ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2017-ENST-0034

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Communication et Électronique »

présentée et soutenue publiquement par

Elena LUKASHOVA

le 7 julliet 2017

Récepteurs Interférence-aware

pour les Systèmes LTE MIMO Mono-Utilisateur

Directeur de thèse : **Prof. Florian KALTENBERGER**
Co-encadrement de la thèse : **Prof. Christian BONNET**

Jury

M. Luc Deneire, Professeur, Université de Nice	Rapporteur
M. Markus Rupp, Professeur, Université technique de Vienne	Rapporteur
M. Raymond Knopp, Professeur, EURECOM	Examinateur
Mme Mari Kobayashi, Professeur, CentraleSupélec	Examinateur
M. Tommy Svensson, Professeur, Université de Technologie de Chalmers	Examinateur
M. Sebastian Wagner, Expert en chef des systèmes sans fil , TCT	Examinateur

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

THÈSE

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

EURECOM
Sophia Antipolis

DISSERTATION

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

Specialization

Communication and Electronics

École doctorale Informatique, Télécommunications et Électronique

Interference-Aware Receivers for Single-User MIMO LTE Systems

by

Elena Lukashova

07 July 2017

Composition of the Jury

Reporters :	Prof. Luc Deneire	University of Nice, France
	Prof. Markus Rupp	Technical University of Vienna, Austria
Examiners:	Prof. Raymond Knopp	EURECOM, France
	Prof. Mari Kobayashi	CentraleSupélec, France
	Prof. Tommy Svensson	Chalmers University of Technology, Sweden
	Dr. Sebastian Wagner	TCL Corporation, France
Advisors:	Prof. Florian Kaltenberger	EURECOM, France
	Prof. Christian Bonnet	EURECOM, France

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

EURECOM
Sophia Antipolis

DISSERTATION

Présentée pour obtenir le grade de
Docteur de
Telecom ParisTech

Spécialité

Communication et Électronique

École doctorale Informatique, Télécommunications et Électronique

Récepteurs Interférence-aware pour les Systèmes LTE MIMO Mono-Utilisateur

par

Elena Lukashova

07 juillet 2017

Composition du Jury

Rapporteurs :	Prof. Luc Deneire	Université de Nice, France
	Prof. Markus Rupp	Université technique de Vienne, Autriche
Examinateurs :	Prof. Raymond Knopp	EURECOM, France
	Prof. Mari Kobayashi	CentraleSupélec, France
	Prof. Tommy Svensson	Université de Technologie de Chalmers, Suède
	Dr. Sebastian Wagner	TCT, France
Directeurs de thése :	Prof. Florian Kaltenberger	EURECOM, France
	Prof. Christian Bonnet	EURECOM, France

Acknowledgements

I would like to express my gratitude to my advisor, Prof. Florian Kaltenberger. His curiosity, competence and confidence were great motivation for me; his ability to give a direction and help me find a solution through the alternative paths when all the intuitive paths were blocked was precious. During this journey, I was given a rare opportunity to apply both theoretical and practical approaches to the research, as we always kept in mind the real-world applications of our ideas. I would also like to remark his honesty, flexibility, his support and protection, which made this adventure possible.

I would also like to thank my co-advisor, Prof. Christian Bonnet who started working with me during my master studies at Eurecom, for his valuable suggestions and for teaching me how to turn any scientific deliverable into a beautiful story.

I am deeply grateful to Prof. Raymond Knopp for his recommendations and support, as well as for the encouraging and helpful discussions, which saved me many hours of wondering in the limbo of wireless communications.

I am very thankful to Robin Thomas, Xiwen Jiang, and Konstantinos Alexandris: we all were working with OpenAirInterface, and shared the same moments of frustration (debugging period) and happiness (the bug is solved, "the UE is decoding!"). Small achievements can definitely make a day, and receiving an expected value of throughput at the end of a new experiment or a giant merge were among of those ones. Special thanks to Kalyana Gopala for his strong knowledge of fundamentals and a talent to teach, and to Sumit Kumar for everyday discussions and co-supervisions of the lab sessions.

I am very happy that I met Cedric Roux and Melek Önen, whose help and support during this period was vital and precious.

I find myself on this place thanks to two wonderful ladies, my grandmothers: one of them was in love with France and taught me to dream, and the second one taught me to push for an ambitious goal. I would like to thank my family on both sides for their faith in me, deep love, patience and infinite support.

To my grandmothers

Abstract

Receiver design has been intensively studied over the years and has evolved together with the standardization process of wireless communication systems. Initially conceived as a simple linear processing, it has since then transformed into advanced non-linear detection mechanisms. Advanced architectures allow to achieve high levels of performance even in the harmful presence of fading, multipath propagation, and interference. However, the real-world feasibility remains the prime criterion when trying to balance the complexity and the performance of receivers.

In this thesis we present a pragmatic evaluation of the low complexity alternatives to the optimum joint Maximum Likelihood (ML) detection in Single-User Multiple-Input-Multiple-Output (MIMO) systems with spatially multiplexed transmissions. The complexity reduction is achieved thanks to the novel interference-aware log-likelihood metrics introduced by Ghaffar and Knopp. We focus on two low complexity receiver architectures: the ML Parallel Interference Aware (PIA) detection and the ML Successive Interference Aware Canceling (SIC). To be eligible for deployment in practical modems, the receiver design should fulfill computational efficiency requirements and be evaluated in simulators or emulators that are compliant with real-world wireless standards. The simulations and experiments are performed in the downlink simulator of OpenAirInterface, which is extensively developed during this thesis.

In the first part of this thesis, we present a complete comparative study of the PIA and the SIC receivers, focusing on the implementation, optimal SNR regimes and levels of the Modulation and Coding Schemes (MCS), and on the computational effort. A theoretical analysis of the potential performance, supported by exhaustive empirical results, demonstrates that our practical SIC receiver outperforms the PIA receiver in terms of throughput and computational time. The latter is achieved by replacing the time consuming interference-aware LLR metric of one of the spatial streams with the light-weight SIC block with interference-free metrics. This result is particularly important since the time consumption for downlink signal processing is one of the crucial performance indicators and criteria for deployment in real-time modems.

To make our receivers compatible with practical LTE MIMO systems, we revisit in the second part of this thesis essential LTE protocols such as the Incremental Redundancy Hybrid Automatic Repeat Request (HARQ) protocol and the Link Adaptation. Exploiting the specific structure of our SIC receiver, we show how the multi-round SIC procedure can be used to handle retransmission processing at the UE.

Real time-feedback estimation for adaptive transmissions is a challenging aspect of the receiver procedures. The main limitations are brought by signaling overhead, channel

adaption thresholds, and non-perfect feedback channels. Even though significant research efforts have been made to circumvent these challenges, the complexity-performance trade-off is still an open question. We propose a low complexity estimation of the three feedback components: the Rank Indicator (RI), the Precoder Matrix Indicator (PMI), and the Channel Quality Indicator (CQI). The empirical results of the proposed RI estimation methodology reveal that both our receivers achieve maximum throughput if the base station performs spatial multiplexing transmission even in poorly conditioned channels. The gap between the performance levels of our PMI estimation and the optimal estimation strategy based on the maximum mutual information criterion, is negligible. The proposed CQI estimation is based on Exponential Effective SINR mapping techniques does not provide the optimal performance. However, the CQI errors can be recovered with the HARQ retransmissions.

Contents

Acronyms

Here are the main acronyms used in this document. The acronyms for logical, physical and transport channels of Long Term Evolution are given in tables in Section 2.2.

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
ACK	Acknowledge
AMC	Adaptive Modulation and Coding Scheme
AoA	Angle of Arrival
AWGN	Additive White Gaussian Noise
BLER	Block Error Rate
CC	Chase Combining
CDF	Cumulative Density Function
CDMA	Code Division Multiple Access
CE	Channel Estimation
CLSM	Closed Loop Spatial Multiplexing
COMP	Coordinated Multi-Point
CQI	Channel Quality Indicator
CRC	Cyclic Redundancy Check
CSI	Channel State Information

CW	Codeword
DCI	Downlink Control Information
DL	Downlink
DLSCH	Downlink Shared Channel
ECM	Equivalent Channel Matrix
EDGE	Enhanced Data GSM Evolution
EESM	Exponential Effective SNR mapping
eICIC	Inter-Cell Interference Coordination
eNodeB	Enhanced NodeB
EPA	Extended Pedestrian A
EPAL	Extended Pedestrian A with Low Correlation
EPAM	Extended Pedestrian A with Moderate Correlation
EPC	Evolved Packet Core
EVA	Extended Vehicular A
FDD	Frequency Division Duplex
GSM	Global System for Mobile Communications
HARQ	Hybrid Automatic Repeat Request
HSPA	High Speed Packet Access
HSPA+	Evolved High Speed Packet Access
IA	Interference-Aware
IF	Interference-free
i.i.d.	independent and identically distributed
IR	Incremental Redundancy
ITU	International Telecommunication Union
LLR	Log-Likelihood Ratio
LOS	Line-of-Sight
LS	Least Squares
LTE	Long Term Evolution

LUT	Look-Up Table
LQM	Link Quality Metric
NACK	Not-Acknowledge
NAIC	Network-Aided Assisted Interference Cancellation
NAS	Non-Access Stratum
NDI	New Data Indicator
NOMA	Non-Orthogonal Multiple Access
MAC	Medium Access Control
MAP	Maximum A Posteriori
MCS	Modulation and Coding Scheme
MF	Matched Filter
MI	Mutual Information
MIESM	Mutual Information Effective SNR mapping
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
ML	Maximum Likelihood
MMSE	Minimum Mean Squared Error
MRC	Maximal Ratio Combining
MSE	Mean Squared Error
MUSIC	MultiUser Superposition Transmission
OAI	OpenAirInterface
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiplexing Access
PDCP	Packet Data Convergence Protocol
PDP	Power Delay Profile
PDU	Packet Data Unit
PHY	Physical Layer
PIA	Parallel Interference Aware

PMI	Precoder Matrix Indicator
QAM	Quadrature Amplitude Modulation
R-ML	Reduced Complexity ML
RE	Recourse Element
RI	Rank Indicator
RLC	Radio Link Control
RRC	Radio Resource Control
RV	Redundancy Version
SDMA	Spatial Division Multiple Access
SDU	Service Data Unit
SIC	Successive Interference Canceling
SINR	Signal-to-Interference-plus-Noise Ratio
SIMO	Single Input Multiple Output
SISO	Single Input Single Output
SNR	Signal-to-Noise Ratio
SVD	Singular Value Decomposition
TB	Transport Block
TD-SDMA	Time Division Synchronous Code Division Multiple Access
TDD	Time Division Duplex
TM	Transmission Mode
TPMI	Temporarily PMI
UE	User Equipment
UL	Uplink
UMTC	Universal Mobile Telecommunications Service
ZF	Zero Forcing
ZMCSCG	Zero Mean Circularly Symmetric Complex Gaussian

Notations

j	The complex number verifying $j^2 = -1$.
$\Re(a)$	Real part of complex number a
$\Im(a)$	Imaginary part of complex number a
$(.)^*$	Conjugate operator
$\mathbb{C}$	The set of all complex numbers
$\mathbb{C}^{m\times n}$	The set of $m \times n$ matrices with complex-valued entries
$\mathbf{I}_n$	Identity matrix of size $n \times n$
$\lvert a\rvert$	Absolute value of complex number a
$\lVert\mathbf{a}\rVert$	Euclidean norm of vector $\mathbf{a}$
$\lVert\mathbf{A}\rVert$	Euclidean norm of matrix $\mathbf{A}$
$\lVert\mathbf{A}\rVert_{\mathrm{F}}$	Frobenius norm of matrix $\mathbf{A}$
$\mathcal{K}(\mathbf{A})$	Condition number of matrix $\mathbf{A}$
$\mathrm{Tr}(\mathbf{A})$	Trace of matrix $\mathbf{A}$
$\mathbf{A}^{\frac{1}{2}}$	Square root of matrix $\mathbf{A}$
$[\mathbf{A}]_{nn}$	n-th diagonal element of matrix $\mathbf{A}$
$\mathrm{vec}[\mathbf{A}]$	Vectorization operator which stacks the columns of matrix $\mathbf{A}$
$I(.;.)$	Mutual Information between the two arguments
$p(.)$	Probability density function of the argument
$\mathcal{H}(.)$	Entropy of the argument
$\mathbb{E}(.)$	Expected value of the argument
$\lvert\mathcal{N}\rvert$	Cardinality of a set $n \in \mathcal{N}$
$\mathcal{I}(.)$	Mapping function of the argument
$\mathcal{I}^{-1}(.)$	Reverse mapping function of the argument
$\mathcal{P}(.)$	Projection onto constellation
$y, \mathbf{y}, \mathbf{Y}$	Received signal (scalar, vector, matrix)
$x, \mathbf{x}, \mathbf{X}$	Transmitted signal (scalar, vector, matrix)
$h, \mathbf{h}, \mathbf{H}$	Channel (scalar, vector, matrix)
$n, \mathbf{n}$	AWGN noise (scalar, vector)
$y_{\mathrm{MF}}, \mathbf{y}_{\mathrm{MF}}$	Matched Filter output, (scalar, vector)
M_0	Modulation order of the first codeword
M_1	Modulation order of the second codeword
$\mathcal{Q}^M$	Modulation alphabet of order M

Chapter 1

General Introduction

1.1 Introduction

Wireless communications have become an integral part of people's lives. Indeed, usage of radio links between the base station and the mobile station withdrew the necessity to be in geographical proximity of the stationary terminal to access services by introducing *mobility*. In the modern world, the user needs only a pocket-size device to have permanent access to the network. Mobile services provide voice connectivity (conventional call) and data connectivity (Web browsing, audio and video streaming, file sharing, social networking, software download, etc). A recent statistic report delivered by [Ericsson, 2016] reveals that while voice traffic remains at approximately the same level since 2012, data traffic is constantly growing, and its contribution between $2015 - 2016$ has increased by 50%, as can be seen in Fig. 1.1.

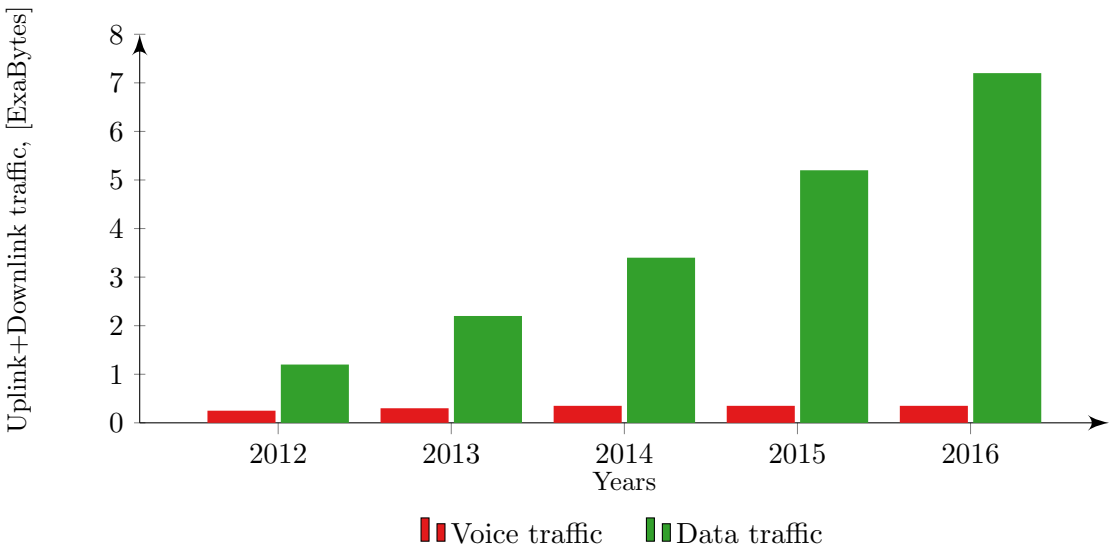


Figure 1.1: Monthly uplink and downlink traffic in $2012 - 2016$ reported in [Ericsson, 2016]. Traffic does not include Wi-Fi, but VoIP is taken into account in data traffic.

The increase of mobile traffic demand has been and remains one of the main drivers of the constant development of wireless communication technologies. Its evolution is often seen as a sequence of generations following one another, separated by a revolutionary step forward. The first analog mobile communications, known as 1G (G standing for Generation), were introduced in the 1980s. Within a decade, they were replaced by the 2G standard, which included the Global System for Mobile Communications (GSM) and the Enhanced Data GSM Evolution (EDGE) systems. The second generation of wireless communicational technology introduced digital encryption and mobile data services, such as text and image messaging. By the beginning of 2008, users, connected to 3G and 3G+ technologies with their smartphones and portable modems for laptops, were able to enjoy the benefits of the mobile broadband access. To make a system more flexible and less vulnerable to the quality of the channel, one of the 3G protocols – High Speed Packet Access (HSPA) – introduced adaptive modulation and coding, fast dynamic scheduling and retransmission protocols. As a next step forward, HSPA+ standard provided a significant increase of transmission data rates by incorporating *multiple-input-multiple-output* (MIMO) systems. Such systems employ multiple antennas at the transmitter and the receiver, and configure multiple data stream transmission. In 2009, the Third Generation Partnership Project (3GPP) introduced the fourth generation (4G) of mobile technology with the packet-switched radio interface — *Long Term Evolution* (LTE). LTE technology has gained popularity thanks to its revolutionary high peak data rates, low latency and improved quality of service. The new fifth generation (5G) of wireless technologies has just emerged and the standardization has not been finished yet. Although, it is already clear that this standard will use new waveforms, high order MIMO and massive MIMO systems. 5G technology is promised to support a large number of densely deployed interconnected devices, to allow remote manufacturing and to bring reliable on-road wireless coordination.

According to [Ericsson, 2016], LTE is forecast to become by 2019 the leading mobile access technology with 4.6 billions clients (see Fig. 1.2). The number of EGDE and HSPA subscriptions are expected to gradually reduce and to fall to 5 times and 3 times lower than LTE values by 2020, respectively.

In LTE, a wireless link between the eNodeB (standing for the base station) and the UE (standing for the user equipment), also known as an *air interface*, is based on the fusion of the *orthogonal frequency division multiplexing* (OFDM) technology and MIMO systems, which allow to combine the high data rates of multi-stream transmission [Alouini and Goldsmith, 1999; Telatar, 1999] with impressive spectral efficiency and robustness to fading and to inter-symbol interference [Stuber et al., 2004].

The OFDM technology allows to transform a frequency selective channel into a set of parallel frequency flat subchannels, each constituting a bandwidth of one subcarrier. These subcarriers are designed with the minimum spacing insuring the orthogonality property of the signals in the time domain. Each subcarrier is mapped onto one of constellations of the Quadrature Amplitude Modulation (QAM) family. The initial encoded bit sequence is thus distributed over parallel data streams, which increases the transmission data rates.

MIMO technology, on the other hand, allows to exploit a new dimension of wireless

systems — spatial layers. The wireless links between the pairs of the antennas at the base station and at the UE can be used to transmit multiple copies of the same signal (*diversity*) ensuring a high signal recovery probability; or to transmit multiple independent signals on available spatial layers (*spatial multiplexing*), thus increasing the system throughput. Both diversity and spatial multiplexing gains can be obtained simultaneously, but not at the maximum level. This phenomenon is known as the *diversity-multiplexing tradeoff* [Zheng and Tse, 2003]. If the knowledge of the *channel state information* (CSI) is available, the MIMO system may tune the transmission scheme to act optimally in the current channel conditions. In LTE, we use so-called *transmission modes* (with 10 modifications currently available) which define the degree of diversity and spatial multiplexing to be used; and the system is able to switch dynamically between different transmission strategies.

MIMO networks include Multi-User MIMO systems, where the base station equipped with the multiple antennas transmits signals to multiple users with one or multiple receive antennas; Single-User MIMO systems, where all simultaneous data streams are dedicated to one user with one or multiple receive antennas; and the degenerate case of *single-input-single-output* (SISO) systems, where both the base station and the UE are equipped with only a single antenna.

In LTE, the Single-User MIMO system can be configured with at maximum two codewords which are mapped onto multiple spatial layers (up to four layers). As a result, the signals — *both dedicated to the same user* — suffer from cross-layer interference, which

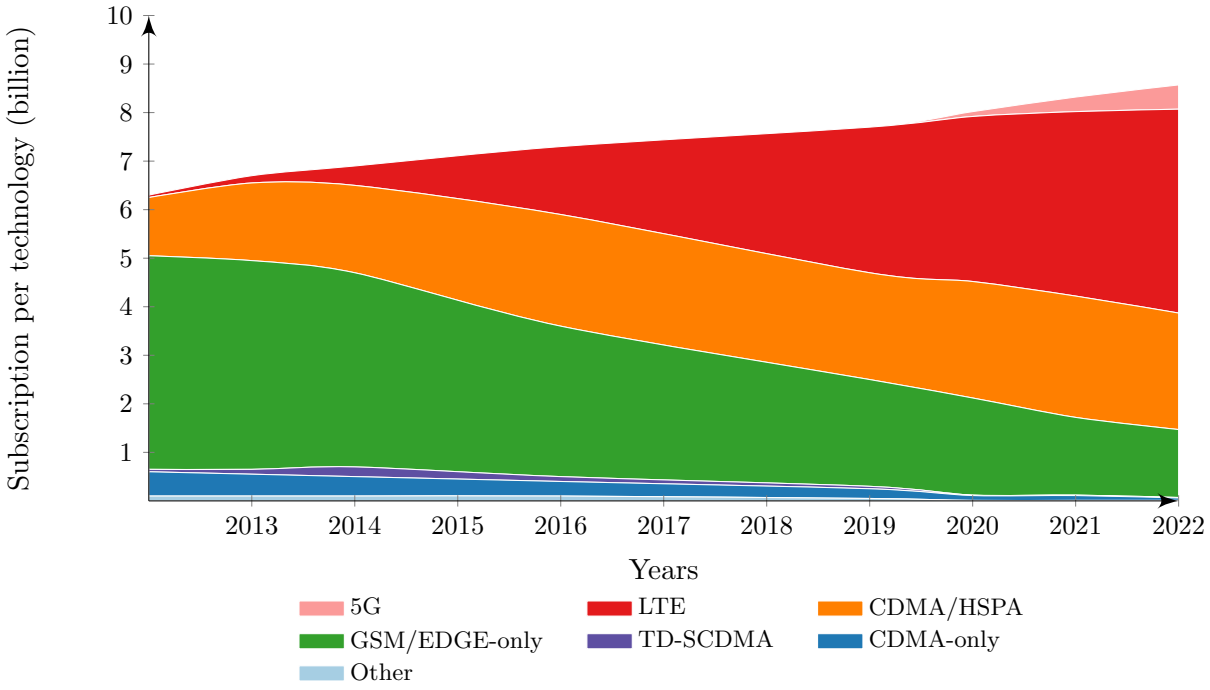


Figure 1.2: Mobile subscription per technology in billions according to [Ericsson, 2016]. Terms UMTS, TD-SCDMA, HSPA+, CDMA stand for Universal Mobile Telecommunications Service, Time Division Synchronous Code Division Multiple Access, Evolved High Speed Packet Access, and Code Division Multiple Access, and belong to the Third Generation of wireless mobile technology.

significantly degrades the performance. Interference management can be applied at the transmitter using *signal preprocessing* techniques: advanced receiver design at the UE and joint optimization at both sides of the MIMO system. Preprocessing techniques, such as precoding, can be codebooks-based and non-codebook based. In the first case, the precoding codebooks are standardized and shared between the base station and the UE [3GPP, 2015c]. In the latter case, so-called Demodulation Reference Signals are added before precoding and provide information about joint influence of the channel and precoding. The main role in the Single-User MIMO interference management is thus given to the receiver architecture.

