dans ce chapitre, nous montrons les résultats de la implémentation de l'ordonnanceur dans le logiciel défini OAI radio (SDR) plate-forme [3] afin de tester les performances et la conformité de nos stratégies d'allocation de ressources dans le LTE. Nous montrons que la implémentation de ces politiques dans un système réel est possible sans une coordination ou complexe et fastidieux processus d'optimisation. Nous montrons que nos techniques d'ordonnancement travaillent pour différents environnements et surtout, nous montrent que les résultats sont en accord avec les résultats théoriques présentés dans le chapitre 4. Les résultats seront présentés comme une partie de

- Villa, Tania; Knopp, Raymond; Merz, Ruben, "**Dynamic resource allocation for time-varying channels in next generation cellular networks, Part II: applications in LTE**", under preparation.

A.3 Résumé du Chapitre 2

A.3.1 Evolution des systemes san fils

L'évolution des systèmes de communication sans fil est principalement tirée par l'introduction de nouveaux services et la disponibilité de technologies plus avancées. Au cours des deux dernières décennies, les réseaux cellulaires ont augmenté de façon exponentielle et la demande de services nouveaux et améliorés est devenue un enjeu important pour les opérateurs. Il existe un besoin pour de nouvelles technologies permettant de pallier les limitations de capacité du réseau et de maintenir la qualité de service demandés par les utilisateurs. Cela a motivé le développement de nouvelles normes comme (3GPP) la norme LTE [6] le projet de partenariat de troisième génération afin de fournir des débits plus élevés et une qualité de service améliorées dans

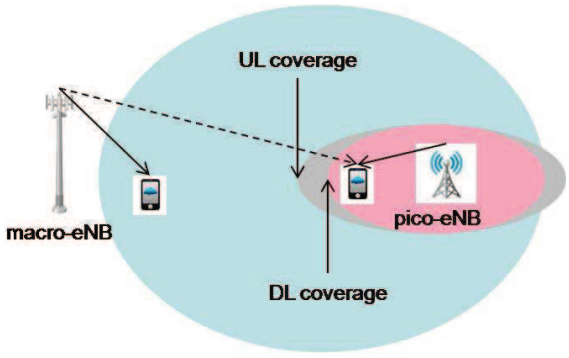


Figure A.3: Macro-eNB vs pico-eNB.

les réseaux sans fil.

Le trafic de données va continuer à croître, avec des abonnements de données mobiles et une augmentation du volume de données moyenne par abonnement. En fait, on s'attend à l'ensemble du trafic de données mobiles de poursuivre la tendance de doubler chaque année [1]. Cette croissance du trafic et des services apportera de nouveaux défis techniques pour les opérateurs, les interférences étant l'une des meilleures performances de limitation.

A.3.2 Interférence dans les reseaux 4G

Pendant les 20 dernières années, il ya eu une augmentation massive du volume de trafic, le nombre de périphériques connectés, et une demande accrue pour les données vidéo. Les réseaux cellulaires futures devraient être en mesure de faire face à cette demande accrue et gérer tout le trafic d'une manière efficace.

Il y a des nouveaux défis techniques et les scénarios possibles avec interférence qui varient selon le type de déploiements, les exigences, débit de données élevé et les niveaux de qualité de service. Ces nouveaux scénarios d'interférence ont été considérés lors de la LTE sortie 10 normalisation:

- Interférence macro-picocell(voir la figure A.3).
- Interférence macro-Home-eNodeB (HeNB) (voir la figure A.4).

A.3.3 Gestion et adaptation de liaison pour LTE

Dans LTE, les algorithmes d'ordonnancement et de l'adaptation du taux de la couche MAC peuvent être combinés avec la gestion des ressources radio à la couche RLC pour obtenir un débit plus élevé, plus la bande passante et donc des réseaux plus efficaces. La implémentation des algorithmes efficaces

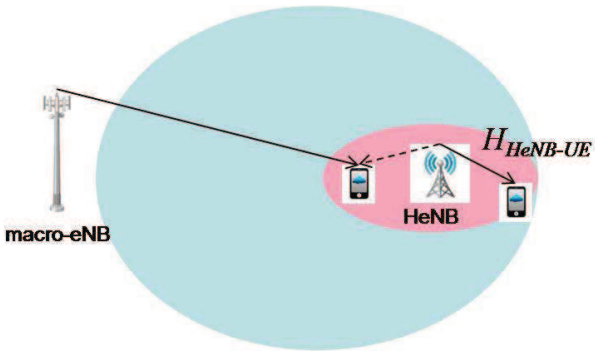


Figure A.4: Macro-eNB vs HeNB.

pour la gestion des ressources radio, ordonnancement de paquets, le contrôle d'admission ou de la puissance et de contrôle d'interférence sont importants pour optimiser la capacité et la performance.

La planification consiste à allouer les ressources de transmission, blocs de ressources physiques (PRB) en LTE, pour les utilisateurs, toutes les possibilités de transmission. Compte tenu des variations rencontrées dans la qualité d'un canal sans fils, le choix d'autres paramètres tels que la modulation et de codage (MCS) peut être adapté dans le but de maximiser la capacité de la cellule, tout en satisfaisant aux exigences en matière de qualité de service de chaque utilisateur. De cette manière, le caractère aléatoire de la liaison radio peut être prise en compte et exploité pour utiliser les ressources de la manière la plus efficace. Le planificateur interagit étroitement avec le gestionnaire HARQ qui est responsable de la programmation des retransmissions en cas de mauvaise réception. La norme LTE prend en charge la programmation dynamique, canal dépendant de renforcer la capacité globale du système.

Dans LTE, le planificateur réside au eNB. La capacité est partagée entre plusieurs utilisateurs sur une base à la demande. Le but de l'ordonnanceur est de décider quel terminal ou une station de base d'émission et sur lequel ensemble de ressources.

Semblable à des ordonnanceurs OFDMA utilisés sur le DL, les planificateurs SC-FDMA pour l'UL peuvent être à la fois du temps et de fréquence opportuniste. Une différence importante entre DL et UL est que les rapports CQI n'est pas nécessaire car le planificateur est situé à l'eNB qui permet de mesurer la qualité du canal UL par signaux sonores de référence (SRS) [61].

En LTE, la bande passante disponible est divisée en N sous-porteuses. De N sous-porteuses, 12 ou 24 sous-porteuses adjacentes sont regroupées formant ce qu'on appelle un bloc de ressources (RB), qui représente la ressource

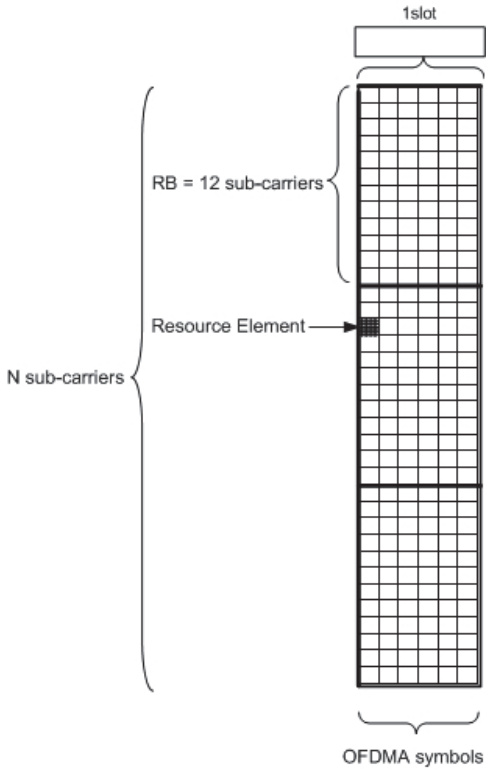


Figure A.5: LTE resource grid [9].

de planification minimale pour UL et transmissions de DL et correspond à 180 KHz de spectre (voir figure A.5). Les cadres LTE sont divisés en deux fentes de durée T_{slot} = 0.5ms. Une fente est formée par N_{RB} RB dans le domaine de fréquence pour la durée de 6 ou 7 symboles OFDMA dans le domaine du temps, en fonction de la longueur du préfixe cyclique (CP) utilisé. Le CP est utilisé à des fins de synchronisation et est fixée à chaque fente. Un sous-porteuse spécifique à l'intérieur de la RB est appelé un élément de ressources (RE). Depuis les sous-porteuses en OFDMA sont orthogonales n'y a aucune interférence de l'intérieur de la cellule, mais des interférences surviennent dans les cellules voisines.