The specificity of the Single-User MIMO detection is that since both codewords carry useful data, the efficient receiver must apply advanced techniques to recover both of them with the highest probability. Receiver design for MIMO systems can be based on linear filters, such as Zero Forcing (ZF) or Minimum Mean Square Error (MMSE), these receivers are called *linear receivers*; or apply non-linear techniques such as maximum-likelihood (ML) metrics or successive interference canceling (SIC), receivers of this type are called *non-linear*. Linear architectures, despite being attractive thanks to their low complexity, are significantly outperformed by the non-linear design. The non-linear ML-based receivers may perform joint, parallel or successive detection. Joint ML detection is optimum, but its complexity proliferates with the number of spatial layers and the modulation order since an exhaustive search among all possible transmitted vector candidates must be performed. The real-world feasibility remains thus the prime criterion when balancing the complexity and the performance of receivers.

To perform parallel or successive detection, the receiver has to decompose the received signal vector into a set of parallel streams. In these parallel streams, apart from treating interference as Gaussian, the advanced receivers may exploit the actual interference structure, as in LTE systems all the signals are mapped onto modulation alphabets from the QAM family. [Lee et al., 2015] showed that even minimum knowledge of interference significantly contributes to system performance. Furthermore, since in Single-User systems both codewords are expected at the UE, the receiver possesses the knowledge of the exact modulation constellation for both of them, which it may use in *log-likelihood metric* (LLR) computations to reduce the search space dimension [Ghaffar and Knopp, 2010a]. The receivers based on this concept are called *interference-aware* (IA) receivers.

1.2 Thesis Outline and Contributions

Although the base station processing and the majority of the protocols of the wireless systems are now standardized by Third Generation Partnership Project (3GPP), the receiver architecture and implementation remain vendor-dependent. The main focus of this thesis is on low complexity alternatives to the joint ML detection in Single-User MIMO systems. Precisely, we have studied two reduced complexity receiver architectures: the ML Parallel Interference Aware (PIA) detection and the ML Successive Interference Aware Canceling (SIC). The receiver design should fulfill computational efficiency requirements and be evaluated in practical simulators or emulators that are compliant with real-world wireless standards.

To make our receivers compatible with practical LTE MIMO systems, we revisit essential LTE protocols such as the Incremental Redundancy Hybrid Automatic Repeat Request (HARQ) protocol and the Link Adaptation. We seek to define methodologies that would reflect the receiver design and thus would allow to achieve maximum performance gains.

Chapter 2: Background This chapter covers the essential concepts required to carry out this thesis. First, we introduce the basics of MIMO systems and discuss the MIMO transmission techniques and the capacity of the multi-antenna channels. We then deliver insights on LTE, including a standardization overview, physical layer structure and signal processing. We illustrate the classical signal model for MIMO systems which is used throughout this thesis, review the receiver architectures based on different detection criteria, and discusses interference handling and complexity reduction techniques. Finally, we conclude with a brief introduction to HARQ protocols and Link Adaptation technique.

Chapter 3: Signal Model and Simulation Framework This chapter presents the downlink simulator of OpenAirInterface (OAI), which was extensively developed and used for the link-level simulations throughout this thesis. We describe building blocks of the simulator, the key procedures and the available scenarios. Special attention is paid to the analytical models that are used in the transmission schemes and for channel modeling purposes.

Chapter 4: Architecture of the R-ML PIA and SIC Receivers This chapter provides a comparative study of two low complexity Maximum-Likelihood receivers with the Parallel Interference-Aware detection and with the Successive Interference Aware Canceling detection. Empirical simulations deliver insights on the optimal SNR regimes and Modulation and Coding Scheme (MCS) combination for both codewords to maintain long-term throughput at high level. Furthermore, we provide an information-theoretic analysis in order to define potential performance bounds based on the concept of outage mutual information. We build upon these results and investigate how close the practical outcomes of the simulations are to the theoretical expectations. Finally, we quantify the computational effort required by the receivers and evaluate their feasibility to be used in the real modem.

The main contributions of this chapter are presented in the following paper:

- E. Lukashova, F. Kaltenberger, and R. Knopp, "Reduced Complexity ML Interference-Aware Parallel and SIC Receivers for LTE SU-MIMO," in 14th International Symposium on Wireless Communication Systems (ISWCS 2017), Bologna, Italy, August 2017.

Chapter 5: HARQ Protocol for the R-ML PIA and SIC Receivers This chapter first introduces the necessary methodology for Incremental Redundancy HARQ support for the Single-User MIMO system with the Parallel Interference Aware and the Successive

Interference Canceling receiver architectures that we implemented in our downlink simulator in Chapter 4. This allows us to evaluate the throughput and reliability gains coming from the retransmission rounds. In our study, we model the full set of retransmission scenarios: the UE successfully decodes both codewords and no retransmission is needed; only one out of two codewords is decoded and one codeword is thus retransmitted; decoding of both codewords fails, and two spatially multiplexed codewords are retransmitted. The UE signal processing is discussed for each receiver in particular. Special care is given to the multi-round SIC procedure available for our SIC receiver.

Secondly, we quantify the performance of our receivers with the HARQ protocol in terms of throughput and reliability in various frequency-selective environments.

Finally, we test different retransmission schemes with respect to available precoding options in scenarios with one successfully decoded codeword while another one is in error. The comparison is done by applying the throughput metric in the situations with actual and outdated CSI in frequency-selective fading channels.

Some parts of this work were recently published in:

- E. Lukashova, F. Kaltenberger, and R. Knopp, "Single-User MIMO ML successive interference canceling receiver with HARQ-IR protocol," in 21st International ITG Workshop on Smart Antennas (WSA 2017), Berlin, Germany, March 2017.

Chapter 6: PHY Abstraction for the R-ML PIA and SIC Receivers To be eligible for real-time devices, the abstraction algorithms should satisfy certain requirements, such as low memory consumption, time efficiency, and feasible accuracy. This chapter investigates different abstraction methodologies for physical layer of Single-User MIMO LTE systems employing our R-ML Parallel Interference Aware and Successive Interference Canceling receivers, and evaluates their feasibility. A Mutual Information Effective SNR Mapping (MIESM) methodology based on look up tables is developed and validated. A look-up table quantization analysis is performed and abstraction results are compared with the ones obtained from direct precise computations of the mutual information levels. The advantages and drawbacks of this methodology as well as the feasibility of its deployment for real-time transmissions are extensively discussed. As an alternative, this chapter considers an Exponential Effective SNR (EESM) mapping based on the approximation of the PIA and SIC detection mechanisms by the MMSE and the Interference-free receivers and shows that even this simple and light-weight solution provides sufficient accuracy to perform the link adaptation with practically feasible accuracy.

The parts of this chapter which are dedicated to MIESM abstraction for our PIA receiver are published in:

- E. Lukashova, F. Kaltenberger, R. Knopp, and C. Bonnet, "PHY layer abstraction for SU-MIMO LTE system employing parallel Interference-Aware detection," in International Workshop on Link and System Level Simulations 2016 (IWSLS2'16), Vienna, Austria, June 2016.

Chapter 7: Link Adaptation for the R-ML PIA and SIC Receivers In this chapter, we study a low-complexity Link Adaptation strategy applied to our PIA and SIC

receivers. This concept includes the estimation of the Rank Indicator (RI), the Precoder Matrix Indicator (PMI), and the Channel Quality Indicator (CQI). After detailing the existing solutions, we present a RI estimation strategy based on the condition number of the channel matrix. To evaluate performance gains, we perform link level simulations for different values of condition number threshold. The proposed RI estimation is followed by the introduction of a light-weight precoder calculation methodology based on the evaluation of the correlation coefficients of the channel matrix. Our methodology is then numerically compared with the optimal maximum mutual information based criteria. Our CQI estimation is based on the proposed EESM abstraction methodology studied in Chapter 6.

Chapter 8: Conclusion and Future Research This chapter summarizes the contributions developed in the previous chapters and highlights the most valuable findings. It also provides perspectives and discusses ideas and methodologies to extends current achievements to 5G receivers using Block QR decomposition techniques, existing low-complexity LLR metrics and slicing with intelligent ordering.

Chapter 9: Résumé en Français This chapter provides a short summary in French language of the motivation and the most important findings obtained in the scope of this thesis.

Appendix A: Measurement Campaign Over the course of these doctoral studies, we have made a few contributions to the French FUI FAPIS project "4G in Vitro". The goal of this project was to bridge the gap between simulations and live experiments by developing a tool that emulates field conditions within a lab. To do so, we performed multiple drive test campaigns which aimed to collect field measurements (such as GPS positions), Key Performance Indicators (such as uplink and downlink throughput), Reference Signal Received Power, Reference Signal Received Quality, and CQI. The obtained data is stored in the traces in a way that the user is able to easily reproduce the results in a virtual mode.

Chapter 2

Background

Contents

2.1 Basics of MIMO Systems

Reliability and increasingly higher data rates have been a constant motivation for the development of wireless communicational systems. The challenging factors impacting these two key performance indicators are practical coding schemes, spectrum and power constrains, the time-varying multipath nature of the channel, and the complexity of the transmitter/receiver chain. Conventional single antenna point-to-point systems operate in time and frequency domains. Multi-antenna systems, widely known as Multiple-Input-Multiple-Output (MIMO), utilize an extra dimension — space. The multiple antenna

techniques were first developed in the middle of the twentieth century, for military purposes. Civil applications of the MIMO systems followed in 90s, when Spatial Division Multiple Access (SDMA) was introduced by [Roy, 1997] and allowed to improve cell coverage and capacity. The proposed SDMA scheme was illustrated with the antenna array deployed at the base station, combined with adaptive signal processing techniques. At the same time a transmission scheme with spatially separated transmitters (tunned to the same frequency) and multiple receivers with the antenna arrays was shown to increase the broadcast information capacity [Paulraj and Kailath, 1994]. These concepts were further supported by a fundamental information-theoretic analysis performed by [Foschini and Gans, 1998; Telatar, 1999], which dramatically extended the performance boundaries. A variety of antenna arrays techniques were developed aiming to provide reliable radio communications or supply high data rates to users.

2.1.1 MIMO Transmission Techniques

Consider an arbitrary wireless communicational system with n_{tx} transmit and n_{rx} receive antennas. Conventional point-to-point communication system consists of the transmitter and the receiver both equipped with a single antenna and is called *single-input-single-output* (SISO) system. More advanced systems are equipped with multiple antennas at the transmitter and/or the receiver side. Depending on the configuration, such systems are called *single-input-multiple-output* (SIMO), *multiple-input-single-output* (MISO), or *multiple-input-multiple-output* (MIMO) systems. MIMO systems are based on the idea of combining intelligently the multiple received copies of the same signal or independent streams in a complementary way to achieve throughput gains when the medium is favorable, or satisfy reliability requirements when the channel is ill-conditioned. MIMO transmission methods can be classified as *diversity*, *spatial multiplexing*, and *beamforming*.

Diversity. If the transmit or receive antennas are separated by a sufficient interval inside the antenna array, fading can be considered uncorrelated. The probability of multiple copies of the same signal being simultaneously in deep fade is low, and the communication link is more reliable. The received copies are complementary added up at the receiver, using maximum ratio combining [Shah and Haimovich, 2000]. Diversity design can be applied in SIMO systems (receive diversity), MISO, or MIMO systems (transmit diversity in the last two cases). The transmit diversity is achieved thanks to intelligent *space-time* code design [Alamouti, 1998; Guey et al., 1999; Tarokh et al., 1998]. The maximum diversity gain that can be achieved in arbitrary MIMO system equals the number of independent paths between antenna pairs: $n_{\text{tx}}n_{\text{rx}}$.

Multiplexing. To increase the overall system throughput, the paths between the transmit and the receive antenna pairs can be used to send independent signals in favorable channel conditions using the same time-frequency resources. Such signals use an independent spatial resource, and thus are called *spatially multiplexed*. The success of the scheme also depends on the receiver architecture and interference handling strategy. An example of the receiver supporting this transmission scheme is the VBLAST scheme [Hassibi and Hochwald, 2002] where the transmitter sends spatially multiplexed signal to one or multiple users. In this case we talk about *Single-User* MIMO or *Multi-User* MIMO respectively [Gesbert et al., 2007].

In an arbitrary MIMO system, both the diversity and the multiplexing gains can be experienced simultaneously, although not at the maximum level, but rather with so called *diversity-multiplexing trade-off* [Zheng and Tse, 2003].

Beamforming. By manipulating the weights on each transmit antenna, the transmitter focuses power and constructs the beam in the required directions to provide a particular user with throughput benefits and reduce co-channel interference. An exhaustive survey of the beamforming schemes was presented by [Godara, 1997].

2.1.2 Capacity of MIMO channels

Knowledge of information-theoretic limits is essential to understand the potential of wireless systems. The seminal work of Claude Shannon [Shannon, 1948] introduces fundamental results on the capacity limits for Additive White Gaussian Noise (AWGN) channels. The channel capacity C is a maximum rate at which a reliable communication can be performed, without constraints on the complexity of the transmitter or the receiver. It means that for any rate $R < C$ there exists a code that guarantees the signal reception with an arbitrary small probability of error P_e. Shannon proved that the channel capacity equals the mutual information $I(X;Y)$ between the transmitted signal X and the received signal Y maximized over all possible input distributions:

$$C = \max I(X;Y) = \max_{p(x)} \int_{x,y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}, \tag{2.1}$$

where $p(.)$ is a probability density function. Applying Jensen's inequality to the definition of the mutual information (2.1) and replacing integrals with sums, we obtain the representation of the mutual information through entropy (2.2). The entropy $\mathcal{H}(.)$ is defined as an expected value of mutual information: $\mathcal{H}(.) = \mathbb{E}[I(.)] = -\mathbb{E}\log p(.)$.

$$\begin{aligned}
I(X;Y) &= \sum_{x,y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)} \\
&= \sum_{x,y} p(x,y) \log_2 \frac{p(x,y)}{p(y)} - \sum_{x,y} p(x,y) \log_2 p(x) \\
&= \sum_{x,y} p(y)p(x|y) \log_2 p(x|y) - \sum_{x,y} p(x,y) \log_2 p(x) \\
&= \sum_{y} p(y) \left(\sum_{x} p(x|y) \log_2 p(x|y) \right) - \sum_{x} \log_2 p(x) \left(\sum_{y} p(x,y) \right) \\
&= -\sum_{y} p(y) \mathcal{H}(X|Y) - \sum_{x} \log_2 p(x)p(x) \\
&= -\mathcal{H}(X|Y) + \mathcal{H}(X) \\
&= \mathcal{H}(X) - \mathcal{H}(X|Y).
\end{aligned} \tag{2.2}$$

Given an AWGN channel with SNR γ and bandwidth B, the optimal input distribution is Gaussian, and the corresponding capacity can be computed as follows:

$$C = B \log_2 (1 + \gamma) \quad \text{bps}, \tag{2.3}$$

assuming

$$\gamma = \frac{E_s}{N_0 B}, \tag{2.4}$$

where E_s is the total average energy per symbol and N_0 is the noise variance.

In the literature, we often meet the terms *ergodic capacity* and *outage capacity*. The ergodic capacity is also called the Shannon capacity and is used to obtain the averaged values of long-term throughput. This metric provides a valid performance representation when the channel varies fast and the realizations follow the a certain channel distribution.

In contrast, the outage capacity is an instantaneous channel characterization. The probability of outage P_{out}, given by

$$P_{\text{out}} = \Pr(C < R), \tag{2.5}$$

is a measure of the reliability of communication: an *outage event* occurs when the rate R assigned to the codeword is higher than the capacity C that the radio link can support under the current channel conditions. This metric is often used when we wish to adapt the transmission rate to the current channel conditions. This requires the presence of *Channel State Information* (CSI) at the transmitter. Generally, CSI is obtained by the UE based on the channel estimates from pilot symbols and is then fed back to the transmitter.

MIMO systems exploit various space-time coding techniques to improve the reliability or increase the data rates. Unfortunately, the existing codes do not provide both reliability and throughput gains at the maximum level simultaneously. For example, a spatial multiplexing transmission scheme allows MIMO systems to potentially reach a capacity gain of the factor $\min\{n_{\text{tx}}, n_{\text{rx}}\}$ compared to traditional SISO architecture, given the same bandwidth and total transmit power. In contrast, diversity techniques, such as Alamouti, reduce achievable rates [Sandhu and Paulraj, 2000] by not exploiting the spatial multiplexing potential of multiple layers. The fundamental works on the MIMO capacity for fading channels were performed by [Foschini and Gans, 1998] and [Telatar, 1999] in the late 1990s. Telatar showed that the capacity of the deterministic MIMO channel can be computed as

$$C = \max_{\text{Tr}(\mathbf{R}_{xx})=n_{tx}} \log_2 \det \left(\mathbf{I}_{n_{\text{rx}}} + \frac{E_x}{N_0 n_{\text{tx}}} \mathbf{H} \mathbf{R}_{xx} \mathbf{H}^H \right), \tag{2.6}$$

with n_{tx} transmit and n_{rx} receive antennas in fading channel $\mathbf{H} \in \mathbf{C}^{n_{\text{rx}} \times n_{\text{tx}}}$.

In practice, a MIMO channel varies as a random ergodic process, and to compute the ergodic capacity of such a channel, we thus take an expectation:

$$\overline{C} = \mathbb{E}\{C(\mathbf{H})\} = \mathbb{E}\left\{ \max_{\text{Tr}(\mathbf{R}_{xx})=n_{\text{tx}}} \log_2 \det \left(\mathbf{I}_{n_{\text{rx}}} + \frac{E_x}{N_0 n_{\text{tx}}} \mathbf{H} \mathbf{R}_{xx} \mathbf{H}^H \right) \right\}. \tag{2.7}$$

Depending on the availability of the CSI at the transmitter, the ergodic capacity (2.7) can be further developed. The presence of the CSI allows to perform an optimum

power allocation, such as *waterfilling*. Following the waterfilling concept, the weighting coefficients λ_i and μ_i are obtained through Singular Value Decomposition (SVD) and Eigen Value Decomposition (EVD) which transform the MIMO channel into r equivalent SISO channels.

$$\overline{C}_{\text{CSI}} = \mathbb{E}\left\{ \max_{\sum_{i=1}^{r} \mu_i = n_{\text{tx}}} \sum_{i=1}^{r} \log_2\left(1 + \frac{E_x}{N_0 n_{\text{tx}}} \mu_i \lambda_i\right) \right\}, \tag{2.8}$$

where $i = \{1, 2, \ldots, r\}$ is the stream index and $r = \min\{n_{\text{tx}}, n_{\text{rx}}\}$ is the total number of equivalent SISO streams. If the CSI is not available, the transmitter distributes power equally. In this case the capacity can be computed as follows:

$$\overline{C}_{\text{NOCSI}} = \sum_{i=1}^{r} \log_2\left(1 + \frac{E_x}{N_0 n_{\text{tx}}} \lambda_i\right). \tag{2.9}$$

The probability of outage in case of a random MIMO channel is given by:

$$P_{\text{out}} = \Pr(C(\mathbf{H}) < R), \tag{2.10}$$

and the outage capacity of MIMO channel C_ε is the maximum rate that guarantees that the probability of outage (2.11) is less than ε:

$$P(C(\mathbf{H}) \leq C_\varepsilon) \leq \varepsilon, \, \forall C \leq C_\varepsilon. \tag{2.11}$$

In Long Term Evolution (LTE) systems, the initial bit sequence is mapped onto constrained M-QAM symbols using Bit Interleaved Coded Modulation (BICM). BICM was initially proposed by [Zehavi, 1992] and the theoretical analysis was developed by [Caire et al., 1998]. The most widely applied alphabets are 4QAM, 16QAM, 64QAM, and 256QAM. The code cardinality is 2^M, where M denotes the modulation order which defines how many bits are carried by a single symbol. We assume that all the symbols of the transmitted vector $\mathbf{x}$ belong to the same modulation alphabet $\mathcal{Q}^M$. As follows from (2.2), the capacity of random MIMO channels with finite alphabets can be expressed through the mutual information $I(\mathbf{X}; \mathbf{Y}|\mathbf{H})$:

$$I(\mathbf{X}; \mathbf{Y}|\mathbf{H}) = \mathcal{H}(\mathbf{X}|\mathbf{H}) - \mathcal{H}(\mathbf{X}|\mathbf{Y}, \mathbf{H}) = \log_2 M - \mathcal{H}(\mathbf{X}|\mathbf{Y}, \mathbf{H}). \tag{2.12}$$

The second term of (2.12) is expressed as follows:

$$\mathcal{H}(\mathbf{X}|\mathbf{Y}, \mathbf{H}) = \sum_{\mathbf{x} \in \mathcal{Q}^M} \int_{\mathbf{y}} p(\mathbf{x}, \mathbf{y}|\mathbf{H}) \log_2 \frac{1}{p(\mathbf{x}|\mathbf{y}, \mathbf{H})} d\mathbf{y} \tag{2.13}$$

$$= \sum_{\mathbf{x} \in \mathcal{Q}^M} \int_{\mathbf{y}} p(\mathbf{x}, \mathbf{y}|\mathbf{H}) \log_2 \frac{\sum_{\mathbf{x}' \in \mathcal{Q}^M} p(\mathbf{y}|\mathbf{x}', \mathbf{H})}{p(\mathbf{y}|\mathbf{x}, \mathbf{H})} d\mathbf{y}. \tag{2.14}$$

The right side of (2.13) does not have a closed form, but can be approximated using Monte-Carlo simulations, and after averaging across large amount of noise realizations we get:

$$I(\mathbf{X};\mathbf{Y}|\mathbf{H}) = \log_2(M^{n_{\text{tx}}}) - \frac{1}{MN_n n_{\text{tx}}}\left(\sum_{\mathbf{x}\in Q^M}\sum_{n\in\mathcal{N}}\log\sum_{\mathbf{x}'\in\mathcal{Q}^M}\exp\left[\frac{-\|\mathbf{y}-\mathbf{H}\mathbf{x}'\|^2+\|\mathbf{n}\|^2}{N_0}\right]\right), \quad (2.15)$$

where the noise realization is drawn from a set $n \in \mathcal{N}$ with the maximum number of noise realizations $|\mathcal{N}| = N_n$.

2.2 Basics of LTE

2.2.1 Standardization Overview

With time, the role of standard developing organizations and units such as *Third Generation Partnership Project* (3GPP) and *International Telecommunication Union* (ITU) in the mobile technology design has become more prominent. ITU defines the requirements that the new standards should target (for example, throughput, latency, etc), performs active research studies and issues recommendation documents. 3GPP, being a group of telecommunications associations known as the Organizational Partners, covers in the fundamental specifications the core network, the radio access, and the service capabilities.

3GPP experts have developed multiple advanced features in consequent Releases 8–13. In particular, Release 8 introduced LTE for the fist time and provided a fundamental description of a flat network architecture with Orthogonal Frequency Division Multiplexing Access (OFDMA) for downlink and Single Carrier Frequency Division Multiplexing Access (SC-FDMA) for uplink, functionalities of the base station and baseline MIMO Transmission Modes. The promised peak data rates of a 2×2 MIMO system given 20 MHz of LTE bandwidth were 300 Mbit/s and 75 Mbit/s for downlink and uplink respectively. To improve the network coverage, flat network architecture was extended to the heterogeneous with LTE-compatible femtocells [3GPP, 2013] in Release 9. To reduce the cost of maintenance, the networks gained self-organizing features [3GPP, 2010] and now became capable of self installation, optimization and healing.