Le nombre de RBs disponibles dépend de la largeur de bande du canal (voir figure A.6), et en fonction de la longueur de la CP, un nombre différent de symboles OFDMA est logé dans une rainure. La table A.1 donne le nombre différent de RBs disponibles pour chacune des largeurs de bande spécifiées dans la norme 3GPP avec le nombre correspondant de RB.

La gestion des ressources radio vise à la programmation des ressources disponibles de la meilleure façon de permettre aux utilisateurs d'obtenir une qualité de service spécifiques. Un mécanisme intelligent doit considérer l'interférence créée avec des ressources physiques déjà attribués.

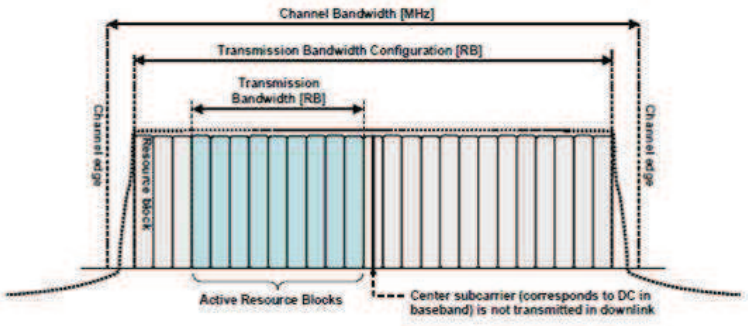


Figure A.6: LTE bandwidth et RBs [28].

Table A.1: N_{RB} vs Downlink System Bandwidth

Bandwidth	1.4 MHz	3 MHz	5 MHz	10 MHz	15 MHz	20MHz
N_{RB}	6	15	25	50	75	100

A.4 Résumé du Chapitre 3

L'expansion actuelle et croissante des réseaux cellulaires représente un défi que l'aménagement et le déploiement nouvelle infrastructure est extrêmement coûteux. Les cellules de petite taille sont considérées comme une solution rentable pour étendre la couverture et réseaux de capacité de la technologie LTE et post-LTE [4].

Les cellules de petite taille sont des points d'accès sans fils à faible puissance fonctionnant dans un spectre sous licence, utilisé en plein air pour améliorer la couverture, ou à l'intérieur de l'entreprise ou de l'utilisation à domicile. Le concept de cellules de petite taille comprennent les femtocells, pico et micro-cellules.

Dans le cas d'utilisation à domicile, les femtocells offrent une haute qualité, l'accès cellulaire à haut débit. Ils sont déployés par les utilisateurs finaux et connecté au réseau de l'opérateur en un numérique ligne d'abonné (DSL), un modem câble ou fibre optique [24]. En raison de la nature imprévue de déploiements femtocell, ils peuvent souffrir de fortes interférences inter-cellule avec femtocells voisins dans les déploiements denses [26, 54, 68]. En outre, la coordination est difficilement réalisable en raison des retards induits par le infrastructure de backhaul de ces réseaux à domicile femtocell.

Dans ce chapitre, nous étudions un régime décentralisée d'atténuation d'interférence que combine IR-HARQ avec une annulation d'interférence décodeur. Notre évaluation de la performance basée sur la modélisation ana-

lytique et expériences de Monte Carlo montre que notre système est efficace pour la lutte contre l'interférence sans nécessiter de coordination.

A.4.1 Interférence dans les reseaux small cells

Nous nous concentrons sur les technologies LTE et LTE-Advanced (dite 4G) avec des couches physiques OFDMA et explorer des stratégies alternatives pour limiter les interférences. OFDMA assure l'orthogonalité des sous-porteuses et, par conséquent, il n'y a pas de interférence intra-cellule. Cependant, l'interférence peut être vécue par les utilisateurs dans les cellules adjacentes.

Contrairement au travail précédent, nous profitons de la nature non-gaussien d'ingérence dans les déploiements femtocell où il ya généralement un seul ou deux fortes brouilleurs dominantes [60]. Nous considérons des signaux provenant d'alphabets discrètes afin de pouvoir bénéficier de la structure de l'interférence. Les signaux gaussiens atteindre le maximum d'efficacité spectrale. Cependant, les systèmes pratiques utilisent des petits entrée alphabets finis de taille.

Sous ces hypothèses, nous proposons une stratégie décentralisée qui combine l'annulation d'interférence décodage [32] avec une politique de redondance HARQ incrémentale [20].

Pour évaluer la performance de cette stratégie, nous développons une modèle analytique du débit obtenu par un protocole HARQ avec une annulation d'interférence décodeur. En particulier, notre modèle s'appuie sur une caractérisation de l'information-théorique du taux réalisable avec une annulation d'interférence.

A.4.2 Modèle du system

Nous nous concentrons sur un scénario de liaison descendante. Sans perte de généralité, nous considérons actuellement de l'antenne unique. Par ailleurs, les transmissions sont fendue et parfaitement synchronisés

Nous avons N_u émetteurs, où noeud 0 est l'émetteur de intérêt et les autres $N_u - 1$ émetteurs sont brouilleurs. Nous laissons d_k soit la distance entre le noeud k et le récepteur. Les systèmes 4G sont basées sur une couche physique OFDMA [61]. Nous laissons $y[m]$ le signal reçu dans un RB particulier au moment de m. Dans LTE, chaque RB est défini comme un groupe de sous-porteuses K. Au sein d'une cellule donnée, on suppose que les RBs sont orthogonales entre elles. Par conséquent, nous pouvons écrire le signal

reçu $y[m]$ au temps m comme

$$y[m]=\sum_{j=0}^{K-1}\sum_{k=0}^{N_u-1}\sqrt{P_k d_k^{-\alpha}}h_{k,j}[m]\mu_{k,j}[m]x_{k,j}[m]+z_j[m]. \tag{A.1}$$

où $x_{k,j}[m]$ est le signal transmis à partir du noeud k dans le $j$ième sous-porteuse d'un RB particulier, $\mu k, j[m]$ est un soi-disant *de facteur d'activité*, $z_j[m]$ est le bruit thermique, P_k est la puissance de transmission, α est l'exposant de perte de trajet et $h_{k,j}[m]$ est le coefficient du canal. Nous modélisons $z_j[m]$ comme le bruit, un processus indépendamment et identiquement distribuées (iid) de moyenne nulle blanc gaussien additif (AWGN) de variance σ^2.

Si nous nous concentrons sur une sous-porteuse particulier, le signal reçu dans le $j$ième sous-porteuse au temps m est

$$y_j[m] = \sum_{k=0}^{N_u-1}\sqrt{P_k d_k^{-\alpha}}h_{k,j}[m]\mu_{k,j}x_{k,j}[m]+z_j[m] \tag{A.2}$$

La variable aléatoire $h_{k,j}[m]$ est iid pour chaque emplacement avec une distribution de Rayleigh. Par conséquent, le coefficient de canal reste constante pendant la durée d'un créneau. Les facteur d'activité modèlisent la charge de trafic et/ou de transmission discontinue (DTX) des caractéristiques de la technologie LTE [61]. Nous modélisons $\mu_{k,j}$ avec un Bernoulli iid, une distribution avec le paramètre p.

Ces caractéristiques sont prises en compte car elles ont un effet direct sur la distribution d'interférence. Le protocole de retransmission est un schéma HARQ utilisant IR [20]. Pour des fins de comparaison, nous considérons également une schéma ARQ simple qui retransmet le même bloc de données dans le cas de transmission infructueuse. Le paramètre $\mathbb{M}_{max}$ est le nombre maximum de tours ARQ. Par conséquent, une trame donnée peut être retransmis à la plupart des $\mathbb{M}_{max}$ fois et est jeté si $\mathbb{M}_{max}$ est atteint. Nous supposons parfaite information de l'état du canal(CSI) et des signaux d'interférence désirés au niveau du récepteur et nous permettent R de définir la vitesse de transmission donnée par un MCS particulier.

A.5 Résumé du Chapitre 4

Les performances de la technologie LTE en termes d'efficacité spectrale et des débits de données disponibles est, relativement parlant, plus limitée par

l'interférence de cellules adjacentes que par rapport aux normes de communication précédentes [27]. Moyens pour réduire ou contrôler l'interférence inter-cellules pourrait avoir des retombées importantes pour les performances LTE, notamment en termes de la qualité de service fourni à chaque utilisateur.

Les algorithmes d'allocation des ressources efficaces sont l'un des éléments clés dans la fourniture de haute efficacité du spectre LTE [34]. Un algorithme d'ordonnancement intelligent peut, en même temps, contribuer à réduire l'impact des interférences.

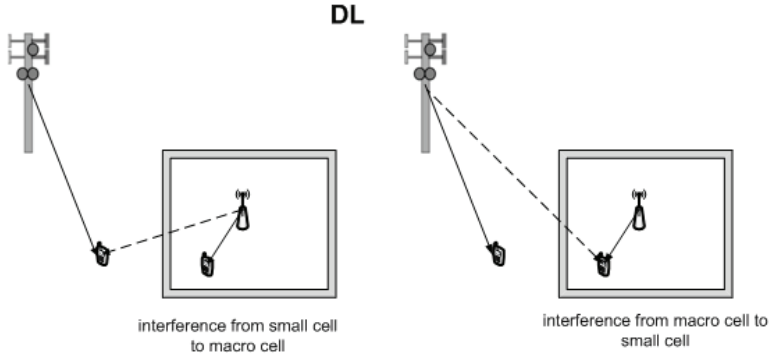
La performance d'un algorithme d'ordonnancement particulier peut être évaluée à l'aide de simulations exhaustives, mais il est temps et a un coût de calcul élevé. En utilisant la théorie de l'information, nous pouvons analyser le débit possible, ou de l'efficacité spectrale, sous différentes hypothèses du modèle de système.

Plutôt que des simulations, nous adoptons une approche théorique de l'information pour obtenir des expressions analytiques qui représentent le débit à long terme du réseau et d'envisager des cas pratiques où il ya une contrainte sur la probabilité d'interruption représentant la latence du protocole. Nous considérons les signaux provenant des alphabets discrets pour modéliser des systèmes pratiques et nous considérons le cas où les informations de CQI sont indisponibles, ou pas à jour.

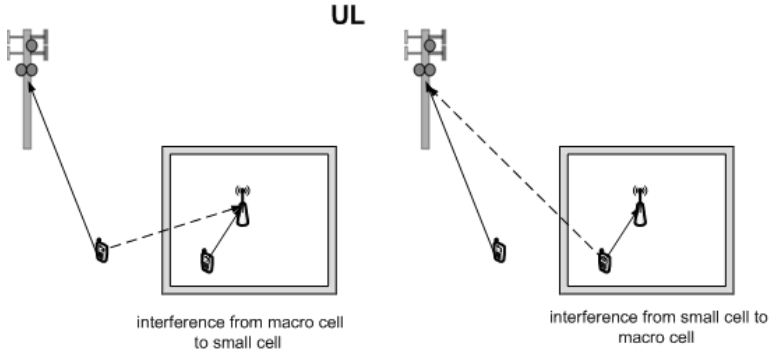
A.5.1 Applications clés

Les réseaux sans fils cellulaires peuvent être divisés en homogènes et HetNets. Hétérogène implique qu'il existe différents types de cellules dans le réseau, à savoir les stations de base de faible puissance sont distribués à travers d'un réseau de cellules macro. Ces stations de base de faible puissance peuvent être microcells, les picocells, relais, femtocells ou systèmes d'antennes distribuées [33]. D'une part, les microcells, les picocells et les relais sont déployées par l'opérateur pour augmenter la capacité et la couverture dans les lieux publics, les entreprises de bâtiments, etc. D'autre part, les femtocells sont déployées à la maison pour améliorer la capacité déployée par l'utilisateur. On note généralement les stations de base de faible puissance par de cellules de petite taille.

Les HetNets cellulaires fonctionnent généralement sur spectre sous licence détenue par l'opérateur de réseau. L'ingérence plus grave est connu lorsque les cellules de petite taille sont déployés sur la même fréquence porteuse que les macro-cells [47]. Scénarios de brouillage plus difficiles sont identifiés dont l'interférence peut venir à travers les couches (macro–à cellules de pe-



(a) Interference scenarios DL.



(b) Interference scenarios UL.

Figure A.7: Figure (a) montre les scénarios d'interférence pour HetNets dans le DL, la figure (b) montre les scénarios d'interférence pour HetNets dans le UL

tite taille, cellules de petite taille–macrocells), par exemple, un utilisateur macrocellulaire loin de la station de base transmet à un très haut pouvoir de blesser les femtocells dans les environs . Les interférences peuvent aussi être vécue entre les petites cellules à la fois dans l'UL et canaux de DL (voir la figure A.7). Dans le cas d'interférences inter-couches, le planificateur de la macro-cell doit prendre en compte l'interférence sporadique des femtocells, car ils seront servent seulement quelques utilisateurs.

La communication M2M, une partie de la révolution d'internet des objets (IoT), devrait permettre de créer un nombre croissant d'appareils connectés, qui dépassera les communications homme-à-homme au cours des années suivantes (50 milliards de machines contre sept milliards de personnes pour 2011) [25, 53].

Dans les scénarios de trafic épars et latence limitées, les arrivées de pa-

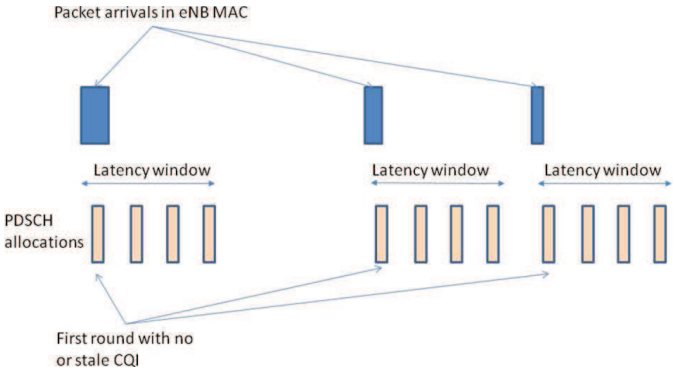


Figure A.8: Le trafic épars dans un scénario de retard limité. Les arrivées du traffic dans la couche MAC eNB sont rares comme représenté en bleu (il y a trois d'entre eux). La contrainte de latence est de quatre emplacements, à savoir il ăăăă sont jusqu'à quatre attributions de canaux PDSCH possibles. En raison de la circulation clairsemée, le CQI n'est pas à jour ou n'est pas disponible sur la première tranche.

quets sont sporadiques et doivent être programmés sous une contrainte de latence (voir la figure A.8). Dans ce contexte, le CQI est généralement obsolète ou indisponible. Notez que le CQI pas à jour se produit aussi à cause de mobilité modérée à élevée, de l'insuffisance de liaison montante CQI périodicité ou de non-stationnaire interférence inter-cellules. Celui-ci deviendra de plus en plus importante avec LTE Release 10 réseaux et leur hétérogénéité inhérente. Par conséquent, le planificateur doit fonctionner à l'aveuglette pour AMC et ne peut bénéficier de feedback après le premier tour de transmission HARQ sous forme de signalisation ACK/NACK.