An idea to circumvent the limited spectrum by allowing user to transmit on different carriers — fragmented spectrum of several base stations was implemented in Release 10, which extended the maximum bandwidth to 100 MHz. Further throughput boost was enabled with high order MIMO transmission schemes (8×8 for downlink and 4×4 for uplink) and support of relay nodes. However, the problem of increased levels of interference became more significant, which was only partially solved with enhanced Inter-Cell Interference Coordination (eICIC). The advanced interference management came with *Coordinated Multi-Point* (COMP) transmission and with advanced receivers [3GPP, 2012] in Release 11, and Network-Aided Assisted Interference Cancellation (NAIC) [3GPP, 2014] technique in Release 12. The NAIC receivers benefit from the exchange of cell configuration information between the neighboring eNodeBs through X2 interface, and neighboring cell configuration parameters between the eNodeB and the UE. The extensive research on physical layer advances introduced *MultiUser Superposition Transmission* (MUST)

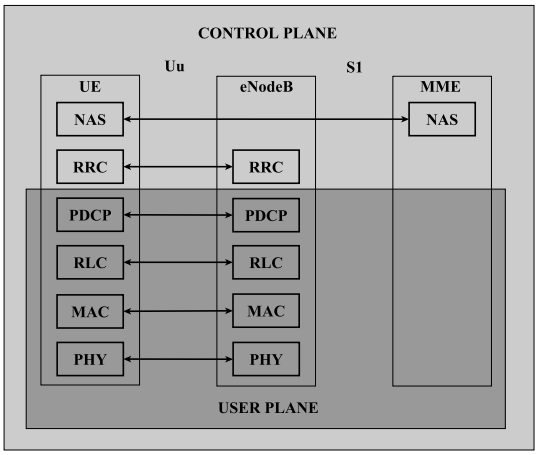


Figure 2.1: LTE air interface protocol stack from the UE point of view.

[3GPP, 2015e; Vanka et al., 2012] and *Non-Orthogonal Multiple Access* (NOMA) [Lan et al., 2014; Yan et al., 2015] in Release 13. Additionally to the enhancements of existing services and features, Release 13 considered 5G-specific features such as machine-type communications, public safety features and small cell dual-connectivity and architecture. Releases 14 and 15 are still ongoing standardization processes and are expected to come at the end in 2017 and 2018 respectively. Release 14 is dedicated to Mission Critical enhancements, LTE support for V2x services, and inter-band Carrier Aggregation, while Release 15 is focused on the importance of forward compatibility in both radio and protocol design and releasing 5G specifications.

2.2.2 Protocol Stack and Data Flow

LTE design incorporates a layered protocol stack consisting of six main components: the Non-Access Stratum (NAS), the Radio Resource Control (RRC), the Packet Data Convergence Protocol (PDCP), the Radio Link Control (RLC), the Medium access control (MAC) and the Physical Layer (PHY), as illustrated in Fig. 2.1. The majority of blocks are common between the user plane and the control plane at the UE side, with only RRC being control-plane specific. Each layer receives Service Data Units (SDU) from the preceding layer and passes Packet Data Units (PDU) to the layers below. Fig. 2.2 illustrates an example of a downlink data flow.

The NAS is a functional layer between the core network and the UE. It is responsible for the establishment of communication sessions, identity management, call control and support of continuous communication in mobility scenarios. The next protocol layer, the RRC, manages radio bearers and low layer configuration in a dialog between the UE and the base station. The PDCP deals with header compression and decompression, security processes such as ciphering or deciphering of the IP packets and verification. The RLC is located between the PDCP and the MAC layers. It reformats SDUs by segmentation of the transport blocks passed by the PDCP with excess of the length or concatenation, if required. It is also responsible for Hybrid Automatic Repeat Request (HARQ) reordering. The MAC layer is connected to the RLC and the PHY layer through the logical and

Table 2.1: LTE Logical and Transport Channels

Channel	Acronym	Purpose
Logical Channels:		
Broadcast Control Channel	BCCH	Conveys system information.
Paging Control Channel	PCCH	Carries paging information to registered terminals.
Common Control Channel	CCCH	Transfers control information to/from idle UEs.
Dedicated Control Channel	DCCH	Transmit dedicated control information to/from the UE.
Dedicated Traffic channel	DTCH	Used for user-plane traffic.
Multicast Traffic Channel	MTCH	Associated with multicast/broadcast services.
Multicast Control Channel	MCCH	Control information for reception of the MTCH.
Transport Channels:		
Broadcast Channel	BCH	Transmits the Maser Information Block.
Paging channel	PCH	Conveys paging information from PCCH.
Downlink Shared Channel	DLSCH	Main transport channel for downlink data.
Multicast Channel	MCH	Used for multicast/broadcast services.
Uplink Shared Channel	ULSCH	Main transport channel for uplink data.
Random-Access Channel	RACH	Transmits data fro initial access or state changes.

transport channels respectively, which are summarized together with their main functions in Table 2.1. The MAC layer performs reception of the transport blocks, combining the data and generation of acknowledgment (ACK) and non-acknowledgment (NACK) messages, as well as scheduling and multiplexing of users. The PHY layer, which is the focus of this thesis, includes channel coding, modulation, precoding, resource mapping, etc. A detailed description of the physical layer is given in Section 2.2.6.

2.2.3 Transmission Modes

The LTE has been designed to allow maximum flexibility in the exploitation of the benefits of MIMO channels. The eNodeB can be configured to serve the UEs in different downlink Transmission Modes (TM) (ten, according to Release 13 [3GPP, 2016]), which range from transmit diversity, over beamforming to spatial multiplexing. The Transmission Mode is set for every scheduled UE and depends on capabilities of the UE and eNodeB, on engineering of the eNodeB cell site and on propagation environment. The Transmission Mode is configured in the RRC message during the initial connection to the network or during the reestablishment of the RRC connection after a radio link failure. To avoid signaling overhead, each Transmission Mode supports at least two transmission schemes, depending on the received Downlink Control Information (DCI) and the Search Space for PDCCH candidate (Table 2.2, Fig. 2.4). The Transmission Modes above TM2 are applicable only to the DLSCH channel, while the BCH and PCH are transmitted using single antenna or transmit diversity schemes only.

The essential part of Transmission Modes $3-6$ is the Channel State Information from the UE which characterizes the quality of a radio link and the UE capabilities. The CSI feedback may have up to three components: the Rank Indicator (RI), the Precoder Matrix Indicator (PMI) and the Channel Quality Indicator (CQI). The RI defines the amount of the spatial layers that the UE can support, and may take values in $\mathcal{RI} \in \{1, 2, 3, 4\}$. However, only two codewords are allowed for simultaneous transmission in the current

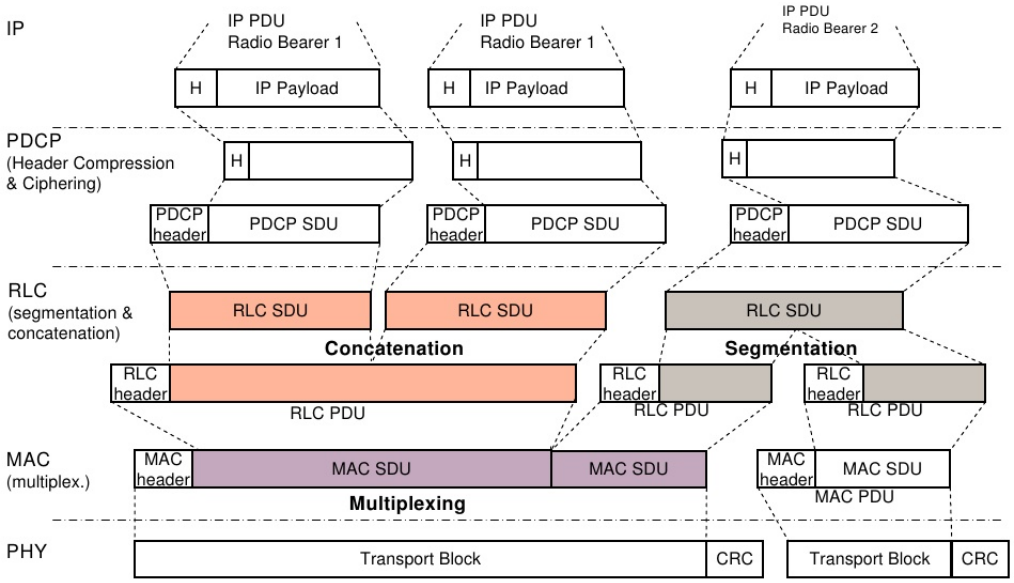


Figure 2.2: Illustration of the LTE downlink data flow [Lindström, 2009]. Each layer receives a SDU from the preceding layer and passes a PDU to the layers below.

standard. The PMI is used to signal the weights to be applied during the precoding process, in order to maximize the downlink SINR or maximize system capacity and reduce interference. The precoder is chosen from the set of predefined codebooks [3GPP, 2015c]. The CQI is used to indicate the maximum MCS that the UE can support with less than 10% Block Error Rate (BLER), but it is up to the eNodeB to take it into account. In LTE, the base station is intelligent enough to track the Block Error Rate based on the ACK/NACK rates.

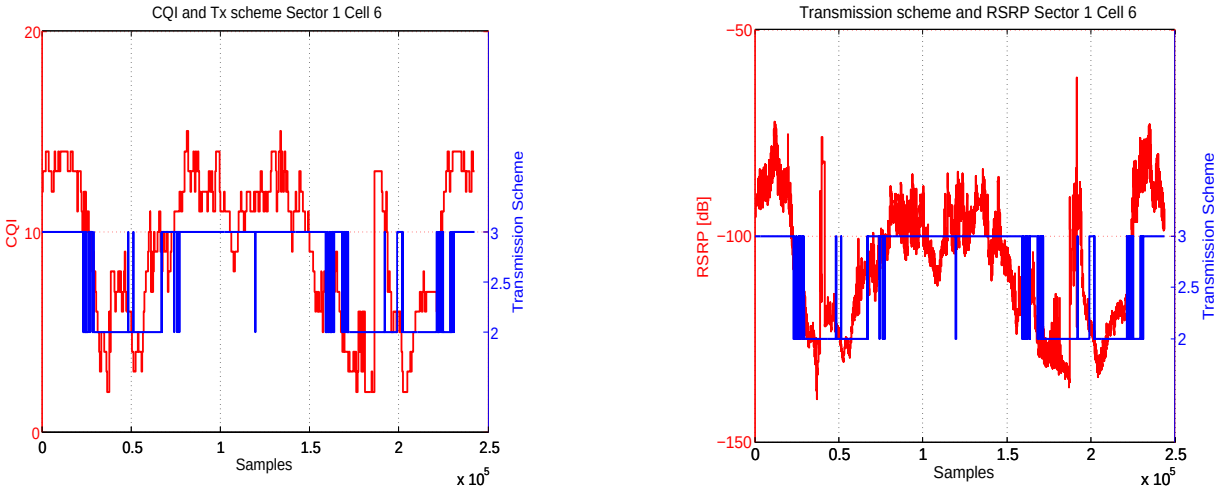


Figure 2.3: Correlation between the transmission scheme, the CQI and the RSRP tracked during a measurement campaign in Sophia Antipolis, France. When the link quality is below certain threshold, the Transmit diversity is preferred over Open Loop spatial Multiplexing.

During one of our measurements campaigns, performed in Sophia Antipolis, France, in the real network with Ericsson's eNodeB configured in TM $1-3$, we tracked the correlation between the quality of radio link (CQI, Reference Signal Received Power) and the transmission scheme (Fig. 2.3): when the link quality is below a certain threshold, the Transmit diversity is preferred over the Open Loop Spatial Multiplexing.

TM1 is the most primitive Transmission Mode that supports SISO and SIMO combinations and is used in single antenna cells (femtocells in practice) and for some control channels.

TM2, also known as Single Layer Transmit Diversity, is a default Transmission Mode used to increase robustnesses since the same information is transmitted over multiple antennas. Alamouti coding is used in the setting of two transmit antennas, and is combined with Frequency Shift Time Diversity codes for four transmit antennas.

A group of the Transmission Modes, supporting spatial multiplexing, includes Multi-User MIMO TM5, and Single-User MIMO TM3 Open Loop Spatial Multiplexing (OLSM), TM4 Single-User MIMO Closed Loop Spatial Multiplexing (CLSM) with two codewords and Single Layer Beamforming TM6. All spatial multiplexing modes, introduced in Release 8, support Channel State Information feedback to a certain extent. Precisely, the UE performs the channel estimation in TM3-6 based on reference symbols and feedbacks the CQI and the LTE codebook based PMI. In addition to the PMI and the CQI, TM4 also feedbacks the RI.

A group of advanced Transmission Modes (TM7-10), supporting non-codebook based precoding, was recently introduced in Release 10 and 11. These Transmission Modes handle up to 8 transmission layers.

2.2.4 Downlink Control Information

The DCI carries scheduling assignments, requests for the aperiodic CQI and uplink power control command. To proceed with the decoding of the real data, the UE first has to decode the DCI and extract the critical information about the modulation scheme, the resource allocation, etc. Various DCI formats have common information fields, such as the MCS, the redundancy version, the resource allocation type, while some types of the fields are valid only for certain formats. The detailed DCI field description is provided in [3GPP, 2015b]. Every DCI format is assigned to a particular use case. The most frequently used DCI formats were introduced in LTE Release 8 aiming to serve Transmission Modes $1-6$.

Not every DCI format is used carry the scheduling signaling for the physical downlink channels. For example, format 0 and format 1C convey assignments for Physical Uplink Shared Channel (PUSCH) and paging or system information, and formats 3 and 3A are used to transfer power control commands for Physical Uplink Control Channel (PUCCH) and PUSCH for a group of UEs.

DCI format 1 is a baseline format used to carry one Physical Downlink Shared Channel (PDSCH) codeword in SISO and SIMO antenna configuration. For systems with spatial multiplexing, format 1D is used in Multi-MIMO case, while formats 1B (rank one only), 2 and 2A are dedicated to Single-User MIMO systems. In addition to the standard fields (modulation and coding scheme, redundancy version, etc), formats 1B, 1D and 2 carry precoding information shared between the base station and the UE. Format 2A has

Table 2.2: LTE Transmission Modes, DCI formats and Transmission schemes [3GPP, 2016]

TM	DCI format	Search Space	Tx Scheme
TM1	1A	Common and UE specific by C-RNTI	Single antenna port 0
	1	UE specific by C-RNTI	
TM2	1A	Common and UE specific by C-RNTI	Transmit diversity
	1	UE specific by C-RNTI	
TM3	1A	Common and UE specific by C-RNTI	Transmit diversity
	2A	UE specific by C-RNTI	Large delay CDD or transmit diversity
TM4	1A	Common and UE specific by C-RNTI	Transmit diversity
	2	UE specific by C-RNTI	CLSM or transmit diversity
TM5	1A	Common and UE specific by C-RNTI	Transmit diversity
	1D	UE specific by C-RNTI	Multi-user MIMO
TM6	1A	Common and UE specific by C-RNTI	Transmit diversity
	1B	UE specific by C-RNTI	CLSM using a single transmission layer
TM7	1A	Common and UE specific by C-RNTI	Single antenna port 0, if the number of PBCH antenna port is one, otherwise transmit diversity
	1	UE specific by C-RNTI	Single antenna port 5
TM8	1A	Common and UE specific by C-RNTI	Single antenna port 0, if the number of PBCH antenna port is one, otherwise transmit diversity
	2	UE specific by C-RNTI	Dual layer transmission ports 7-8
TM9	1A	Common and UE specific by C-RNTI	Non-MBSFN subframe: Single antenna port 0, if the number of PBCH antenna port is one, otherwise transmit diversity. MBSFN subframe: Single antenna port 7
	2C	UE specific by C-RNTI	Up to 8 layer transmission ports 7-14 or single antenna ports 7, 8, 11 or 13
TM10	1A	Common and UE specific by C-RNTI	Non-MBSFN subframe: Single antenna port 0, if the number of PBCH antenna port is one, otherwise transmit diversity. MBSFN subframe: Single antenna port 7
	2D	UE specific by C-RNTI	Up to 8 layer transmission ports 7-14 or single antenna ports 7, 8, 11 or 13

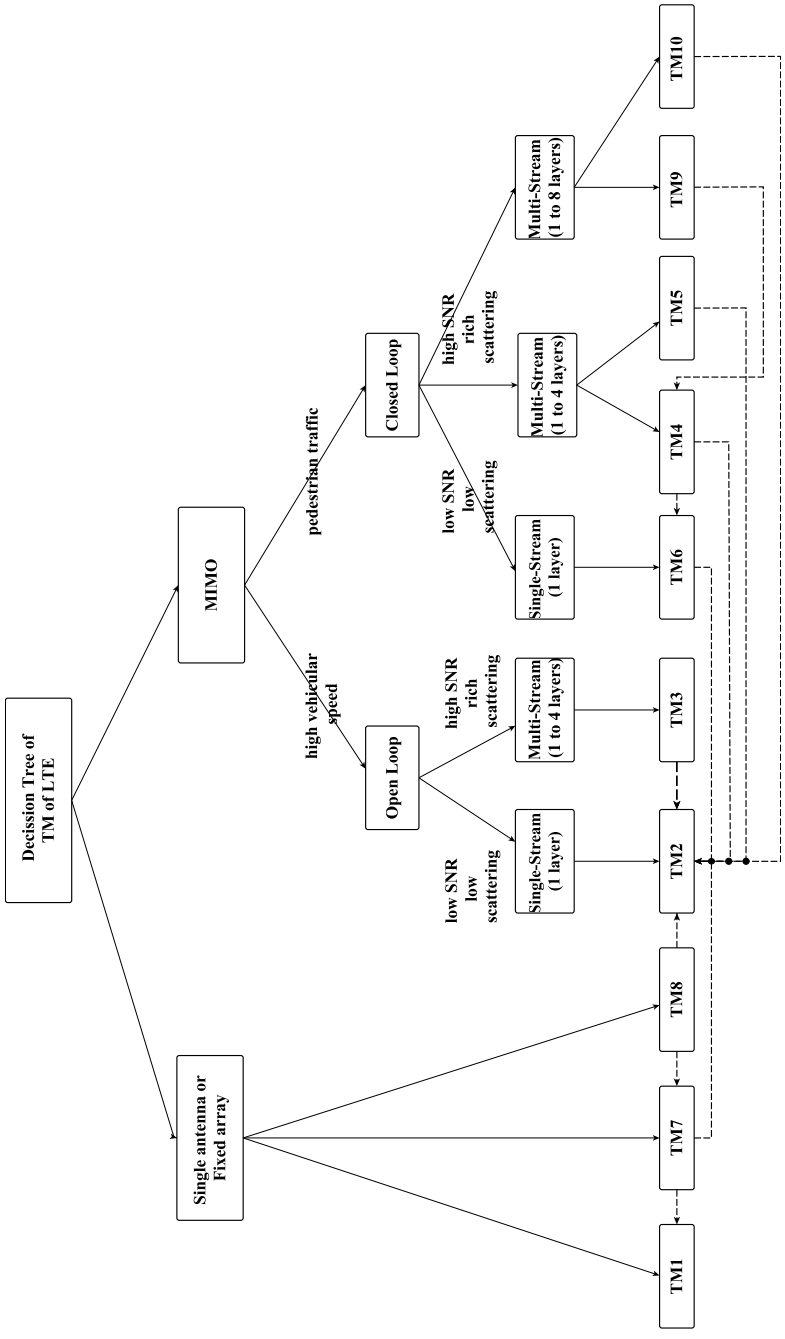


Figure 2.4: Decision Tree of LTE Transmission Modes based on [4GAmericas, 2013]. The dashed lines symbolize the fallback links between the transmission schemes.

an active precoding information field only if 4 antenna ports are enabled at the eNodeB.

Recent LTE releases introduced new Transmission Modes which are based on non-codebook precoding and thus require corresponding DCI formats: formats 2C and 2D, that support up to 8 transmission layers.

2.2.5 LTE Frame Structure

The LTE physical layer is built from so-called *Physical Resource Blocks* (PRB). The PRB are designed in both time and frequency domain. One PRB occupies 12 subcarriers on the frequency domain (15 kHz of bandwidth for each subcarrier). On the time domain, one PRB is granted one *time slot*, which lasts 0.5 ms. Two consequent times slots form a *subframe*, which are similarly grouped into *frames* with a total duration of 10 ms (see Fig. 2.5).

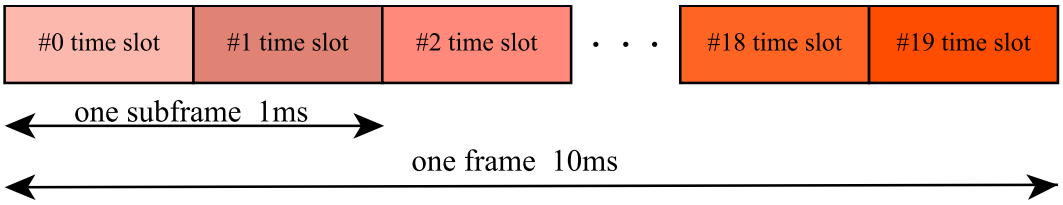


Figure 2.5: Frame subdivision in LTE.

LTE Releases 13 and 14 support three Frame Structures: Type 1, which is used for the Frequency Division Duplex (FDD), Type 2 which is applied for the Time Division Duplex (TDD), and a new Type 3 which is designed for the LTE Augmented Access (LAA). During a FDD transmission, the uplink and downlink frames are separated in frequency, as it is shown in Fig. 2.6. The frequency separation must be sufficient for the uplink and downlink frames not to interfere with each other. In contrast, a TTD transmission does not require a paired spectrum, since both uplink and downlink frames are sent through the same channel (Fig. 2.7). The TTD Frame structure introduces Special Subframe (SS), which is used to switch between the uplink and the downlink transmissions, and contains three fields: Downlink Pilot Time Slot (DwPTS), Uplink Pilot Time Slot (UpPTS) and Guard Period (GP). The proportion of uplink and downlink subframes is determined by one of the seven uplink-downlink configurations. A TTD transmission allows to exploit channel reciprocity, as the transmitter and the receiver may use the same set of parameters which is crucial in Massive MIMO [Rogalin et al., 2014; Shepard et al., 2012]. It also allows to dynamically change the uplink/downlink ratio based on the current user demand [ElBamby et al., 2014; Zhang and Fang, 2015]. The LAA transmission is built upon the TDD frame structure, but the downlink transmission may start anywhere within a subframe and finish either at the end of the last subframe, or follow one of the DwPTS durations.

The minimum amount of resources that can be scheduled to one user is two consecutive PRB. The total number of PRB available at the base station for scheduling depends on the supported bandwidth. A typical base station supports 1.4, 5, 10 and 20 MHz transmission, which is equivalent to allocating 6, 25, 50 and 100 PRBs. Every time slot is

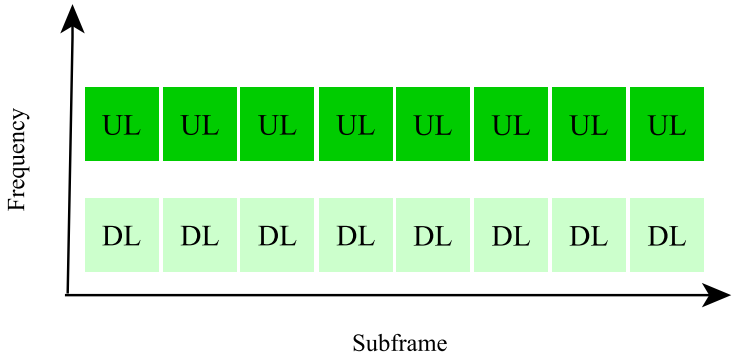


Figure 2.6: Example of an FDD transmission, where the uplink and downlink subframes are sent via paired spectrum.

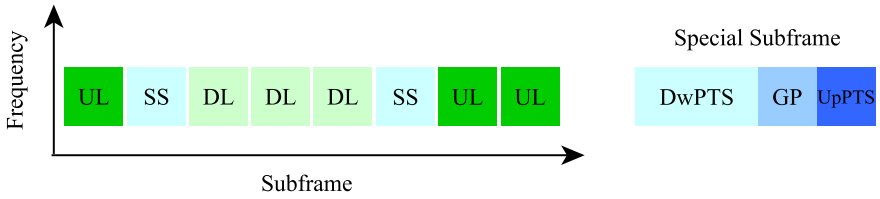


Figure 2.7: Example of a TDD transmission. The Special Subframe includes DwPTS, UpPTS and GP. The DwPTS contains Primary Synchronization Signal, while UwPTS may carry Physical Random Access Channel or Signaling Reference Signal.

filled with 6 or 7 OFDM symbols. The exact amount of the OFDM symbols depends on whether Normal or Extended Cyclic Prefix are applied. This structure produces a two-dimensional unit (symbol duration × subcarrier bandwidth) — Resource Element, placed in the time-frequency resource grid. Every Resource Element is mapped onto one of the physical channels, such as control channels, broadcast channels, channels for conveying user-specific data. The mapping is defined by [3GPP, 2015c] and illustrated in Fig. 2.8 in configuration for a 1.4 MHz LTE bandwidth. The resource grid scales with the number of PRBs, and may slightly vary based on allocation parameters, for example, the amount of PDCCH symbols, the number of antenna ports, etc. The remaining sections shall provide more insights on the physical layer procedures and detail the signal processing.

2.2.6 Physical Layer Processing

The physical layer provides services to the MAC layer in the form of transport channels. The majority of the transport channels are mapped onto physical channels, which can be seen as a set of time-frequency resources. Only physical channels, which carry Downlink or Uplink Control Information do not have a corresponding pair. The central physical channels are summarized in Table 2.3.

The primary goal of the LTE physical layer is to prepare the data received from the upper layers for a reliable transmission between the base station and the UE. Fig. 2.9 illustrates the downlink physical layer signaling for the LTE system with two spatially multiplexed codewords. As an input to signal processing, the physical layer receives the

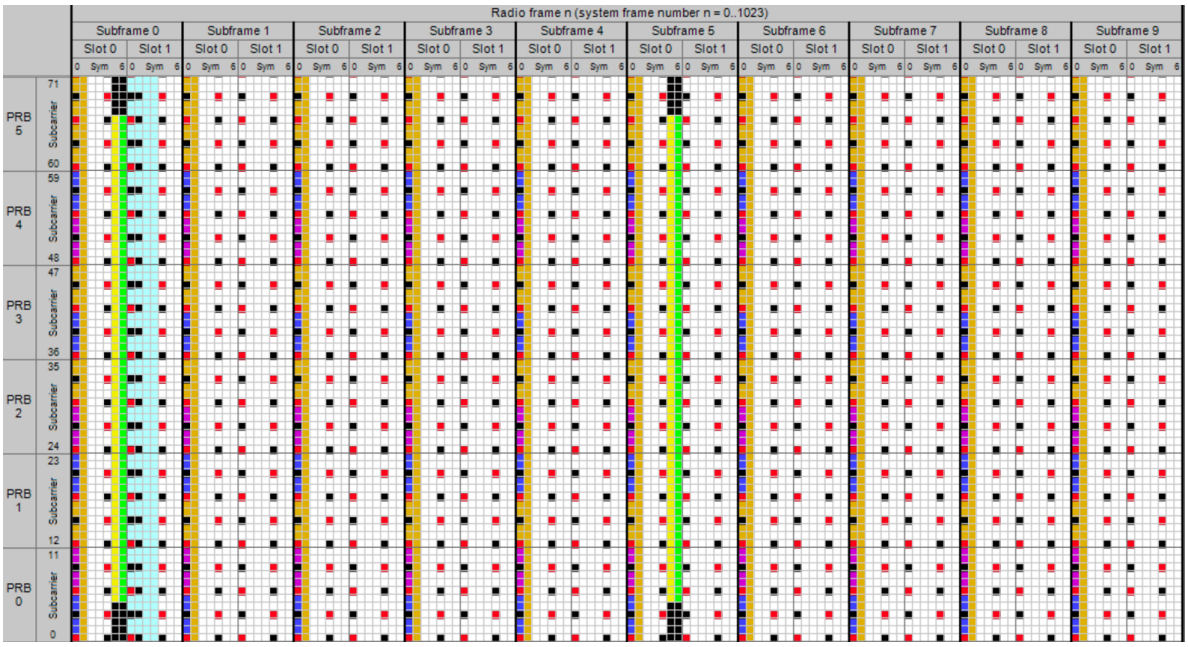


Figure 2.8: Example of the FDD LTE resource grid [Dhagle, 2009] on the antenna port 0 of the base station in the 2×2 MIMO system. The LTE bandwidth is set to 1.4MHz (6 PRBs and 72 subcarriers) with normal Cyclic Prefix and 2 PDCCH symbols. The green and yellow cells correspond to the Primary and Secondary Synchronization Channels. The light blue cells depict the Physical Broadcast Channel. The red pixels are the Reference Signals for the selected antenna port, while the black pixels are the Reference signals for the second antenna port. The dark blue, purple and brown cells mark three control channels: Physical Control Format Indicator Channel, Physical HARQ Indicator Channel, Physical Downlink Control Channel. The resources in the white area are dedicated to the PDSCH transmission.

Table 2.3: LTE Physical Channels

Channel	Acronym	Purpose
Logical Channels:		
Physical Broadcast Channel	PBCH	Conveys essential system information to the UE.
Physical Downlink Control Channel	PDCCH	Carries DCI information.
Physical Uplink Control Channel	PUCCH	Carries Uplink Control Information.
Physical Downlink Shared Channel	PDSCH	Main channel to convey downlink traffic.
Physical Uplink Shared Channel	PUSCH	Main channel to convey uplink traffic.

data organized in two transport blocks from the MAC layer and outputs two codewords to the propagation medium. In the case of an HARQ retransmission, it is possible that one of the transport blocks is disabled and the MAC layer passes only one transport block even though the system is configured in CLSM. The transport block size can be obtained from the look-up table with the input parameters of I_{mcs} modulation scheme index and N_{prb} the number of the physical radio blocks, dedicated to the UE [3GPP, 2016].

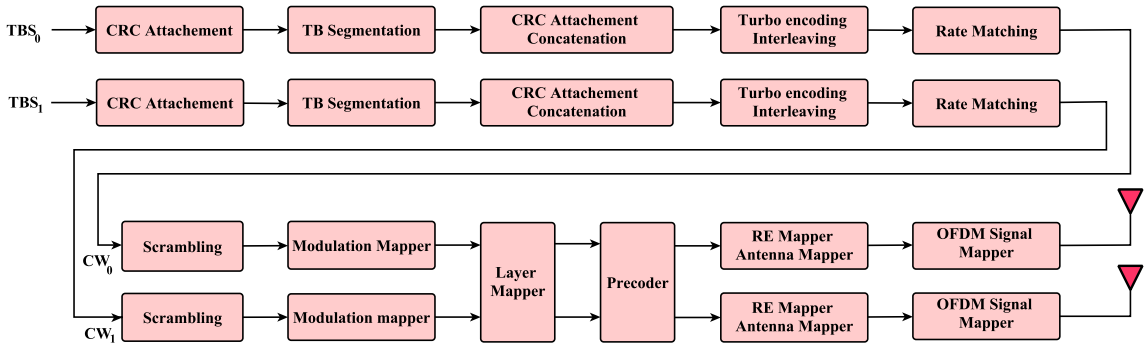


Figure 2.9: LTE Physical Layer signal processing at the eNodeB in 2×2 MIMO system with 2 spatially multiplexed codewords.

2.2.6.1 CRC attachment

To prevent the UE from spending time on decoding an erroneous packet, early stopping algorithms were developed. One of these algorithms is a *Cyclic Redundancy Check* (CRC) — 24-bit are appended to each transport block, computed from all the bits using CRC generator polynomial $g_A(D)$. The CRC check results are used in HARQ procedures to indicate decoding success or failure and generate ACK/NACK message. Furthermore, it can be used in Successive Interference Canceling receivers to avoid error propagation.

2.2.6.2 Segmentation

The transport block size is obtained from the look-up table with the modulation and coding scheme index $\mathrm{I_{mcs}}$ and the number of physical radio blocks $\mathrm{N_{prb}}$, dedicated to the UE, as the input parameters. The greater the MCS, the higher the spectral efficiency, at the price of higher SNR requirements. If the size of the transport block including the CRC bits exceeds 6144 bits (e.g., in high mobility scenarios), the transport block is divided into shorter segments, called code blocks, which are easier to process by the turbo encoder. After segmentation, each block is appended with the CRC bits based on the polynomial generator $g_B(D)$. If the Code Block level CRC check passes, the UE checks for the CRC at the Transport Block level to increase the probability of potential error detection as the Code block level CRC is different from the Transport Block level CRC.

If required, zero-padding bits are appended at the beginning of the transport block to match the discrete transport block sizes. The code blocks are then concatenated.

2.2.6.3 Turbo encoding and Interleaving

LTE channel coding is based on Turbo codes [Berrou and Glavieux, 1996]. Turbo codes belong to the family of forward error-correction codes and are famous for their near-capacity-achieving qualities [Berrou et al., 1993]. A turbo-encoder consist of a concatenation of two convolutional encoders linked by an interleaver [3GPP, 2015b]. The bits are reordered in the interleaver to spread the possible burst error caused by the channel imperfection and noise distortion, and convert it from the burst error to discrete errors.

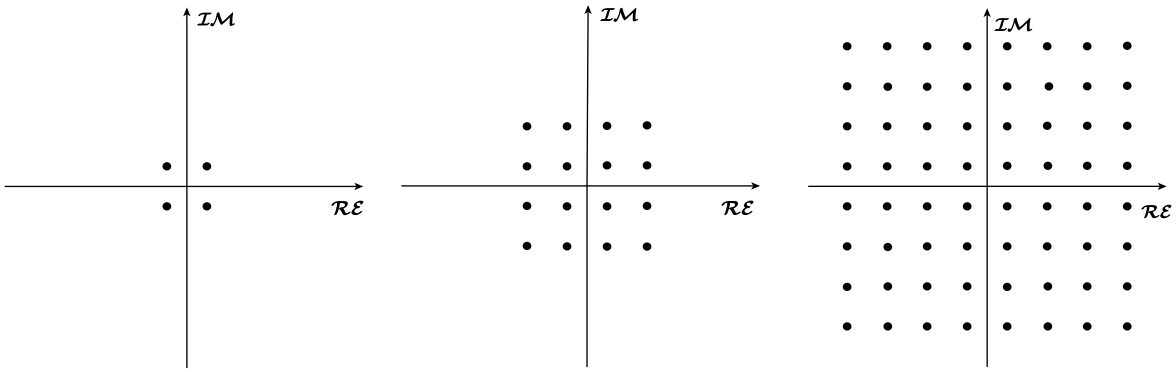


Figure 2.10: Constellations QPSK, 16QAM and 64QAM used for modulation mapping with Grey labeling in LTE systems. Every symbol conveys 2, 4 and 6 bits respectively. QPSK is used for poor channel conditions, while 64QAM provides gains in high SNR regime.

The code rate of the turbo encoder is 1/3, where one input source bit is converted into one systematic and two parity-check bits. The encoded bits are written into a circular buffer.

2.2.6.4 Rate Matching

The purpose of the Rate Matching block is to form the exact set of the encoded bits to be transmitted from the bits written in a circular buffer. This exact set depends on the modulation and coding scheme (defines the rate), as well as on the redundancy version (points to the starting point to read from in a circular buffer [Cheng et al., 2008]).

2.2.6.5 Scrambling

The scrambling approach is introduced to combat the interference from neighboring cells. Scrambling operations are performed at the bit level and consist in the multiplication of the bits delivered by the rate matching unit by the scrambling sequence. The scrambling sequence uses order-31 Gold code, providing 2^{31} unique linearly independent sequences. The scrambling sequence generator is reinitialized every subframe based on the cell identity, subframe number, and the UE identity. This prevents the desired signal decoded at the receiver from being matched to the interfering signals from the neighboring cells.

2.2.6.6 Modulation Mapping

The modulation block maps the scrambled bits onto modulation symbols of the constrained constellation. Current LTE systems use the 2^MQAM constellation, where the modulation order $M \in \{2, 4, 6\}$. Release 12 allowed for 256QAM modulation. Modulation order indicates how many bits are carried by a symbol: 2 bits for the QPSK, 4 bits for 16QAM and 6 bits for 64QAM Fig. 2.10).

2.2.6.7 Layer Mapping and Precoding

The input to the layer mapping unit are the modulation symbols corresponding to one or two transport blocks. Depending on the Transmission Mode and the transmission scheme the UE operates in, various number of spatial layers is configured for the transmission. The standard restricts the maximum transport block number to two, while the maximum number of spatial layers supported in LTE transmission is eight. Even through the DCI format does not have an independent field to signal the active number of layers, the UE may extract the number of layers from a set of DCI parameters.

LTE precoding can be classified into *codebook-based precoding* and *non-codebook-based precoding*, depending on the reference signals. If only cell-specific reference signals are present, the codebook-based precoding is applied (Transmission Modes $3-6$), while the non-codebook-based precoder requires a presence of the demodulation reference signals (Transmission Modes $7-9$).

The codebook-based precoding is characterized by the following properties: constant modulus, nested property, and minimal complex multiplications [Sesia, 2009]. The constant modulus means that only the phase of the signal changes after precoding, while the amplitude remains unchanged. The nested property dictates that the low rank precoder represent a subset of the higher rank precoder matrices or vectors. The LTE codebook is constructed from the QPSK alphabet and all the symbols are multiplied by ± 1 and $\pm j$ to minimize the amount of complex multiplications and simplify the implementation. In Transmission Modes 4 and 6, if the DCI the CSI reporting is configured in PUSCH, the UE selects the desired precoder based on throughput maximization or SINR maximization principle and feedbacks it to the eNodeB as the Precoder Matrix Indicator (PMI). Similar with the CQI, it is the eNodeB who makes the final decision. More details about codebook-based precoding can be found in Section 3.2.3.

2.2.6.8 Resource Element and Antenna Mapping and OFDM Signal Generation

The Resource Element (RE) mapping puts the precoded symbols in the specific place on the LTE Resource Grid, assigned by the MAC layer. The resource elements are grouped into resource blocks, 84 RE per block with 12 subcarriers and 7 OFDM symbols for a normal Cyclic Prefix configuration. Apart from the RE, which carries the data, the radio block may also include control signals (PDCCH) and reference signals (Cell-specific or UE-specific).

Depending on the transmission scheme, the antenna mapping maps the modulated symbols to different antenna ports, specified in the standard [3GPP, 2016].

As the last step before the transmission, the OFDM signals are generated through the IFFT and the Cyclic Prefix is inserted.

2.2.7 Hybrid Automatic Repeat Request Protocol

The Hybrid Automatic Repeat Request (HARQ) retransmission protocols were developed to reduce the amount of the lost packages over fading channels through multiple retransmissions. If the codeword is successfully decoded, the UE sends an acknowledgment message (ACK) on one of the uplink channels and the eNodeB proceeds to the transmission of

the next data packet. In the opposite case, the UE sends a non-acknowledgment message (NACK) and the eNodeB retransmits the package. At the UE side, the log-likelihood ratios (LLR) of the corresponding bits will be added up and the new decoding attempt will be performed.

The HARQ protocol has two modifications: Chase Combining (CC) and Incremental Redundancy (IR)[Onggosanusi et al., 2003]. HARQ-CC can be thought of as a repetition code: upon reception of the NACK message, the eNodeB resends exactly the same copy of the previously transmitted message. Thus, there is no coding gain provided by this method. On contrary, using more complex IR type of the protocol, the eNodeB sends a different redundancy version (RV) of the message at each retransmission round. The RV is generated from the different fractions of the systematic and parity check bits, delivering coding gain with every new retransmission.

2.2.8 Link Adaptation

The LTE system architecture supports the instantaneous adaptation of the transmission to the current channel conditions. The adaptation includes a set of techniques and protocols, regrouped under the term of *Channel State Information*, which provides the eNodeB with the optimal transmission parameters to achieve maximum throughput in a current propagation environment.

Link adaptation techniques fall into two groups: open loop adaptation and closed loop adaptation. Open loop link adaptation does not involve the UE directly and relies on the statistics of the ACK/NACK confirmations at the eNodeB. It is used when the closed loop method is not available or is not reliable. The closed loop strategy is based on the *Channel State Information* feedback: the set of optimal transmission parameters is calculated at the UE taking into account channel estimates and, depending on the reporting configuration, is periodically or aperiodically sent to the eNodeB via one of the uplink channels. The reporting configuration is conveyed by the *RRC Connection Reconfiguration* or the *RRC Connection Setup* messages.

The channel state information includes three informational items that are sent to the eNodeB via one of the uplink channels: the CQI, the RI and the PMI. During a single antenna transmission, only the CQI is required, while the PMI and the RI are necessary for optimization of transmissions using multiple antennas.

2.3 Overview of MIMO Detection Strategies

MIMO systems with spatially multiplexed transmission achieve throughput gains compared to the conventional SISO systems, and these gains scale with the number of spatial layers [Sandhu and Paulraj, 2000; Telatar, 1999]. However, multi-layer transmissions make the performance vulnerable to interference created by different streams. In order to reduce the interference, pre-transmission signal processing at the base station was introduced. This preprocessing includes a set of techniques such as scheduling users in orthogonalized manner in terms of time-frequency-space resources, beamforming and precoding [Chen et al., 2014; Ghaffar, 2012; Wang et al., 2014]. In this thesis, we focus on the interference

management techniques at the receiver, which are reviewed in the following sections.

The process of extracting the data from a received signal corrupted by noise distortions and interference is called *detection*. The detector is given an observation of the received signal and, following a certain criterion, has to compute an estimate of the transmitted signal. The detector may or may not have some prior knowledge about the transmitted signal, such as the modulation alphabet that the symbols are mapped on. The detection criterion plays the leading role in the receiver architecture: it defines the system behaviors, the sensitivity to interference and noise distortions, and the computational complexity. Linear criteria allow to achieve low computational and implementation complexity at the price of performance degradation, while advanced non-linear estimators are optimal or near-optimal, but their complexity proliferates with the dimensionality of the MIMO system and the modulation order of the codewords.

First, we describe the classical signal model for a MIMO system, which is used throughout this chapter. We then provide an overview of the receiver architectures based on different detection criteria and discuss the interference handling and the complexity reduction techniques.

2.3.1 Classical MIMO Signal Model and General Receiver Architecture

Consider a classical downlink MIMO signal model:

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n_{\text{rx}}} \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} & \cdots & h_{0n_{\text{tx}}} \\ h_{10} & h_{11} & \cdots & h_{1n_{\text{tx}}} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_{\text{rx}}0} & h_{n_{\text{rx}}1} & \cdots & h_{n_{\text{rx}}n_{\text{tx}}} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n_{\text{tx}}} \end{bmatrix} + \begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_{n_{\text{rx}}} \end{bmatrix}, \tag{2.16}$$

where $\mathbf{y}$ is the received signal vector on n_{rx} antennas of the UE, $\mathbf{x}$ is the signal vector transmitted from n_{tx} antennas of the base station, $\mathbf{H}$ is a channel matrix and $\mathbf{n}$ is Gaussian noise. We discuss the configuration where the transmitted symbols $\mathbf{x}$ is mapped onto some known alphabet $\mathcal{Q}^{\text{M}}$ of modulation order $M \in \{2, 4, 6\}$ for 4QAM, 16QAM and 64QAM respectively. This signal model is used throughout this chapter to illustrate the different detection techniques.

Figure 2.11: Main functional blocks of a general receiver architecture. For simplicity, LTE-specific blocks (such as unscrambling) are omitted, and the turbo-decoder is replaced with the more general term "decoder".

Signal detection in MIMO receivers (see Fig. 2.11) often begins with linear processing. The linear filters, used at this phase, aim to suppress interference and to decompose the MIMO channel into parallel subchannels to simplify detection. The compensated

signals are passed to the decision block, which deduces the transmitted symbols. The decision logic can be based on the linear estimates, their projections onto constellation (if the constellation knowledge is available), or on non-linear metrics. The outputs of the decision block are then decoded and the sequence of bits is obtained.

In the next sections, we review major detection techniques that are used in MIMO receivers.

2.3.2 Linear Receivers

At the UE, the signal can be processed by linear or non-linear techniques, allowing for significant interference reduction. The idea behind *linear equalization* is to transform the MIMO detection into a set of parallel SISO detectors by eliminating the interference among the layers. This can be done using linear filters, such as Zero-forcing (ZF) and Minimum Mean Square Error (MMSE) [Onggosanusi et al., 2002]. The ZF filter $\mathbf{W}_{\mathrm{ZF}}$ applies a pseudo-inverse of the channel matrix to the received signal vector:

$$\mathbf{W}_{\mathrm{ZF}} = \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H. \tag{2.17}$$

Similarly, the MMSE filter $\mathbf{W}_{\mathrm{MMSE}}$ applies the channel pseudo-inverse, but also takes into account the noise variance σ_{n}^2, and thus is less sensitive to noise-related distortions [Onggosanusi et al., 2002]. [Hedayat and Nosratinia, 2007] demonstrated that MMSE receivers enjoy full spatial diversity at low rates, but do not possess these benefits at high rates.

$$\mathbf{W}_{\mathrm{MMSE}} = \left(\mathbf{H}^H\mathbf{H} + \sigma_{\mathrm{n}}^2\mathbf{I}\right)^{-1}\mathbf{H}^H. \tag{2.18}$$

The symbols estimates are obtained following

$$\widehat{\mathbf{x}}_{\mathrm{ZF}} = \mathbf{W}_{\mathrm{ZF}}\mathbf{y},$$
$$\widehat{\mathbf{x}}_{\mathrm{MMSE}} = \mathbf{W}_{\mathrm{MMSE}}\mathbf{y}.$$

To make the MMSE and the ZF receives compatible with the BICM that is used in LTE, the decision logic must follow soft-decision metrics that are required for each bit of the modulated symbol. These metrics are based on log-likelihood ratios and are derived with application to the MMSE and ZF receivers by [Butler and Collings, 2004; Seethaler et al., 2004]

Linear receivers are characterized by low computational complexity and acceptable performance level compared to optimum receivers. Furthermore, linear filters can compose a part of signal processing in advanced (non-linear) receivers, such as the MMSE or the ZF-based V-BLAST developed at the Bell Labs research center [Bai et al., 2009; Wolniansky et al., 1998]. Traditional ZF and MMSE detectors are interference-unaware, meaning that after the layers are decoupled, the SISO detector assumes that the interference is absorbed into the Gaussian noise. Recent works have shown that they can be followed by the interference-aware log-likelihood metrics, as it is done in LTE-compatible comparison of Block QR technique with the Block MMSE [Thomas et al., 2014].

2.3.3 Non-linear Receivers

Optimum or near-optimum non-linear MIMO receivers deliver high performance gain, but often require a prohibitive computational complexity. The detection can be joint, when all components of the received signal are processed together; parallel, when the MIMO channel is decomposed into streams which are treated in parallel; and successive, when the symbols are detected one after the other. The main representatives of non-linear detectors are receivers based on the minimum-distance metric, such as optimum Maximum-Likelihood receivers, their approximations and variations, and Successive Interference Canceling receivers.

2.3.3.1 Maximum Likelihood Detection

Conventional Maximum-Likelihood (ML) MIMO detectors perform an exhaustive search over all possible vector candidates, and the *most likely candidate* is located at the minimum distance from the received vector:

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}\in\mathcal{Q}^{\mathrm{M}}}{\arg\min} \left\|\mathbf{y} - \mathbf{H}\mathbf{x}\right\|^2. \tag{2.19}$$

In practice, the ML detection is replaced with the *Maximum a posteriori probability* (MAP) detection [Caire et al., 1998] which aims to maximize the probability of correct detection. The ML and MAP criteria are equivalent if the transmitted signals are equally likely. This assumption holds on modern digital systems, as the bits are interleaved during the encoding process. The distance metric is defined as follows:

$$D(\mathbf{y}|\mathbf{H}\mathbf{x}) = \left\|\mathbf{y} - \mathbf{H}\mathbf{x}\right\|^2. \tag{2.20}$$

Similarly, the minimum distance is defined as

$$\lambda\left(\mathbf{y}, \mathbf{H}, \mathbf{x}\right) = \min_{\mathbf{x}\in\mathcal{Q}^{\mathrm{M}}} \{D(\mathbf{y}|\mathbf{H}\mathbf{x})\} = \max_{\mathbf{x}\in\mathcal{Q}^{\mathrm{M}}} \{-D(\mathbf{y}|\mathbf{H}\mathbf{x})\}. \tag{2.21}$$

The complexity of the ML and MAP detector proliferates with the dimensionality of MIMO systems and the modulation order of the codewords since it elevates the number of candidate symbols. The sub-optimal Reduced Complexity Maximum Likelihood (R-ML) receiver design was introduced to circumvent the extremely high computational complexity of the ML and MAP receivers. The performance of the joint ML detector is an upper-bound for the performance of the MIMO systems relying on the linear filters [Moon et al., 2012].