A.5.2 Analysis pour les réseaux avec interférence

Nous considérons un système de transmission à fentes et nous adoptons une approche de l'information-théorique pour analyser les performances de débit. Quand il y a plus d'un utilisateur, nous supposons que toutes les transmissions à chaque emplacement sont synchronisées et nous faisons le processus d'interférence aléatoire avec l'utilisation de facteurs d'activité. Les derniers modèlisent des figures d'interférence sporadiques caractéristiques des futurs déploiements des réseaux hétérogènes, en particulier l'ingérence vu des stations de base à petites cellules avec du trafic en rafales dans le récepteur d'un utilisateur macrocellulaire. Ils peuvent aussi modéliser les réseaux bi-porteuses avec la programmation "cross-carrier". Dans ce type de réseau, nous pouvons parler de transporteurs propres et sales. D'une part, les transporteurs propres sont utilisés par la macro-cell pour transporter leurs données ainsi que la signalisation pour les petites cellules en raison de leur propriété d'interférence contrôlée. D'autre part, les transporteurs sales

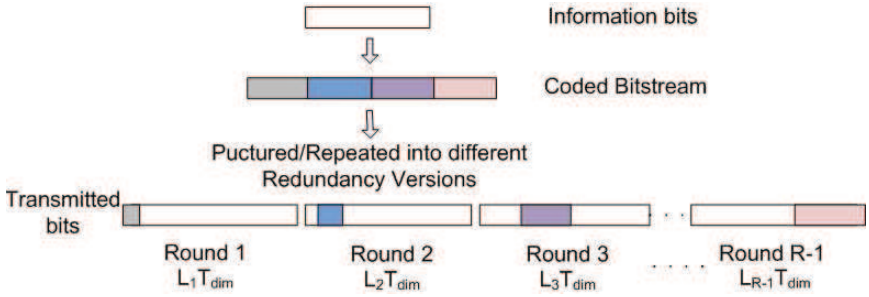


Figure A.9: Modèle de codage

interfèrent transporteurs où la "nettoyage" se fait avec l'utilisation de HARQ.

Nous considérons un maximum de $\mathbb{M}_{max}$ tours de transmission HARQ et le canal est iid ou constante sur toutes les manches du protocole de transmission. Après chaque transmission, nous recevons un accusé de réception sans erreur (ACK ou NACK) indiquant une transmission réussie ou non. Nous définissons la probabilité de panne comme étant infructueuse de recevoir correctement les informations à la fin du protocole HARQ. Cette probabilité se traduit par la latence du protocole et la qualité de service dans notre système.

En général, nous définissons R_r comme le taux de code à l'$r$eme round. Pour un utilisateur particulier, nous définissons le nombre de dimensions dans le temps comme T_{dim} et le nombre de dimensions de la fréquence comme L_r. Soit L'_r le nombre de dimensions de la fréquence à la tour r. Ensuite, à chaque tour de transmission, le nombre total de dimensions est $L'_r T_{dim}$. En supposant que le canal ne varie pas au cours de T_{dim} dimensions de temps et pour une longueur de paquet d'information de B bits, le taux R_r au $r$eme ronde, en bits/dim est donnée par :

$$R_r = \frac{\log_2 B}{L'_r T_{dim}} \quad \text{bits/dim.} \tag{A.3}$$

Dans IR-HARQ, la retransmission comprend le même ensemble de bits d'information que l'original, cependant, l'ensemble de bits codés sont choisis différemment, et ils peuvent contenir des bits de parité supplémentaires. Dans chacun de la transmission arrondit il y a $L_r T_{dim}$ dimensions, cependant, ce nombre n'est pas nécessairement le même partout tours selon la norme LTE [7] (voir la figure A.9).

Dans le contexte de LTE, le nombre de dimensions physiques $L_r T_{dim}$ désigne le nombre de blocs de ressources attribués à un utilisateur dans une sous-trame de durée 1ms soit un temps de transmission d'intervalle (TTI). Il existe dans la plupart des deux blocs de transport fournis à la

couche physique dans le cas de multiplexage spatial [28]. Dans un système LTE mono-utilisateur, il y a seulement un bloc de transport dans un TTI, représentant un seul mot de code "dans l'air" en même temps. Chaque bloc de transport est effectuée par un processus HARQ, et chaque processus est associé à une sous-trame (nombre de processus est fixe). Dans notre modèle, si le nombre de dimensions d'un utilisateur est inférieur au nombre maximum de ressources disponibles N_T,($L_r T_{dim} < N_T$), alors le reste ne sera pas utilisée. Bien que n'étant pas possible dans la norme LTE actuelle, nous pourrions proposer d'utiliser les ressources non utilisées pour transmettre de multiples mots de code en parallèle (en même temps), pour augmenter le débit. Dans un système multi-utilisateur, les dimensions restantes seraient affectées à d'autres utilisateurs, et donc l'efficacité du protocole doivent être choisis afin de maximiser l'efficacité spectrale globale de la cellule.

Pour modéliser des systèmes pratiques, nous dérivons des expressions pour l'information mutuelle en supposant constellations discrètes. Nous ciblons des réseaux LTE Release 10 d'une couche physique OFDMA, et nous étudions la fois la mono-utilisateur et un cas de brouilleurs dominantes.

Dans ce chapitre, nous avons démontré les avantages de l'adaptation du taux et les dimensions physiques travers tours de protocoles HARQ de transmission. Nous avons obtenu un débit plus élevé que la capacité ergodique dans le cas d'indisponibilité de débit zéro et nous avons montré que la présence d'un message ARQ de la couche supérieure dans le cas d' une probabilité résiduel de panne il y a un débit inférieur. Dans les scénarios pratiques sans contrôle de puissance et de l'information d'état de canal, il n'est pas possible d'obtenir zéro panne débit. Toutefois, nous bénéficions de l'allocation dynamique des ressources et en imposant une contrainte sur la probabilité d'interruption, nous pouvons améliorer le débit en faisant varier le temps de latence du protocole. En motivant l'utilisation de l'allocation des ressources inter-ronde, nous avons dérivé politiques distribués d'allocation des ressources qui sont applicables pour l'UL et canaux de DL. Nous avons considéré les interférences qui peuvent être sporadique en raison des caractéristiques de déploiements de HetNets. Plutôt que d'effectuer des simulations, nous avons dérivé les expressions analytiques, basée sur la modélisation de l'information mutuelle, qui captent la performance de débit du réseau, avec ou sans contrainte de latence.

A.6 Résumé du Chapitre 5

Un élément clé de la technologie LTE est l'adoption de procédures de gestion des ressources radio de pointe afin d'accroître les performances du système. L'utilisation efficace des ressources radio est essentielle pour atteindre les

objectifs de performance du système [21]. Dans ce chapitre, nous étudions les performances de nos politiques d'allocation de ressources pour la planification des transmissions HARQ sous les contraintes de la modulation codé de LTE.

L'algorithme HARQ et l'adaptation de débit est mis en place et géré par le programmateur à la couche MAC. Le mécanisme de planification joue un rôle fondamental, car il est responsable du choix, avec bien du temps et de la fréquence granularité, comment distribuer les ressources radio entre les différents utilisateurs, en tenant compte des conditions de canal actuels et exigences de QoS. L'algorithme d'ordonnancement est une question de la station de base, il n'est pas spécifié dans la norme. La norme ne donne pas des mesures/rapports de la qualité du canal et des procédures spécifiques pour l'allocation dynamique des ressources [28] nécessaire pour la implémentation du terminal (et par conséquent de la station de base ainsi). Ceci permet d'avoir des algorithmes spécifiques au fournisseur qui peuvent être optimisés pour des scénarios spécifiques.

La conception des mécanismes d'allocation des ressources efficaces devient cruciale pour un fonctionnement efficace du réseau. Une complexité faible et l'évolutivité sont des exigences fondamentales pour trouver la meilleure décision d'attribution. Problèmes d'optimisation complexes et non linéaires ou de recherche exhaustive sur toutes les combinaisons possibles serait trop coûteux en termes de coût et de temps de calcul [45]. Un algorithme doit être facilement mis en place et, surtout, devrait exiger très faible coût de calcul. Plusieurs solutions théoriques peuvent être trouvées dans la littérature, mais quand on les étudie de près, ils ne peuvent être déployés dans des systèmes réels compte tenu de la difficulté à mis en place dans des dispositifs réels et le coût de calcul élevé requis. Stratégies robustes devraient garantir la possibilité de travailler dans des scénarios très différents. Il ne devrait pas nécessiter des réglages de paramètres forts, ou il devrait au moins dynamiquement adapter ces paramètres aux changements environnementaux [21].

A.6.1 Implémentation sur OpenAirInterface

OAI est une plate-forme de développement matériel/logiciel open-source et open-forum pour l'innovation dans le domaine des communications radio numériques [3]. Il a été créé par le ministère des communications mobiles à EURECOM sur la base de son expérience en financement public de R&D effectuée dans le cadre de projets de recherche en collaboration [15]. Un des avantages les plus importants d'une plate-forme complète de SDR est que le même code peut être utilisé pour l'émulation comme dans une implémentation réelle, d'assurer une transition en douceur entre les simulations pour les

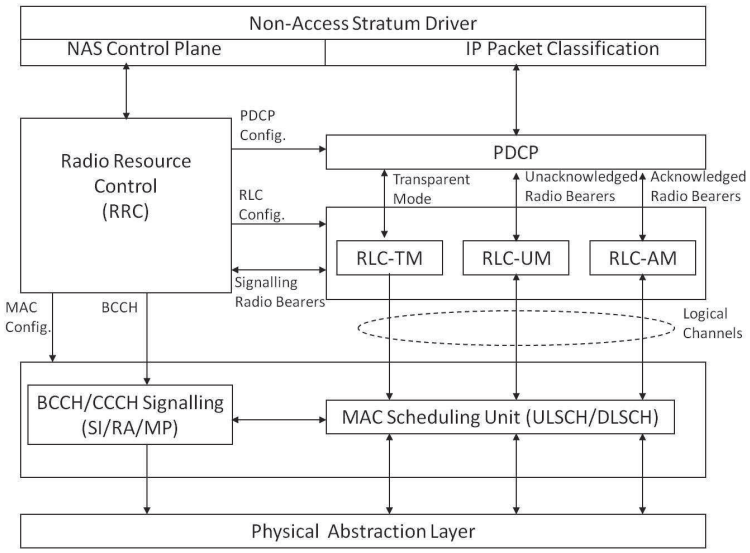


Figure A.10: Pile de protocole pour l'emulation Openair.

tests réels.

La Plate-forme OAI fournit une pile de protocole sans fil complète et matériel de radio (voir la figure A.10) et comprend les implémentations en langage C de la pile de protocole et l'unité PHY d'abstraction correspondant chacune à un node particulier dans le réseau. À la couche réseau, l'OAI mis en place la gestion des ressources radio, le routage multi-casting, contrôle de topologie, et proxy IP mobile.

Pour mettre en place des politiques d'allocation dynamique des ressources telles que nous l'avons décrit dans le chapitre 4 dans les systèmes réels, des changements doivent être introduits dans l'unité de planification MAC. Ces changements nous permettent de modifier le MCS, et le nombre de RBs utilisés dans les tours HARQ. L'allocation des ressources doit être signalé sur le PDCCH avec l'utilisation de la DCI chaque retransmission (les retransmissions sont toujours prévues par PDCCH), au lieu d'une retransmission automatique. La version de redondance doit être explicitement signalé, à savoir l'indice de version de redondance (RVI) correspondant à la version de redondance dans les informations fournies HARQ. En ce qui concerne la figure A.10, il s'agissait de mettre à jour la politique d'ordonnancement de liaison descendante dans l'unité de planification MAC (DLSCH). En outre, les mécanismes de LTE pour l'adaptation de l'ordre de modulation n'étaient pas disponibles dans la implémentation de l'OAI ni dans la eNB ou UE. Ceci a été ajouté à la fois.

A.6.2 Application des techniques pour les modems LTE

Un algorithme d'ordonnancement répond à une métrique pré-formulé relatif à la capacité qui est optimisée dans toutes les solutions d'allocation de ressources possibles satisfaisant un ensemble d'exigences prédéterminées telles que la qualité de service, efficacité spectrale, ou de latence. Pour calculer l'efficacité spectrale, nous devons tenir compte de la TBS, l'ordre de modulation et la probabilité de panne:

$$\mathrm{Spec}_{\mathrm{eff}} = \frac{Q_{m,1}}{\mathrm{G}_1}\mathrm{TBS}\Big(1-P_{\mathrm{out},1}\Big) + \left(\frac{Q_{m,1}}{\mathrm{G}_1}+\frac{Q_{m,2}}{\mathrm{G}_2}\right)\mathrm{TBS}\Big(P_{\mathrm{out},1}(1-P_{\mathrm{out},2})\Big) \tag{A.4}$$

où G_i est le nombre de bits codés par mot de code, et $Q_{m,i}$ l'ordre de modulation utilisé à l'$i$eme ronde. $P_{\mathrm{out},1}$ est la probabilité d'interruption après le premier tour, et $P_{\mathrm{out},2}$ est la probabilité d'interruption après le deuxième tour.

A titre d'exemple de l'optimisation de (A.4), nous fixons le TBS pour toutes les allocations possibles à environ 1000 bits et de la table 7.1.7.2.1-1 dans [7], nous obtenons l'indice du TBS en fonction du nombre de PRBs utilisés dans le premier tour et le TBS plus proche de 1000. Enfin, le tableau 7.1.7.1-1 (voir [7]) nous donne l'indice et de modulation afin MCS correspondante.

Table A.2: Allocatin des ressources vs Modulation et tailles de bloc de transport

$\left(N_{\mathrm{RB}}^{\mathrm{DL}}\right)_1$	$\left(N_{\mathrm{RB}}^{\mathrm{DL}}\right)_2$	TBS	$\mathrm{I}_{\mathrm{TBS}}$	$\mathrm{I}_{\mathrm{MCS}}$	Q_{m}
21	4	936	2	2	2
19	6	1096	3	3	2
17	8	968	3	3	2
15	10	1064	4	4	2
13	12	904	4	4	2
12	13	1032	5	5	2
10	15	1032	6	6	2
8	17	968	7	7	2
6	19	1032	10	11	4
4	21	1000	13	14	4
2	23	1000	21	23	6

La table A.2 montre les différentes allocations utilisées pour générer les résultats en chiffres 5.2 et 5.3, où $\left(N_{\mathrm{RB}}^{\mathrm{DL}}\right)_1$ représente le nombre de PRBs allouées dans le premier tour. Nous utilisons ce tableau pour tester différentes allocations et étudier la performance des codes LTE. En utilisant la DCI,

nous pouvons signaler les nouvelles informations par rapport au nombre de PRBs dans les tours du protocole HARQ consécutives. Le MCS peut également être adapté, mais le TBS reste fixe.