2.3.3.2 Tree Search and Sphere Decoding

Sphere decoding [Viterbo and Boutros, 1999] and tree search detection belongs to the class of search space reduction techniques, where the likely to be transmitted symbols are identified by the same ML metrics (2.19), but the search space is limited according to a particular principle. In tree-search based algorithms, the complexity reduction is achieved via elimination of unreliable paths in the *search tree*.

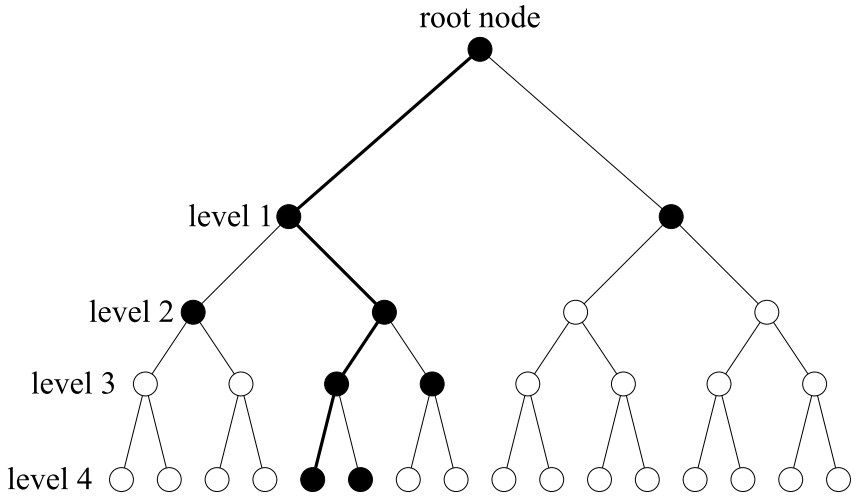


Figure 2.12: Example of traversing through the search tree for 2×2 MIMO system with QPSK modulation. The depth of the tree for the complex-valued alphabet with modulation oder M is $2M$ due to decoupling the real and imaginary parts of the modulated symbol (4 for QPSK). The black nodes are nodes that have been examined, and thick lines symbolizes an example of the path to the most likely symbol.

The idea behind this family of algorithms consists in building and traversing the search tree starting from the top root node down through the branches in a pruning manner, as illustrated in Fig. 2.12, until the ML metric is satisfied (2.19). For the complex symbols mapped onto QAM alphabets, the common approach to construct the search tree search is to decompose the real and imaginary parts of the symbol, converting the M-dimensional complex-valued problem into a $2M$-dimensional real-valued problem [Azzam and Ayanoglu, 2008].

Sphere decoding applies the tree search algorithm, but the most likely candidates are now assumed to be located inside a sphere of radius r (Fig. 2.13). Determining the correct value of d is a challenging task: a radius chosen too small would lead to unreliable

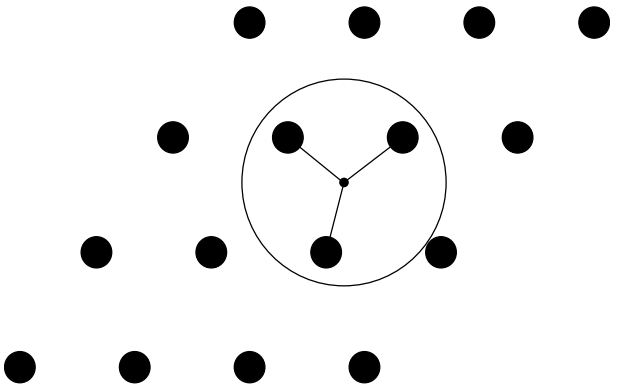


Figure 2.13: Geometrical representation of the sphere decoding algorithm [Viterbo and Boutros, 1999]. The ML metrics are applied to the symbols located inside the sphere with radius r.

estimation of the candidates, while a radius chosen too large would cause a significant complexity increase. One might think of using *covering radius* of the lattice, but this problem has NP hard complexity [Conway et al., 1987]. The complexity of classical sphere decoding is however polynomial [Fincke and Pohst, 1985], but it was proven that the rate of the exponential function depends on the SNR, and is acceptably small for the high SNR values [Jalden and Ottersten, 2005]. Reduced complexity methods based on Pohst [Pohst, 1981] and Schnorr–Euchner enumeration strategies [Schnorr and Euchner, 1993] were proposed by [Damen et al., 2003]. Other representatives of this group are iterative tree-search detection [de Jong and Willink, 2005] and K-best algorithm [Shabany and Gulak, 2008].

LTE compatible implementation of the sphere decoding with QR preprocessing denoted as Fixed-Complexity Sphere Decoder with complexity proportional to the size of QAM constellation was proposed by [Aubert et al., 2009].

2.3.3.3 Reduced-complexity Interference-Aware Detection

Conventional receivers tend to treat interference as Gaussian and apply interference whitening techniques in order to recover the transmitted signal [Winters, 1984]. The interference-aware architecture of the advanced receivers brings significant gains to the system performance, even if only blind interference detection is applied [Lee et al., 2015]. If the interfering symbols come from a fixed alphabet, as it is the case of LTE systems, the receiver may guess the modulation order of the interfering codeword (or acquire this knowledge) and apply it during the detection process. The knowledge of the interferer can be gained from the decoding of DCI.

The search space reduction is achieved thanks to the exploitation of the interference nature and the manipulation of the LLR metrics, which allows to define a single interfering symbol for each desired symbol instead of searching over the complete alphabet. We illustrate this detection strategy, proposed by [Ghaffar and Knopp, 2010a], using an example of a Single-User 2×2 MIMO system, where the decoding of the first codeword with the symbol x_0 is performed in the presence of an interfering codeword with the symbol x_1, and vice versa. Let the first and the second codewords be mapped onto the modulation alphabet $\mathcal{Q}^{M_0}$ and $\mathcal{Q}^{M_1}$, then the full transmitted vector is mapped onto $\mathcal{Q}^{M_0+M_1}$ modulation alphabet.

The detection starts from the signal preprocessing with Matched Filter (MF), and outputs the LLR values obtained via interference-aware soft bit metrics. The distance metric $D(\mathbf{y}|\mathbf{H}_{\text{eff}}\mathbf{x})$ (2.20) can be rewritten as follows:

$$\begin{aligned} D(\mathbf{y}|\mathbf{H}_{\text{eff}}\mathbf{x}) &= \underset{\mathbf{x}\in 2^{M_0+M_1}}{\arg\min} \left\|\mathbf{y} - \mathbf{h}_{\text{eff}0}x_0 - \mathbf{h}_{\text{eff}1}x_1\right\|^2 \\ &= \underset{\mathbf{x}\in 2^{M_0+M_1}}{\arg\min} \left\{\left(\mathbf{y} - \mathbf{h}_{\text{eff}0}x_0 - \mathbf{h}_{\text{eff}1}x_1\right)\left(\mathbf{y}^H - \mathbf{h}_{\text{eff}0}^H x_0^* - \mathbf{h}_{\text{eff}1}^H x_1^*\right)\right\} \\ &= \underset{\mathbf{x}\in \mathcal{Q}^{M_0,M_1}}{\arg\min} \left\{\|\mathbf{y}\|^2 + \|\mathbf{h}_{\text{eff}0}\|^2 |x_0|^2 + \|\mathbf{h}_{\text{eff}1}\|^2 |x_1|^2 \right. \\ &\quad \left. - 2\Re\{y_{\text{MF}_0}x_0^*\} - 2\Re\{y_{\text{MF}_1}x_1^*\} + 2\Re\{\mathbf{h}_{\text{eff}0}^H\mathbf{h}_{\text{eff}1}x_0^*x_1\}\right\}, \end{aligned} \tag{2.22}$$

where y_{MF_0} and y_{MF_1} are the Matched Filter outputs of the received vector $\mathbf{y}$. Omitting

the common term $\|\mathbf{y}\|^2$ and decoupling real and imaginary parts, we obtain

$$\begin{aligned}\lambda = \max_{\mathbf{x}\in\mathcal{Q}^{M_0,M_1}} \{ &- \|\mathbf{h}_{\text{eff0}}\|^2 |x_0|^2 - \|\mathbf{h}_{\text{eff1}}\|^2 |x_1|^2 + 2\left[\Re(y_{\text{MF}_0})\Re(x_0) + \Im(y_{\text{MF}_0})\Im(x_0)\right] \\ &- 2\eta_0\ Re(x_1) - 2\eta_1\Re(x_1)\},\end{aligned} \tag{2.23}$$

where the parameters η_0 and η_1 are defined as

$$\eta_0 = \Re(\rho)\Re(x_0) + \Im(\rho)\Im(x_0) - \Re(y_{\text{MF}_1}), \tag{2.24}$$

$$\eta_1 = \Re(\rho)\Im(x_0) - \Im(\rho)\Re(x_0) - \Im(y_{\text{MF}_1}), \tag{2.25}$$

and the correlation coefficient ρ of the effective channel matrix is $\rho = \mathbf{h}_{\text{eff1}}^H \mathbf{h}_{\text{eff0}}$.

Recall that the minimum distance λ is defined as

$$\lambda(\mathbf{y}, \mathbf{H}, \mathbf{x}) = \min_{\mathbf{x}\in\mathcal{Q}^{\text{M}}} \{D(\mathbf{y}|\mathbf{H}\mathbf{x})\} = \max_{\mathbf{x}\in\mathcal{Q}^{\text{M}}} \{-D(\mathbf{y}|\mathbf{H}\mathbf{x})\}. \tag{2.26}$$

The distance metric $D(\mathbf{y}|\mathbf{H}_{\text{eff}}\mathbf{x})$ is minimized when λ (2.23) is maximized. To achieve so, the terms $\eta_0\Re(x_0)$ and $\eta_1\Re(x_1)$ should be negative, which can be achieved by ensuring that $\Re(x_0)$ is of the opposite sign compared to η_0, and $\Re(x_1)$ is of the opposite sign compared to η_1. With this, we can simplify (2.23) to

$$\begin{aligned}\lambda = \max_{\mathbf{x}\in\mathcal{Q}^{M_0,M_1}} \{ &- \|\mathbf{h}_{\text{eff0}}\|^2 |x_0|^2 - \|\mathbf{h}_{\text{eff1}}\|^2 |x_1|^2 + 2\left[\Re(y_{\text{MF}_0})\Re(x_0) + \Im(y_{\text{MF}_0})\Im(x_0)\right] \\ &- 2|\eta_0||\Re(x_0)| - 2|\eta_1||\Re(x_1)|\}.\end{aligned} \tag{2.27}$$

However, to compute λ metric (2.27) when symbols $x_0 \in 2^{M_0}$ and $x_1 \in 2^{M_1}$ we would need to calculate $2^{M_0+M_1}$ distances. If both codewords are mapped onto 64QAM, this would result in 4096 computations of (2.27). In the setting of a Single-User MIMO, the same amount of calculations would be required for detections of the second stream, which is prohibitive. [Ghaffar and Knopp, 2010a] observed that the terms η_0 and η_1 are independent from x_1, and thus differentiating (2.23) along $\Re(x_1)$ and $\Im(x_1)$ and equating to zero, we obtain:

$$|\Re(x_1)|_{\text{opt}} = \frac{|\eta_0|}{\|\mathbf{h}_{\text{eff1}}\|^2}, \qquad |\Im(x_1)|_{\text{opt}} = \frac{|\eta_1|}{\|\mathbf{h}_{\text{eff1}}\|^2}.$$

The terms $|\Re(x_1)|_{\text{opt}}$ and $|\Im(x_1)|_{\text{opt}}$ respectively represent the absolute values of the real and imaginary part of the optimal symbol x_1 that minimizes the distance metric. Thus, for every possible symbol $x_0 \in \mathcal{Q}^{M_0}$ the receiver needs to define only one interference symbol x_1, which significantly reduces the number of computations of metric (2.26). This derivation was initially developed in [Ghaffar, 2010] and further applied by [Wagner and Kaltenberger, 2014].

2.3.3.4 Successive Interference Canceling

Another representative of the family of non-linear receivers is the Successive Interference Canceling (SIC) receiver. This technique can be combined with preprocessing techniques,

such as the QR decomposition, the MMSE and the ZF filters, as it is done in the famous V-BLAST receiver [Wolniansky et al., 1998]. Each filter from the receiver bank detects one of the data streams, which is then subtracted from the remaining signal stage by stage resulting interference-free signal on the output of the final step. The SIC receiver can also compose a part of the ML receiver (as we will see in Chapter 4) by consequent subtraction of the decoded symbols, finally achieving interference-free signal at the last stage.

The drawback of the SIC technique is its high sensitivity to inaccurate channel estimates [Andrews et al., 2007]. This may cause error propagation: if one of the streams is wrongly detected at the early stages, the wrong information is subtracted and thus error propagates, resulting in multiple wrong detection events. To reduce the risk of occurrence of such an event, the Ordered SIC approach was proposed, where streams with higher probability to be detected correctly are subtracted first based on SNR, SINR or LLR criteria. The LLR-based ordering is shown to provide a 1-3dB gain in comparison to other criteria [Chun and Kim, 2008]. The family of SIC receivers can be categorized as symbol-level SIC receivers, where the detection and interference cancellation are done on a per-subcarrier basis, and codeword-based SIC, where detection and interference cancellation is done for each independently-coded stream due to embedded turbo-decoder inside the feedback loop. Codeword-SIC performs ML or R-ML detection, decoding, re-encoding, and cancellation, while symbol-SIC utilizes linear detection, reconstruction, and cancellation [3GPP, 2014]. Codeword-based approaches outperform symbol-SIC methods due to the error correction capabilities provided by the turbo-decoder [Manchon et al., 2008].

2.3.3.5 QR-Decomposition Techniques

QR matrix decomposition is a decomposition of a real or complex matrix $\mathbf{H} \in \mathbb{R}^{n_{\text{rx}} \times n_{\text{tx}}}$ into a product of the orthogonal (in case of real-valued and squared structure of a matrix $\mathbf{H}$) or unitary matrix $\mathbf{Q} \in \mathbb{R}^{n_{\text{rx}} \times n_{\text{rx}}}$ and an upper triangular matrix $\mathbf{R} \in \mathbb{R}^{n_{\text{tx}} \times n_{\text{tx}}}$:

$$\mathbf{H} = \mathbf{QR}. \tag{2.28}$$

This technique significantly reduces interference on the lower streams, however, in practice it requires high computational complexity. There are several methods to perform a decomposition: Gram-Schmidt process, Givens Rotation and Householder method [Golub and Van Loan, 2012].

Gram-Schmidt process The classical version of the Gram-Schmidt (GS) algorithm [Schmidt, 1908] constructs an orthogonal basis $\{\mathbf{v_1}, \mathbf{v_1}, \ldots, \mathbf{v_n}\}$ given an arbitrary basis $\{\mathbf{u_1}, \mathbf{u_1}, \ldots, \mathbf{u_n}\}$ for an n-dimensional inner product space V. After an orthogonal basis is obtained, any vector in V can be expressed as a linear combination of the vectors of the orthogonal basis. To overcome the loss of orthogonality and proneness to roundoff error, that the traditional version may lead to, the Modified Gram-Schmidt (MGS) orthogonalization was introduced. MGS is superior to GS and introduces the orthogonality loss in a predictable manner [Bjorck, 1994].

An improved lattice-reduction aided MMSE detection with GS process was proposed for 4×4 and 8×8 MIMO systems by [Fujino et al., 2009]. The authors showed that the GS algorithm brought the proposed scheme close to near-ML performance for all constellations, and the decision boundaries became the same as for ML receiver.

Givens Rotations Givens rotations is another method to perform QR decomposition. Originally the rotations were derived by Jacoby to resolve symmetric eigenvalue problem in 1846 and then were applied by Givens to perform matrix decompositions. The plain rotations are represented by the matrix $\mathbf{G}$ constructed from trigonometric functions $\sin\theta$ and $\cos\theta$ that allows to introduce zeros by rotating the vectors of the original matrix $\mathbf{H}$ through a certain angle $\theta \in [0, 2\pi]$. This method has a low hardware complexity but also significantly long latency. An efficient triangular systolic processor array realization based on CORDIC was proposed in [Zhong et al., 2003]. In order to reduce the latency, this technique was then combined with the Three Angle Complex Rotation (TACR) [Maltsev et al., 2006]. A considerable latency reduction was achieved in high speed Tournament Givens Rotation-Based QR decomposition [Lee et al., 2012] thanks to parallelization of the multiple zero inserting steps at every stage of the TACR algorithm.

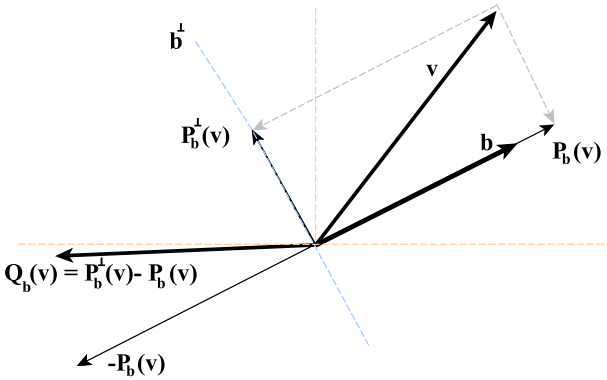


Figure 2.14: Geometrical interpretation of the Householder transform of the vector [Steinhardt, 1988]. $Q_{\mathbf{b}}(\mathbf{v})$ is a length preserving reflection of vector $\mathbf{v}$ about $\mathbf{b}^{\perp}$. $Q_{\mathbf{b}}(\mathbf{v})$ and $\mathbf{v}$ differ only by a sign-reversal in their projections onto $\mathbf{b}$.

Householder Decomposition Housholder transformations are orthogonal transformations which perform reflections through a plane.

Let $\mathbf{b}$ be a normal unit vector, then the reflection across the plane orthogonal to $\mathbf{b}$ can be expressed in a matrix form as

$$\mathbf{Hv} = \mathbf{I} - 2\frac{\mathbf{bb}^T}{\mathbf{b}^T\mathbf{b}}, \tag{2.29}$$

where $\mathbf{I}$ is an identity matrix.

To reflect the vector $\mathbf{v}$ in the hyperplane spine $\{b\}^{\perp}$, one has to multiply $\mathbf{v}$ with the matrix $\mathbf{H}$. The geometrical interpretation of the Householder transformation of the vector

is presented on Fig. 2.14.

All discussed QR decomposition techniques provide the same level of performance, and differ only in the complexity of the algorithms.

The classic ZF-QR decomposition is applied in the following way:

$$\mathbf{H} = \mathbf{QR}, \tag{2.30}$$

and the received signal (2.16) transforms into

$$\mathbf{Q}^H\mathbf{y} = \mathbf{Rx} + \mathbf{Q}^H\mathbf{n}. \tag{2.31}$$

The MMSE-QR algorithms operate on *augmented* channel matrix $\underline{\mathbf{H}}$ and *augmented* received signal $\underline{\mathbf{y}}$ such that

$$\underline{\mathbf{H}} = \begin{bmatrix} \mathbf{H} \\ \sigma_n \mathbf{I}_{n_{\text{tx}}} \end{bmatrix} = \underline{\mathbf{Q}}\underline{\mathbf{R}} = \begin{bmatrix} \mathbf{Q}_0 \\ \mathbf{Q}_1 \end{bmatrix} \underline{\mathbf{R}}, \quad \underline{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix}, \tag{2.32}$$

where the size of matrix $\mathbf{Q}$ is $(n_{\text{tx}} + n_{\text{rx}}) \times n_{\text{tx}}$, of matrix $\mathbf{Q}_0$ is $n_{\text{rx}} \times n_{\text{tx}}$, of $\mathbf{Q}_1$ is $n_{\text{tx}} \times n_{\text{tx}}$, and of matrix $\mathbf{R}$ is $n_{\text{rx}} \times n_{\text{tx}}$.

The augmented received signal is then transformed as following:

$$\underline{\mathbf{Q}}^H\underline{\mathbf{y}} = \mathbf{Q}_0^H\mathbf{y} = \underline{\mathbf{R}}\mathbf{x} - \sigma_n\mathbf{Q}_0^H\mathbf{x} + \mathbf{Q}_1^H\mathbf{n}. \tag{2.33}$$

The term $\sigma_n\mathbf{Q}_1^H\mathbf{x}$ is the remaining interference that cannot be nullified, which is a compromise between avoiding the noise enhancement typical for ZF filters and allowing some interference. This term is not Gaussian, and together with the term $\mathbf{Q}_1^H\mathbf{n}$ can be considered as a *colored* noise. Nevertheless, the MMSE-QR scheme significantly outperforms ZF-QR detection schemes [Wubben et al., 2003].

After QR transformation, the symbol on the lowest layer can be detected in the interference-free environment, and subtracted from the previous layer using SIC detection and so on.

2.4 PHY Abstraction Methodologies

Physical Layer Abstraction is a useful tool for system performance prediction. The idea behind this concept is to map the instantaneous channel conditions to the performance indicators. The instantaneous channel is commonly represented by SNR and channel gains, and the performance is often measured in terms of *Block Error Rate* (BLER).

The mapping technique is the core of the abstraction since it defines the level of accuracy that the performance is quantified with. In frequency-selective environments each OFDM subcarrier achieves different channel gains, meaning that the SNR value varies from subcarrier to subcarrier. It would be time consuming and statistically not feasible to consider performance indicators estimates (for example, throughput) on a per-subcarrier basis. To circumvent this challenge, the concept of *effective Signal-to-Noise Ratio* SNR_{eff} was introduced [Nanda and Rege, 1998]. The effective SNR is an equivalent SNR that the signal would enjoy in AWGN channel, and causes the same block error

rate. Thanks to this property, the mapping process is often called *compressing function*, where the vector of instantaneous SNR_k per subcarrier k is condensed into a single value $\mathrm{SNR}_{\mathrm{eff}}$. The principle of mapping consists in calculating the value of $\mathrm{SNR}_{\mathrm{eff}}$ in a way that it minimizes the square error between $\mathrm{SNR}_{\mathrm{eff}}$ and the equivalent $\mathrm{SNR}_{\mathrm{AWGN}}$ in AWGN channel generated for the same modulation scheme and coding rate.

The two main approaches for the mapping are *Exponential Effective Signal-to-Noise Ratio mapping* (EESM) [Kim et al., 2011; Sandanalakshmi et al., 2007; Stancanelli et al., 2011] and *Mutual Information Effective Signal-to-Noise Ratio mapping* (MIESM) [Olmos et al., 2010; Srinivasan et al., 2008]. The common formula to calculate $\mathrm{SNR}_{\mathrm{eff}}$ for both approaches is

$$\mathrm{SNR}_{\mathrm{eff}}(\beta_1, \beta_2) = \beta_1 I^{-1}\left[\frac{1}{K}\sum_{k=1}^{K}\mathcal{I}_{\mathrm{map}}\left(\frac{\mathrm{SNR}_k}{\beta_2}\right)\right],$$

where β_1, β_2 are rate-dependent adjustment parameters to be discussed in detail in Chapter 6, and $\mathcal{I}_{\mathrm{map}}$ is a mapping function that distinguishes the EESM and MIESM approaches. The EESM mapping function is calculated using the Chernoff union bound for error probabilities and is straightforward to implement:

$$\mathcal{I}_{\mathrm{map}}(\mathrm{SNR}_k) = 1 - \exp(-\mathrm{SNR}_k).$$

The MIESM approach is more challenging as it involves per subcarrier Mutual Information computations. As we mentioned in Section 2.1.2, there is no closed form analytical expression to estimate Mutual Information for a system with finite alphabet and one is restricted to operate with the numerical approximations computed in Monte-Carlo simulations (2.12), which makes this method cumbersome to use in real-time systems. However, both approaches were compared [Brueninghaus et al., 2005] in the scope of the IST project WINNER (Wireless World Initiative New Radio), and the MIESM approach was shown to offer high prediction accuracy for different spatial processing techniques.