Pour tester la faisabilité de nos stratégies d'allocation de ressources sur de véritables modems LTE, nous avons évalué l'efficacité spectrale de ces sytèmes dans une implémentation complète 3GPP PHY/L2 et comparé les résultats obtenus à partir de la simulation unitaire de liaison PHY OAI. Comme un test de validation, nous avons considéré le cas sans interférence et le canal AWGN dans la section 5.4.1.

Pour le scénario avec interférence de la figure A.11 où il y a un seul brouilleur dominant avec un trafic sporadique (eNB_2) qui crée des interférences sur la liaison descendante d'un utilisateur macrocellulaire ($UE_{1,1}$) qui est également la génération de trafic sporadique. Dans ce cas, l'interférence peut modéliser une petite cellule (pico/femtocell) qui transmet seulement une partie du temps.

Pour simuler la macro-cellule entièrement chargée (eNB_1), nous ajoutons un deuxième utilisateur ($UE_{1,2}$) connecté à la même eNB que l'utilisateur d'intérêt qui transmet le trafic constant. Cet utilisateur peut être considéré comme un générateur de trafic utilisé pour "saturer" le flux de eNB_1 de circulation. Dans un scénario avec deux petites cellules, le second UE ($UE_{1,2}$) devrait être inactif.

Dans ce scénario, l'interférence vu à l'utilisateur d'intérêt ($UE_{1,1}$) est solide pendant l'interférence provenant de la macro-cellule sur l'utilisateur de la petite cellule est faible, ce qui peut être dû à la position de l 'utilisateur de la petite cellule (à l'intérieur) ou la distance à la macro-cellule, etc. Dans notre simulation, les UE interférents occupent artificiellement (par conception) de la même sous-trame qui nous permet de contrôler le facteur d'activité.

Une cible est la comparaison des techniques distribués avec la fonction ABS (voir la section 2.2.1) mis en place en LTE sortie 10.

Nous choisissons l'allocation des ressources basée sur les statistiques de l'CQI que le eNB peut garder au fil du temps et à partir de cette information, nous pouvons déduire, dans le temps, si l'utilisateur connaît les interférences ou non. Une autre façon d'obtenir des informations sur l'état de l'interférence, serait d'utiliser un retour parmi les meilleurs en supposant deux valeurs possibles de SNR à être réinjectées. Un niveau de SNR correspondrait à l'état de l'interférence et l'autre à l'état de non-interférence. Enfin, nous pouvons utiliser les expressions dans la section 4.4.2 pour choisir le taux et les dimensions à utiliser dans chacun des tours de transmission.

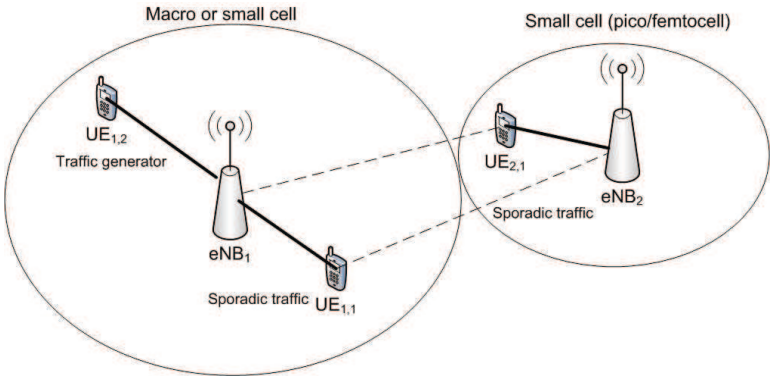


Figure A.11: Scénario avec interférence.

Pour simuler un scénario d'interférence, un facteur d'activité ne peut pas être simulée de manière explicite, toutefois, nous choisissons la charge sur l'interférence d'être faible par rapport à l'utilisateur d'intérêt. Nous sommes actuellement dans le processus de réalisation d'une simulation de protocole complet avec les stratégies de répartition dans la section 5.4.2 avec deux UE et deux eNB qui correspond à la liaison descendante de l'utilisateur d'intérêt avec un brouilleur.

A.7 Conclusion

Nous résumons maintenant les principaux résultats de cette thèse. Dans le chapitre 3, nous avons analysé la performance des réseaux à petites cellules, en particulier les réseaux femtocell avec interférence inter-cellules. Nous avons proposé un protocole de retransmission HARQ décentralisée qui emploie redondance incrémentale combiné avec un récepteur qui peut annuler deux de fortes interférences. Avec l'utilisation de simulations de Monte Carlo pour analyser le débit, nous avons montré que notre système est efficace à la lutte contre l'interférence sans nécessiter de coordination.

Dans le chapitre 4, nous avons d'abord étudié le débit d'une liaison point-à-point pour les canaux variant dans le temps. En adoptant une approche de l'information théorique, nous avons calculé le débit de l'IR-HARQ avec les mécanismes d'allocation dynamique de ressources physiques. Nous avons présenté les politiques taux d'adaptation dans le cas du trafic sporadique et la latence limitées. Ces stratégies peuvent être appliquées à la fois pour la liaison descendante (DL) et la liaison montante (UL) des données. Nous avons ensuite traité le cas où CSI à jour est disponible à l'émetteur et nous avons montré que, même dans le cas de l'information pas à jour avec une faible corrélation avec le canal actuel permet un gain de débit.

Nous avons montré que, en général, en adaptant le nombre de dimensions physiques à travers des tours, nous pouvons exploiter les effets d'atténuation d'interférence de HARQ en l'utilisant non seulement pour récupérer des erreurs mais pour annulation d'interférence. Nous avons proposé des algorithmes d'allocation des ressources efficaces pour augmenter le débit, qui pourrait venir très proche de la performance optimale. Plus tard dans ce chapitre, nous avons étudié le cas des réseaux d'interférence, et nous avons démontré les avantages de l'adaptation du taux et les dimensions physiques a travers des tours de protocoles HARQ de transmission.

Nous motivés à l'utilisation de l'allocation des ressources inter-ronde avec un exemple en utilisant des signaux d'entrée de Gauss et nous avons obtenu un débit plus élevé que la capacité ergodique dans le cas de zéro panne débit et nous avons montré que la présence d'une couche ARQ supérieur en cas de panne résiduelle résultats de probabilité à un débit inférieur. Nous avons étudié des scénarios pratiques (avec des signaux provenant de constellations discrètes) sans contrôle de puissance et CSI. Dans ce cas, il n'est pas possible d'obtenir zéro panne débit, cependant, nous bénéficions de l'allocation dynamique des ressources et en imposant une contrainte sur la probabilité d'interruption, nous avons montré que le débit peut être améliorée en faisant varier le temps de latence du protocole. Nous avons également obtenu des résultats pour le cas de brouillage sporadique dans le canal UL lorsque l'interférence est variable dans le temps en raison vient des autres utilisateurs. Dans ce scénario, on utilise les facteurs d'activité pour représenter la probabilité de l'interférence étant actif.

Dans la dernière partie de ce chapitre, nous avons montré quelques exemples d'applications pratiques pour le cadre analytique. Nous avons montré que l'adaptation des ressources entre les tours HARQ apporte un avantage pour les scénarios comme une topologie Manhattan ou comme une macrocellule ou recouvert par une femtocell. Nous avons terminé le chapitre par une description de la procédure qui doit être suivie dans l'exécution du PHY abstraction de l'utilisation de nos quantités d'information-théorique.