Chapter 3

Signal Model and Simulation Framework

Testbenches constitute a good starting point to run experiments and test new features of communication systems. However, prototype building is often a time and money consuming process. The Communication Systems Department of EURECOM is a founder of OpenAirInterface (OAI) [OAI, 2017] — a unique open source 4G – 5G software/hardware platform for experimental prototyping that provides the flexibility of software implementation while ensuring compliance with 3GPP standards. The OAI strategic areas span a few research axes such as large scale simulations, MODEM development and the Internet of Things.

OAI provides the opportunity to run experiments in simulation, emulation and real-time mode. Several unitary simulations for the physical layer are available to reproduce the behavior of the transport and the physical channels in both uplink and downlink transmissions. One of them, the simulator for DLSCH and PDSCH channels, is used as an experimental platform for the receiver design in this thesis. Our theoretical models are implemented in programming language `C` and the results of the link-level simulations are used to validate the proposed methodologies. The `oaisim` emulator of OAI includes not only the physical layer, but reflects a complete Long Term Evolution (LTE) network with Medium Access Control (MAC) and networking layers. The scenarios may include the OAI or the commercial *Evolved Packet Core* (EPC), the OAI eNodeB, and the OAI UE. Real-time transmission can be run by combining the EURECOM software and hardware units supplied by EURECOM or third party companies.

The simulators of OAI include a variety of available Multiple-Input-Multiple-Output (MIMO) transmission schemes and propagation environment, which conform to 3GPP standards. Users have the flexibility to vary the LTE transmission parameters as well as to customize the propagation environment according to their needs. In this chapter, we first describe the OAI platform, paying a special attention to the downlink simulator that is used to perform extensive link-level simulations throughout this thesis. After that, we detail the channel modeling that was used during our experiments. Finally, we provide analytical models of the transmission schemes that are implemented and make the necessary derivations.

Contents

3.1 Simulation Framework

OpenAirInterface [Kaltenberger et al., 2015] is an open-source software- and hardware-based implementation of 3GPP LTE Release 8 and 9. It also includes some features from LTE-Advanced (Release 10,11 and 12), LTE-Advanced-Pro (Release 13 and 14), and work on 5G New Radio [3GPP, 2017] has just begun.

The multiple use cases of OAI platform include running a classical 3GPP network (Fig. 3.1), simplified network, which may include non-3GPP components, cloud-RAN, which is currently a main target of EURECOM deployment, and simulation/emulation tools.

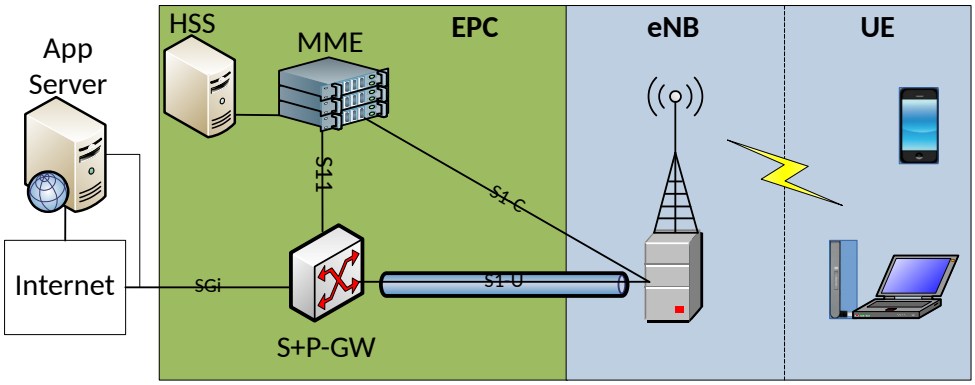


Figure 3.1: The first use case of OAI – running a classical 3GPP network. The EPC, the eNodeB and the UE components are available in OAI. Commercial or third party components are also supported.

Network emulation can be performed using `oaisim` emulator, which incorporates a complete LTE system with message exchange, different transmission schemes, rich propagation and traffic models (Fig. 3.2). The software is equipped with multiple tools for control and monitoring, message analyzer and soft scope [Kaltenberger et al., 2015]. For a detailed analysis, it is also possible to use Wireshark software.

Unitary simulators with a focus on a particular transport or physical LTE channel are listed in Table 3.1. These simulators allow to select the propagation environment from a

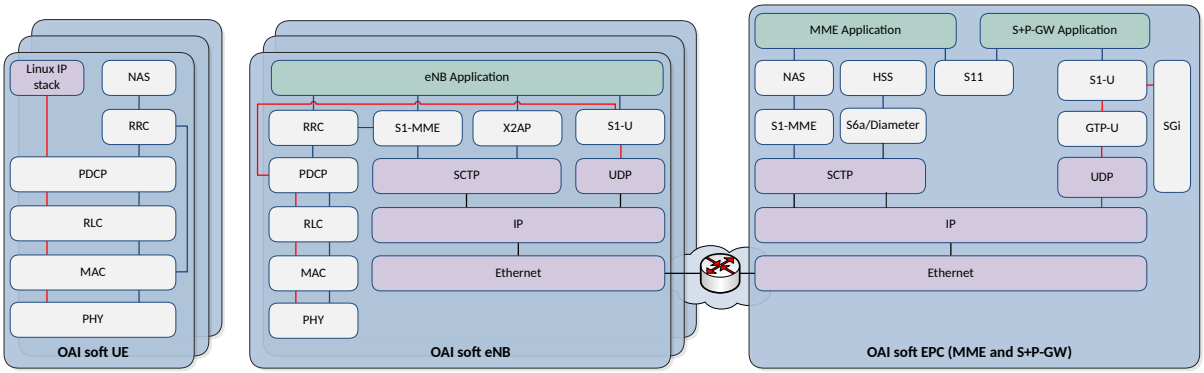


Figure 3.2: OpenAirInterface LTE software stack [Kaltenberger et al., 2015]. The white blocks correspond to the 3GPP layers, the purple blocks symbolize Linux stack, the red lines are responsible for the control plane, while the blue lines are the user plane of LTE.

Table 3.1: OAI Unitary Simulators

OAI DSLSH simulator	`dlsim`
OAI ULSCH simulator	`ulsim`
OAI PUCCH simulator	`pucchsim`
OAI PRACH simulator	`prachsim`
OAI PDCCH simulator	`pdcchsim`
OAI PBCH simulator	`pbchsim`
OAI eMBMS simulator	`mbmssim`

wide set of the channel models, antenna configurations, transmission schemes, and other parameters. As an output, the simulators generate Block Error Rate (BLER) statistics, throughput statistics, Downlink Control Information (DCI) error statistics, and profiling of the code. In the following section, we provide a detailed description of our downlink simulator `dlsim`, which was extensively developed and used to quantify the contributions of this work.

3.1.1 Downlink Simulator

Our downlink simulator `dlsim` models the link-level behaviors of the LTE system during the downlink transmission between the base station and the UE. It is logically split into three functional blocks: the base station procedures, the generation of the propagation medium and the convolution with the transmitted signal, and the UE procedures (see Fig. 3.3).

The physical layer signal processing strictly follows the 3GPP standards [3GPP, 2015b,c, 2016]. In some transmission modes, the feedback link from the UE to the base station is available. In real systems, this link is established through one of the uplink channels.

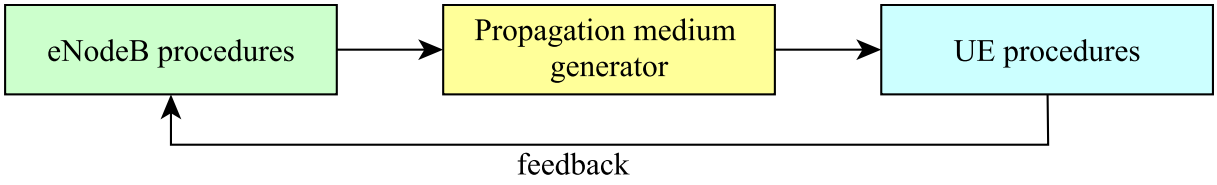


Figure 3.3: General structure of the OAI downlink simulator `dlsim`.

Unlike in `oaisim`, uplink is not implemented in our simulator and is instead replaced with the virtual delay-free and error-free information exchange.

The simulator performs Monte-Carlo simulations for multiple frames over a wide range of the SNR values. By default we assume that every channel realization is drawn from a new seed. For debugging purposes or to generate exact same set of channel realizations in different scenarios the sequence of the seeds can be fixed.

Fig. 3.4 illustrates the main building procedures implemented in our simulator. The simulation starts with the interpretation of the command line arguments. The typical set of the arguments that is used in the simulation in this theses is the following:
`./dlsim -x4 -y2 -z2 -n1000 -S7 -c1 -gN -Y -s5 -f1 -m28 -M28 -R1`.
The interpretation of the essential parameters which define the link-level behavior are summarized in Table 3.2. In this dissertation we included some additional parameters, specific to design of the receivers, implemented in the next chapters of this manuscript.

Several arguments from the command line are used to fill the DCI and generate the corresponding DLSCH parameters for the eNodeB and to encode the DCI structures.

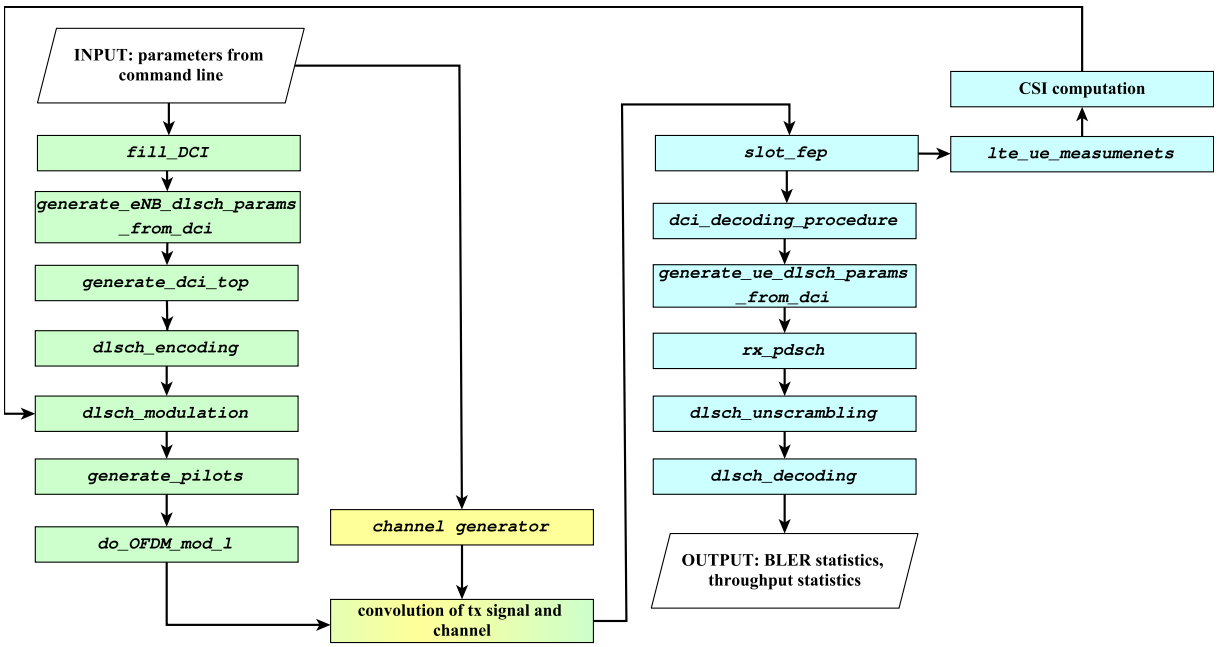


Figure 3.4: Main building blocks of the OAI downlink simulator `dlsim`. The base station related procedures are colored in green, the channel simulation block in yellow, and the UE procedures in blue.

Table 3.2: Essential command line parameters for `dlsim` simulator

Transmission mode	`-x`
Number of transmits antennas	`-y`
Number of receive antennas	`-z`
Number of subrames	`-n`
Subframe index	`-S`
Number of PDCCH symbols	`-c`
Channel	`-g`
Channel estimation	`-Y`
Starting SNR pont	`-s`
SNR step	`-f`
MCS on the fist codeword	`-m`
MCS on the second codeword	`-M`
Number of HARQ rounds	`-R`
Receiver type	`-u`

During the initial transmission round, a random (or predefined) sequence of bits is allocated in the DLSCH buffer, which can be seen as a circular buffer. During the next retransmission rounds, the base station will refer to it again to perform rate-matching with the corresponding redundancy version.

The `dlsch_modulation` functional block performs multiple operations, including mapping the encoded bit sequence onto the modulation symbols, layer mapping and mapping to the Resource Elements and precoding. The output of this block is the transmission data `txdataF` in the frequency domain. After that, the pilots are generated and inserted in the predefined positions, as discussed in Section 2.2.5, and the base station performs the OFDM modulation, preparing the data for the transmission.

The channel generator produces channel coefficients based on the channel model, defined in the command line. The so-called channel descriptor block `fill_channel_desc` generates frequency-flat and frequency-selective channel models with Rayleigh and Rician fading, as well as the channel models, as defined by 3GPP [3GPP, 2015d], such as Extended Pedestrian A (EPA) channel, Extended Vehicular A channel model (EVA), etc.

In practice, the UE periodically performs channel measurements and sends the result to the base station. If the Channel State Information reporting is configured in RRC Connection Reconfiguration or RRC Connection Setup message, the UE also calculates the Channel State Information based on the channel estimates and feeds it back to the base station. In our simulator, the feedback is simulated as delay-free and error-free virtual link. The base station therefore applies the reported information immediately, while the channel changes only during the next retransmission round. Two channel estimation regimes are available: the perfect channel estimation and the Least Square estimation. If the Least Square method is chosen, the channel estimates are obtained by interpolation of the channel coefficients at the Resource Elements that carry Reference Symbols.

After the DCI is decoded and the necessary information is extracted and interpreted, the UE starts the demodulation procedure. The demodulation is implemented in the

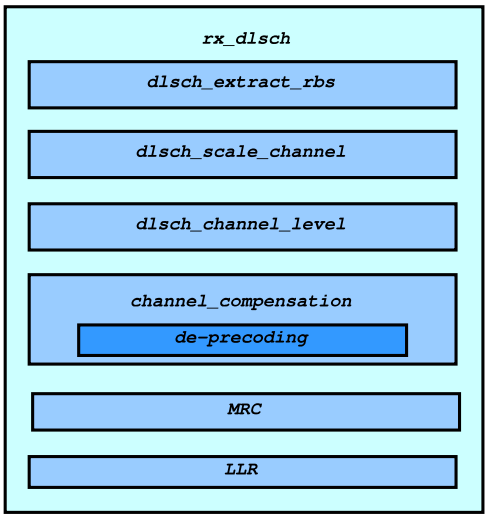


Figure 3.5: Received signal demodulation block at the UE.

`rx_pdsch` procedure (see Fig. 3.5), and includes the extraction of the Radio Blocks, adjusting the power levels of different components of the downlink signal (pilots and data signals), the channel compensation (matched filtering) and the LLR computation. If the Multi-User MIMO or multi-layer Single-User MIMO transmission is activated, depending on the requested receiver type the simulator may use interference-aware LLR metrics described in Section 2.3.3.3. Currently we have available the Interference-Unaware receiver, the Interference-Aware receiver for Multi-User MIMO, the Parallel Interference Aware Receiver and the Successive Interference Canceling Receiver for Single-User MIMO. Both Single-User MIMO receivers were developed in the scope of this dissertation.

The computed LLR values are forwarded to the turbo decoder. If after four turbo iterations the packet is not decoded, the decoding is considered as failed, and the the UE sets NACK message flag, and the packet is retransmitted with Incremental Redundancy HARQ protocol. The default maximum value of the retransmission rounds is set to four. If the packet is successfully decoded, the simulator repeats the described procedure for the rest of the frames, until all the frames undergo transmission. At the end of the experiment, the simulator computes BLER and throughput statistics.

The simulator uses fixed point arithmetic, where all numbers are represented as fractional Q15 numbers. To speed up the computations, some parts of the code are written using streaming SIMD instructions [Skillicorn, 1990], which allow to execute four multiplications simultaneously.

The soft scope is implemented for the visualization purposes. An example with the SIC receiver, developed in the scope of this thesis, is illustrated in Fig. 3.6. The soft scope visualizes the received signal on Time Domain, the Channel Impulse Response, The Channel Frequency Response, and plots the outputs of the Matched Filter with the corresponding LLR Ratios.

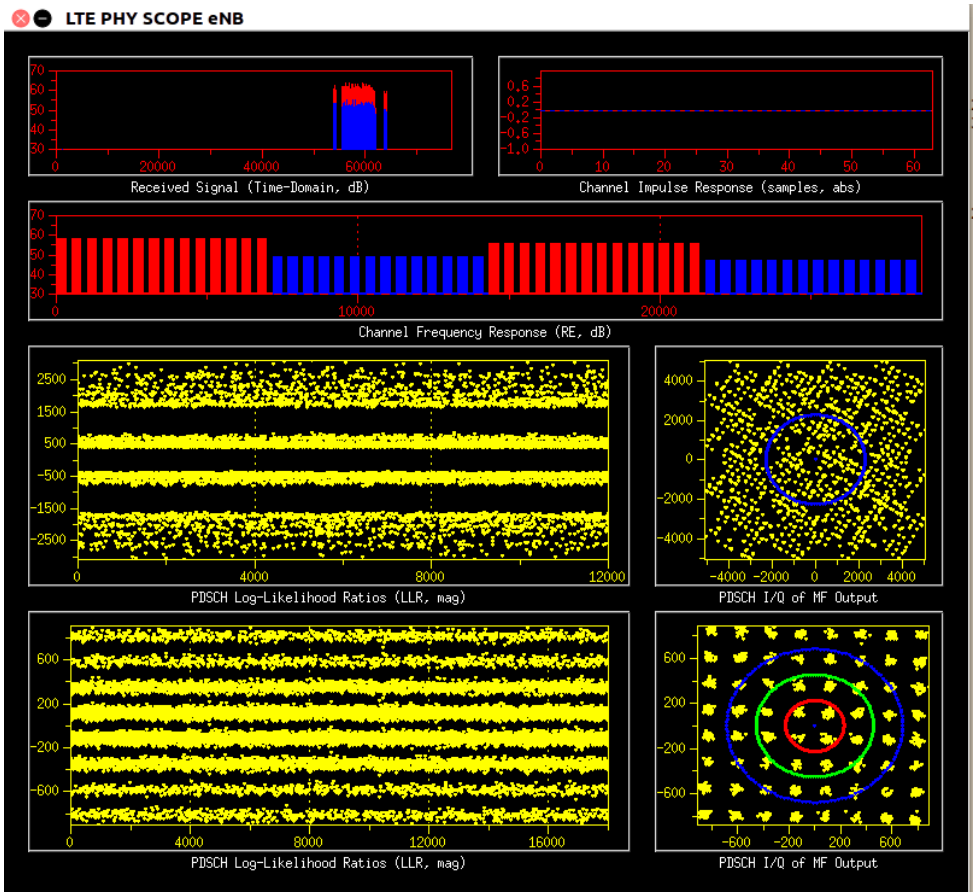


Figure 3.6: Soft Scope for Transmission Mode 4 with SIC receiver.

3.2 Analytical Models

Implemented transmission schemes, as well as simulated channels, are based on analytical models. In the next sections, we shall discuss the channel models available in our `dlsim` and how they are generated. Then, we will provide the signal models for the transmission schemes that are used throughout this thesis.

3.2.1 Channel Modeling

Channel modeling is essential for accurate design of wireless systems and the creation of a realistic representation of the propagation environment. The modeling should be based on the available information characterizing the propagation medium: frequency selectivity, Rician factor, angle of arrival (AoA), spatial correlation level, the distribution followed by the amplitudes of the channel gains, Power Delay Profile (PDP), etc.

The channel matrix $\widetilde{\mathbf{H}}_k = [\widetilde{\mathbf{h}}_{0,k}\ \widetilde{\mathbf{h}}_{1,k}]$ for 2×2 MIMO system given a bandwidth of

$k = 1, \ldots, K$ Resource Elements is constructed in the classical manner:

$$\widetilde{\mathbf{H}}_k = \underbrace{\sqrt{\frac{K_\mathrm{r}}{K_\mathrm{r}+1}}\begin{bmatrix} 1 & e^{-j\phi} \\ e^{j\phi} & 1 \end{bmatrix}}_{\text{LOS component}} + \underbrace{\sqrt{\frac{1}{K_\mathrm{r}+1}}\mathbf{H}_k}_{\text{NLOS component}}, \tag{3.1}$$

where $\mathbf{H}_k$ is a matrix representing a non-Line-of-Sight (NLOS) component and is defined with respect to the level of spatial correlation, as explained below. The Line-of-Sight (LOS) component is defined by two parameters: a phase-shift ϕ, which is strongly related to the antenna array configuration and its orientation towards the Line-of-Sight (LOS) component, and a Rician K-factor K_r. The channel experiences deep fades and can be seen as Rayleigh when $K_\mathrm{r} \simeq 0$. On the other hand, the LOS component becomes more dominant as K_r increases.

The rows of a channel matrix are linked to an index of a transmit antenna and the columns correspond to an index of a receive antenna. Precisely, in a 2×2 MIMO system, the channel coefficient h_{00} represents a channel from a transmit antenna tx_0 to a receive antenna rx_0, h_{01} – a channel from a transmit antenna tx_1 to a receive antenna rx_0, h_{10} – a channel from a transmit antenna tx_0 to a receive antenna rx_1, h_{11} – a channel from a transmit antenna tx_1 to a receive antenna rx_1:

$$\begin{array}{c} \\ \text{rx antennas} \downarrow \end{array} \overset{\text{tx antennas}\rightarrow}{\begin{bmatrix} h_{00} & h_{01} \\ h_{10} & h_{11} \end{bmatrix}}.$$

In this thesis, we consider MIMO channels characterized by a different level of spatial correlation: uncorrelated channels; channels with low, medium, or high correlation; and fully correlation channels.

The entries of spatially uncorrelated channel matrices are modeled as Gaussian *identically and independently distributed* (i.i.d.) variables with a zero mean and a variance of 0.5 per dimension. Their magnitudes thus exhibit the Rayleigh distribution. Although the spatially uncorrelated channel assumption requires sufficient distance between the antennas in the array, it is still often used in research to evaluate the potential performance of the MIMO systems. To gain control over the spatial correlation level, we model correlation matrix $\mathbf{R}$ as follows:

$$\mathbf{h} = \mathrm{vec}[\mathbf{H}] = \mathbf{R}^{\frac{1}{2}}\mathbf{g}, \tag{3.2}$$

where $\mathbf{R} = \mathbb{E}[\mathbf{h}\mathbf{h}^H]$, $\mathbf{g} \in \mathbb{C}^{n_{\mathrm{tx}} n_{\mathrm{rx}} \times 1}$ is a random vector with i.i.d. random entries which have zero mean and unit variance, and vec[.] is a column stacking operator.

Moreover, the design of the correlation matrix is essential at the debugging stage, when the developer has to to calibrate the performance in the interference-free environment. The correlation matrices used in this thesis are summarized in the Table 3.3.

Link-level simulation are carried out using AWGN, frequency-flat, and frequency-selective fading:

- AWGN channel;
- 1-tap Rayleigh frequency-flat channel;
- 8-tap Rayleigh frequency selective channel with Exponential PDP;
- 1-tap Rician frequency-flat channel;
- 8-tap Rician frequency selective channel with Exponential PDP;
- 8-tap Extended Pedestrian A model with low, medium and high correlation, PDP defined in [3GPP, 2015d], and zero Doppler Frequency.

Note that it is possible to set a custom Rician K-factor K_{r} and Angle-of-Arrival α. In our simulations, we set $K_{\mathrm{r}} = 9.5\mathrm{dB}$, and $\alpha = 45^{\circ}$.

3.2.2 Signal Modeling

A deterministic signal model describes an input-output relationship of the downlink transmission between the base station and the UE. It reflects the number of the transmit and receive antennas, character of the transmitted signals (Gaussian inputs, constrained alphabet) and channel statistics (Gaussian, fading, time-variant, etc.). In our simulation scenarios we consider three signal models: the Single-User Closed Loop Spatial Multiplexing (CLSM) with two transmission layers, Single-User CLSM with a single transmission layer and Alamouti precoding.

Table 3.3: Correlation Matrices used in a Simulation

Channel	Correlation Matrix
Interference-free MIMO with real precoder matrix	$\mathbf{R}_{\mathrm{nointerf}} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$
Interference-free MIMO with complex precoder matrix	$\mathbf{R}_{\mathrm{nointerf}} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$
Low correlation	$\mathbf{R}_{\mathrm{low}} = \mathbf{I}_4$
Medium correlation	$\mathbf{R}_{\mathrm{medium}} = \begin{bmatrix} 1 & 0.9 & 0.3 & 0.27 \\ 0.9 & 1 & 0.27 & 0.3 \\ 0.3 & 0.27 & 1 & 0.9 \\ 0.27 & 0.3 & 0.9 & 1 \end{bmatrix}$
High Correlation	$\mathbf{R}_{\mathrm{high}} = \begin{bmatrix} 1 & 0.9 & 0.9 & 0.81 \\ 0.9 & 1 & 0.81 & 0.9 \\ 0.9 & 0.81 & 1 & 0.9 \\ 0.81 & 0.9 & 0.9 & 1 \end{bmatrix}$

3.2.2.1 Transmission Mode 4 Closed Loop Spatial Multiplexing

Consider a 2×2 MIMO system operating in Transmission Mode 4 with CLSM transmission scheme. The base station sends two transport blocks TB_0 and TB_1 mapped onto spatially multiplexed codeword CW_0 and CW_1. The received signal vector $\mathbf{y}_k \in \mathbf{C}^{2\times 1}$ for the k-th Resource Element as seen by the UE is given by

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{P}_k\mathbf{x}_k + \mathbf{n}_k, \quad k = 1, 2..., K, \tag{3.3}$$

where $\mathbf{x}_k$ is the transmitted vector of two complex symbols x_0 and x_1 with variance of σ_0^2 and σ_1^2. The transmitted vector belongs to the $\mathcal{Q}^{M_0,M_1}$ alphabet, such that $\mathcal{Q}^{M_0,M_1} := \mathcal{Q}^{M_0} \times \mathcal{Q}^{M_1}$ is the Cartesian product of two modulation alphabets $\mathcal{Q}^{M_0}$ and $\mathcal{Q}^{M_1}$, and $M_0, M_1 \in \{2, 4, 6\}$ are the modulation orders of the QAM constellations. The vector $\mathbf{n}_k$ is Zero Mean Circularly Symmetric Complex Gaussian (ZMCSCG) white noise of double-sided power spectral density $N_0/2$ at two receive antennas of UE. The matrix $\mathbf{H}_k$ is a 2×2 channel matrix built with respect to one of the channel models, described in the previous section, and $\mathbf{P}_k$ is the precoding matrix employed by the eNodeB at the k-th Resource Element.

Apart from the described scenario, the Transmission Mode 4 can be used in the configuration with a single active codeword. Then:

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{p}_k x_k + \mathbf{n}_k, \tag{3.4}$$

where $\mathbf{p}_k$ is a precoder vector for single-layer transmission, and x_k is a transmission symbol of single codeword CW_0 drawn from $\mathcal{Q}^{M_0}$ alphabet.

Single-User MIMO CLSM is the main focus of this thesis. This transmission scheme was introduced in LTE Release 8. It is used to maximize throughput of a single user, and thus is designed for the high SINR environment and scenarios, where the UE is close to the base station.

3.2.2.2 Alamouti precoding

Alamouti Coding is a two-branch transmit diversity scheme introduced by [Alamouti, 1998] for two transmit and one receive antenna. It belongs to the family of Space Time Block Codes and provides a $2n_{\mathrm{rx}}$ diversity order for the system with two transmit and n_{rx} receive antennas. MIMO systems benefit from the Alamouti precoding under slow fading assumption where the channel stays constant for the duration of two consequent time slots.

In the original scheme of 2×1 antenna configuration the transmitter sends x_0 and x_1 on the first subcarrier (first row of the signal matrix $\mathbf{X}$) from the first and second transmit antennas respectively, while symbols $-x_1^*$ and x_0^* are transmitted on the second subcarrier:

$$\mathbf{X} = \begin{bmatrix} x_0 & x_1 \\ -x_1^* & x_0^* \end{bmatrix}. \tag{3.5}$$

In the canonical case, the channel is represented by the complex vector $\mathbf{h} = \begin{bmatrix} h_0 & h_1 \end{bmatrix}^T$, built with respect to the applied channel model. Being passed through the wireless

medium, the transmitted signal $\mathbf{X}$ is multiplied with the channel gains, resulting, after adding cumulative AWGN noise $\mathbf{n}$, in the received signal $\mathbf{y}$:

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{n}. \tag{3.6}$$

Further developing (3.6), we have:

$$\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 h_0 + x_1 h_1 \\ -x_1^* h_0 + x_0^* h_1 \end{bmatrix} + \begin{bmatrix} n_0 \\ n_1 \end{bmatrix}, \tag{3.7}$$

where y_0 is the received signal during the first time slot, and y_1 is the received signal during the second time slot. Upon reception, the UE performs the conjugation of y_1 (note that conjugation is a linear operation):

$$\begin{bmatrix} y_0 \\ y_1^* \end{bmatrix} = \begin{bmatrix} h_0 & h_1 \\ h_1^* & -h_0^* \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} + \begin{bmatrix} n_0 \\ n_1^* \end{bmatrix}. \tag{3.8}$$

To benefit from the orthogonal structure of Alamouti coding $\left(\mathbf{X}\mathbf{X}^H = (|x_0|^2 + |x_1|^2)\mathbf{I}_2\right)$ and eliminate cross-antenna interference, the receiver applies the Matched Filter $\mathbf{H}^H$ to (3.8):

$$\mathbf{H}^H \begin{bmatrix} y_0 \\ y_1^* \end{bmatrix} = \left(|h_0|^2 + |h_1|^2\right) \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} + \mathbf{H}^H \begin{bmatrix} n_0 \\ n_1^* \end{bmatrix}. \tag{3.9}$$

We now perform an extension to the 2×2 antenna configuration. Let $y_{j,k}$ and $n_{j,k}$ be the members of the received signal vector and noise vector on the j receive antenna at k subcarrier. The received signals seen on the first subcarrier (3.10) and on the second subcarrier (3.11) are:

$$\begin{bmatrix} y_{0,0} \\ y_{1,0} \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} \\ h_{10} & h_{11} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} + \begin{bmatrix} n_{0,0} \\ n_{1,0} \end{bmatrix}, \tag{3.10}$$

$$\begin{bmatrix} y_{0,1} \\ y_{1,1} \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} \\ h_{10} & h_{11} \end{bmatrix} \begin{bmatrix} -x_1^* \\ x_0^* \end{bmatrix} + \begin{bmatrix} n_{0,1} \\ n_{1,1} \end{bmatrix}. \tag{3.11}$$

Complex-conjugated signal on the second slot is seen as:

$$\begin{bmatrix} y_{0,1}^* \\ y_{1,1}^* \end{bmatrix} = \begin{bmatrix} h_{01}^* & -h_{00}^* \\ h_{11}^* & -h_{10}^* \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} + \begin{bmatrix} n_{0,1}^* \\ n_{1,1}^* \end{bmatrix}. \tag{3.12}$$

Combining the received signal from two time slots (3.10) and (3.12), we obtain

$$\begin{bmatrix} y_{0,0} \\ y_{1,0} \\ y_{0,1}^* \\ y_{1,1}^* \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} \\ h_{10} & h_{11} \\ h_{01}^* & -h_{00}^* \\ h_{11}^* & -h_{10}^* \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} + \begin{bmatrix} n_{0,0} \\ n_{1,0} \\ n_{0,1}^* \\ n_{1,1}^* \end{bmatrix}. \tag{3.13}$$

We define the *combined* channel matrix $\mathbf{H}_{\text{comb}}$:

$$\mathbf{H}_{\text{comb}} = \begin{bmatrix} h_{00} & h_{01} \\ h_{10} & h_{11} \\ h_{01}^* & -h_{00}^* \\ h_{11}^* & -h_{10}^* \end{bmatrix}.$$

Similar to the 2×1 case in (3.9), the receiver applies MF $\mathbf{H}_{\text{comb}}^H$ to the combined received signal (3.13):

$$\mathbf{H}_{\text{comb}}^H \begin{bmatrix} y_{0,0} \\ y_{1,0} \\ y_{0,1}^* \\ y_{1,1}^* \end{bmatrix} = \mathbf{H}_{\text{comb}}^H \mathbf{H}_{\text{comb}} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} + \mathbf{H}_{\text{comb}}^H \begin{bmatrix} n_{0,0} \\ n_{1,0} \\ n_{0,1}^* \\ n_{1,1}^* \end{bmatrix}, \tag{3.14}$$

$$\mathbf{H}_{\text{comb}}^H \mathbf{H}_{\text{comb}} = \begin{bmatrix} |h_{00}|^2 + |h_{01}|^2 + |h_{10}|^2 + |h_{11}|^2 & 0 \\ 0 & |h_{00}|^2 + |h_{01}|^2 + |h_{10}|^2 + |h_{11}|^2 \end{bmatrix}. \tag{3.15}$$

For simplicity, we define $d = |h_{00}|^2 + |h_{01}|^2 + |h_{10}|^2 + |h_{11}|^2$. The cross-antenna interference is fully canceled:

$$\mathbf{y}_{\text{MF}} = \begin{bmatrix} d & 0 \\ 0 & d \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} + \mathbf{n}_{\text{MF}}, \tag{3.16}$$

where $\mathbf{y}_{\text{MF}}$ and $\mathbf{n}_{\text{MF}}$ are respectively the received signal and noise vector after Matched Filtering (MF).

In LTE PDSCH channel, Alamouti precoding can be used both as an independent transmission mode – TM2, and as a fall back of the advanced transmission modes, such as TM3 and 4, and is supported by the various DCI formats depending on the Search Space: format 1 and format 2 for UE specific by C-RNTI, format 1A and format 2A for common and UE specific by C-RNTI.

3.2.3 Applied Precoder Selection Strategy

Precoding techniques are designed to increase the spatial diversity gain or the spatial multiplexing gain of MIMO systems.

The LTE standard for Single-User MIMO transmission schemes defines codebook-based [3GPP, 2015b] precoding with the codebook available both at the transmitter and at the receiver. The precoder can be defined via open loop adaptation, where the decision is made solely by the base station, or via closed loop cycle, where the optimal precoder is calculated at the UE based on the estimated channel coefficients, and then is fed back to the eNodeB through one of the uplink channels. The base station decides whether to take the feedback value into account or to impose its own choice of the precoding matrix, and then signals it in the *Temporary Precoder Matrix Indicator* (TPMI) field of the DCI.

The LTE standard allows at maximum two codewords for simultaneous transmission with two possible scenarios: both codewords are active, or one codeword is disabled. There are three unitary matrices available for the precoding in DCI format 2 with two active codewords. The unitary matrices and the corresponding bit fields to PMI interpretation are presented in Table 5.1 [3GPP, 2015b].

The scenario with one disabled codeword is possible when the channel does not support two independent spatial streams simultaneously (high correlation) or if one the codewords is decoded during the preceding retransmission round. A TPMI bit field interpretation for the DCI format 2 for two antenna ports with one active codeword is shown in Table 5.2 [3GPP, 2015b].

Table 3.4: TPMI bit field interpretation for a simultaneous transmission of two codewords

Bit field	TPMI interpretation
0	$\frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$
1	$\frac{1}{2}\begin{bmatrix}1 & 1\\ j & -j\end{bmatrix}$
2	last PMI on PUSCH

Table 3.5: TPMI bit field interpretation for a transmission of a single codeword

Bit field	TPMI interpretation
0	Alamouti
1	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ 1\end{bmatrix}$
2	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ -1\end{bmatrix}$
3	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ j\end{bmatrix}$
4	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ -j\end{bmatrix}$
5	1st column of the last PMI on PUSCH
6	2nd column of the last PMI on PUSCH

The criterion based on which the UE picks up a particular PMI from the set in the LTE codebook, is subject to discussion. A few solutions have been proposed to solve this challenge, including minimum MSE criterion [Bai et al., 2010a], algorithm based on QR decomposition and MSE error [Wagdy et al., 2013], and also a metric-based criteria. The two popular metric-based criteria are the maximum mutual information [Schwarz et al., 2010] and the maximum SNR [Ghaffar and Knopp, 2010b]. However, the receiver architecture should also be taken into account as it may impact the total mutual information of the system (Section 2.1.2). In our simulations for the Parallel Interference Aware and the Successive Interference Aware Canceling receivers, we apply a light-weight solution which maximizes the SINR on the first stream and thus minimizes BLER. The UE selects the precoder matrix $\mathbf{P}$, which ensures that the effective channel of the first stream is stronger than the one of the second stream by evaluating the correlation coefficient $\rho_{10} = \mathbf{h}_{\text{eff1}}^{H}\,\mathbf{h}_{\text{eff0}}$ [Ghaffar and Knopp, 2010b]. Comparing the real and imaginary parts of ρ_{10}, the UE picks up one $\mathbf{P}$ from the following options:

$$\mathbf{P} = \begin{cases} \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}, & \text{for } \Re(\rho_{10}) \geq \Im(\rho_{10}); \\ \\ \frac{1}{2}\begin{bmatrix}1 & 1\\ j & -j\end{bmatrix}, & \text{for } \Re(\rho_{10}) < \Im(\rho_{10}). \end{cases} \tag{3.17}$$

The detail analysis of the proposed precoding strategy in demonstrated in Section 9.6.2.

3.3 Validation of the Downlink Simulator

The starting point of this thesis was to implement the Single-User MIMO TM4 and validate the accuracy of the implementation. At the time, `dlsim` supported SISO mode, Alamouti precoding, as well as TM5 and TM6.

The idea is to decompose the MIMO channel into parallel channels and compare the *per-stream* performance with the performance TM1 and TM6. To do so, we have selected the AWGN channel, where TM4 does not suffer from the cross-layer interference.

Consider the performance TM1 as a reference. TM6 and both streams of TM4 lose 3 dB of performance due to the power split at the transmitter, which is compensated with the precoding gain of 3 dB. TM6 additionally receives a 3 dB benefit from array gain accessed through Maximum Ration Combining. However, TM4 does not have a cross-layer interference thanks to the special channel modeling, and does not benefit from the Maximum Ration Combining. We thus conclude that each stream of the TM4 is expected to demonstrate performance close to the SISO transmission and TM6 should outperform them by 3 dB (Table 3.6).

Table 3.6: Expected performance gains of TM1, TM4 and TM6

	power split at the tx	precoder gain	array gain	accum gain
TM4 MIMO per stream	-3dB	+3dB	0dB	0dB
TM6 MISO	-3dB	+3dB	+3dB	+3dB
TM1 SISO	ref	ref	ref	ref

The empirical performance, illustrated in Fig. 3.7, shows a good match to the expectations. A slight gap between curves for TM1 and TM4 can be explained by the different number of pilots used in these transmission modes.

3.4 Summary

The OAI downlink simulator is a flexible tool for rapid prototyping of new physical layer solutions. The simulator provides a skeleton of LTE-conforming blocks of the signal processing, leaving for a developer or researcher the opportunity to add new features (receiver design, feedback, antenna configuration, abstraction). A variety of the transmission schemes and fading environments is offered based on analytical channel models, where new dependencies and model parameters may be introduced in order to tune the simulations for a particular scenario.

In the scope of this thesis, our contribution to the simulator was done in the following directions. First, we have implemented and validated our new receivers based on the Parallel Interference Aware detection and the Successive Interference Aware Canceling detection. The accuracy of the implementation was validated by a comparison with the performance of TM1 and TM6. We then proceeded with the extension of the HARQ protocol, already implemented for single layer transmissions, to dual codeword transmission, which included the implementation of the multi-round SIC procedure. Finally, the

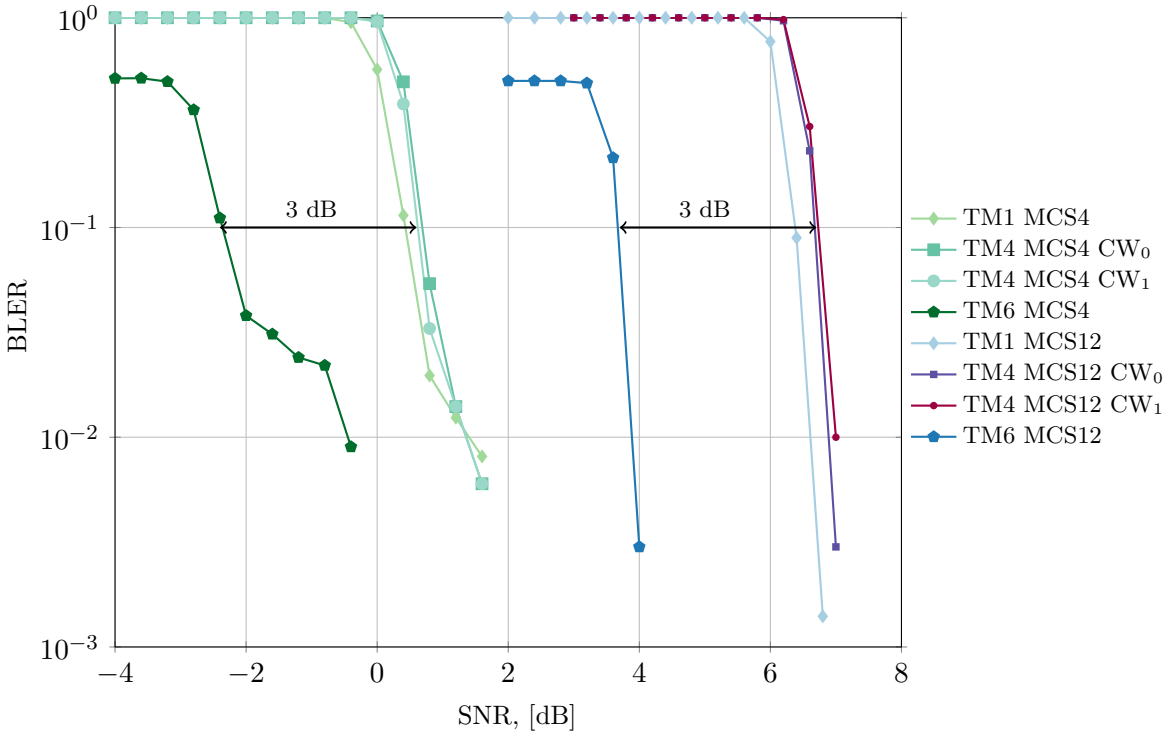


Figure 3.7: Empirical BLER curves demonstrate performance gains of TM1, TM4 and TM6 in AWGN channel.

Link Adaptation, which included the Rank Indicator estimation, the Precoder Matrix estimations and the Channel Quality indicator was added.

Chapter 4

Architecture of the R-ML PIA and SIC Receivers

Over the last few years, enhancing system performance through the design of advanced receiver architectures has been extensively studied [3GPP, 2014]. However, the real-world feasibility remains the prime criterion when trying to balance the complexity and the performance of receivers. To circumvent the extremely high computational complexity of optimum Maximum Likelihood (ML) receivers, several sub-optimal Reduced Complexity ML-based (R-ML) approaches have been introduced for Multiple-Input-Multiple-Output (MIMO) systems.

Single-User MIMO design can be seen as a system with two codewords (CW) that interfere with each other. Since both codewords are dedicated to the same user, the main objective of a receiver is to recover both of them with the highest probability. Conventional receivers absorb interference into noise, a priori limiting the potential performance. In contrast, Interference Aware (IA) receivers exploit the structure of interference and obtain thus significant performance gains [Ghaffar and Knopp, 2010a; Lee et al., 2015]. Closed Loop Spatial Multiplexing (CLSM) transmissions utilize Downlink Control Information (DCI) format 2, which carries the information about two Modulation and Coding Scheme (MCS) values, associated with the two transmitted Transport Blocks (TB). Consequently, the Single-User MIMO receiver can benefit from interference knowledge by applying interference-aware soft log-likelihood ratio (LLR) metrics [Ghaffar and Knopp, 2010a].

Contributions Single-User MIMO multi-stream interference-aware detection falls into three groups: conventional ML, R-ML Parallel Interference-Aware (PIA) and R-ML Successive Interference Aware Canceling (SIC). First, we detail the design and the implementation of our PIA and SIC receivers and quantify the computational effort for signal processing and decoding of a downlink frame. We then provide an information-theoretic analysis in order to define potential performance bounds for these receivers. The theoretical analysis is followed by evaluation of the empirical throughput achieved via link-level simulations, which deliver insights on the optimal SNR regimes and MCS combination for both codewords to maintain throughput at high level. Finally, we investigate how close

the empirical outcomes of the simulations are to the theoretical expectations.

Contents

4.1 State of the Art

Single-User MIMO systems exploit multiple transmit and receive antennas to achieve the capacity and reliability gains, as well as to reduce the sensitivity of the system to the damaging influence of interference. The importance of the receiver architecture grows along with the complexity of the wireless system. The relatively easy to build linear receivers, such as Zero Forcing (ZF) and Minimum Mean Squared Error (MMSE), are sensitive to the number of paths in a multi-path environment and show not capacity-achieving performance in MIMO systems [Onggosanusi et al., 2002]. The outage performances of MMSE and ZF receivers match in high data rates scenarios, while MMSE receiver is more favorable [Hedayat and Nosratinia, 2007] for low rate data transmissions. Thanks to these advantages, the MMSE receivers are widely deployed in practical systems. The V-BLAST receiver, developed at the Bell Labs research center, has a simple and elegant architecture which incorporates the linear ZF filter with the SIC technique [Golden et al., 1999; Wolniansky et al., 1998]. This technique allows the receiver to decode one of the codewords, subtract it from the total received signal and extract another codeword from the residue [Miridakis and Vergados, 2013]. The MMSE-SIC receiver, which uses the MMSE filter instead of ZF, outperforms the V-BLAST scheme, bringing performance closer to the optimum ML detection [Bai et al., 2009].

Although ML receivers are the optimum receivers for the MIMO systems, their complexity grows exponentially with the number of spatial layers and the modulation order of the codeword, as the receiver has to perform an exhaustive search among all possible transmitted vector candidates. Near-optimal non-linear solutions with reduced implementation complexity have become a fair compromise to balance complexity-performance trade-off. In practice, the ML detection is replaced with the maximum *a posteriori* probability (MAP) approximation [Caire et al., 1998]. However, the complexity is still high and further simplifications are required to make it practical.