Dans le chapitre 5, nous avons présenté la conception d'ordonnanceurs pratiques pour les stations de base LTE. Nous avons montré les résultats de la implémentation de nos stratégies d'allocation de ressources dans la plateforme SDR OAI.

Avec l'utilisation d'une application de modem LTE entièrement conforme, nous avons abordé le cas de l'adaptation de l'allocation des ressources dans le cadre des contraintes de LTE code-modulation. En utilisant la DCI, nous avons ajusté le nombre de PRBs à travers des tours de transmission HARQ

et nous avons prouvé que par l'adaptation des PRBs, nous avons obtenu un débit plus élevé par rapport à une allocation statique. De plus, nous avons montré que les résultats sont en accord avec ceux obtenus à partir de l'analyse théorique du chapitre 4. Nous avons également constaté que les performances des codes LTE est différente de la théorie, comme tous les débits de code se comportent de la même façon. À la suite de l'algorithme taux d'appariement, MCS supérieur ont une moins bonne performance par rapport à MCS inférieure. Ce dernier peut être surmonté, par l'abaissement de la MCS dans les tours consécutifs à l'utilisation de la technologie LTE réservé indices de MCS pour le fonctionnement HARQ. Dans la dernière partie, nous avons présenté un implémentation PHY/MAC complète de la pile de protocole des stratégies d'ordonnancement.

Notre cadre d'analyse peut être utilisée pour calculer les paramètres nécessaires pour effectuer PHY abstraction. Afin de procéder à des évaluations de performance des réseaux à grande échelle, il est utile de faire abstraction de la PHY, depuis le temps de simulation peut croître de façon exponentielle et devenir des calculs pas possible. La PHY abstraction aide à réduire le temps de simulation sans coût de calcul élevé.

Bibliography

[1] Ericcson mobility report, june 2013.

[2] http://www.3gpp.org/ftp/specs/archive/ (last visited: 3 october 2009).

[3] http://www.openairinterface.org/ (last visited: 25 july 2013).

[4] http://www.smallcellforum.org/ (last visited: 23 july 2013).

[5] Ericsson white paper: Lte release 12, January 2013.

[6] 3GPP TR 36.201 3rd Generation Partnership Project. Technical specification group radio access network; evolved universal terrestrial radio access (e-utra); lte physical layer; general description, March 2011.

[7] 3GPP TR 36.213 3rd Generation Partnership Project. Technical specification group radio access network; evolved terrestrial radio access (e-utra); physical layer procedures, March 2011.

[8] 3GPP TS 36.101 3rd Generation Partnership Project. Technical specification group radio access network; evolved terrestrial radio access (e-utra); user equipment (ue) radio transmission and reception, March 2009.

[9] 3rd Generation Partnership Project, 3GPP TS 36.211. Technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA); physical channels and modulation, December 2008.

[10] A. Annamalai and C. Tellambura. A simple exponential integral representation of the generalized marcum q-function qm (a, b) for real-order m with applications. In *Military Communications Conference, 2008. MILCOM 2008. IEEE*, pages 1 –7, nov. 2008.

[11] Huseyin Arslan. *Cognitive Radio, Software Defined Radio, and Adaptive Wireless Systems (Signals and Communication Technology)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.

[12] F. Baccelli and P. Brémaud. *Palm Probabilities and Stationary Queues*. Springer Verlag Lecture Notes in Statistics, March 1987.

[13] K. Balachandran, S.R. Kadaba, and S. Nanda. Channel quality estimation and rate adaptation for cellular mobile radio. *Selected Areas in Communications, IEEE Journal on*, 17(7):1244 –1256, jul 1999.

[14] Hari Balakrishnan, Peter Iannucci, Jonathan Perry, and Devavrat Shah. De-randomizing shannon: The design and analysis of a capacity-achieving rateless code. *CoRR*, abs/1206.0418, 2012.

[15] Bilel Ben Romdhanne, Navid, Nikaein, Knopp Raymond, and Bonnet Christian. OpenAirInterface large-scale wireless emulation platform and methodology. In *PM2HW2N 2011, 6th ACM International Workshop on Performance Monitoring, Measurement and Evaluation of Heterogeneous Wireless and Wired Networks, October 31, 2011, Miami, Florida, USA*, Miami, ÉTATS-UNIS, 10 2011.

[16] R. A. Berry and R. G. Gallager. Communication over fading channels with delay constraints. *IEEE Trans. Inf. Theor.*, 48(5):1135–1149, sep 2006.

[17] I. Bettesh and S. Shamai. Optimal power and rate control for minimal average delay: The single-user case. *Information Theory, IEEE Transactions on*, 52(9):4115 –4141, sept. 2006.

[18] Pierre Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. Texts in applied mathematics. Springer, 1999.

[19] G. Caire, G. Taricco, and E. Biglieri. Optimum power control over fading channels. *Information Theory, IEEE Transactions on*, 45(5):1468 –1489, jul 1999.

[20] G. Caire and D. Tuninetti. The throughput of hybrid-ARQ protocols for the Gaussian collision channel. *Information Theory, IEEE Transactions on*, 47(5):1971 –1988, july 2001.

[21] F. Capozzi, G. Piro, L.A. Grieco, G. Boggia, and P. Camarda. Downlink packet scheduling in lte cellular networks: Key design issues and a survey. *Communications Surveys Tutorials, IEEE*, 15(2):678–700, 2013.

[22] C.C. Chai, Tjeng Thiang Tjhung, and Leng Chye Leck. Combined power and rate adaptation for wireless cellular systems. *Wireless Communications, IEEE Transactions on*, 4(1):6 – 13, jan. 2005.

[23] T.V.K. Chaitanya and E.G. Larsson. Outage-optimal power allocation for hybrid arq with incremental redundancy. *Wireless Communications, IEEE Transactions on*, 10(7):2069 –2074, july 2011.

[24] V. Chandrasekhar, J. G. Andrews, and A. Gatherer. Femtocell networks: A survey. *IEEE Communications Magazine*, 46(9):59–67, 2008.

[25] Yu Chen and Wei Wang. Machine-to-machine communication in LTE-A. In *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, pages 1 –4, sept. 2010.

[26] H. Claussen. Performance of macro- and co-channel femtocells in a hierarchical cell structure, Sept. 2007.

[27] E. Dahlman, S. Parkvall, J. Sköld, and Per Beming. *3G Evolution HSPA and LTE for Mobile Broadband*. Academic Press, 1st edition, 2007.

[28] E. Dahlman, S. Parkvall, and J. Sköld. *4G LTE/LTE-Advanced for Mobile Broadband*. Academic Press, 1st edition, 2011.

[29] H El Gamal, G Caire, and M. O. Damen. The mimo arq channel: Diversity-mutliplexing-delay tradeoff. *IEEE Transactions on Information Theory*, pages 3601–3621, August 2006.

[30] U. Erez, M.D. Trott, and Gregory W. Wornell. Rateless coding for gaussian channels. *Information Theory, IEEE Transactions on*, 58(2):530–547, 2012.

[31] EU FP7 Project LOLA (Achieving Low-Latency in Wireless Communications). *D4.3 Adaptive Modulation and Coding Scheme and Hybrid ARQ Mechanism*, v3.0, January 2013.

[32] Rizwan Ghaffar and Raymond Knopp. Interference suppression for next generation wireless systems. In *VTC 2009-Spring. IEEE 69th Vehicular Technology Conference, 2009. April 26-29, 2009, Barcelona, Spain*, April 2009.

[33] Amitabha Ghosh, Nitin Mangalvedhe, Rapeepat Ratasuk, Bishwarup Mondal, Mark Cudak, Eugene Visotsky, Timothy A. Thomas, Jeffrey G. Andrews, Ping Xia, Han-Shin Jo, Harpreet S. Dhillon, and Thomas David Novlan. Heterogeneous cellular networks: From theory to practice. *IEEE Communications Magazine*, 50(6):54–64, 2012.

[34] Arunabha Ghosh, Jun Zhang, Jeffrey G. Andrews, and Rias Muhamed. *Fundamentals of LTE*. Prentice Hall Press, 1st edition, 2010.

[35] A.J. Goldsmith. The capacity of downlink fading channels with variable rate and power. *Vehicular Technology, IEEE Transactions on*, 46(3):569 –580, aug 1997.

[36] N. Gopalakrishnan and S. Gelfand. Rate selection algorithms for ir hybrid arq. In *Sarnoff Symposium, 2008 IEEE*, pages 1 –6, april 2008.

[37] Aditya Gudipati and Sachin Katti. Strider: automatic rate adaptation and collision handling. *SIGCOMM Comput. Commun. Rev.*, 41(4):158–169, aug 2011.

[38] H. Holma and A. Toskala. *LTE for UMTS: Evolution to LTE-Advanced*. John Wiley & Sons, 2011.

[39] Yichao Huang and B.D. Rao. An analytical framework for heterogeneous partial feedback design in heterogeneous multicell ofdma networks. *Signal Processing, IEEE Transactions on*, 61(3):753–769, 2013.

[40] Peter Iannucci, Jonathan Perry, Hari Balakrishnan, and Devavrat Shah. No symbol left behind: A link-layer protocol for rateless codes. In *ACM MobiCom*, Istanbul, Turkey, August 2012.

[41] IST-2003-507581 WINNER D1.3 version 1.0. Final usage scenarios.

[42] A. Khandekar, N. Bhushan, Ji Tingfang, and V. Vanghi. Lte-advanced: Heterogeneous networks. In *Wireless Conference (EW), 2010 European*, pages 978–982, 2010.

[43] Su Min Kim, Wan Choi, Tae Won Ban, and Dan Keun Sung. Optimal rate adaptation for hybrid arq in time-correlated rayleigh fading channels. *Wireless Communications, IEEE Transactions on*, 10(3):968 –979, march 2011.

[44] Georgios P. Koudouridis and Hong Li. Distributed power on-off optimisation for heterogeneous networks - a comparison of autonomous and cooperative optimisation. In *CAMAD*, pages 312–317. IEEE, 2012.

[45] Raymond Kwan, Cyril Leung, and Jie Zhang. Multiuser scheduling on the downlink of an lte cellular system. *Rec. Lett. Commun.*, 2008:3:1–3:4, jan 2008.

[46] A. Larmo, M. Lindström, M. Meyer, G. Pelletier, J. Torsner, and H. Wiemann. The LTE link-layer design. *IEEE Communications Magazine*, pages 52–59, April 2009.

[47] Lars Lindbom, Robert Love, Sandeep Krishnamurthy, Chunhai Yao, Nobuhiko Miki, and Vikram Chandrasekhar. Enhanced inter-cell interference coordination for heterogeneous networks in lte-advanced: A survey. *CoRR*, abs/1112.1344, 2011.

[48] A. Lioumpas, A. Alexiou, C. Anton-Haro, and P. Navaratnam. Expanding LTE for devices: Requirements, deployment phases and target scenarios. *Wireless Conference 2011 - Sustainable Wireless Technologies (European Wireless), 11th European*, pages 1 –6, april 2011.

[49] M. Luby. Lt codes. In *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, pages 271–280, 2002.

[50] Justin Manweiler, Sharad Agarwal, Ming Zhang, Romit Roy Choudhury, and Paramvir Bahl. Switchboard: a matchmaking system for multiplayer mobile games. In Ashok K. Agrawala, Mark D. Corner, and David Wetherall, editors, *MobiSys*, pages 71–84. ACM, 2011.

[51] Ruben Merz and Jean-Yves Le Boudec. Performance evaluation of impulse radio uwb networks using common or private acquisition preambles. *IEEE Transactions on Mobile Computing*, 8:865–879, July 2009.

[52] Hyung G. Myung, Junsung Lim, and David J. Goodman. Single carrier fdma for uplink wireless transmission. *Vehicular Technology Magazine, IEEE*, 1(3):30–38, Sept. 2006.

[53] Navid Nikaein and Srdjan Krea. Latency for real-time machine-to-machine communication in LTE-based system architecture. *Wireless Conference 2011 - Sustainable Wireless Technologies (European Wireless), 11th European*, pages 1 –6, april 2011.

[54] Nokia Siemens Networks. Initial home nodeB coexistence simulation results. *3GPP TSG-RAN WG4 R4-070902*, June 2007.

[55] E. Ohlmer and G. Fettweis. Rate adaptation for interference cancelation receivers in slowly time-variant mimo channels. In *Sarnoff Symposium (SARNOFF), 2012 35th IEEE*, pages 1 –5, may 2012.

[56] R. Palanki and Jonathan S. Yedidia. Rateless codes on noisy channels. In *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*, pages 37–, 2004.

[57] S. Parkvall, A. FuruskaÌĹ andr, and E. Dahlman. Evolution of LTE toward imt-advanced. *Communications Magazine, IEEE*, 49(2):84 –91, february 2011.

[58] Jonathan Perry, Peter A. Iannucci, Kermin E. Fleming, Hari Balakrishnan, and Devavrat Shah. Spinal codes. In *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication*, SIGCOMM '12, pages 49–60, New York, NY, USA, 2012. ACM.

[59] J. G. Proakis. *Digital Communications*. McGraw–Hill, 4th edition, 2001.

[60] Changkyu Seol and Kyungwhoon Cheun. A statistical inter-cell interference model for downlink cellular OFDMA networks under log-normal shadowing and multipath Rayleigh fading. *Communications, IEEE Transactions on*, 57(10):3069 –3077, october 2009.

[61] Stefania Sesia, Issam Toufik, and Matthew Baker. *LTE, The UMTS Long Term Evolution: From Theory to Practice*. Wiley Publishing, 2009.

[62] N. G. Shephard. From characteristic function to distribution function: A simple framework for the theory. *Econometric Theory*, (7):519–529, 1991.

[63] A. Shokrollahi. Raptor codes. *Information Theory, IEEE Transactions on*, 52(6):2551–2567, 2006.

[64] Leszek Szczecinski, Ciro Correa, and Luciano Ahumada. Variable-rate retransmissions for incremental redundancy hybrid arq. *CoRR*, abs/1207.0229, 2012.

[65] T. Tabet and R. Knopp. Cross-layer based analysis of multi-hop wireless networks. *IEEE Transactions on Communications*, 58(7), July 2010.

[66] A. Tajer, N. Prasad, and Xiaodong Wang. Fair rate adaptation in multiuser interference channels. In *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pages 2083 –2087, june 2010.

[67] Daniela Tuninetti. Transmitter channel state information and repetition protocols in block fading channels. In *Information Theory Workshop, 2007. ITW '07. IEEE*, pages 505 –510, sept. 2007.

[68] T. Villa, R. Merz, and P. Vidales. Performance evaluation of OFDMA femtocells link-layer in uncontrolled deployments. In *Proceedings of European Wireless 2010*, April 2010.

[69] Geng Wu, S. Talwar, K. Johnsson, N. Himayat, and K.D. Johnson. M2m: From mobile to embedded internet. *Communications Magazine, IEEE*, 49(4):36 –43, april 2011.

[70] P. Wu and N. Jindal. Performance of hybrid-arq in block-fading channels: A fixed outage probability analysis. *Communications, IEEE Transactions on*, 58(4), April 2010.

[71] K. Zheng, F. Hu, W. Xiangy, M. Dohler, and W. Wang. Radio resource allocation in LTE-Advanced cellular networks with M2M communications. *IEEE Communications Magazine, accepted for publication*, 2012.