4.1.1 Reduced Complexity Receiver Design

The key idea behind the majority of near-optimal solutions with lower complexity consists in reducing the search space by removing unreliable candidates or selecting the most probable symbols that could have been transmitted. The simplified ML iterative detection three-step scheme [Sung et al., 2003] was designed for coded systems. Symbols that are most likely to have been transmitted are identified through the ZF or the MMSE detector and the corresponding interference signals are regenerated and canceled out from the received signal. The symbol combinations corresponding to the most likely symbols are chosen, and the tentative decisions values are obtained by the detection. The likely symbols and tentative decisions are combined, leading to the definite decision on the candidate symbol combination, for which a distance metric is computed and the final decision on the transmitted signal is made. Another reduced complexity scheme is based on separating the MIMO channel into subchannels and removing the less probable candidates from the search space by employing a reliability criteria on normalized likelihood functions [Kim et al., 2010]. The concept of "sensitive" bits was described in [Li et al., 2004]: appropriate sensitive bits are identified among the initially estimated bits and ML detection is applied to further improve the results on the sensitive fraction.

Other representatives of space search reduction group are sphere decoding, tree search as well as techniques based on QR decomposition (QRD). The key principle of sphere decoder with polynomial complexity consists in restricting the search space to the candidate points that are located inside a sphere of radius d [Viterbo and Boutros, 1999]. Determining the sphere radius is a challenging task since a short radius makes the system vulnerable to erroneous detection, while a large radius does not provide significant complexity reduction. The LTE-compatible sphere decoders and QRD-based detectors have been reported by [Aubert et al., 2010, 2009] and [Thomas et al., 2014] respectively.

Another research direction for complexity reduction consists in manipulating the LLR metrics. One of the possible solutions is to simplify the soft decision bit metrics to the minimum distance between the received symbol and $\sqrt{Card}/2$ constellation points on the real line instead of $Card/2$ without compromising the performance [Akay and Ayanoglu, 2004], where $Card = 2^M$ is the cardinality of QAM alphabet and M is the modulation order. Ghaffar and Knopp reduced one complex dimension of search without introducing a loss in performance by decoupling the real and imaginary parts of the LLR metric [Ghaffar and Knopp, 2010a]. Combined with the Matched Filter (MF), metrics are division-free and can therefore be easily deployed in real-time modems.

Single-User MIMO LTE system often experience interference-limited scenarios, where two spatially multiplexed codewords are mapped onto the multiple transmission streams and suffer from cross-layer interference. Interference management can be applied at the transmitter using *signal pre-processing* techniques, at the UE with the help of advanced receiver design, or the optimization can be jointly performed at both sides of the MIMO system. Preprocessing techniques, such as precoding, can be codebook-based and non-codebook based. In the first case, the precoding codebooks are defined by 3GPP and shared between the base station and the UE [3GPP, 2015c]. In the latter case, so-called Demodulation Reference Signals are added before precoding providing information about joint influence of the channel and precoding. The main role in the Single-User MIMO

interference management is thus given to the receiver architecture.

Conventional receivers tend to treat interference as Gaussian and apply interference whitening technique in order to recover transmitted signal [Winters, 1984]. The advanced receiver architecture may allow for exploiting even minimal knowledge of interference structure, bringing significant gains to the system performance [Lee et al., 2015]. Ghaffar and Knopp developed the simplified receiver design, that is capable of interference mitigation through exploiting its structure [Ghaffar and Knopp, 2009]. The detailed analysis of the proposed scheme showed that the complexity in terms of the number of real-valued multiplications and additions is significantly lower that the one of max-log MAP receiver. Real-time measurements confirmed that the proposed receiver offers gains increasing with the modulation order compared to the non-interference architecture [Wagner and Kaltenberger, 2014]. In Multi-User MIMO systems, the gains from interference awareness may further be improved by avoiding blind interference detection and instead using network-aided signaling to deduct the modulation order of the interfering codeword. For Single-User scenarios, the MCSs of both codewords are available at the receiver in the DCI format. Hence, the receiver does not have to guess the modulation order of interference and can apply the matching Interference Aware soft LLR metrics. The interference-aware receiver was further extended to the sophisticated dual-stream scenario [Ghaffar et al., 2015].

While Ghaffar and Knopp considered LLR computation only for the desired codeword, [Kwon et al., 2012] insisted that the interference decoding even in Multi-User MIMO systems is important, since the interference symbols are correlated at the bit level, and the complete knowledge of interference may thus improve the overall system performance. The researchers illustrate the iterative interference-aware successive decoding for Gaussian channels where the second decoding iteration for the desired signal is enriched with the *a priori* information about the desired and interference signals achieved on the first iterations of successive decoding.

4.2 System Model

We consider a Closed Loop Spatial Multiplexing TM4 scenario. The base station is equipped with $n_{\text{tx}} = 2$ antennas and transmits two spatially multiplexed codewords to the UE with $n_{\text{rx}} = 2$ receive antennas. The codewords CW_0 and CW_1 belong to MCS_0 and MCS_1, with rates R_0 and R_1. We refer to the lower-rate R_0 codeword as CW_0, and CW_1 is always provided with an equal or higher rate R_1. Currently, we focus on the scenario without retransmissions; Chapter 5 shall provide results for multiple Hybrid Automatic Repeat Request rounds.

In this Chapter, the signal model and the precoder calculation are identically close to those defined in the Chapter 3. The received signal vector $\mathbf{y}_k$ for the k-th Resource Element observed by the UE is given by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{P}_k \mathbf{x}_k + \mathbf{n}_k, \quad k = 1, 2..., K,$$

where $\mathbf{x}_k$ is the transmitted vector of two complex symbols x_0 and x_1 with variance of σ_0^2 and σ_1^2. The transmitted vector belongs to the $\mathcal{Q}^{M_0,M_1}$ alphabet, such that $\mathcal{Q}^{M_0,M_1} :=$

$\mathcal{Q}^{M_0} \times \mathcal{Q}^{M_1}$ is the Cartesian product of two modulation alphabets $\mathcal{Q}^{M_0}$ and $\mathcal{Q}^{M_1}$, and $M_0, M_1 \in \{2, 4, 6\}$ are the modulation orders of the QAM constellations. The vector $\mathbf{n}_k$ is Zero Mean Circularly Symmetric Complex Gaussian (ZMCSCG) white noise of double-sided power spectral density $N_0/2$ at two receive antennas of UE. The matrix $\mathbf{H}_k$ is a 2×2 channel matrix built with respect to one of the channel models, described in Section 3.2.1, and $\mathbf{P}_k$ is the precoding matrix employed by the eNodeB at the k-th Resource Element.

For the sake of simplicity, we drop the Resource Element index and replace the multiplication of $\mathbf{H}$ and $\mathbf{P}$ with the effective channel $\mathbf{H}_{\text{eff}}$:

$$\mathbf{y} = \mathbf{H}_{\text{eff}}\mathbf{x} + \mathbf{n}, \quad \mathbf{H}_{\text{eff}} = [\mathbf{h}_{\text{eff0}}\ \mathbf{h}_{\text{eff1}}]. \tag{4.1}$$

4.3 Receiver Implementation

The block schemes illustrating the architecture of the R-ML PIA and the R-ML SIC receivers are depicted in Fig. 4.1 and Fig. 4.2 respectively.

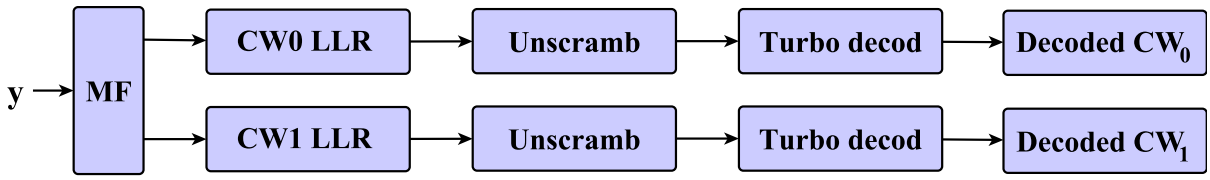


Figure 4.1: Block scheme of R-ML PIA receiver. Both codewords are handled in the same manner using interference-aware low-complexity bit metrics.

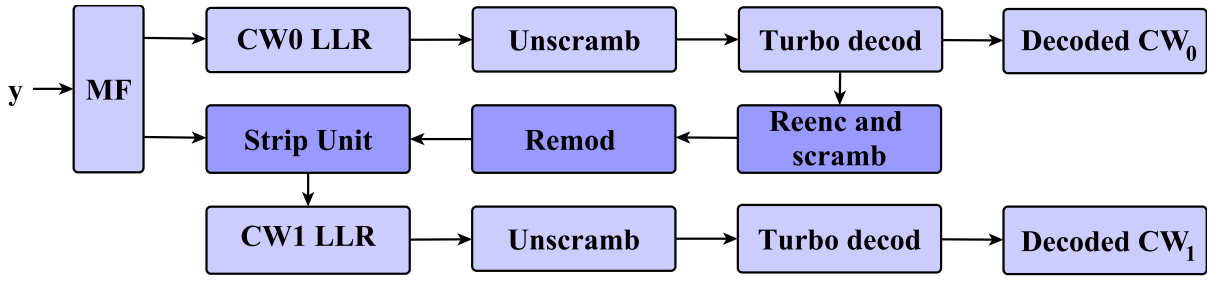


Figure 4.2: Block scheme of R-ML SIC receiver. The SIC block includes re-encoding and scrambling previously decoded interfering signal, remodulation, subtracting unit.

The receivers treat the first codeword in the common way, but handle the second codeword differently. We fist describe the common signal processing blocks and then detail the individual procedures. Both receivers take advantage of low-complexity bit metrics [Ghaffar and Knopp, 2010a].

4.3.1 Common R-ML IA Blocks

The signal processing starts with the linear Matched Filter operation, common for both PIA and SIC receivers. The received signal $\mathbf{y}$ (4.1) is transformed into

$$\mathbf{y}_{\mathrm{MF}} = \frac{1}{\sqrt{|h_{\mathrm{av}}|^2}} \mathbf{H}_{\mathrm{eff}}^H \mathbf{y}, \tag{4.2}$$

where the averaged channel level is

$$|h_{\mathrm{av}}|^2 = \max\left(\|\mathbf{h}_{\mathrm{eff0}}\|^2, \|\mathbf{h}_{\mathrm{eff1}}\|^2\right).$$

In the following derivations, we omit the averaged channel level for the sake of simplicity. The transformed channel matrix can be obtained as follows:

$$\mathbf{H}_{\mathrm{eff}}^H \mathbf{H}_{\mathrm{eff}} = \begin{bmatrix} \|\mathbf{h}_{\mathrm{eff0}}\|^2 & \rho \\ \rho^* & \|\mathbf{h}_{\mathrm{eff1}}\|^2 \end{bmatrix},$$

and the correlation coefficients can be computed as follows:

$$\begin{aligned} \rho &= h_{\mathrm{eff00}}^* h_{\mathrm{eff01}} + h_{\mathrm{eff10}}^* h_{\mathrm{eff11}}, \\ \rho^* &= h_{\mathrm{eff01}}^* h_{\mathrm{eff00}} + h_{\mathrm{eff11}}^* h_{\mathrm{eff10}}. \end{aligned}$$

After matched filtering, both PIA and SIC receivers provide an identical treatment to the lower-rate first codeword by computing values of LLR_0 using the IA bit metric `M0_M1_llr` [Ghaffar and Knopp, 2010a], which treats the second codeword as interference. The soft LLR values are then fed to the unscrambling procedure and consequently to the turbo decoder.

If the first codeword is decoded correctly, the UE forms the acknowledging message ACK_0 to confirm the successful decoding to the base station. Following this scenario, the SIC receiver triggers the SIC procedure. However, no attempt is made to decode the second codeword if the first codewords is in error.

The PIA receiver acts differently (see Fig. 4.1): the LLR values for the first and the second codewords are computed in parallel using the interference aware `M0_M1_llr` and `M1_M0_llr` functions [Ghaffar and Knopp, 2010a], and fed to two independent turbo-decoders. The probability of the second codeword to be successfully decoded thus does not depend on the decodability of the first codeword.

For multiple HARQ rounds, which will be discussed in Chapter 5, the LLR values of undecoded codewords are stored at the receiver and are updated during the next rounds.

4.3.2 SIC Procedure

The idea behind the SIC procedure is to subtract previously decoded interfering signal from the total received signal so that the second codeword is decoded in an interference-free environment. In our receiver, the SIC procedure works with compensated signals — the outputs of the Matched Filter (4.2).

The SIC procedure starts by the re-encoding the recently decoded sequence of bits of the first codeword and mapping them onto modulation symbols x_0 (see Fig. 4.2). The compensated received signal on the second antenna is given by

$$y_{\mathrm{MF}_1} = \left(\rho^* x_0 + \|\mathbf{h}_{\mathrm{eff1}}\|^2 x_1\right) + n_1', \tag{4.3}$$

where the noise term $n_1' = \mathbf{h}_{\mathrm{eff1}}^H n_1$. After multiplication by x_0, obtained from the successful decoding of the first codeword, with correlation coefficient ρ^* and subtraction the result from y_{MF_1} in the subtracting unit, x_1 enjoys interference-free detection:

$$\widetilde{y}_{\mathrm{MF}_1} = \|\mathbf{h}_{\mathrm{eff1}}\|^2 x_1 + n_1'. \tag{4.4}$$

The LLR values for the decoding of the second codeword can now be computed with the light-weight interference-free bit metric with the help of the `M1_llr` function. A single modulation order metric requires less computational effort as it involves a smaller amount of multiplications and summations.

The example of the successful performance of the SIC receiver is illustrated in Fig. 4.3.

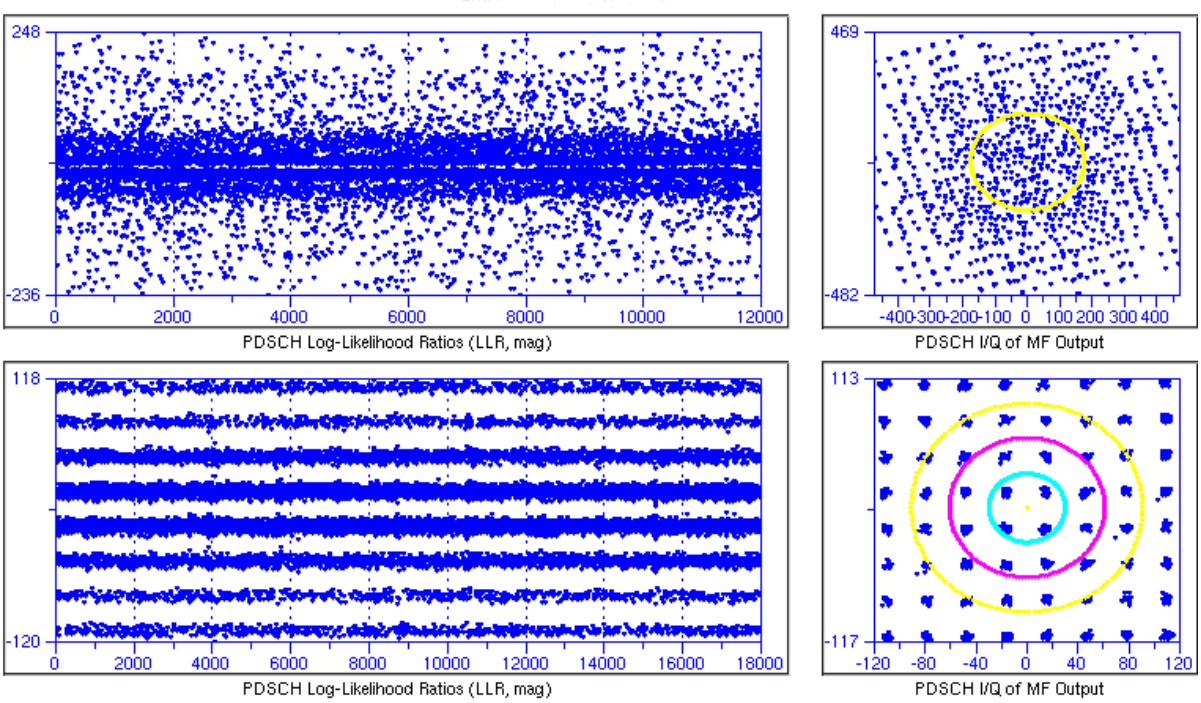


Figure 4.3: LLR and compensated received signal for the first and the second codewords after the SIC procedure. The LLR values for the first and the second codewords, which are mapped onto 64QAM, are shown in the upper left ($\mathrm{LLR}_0 \longmapsto$ `6_6_llr`) and lower corner ($\mathrm{LLR}_1 \longmapsto$ `6_llr`). The right side represents the compensated signals y_{MF_1} and $\widetilde{y}_{\mathrm{MF}_1}$. In the lower right corner, clear 64QAM constellation justifies, that the second stream is detected in the interference-free environment.

We have now discussed the detection procedures of our PIA and SIC receivers. We expect the SIC receiver to outperform the PIA receiver thanks to interference-free detection

of the second stream. Although, it is likely that in the low SNR regime, the PIA receiver may obtain slightly higher throughput, since the detection of the second stream is not blocked by the erroneous detection of the first codeword. To evaluate the computational cost of both receivers and examine the feasibility of their practical implementation, we perform analysis of computational effort using the code profiling techniques available in our downlink simulator.

4.3.3 Computational Effort

To be practically feasible for real-time transmissions, the receiver design has to be computationally efficient. In a real LTE modem, the ACK/NACK report must be generated in the third subframe after the data reception, thus giving a two millisecond window to process the data. Using real-time measurements, we verify that our receivers satisfy this requirement. The measurements are performed using one thread of the 64-bit machine with 2.10 GHz processor and 8 GB of memory.

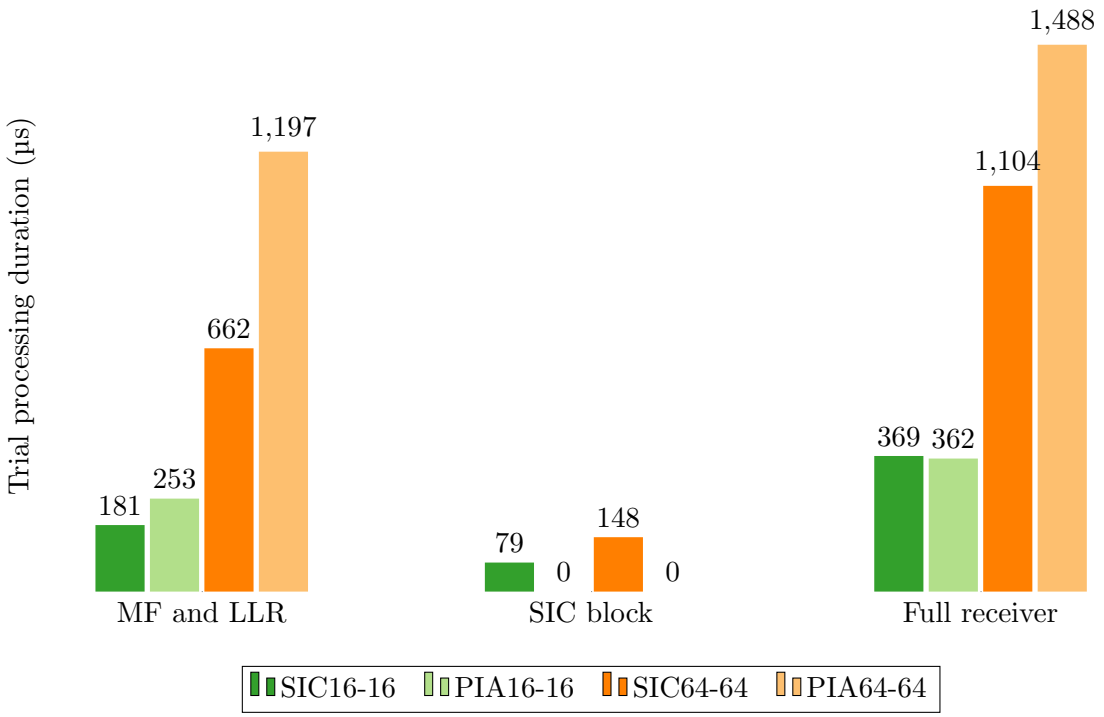


Figure 4.4: Time consumed by our SIC and PIA receivers to process downlink data depending on the constellations that the codewords are mapped onto. When both codewords are mapped onto 64QAM, the SIC receiver is 300 µs faster thanks to the SIC block that takes less time than the `6_6_llr` interference aware LLR metric for the CW_1. For high modulation orders, the SIC receiver is thus 25% more time efficient, and this gain scales with the modulation order.

The measurements take into account the Matched Filtering and LLR computations, unscrambling, SIC block and decoding. The subframe processing duration is averaged over 10000 frames. Both receivers take an equal amount time to process the data when both codewords are mapped onto 16QAM constellation, as illustrated in Fig. 4.4. This

means that the SIC block with the `4_llr` interference-free LLR metric takes approximately the same time as the `4_4_llr` interference aware LLR metric for the second codeword in PIA detection. The computational time is significantly reduced for high modulation orders, when both codewords are mapped into 64QAM. In this settings, the SIC receiver takes 300 µs less to detect one subframe than the PIA receiver since it is significantly faster to perform interference canceling with the `6_llr` interference-free LLR metric, than compute full `6_6_llr` metric. Our SIC receiver is thus 25% more time efficient for high modulation orders. Both our receivers take less than 1.5 ms to process one downlink subframe in 5 MHz bandwidth, and thus can be deployed in real systems. For higher bandwidth we propose to use multi-threading.

4.4 Mutual Information Analysis

Before we present the results of the practical simulations performed in our downlink simulator, we investigate the theoretical potential of our SIC and PIA receivers. The Mutual Information analysis provides the theoretical expectations of the achievable performance under the idealistic assumption of an infinite block length and zero outage.

The works on deriving the Mutual Information probability density function and the outage Mutual Information for MIMO systems are focused on Gaussian signals [McKay et al., 2008; Wang and Giannakis, 2004a]. The fundamental results from non-asymptotic information theory for finite blocklength codes in Gaussian channels were recently developed [Polyanskiy et al., 2010, 2011; Verdu, 2012]. For practical LTE systems where the transmitted vector belongs to a finite discrete alphabet, there is no closed form expression for the mutual information. Instead, it can be numerically approximated via Monte-Carlo simulations. Following the mutual information chain rule, the total mutual information of the MIMO system with joint I_{ML} ML decoding can be decomposed into I_0 and I_1, without compromising the performance:

$$\underbrace{I\left(X_0, X_1; \mathbf{Y}|\mathbf{H}_{\text{eff}}\right)}_{I_{\mathrm{SIC}}=I_{\mathrm{ML}}} = \underbrace{I\left(X_0; \mathbf{Y}_{\mathrm{MF}}|\mathbf{H}_{\text{eff}}\right)}_{I_0} + \underbrace{I\left(X_1; \mathbf{Y}_{\mathrm{MF}}|X_0, \mathbf{H}_{\text{eff}}\right)}_{I_1}. \tag{4.5}$$

The first component I_0 represents the conditional mutual information between the received signal vector $\mathbf{Y}_{\mathrm{MF}}$ and the transmitted symbol X_0, given the knowledge of the estimated channel matrix $\mathbf{H}_{\text{eff}}$. It thus corresponds to the mutual information of the first codeword for both the SIC and the PIA receivers. The second component I_1 represents the conditional mutual information between the received signal vector $\mathbf{Y}_{\mathrm{MF}}$ and the transmitted symbol X_1, given the knowledge of not only the estimated channel matrix, but also of symbol X_0. It thus equals the mutual information on the second stream, detected in the interference-free environment of the receiver with SIC principle. Following this note, the SIC receiver is informationally lossless compared to the joint ML detection [Foschini and Gans, 1998]. However, this is not the case for the PIA principle, where the receiver knows the modulation order of the first codeword from the DCI, but does not possess a precise information about X_0:

$$\underbrace{I\left(X_0, X_1; \mathbf{Y}|\mathbf{H}_{\text{eff}}\right)}_{I_{\mathrm{ML}}} \leq \underbrace{I\left(X_0; \mathbf{Y}_{\mathrm{MF}}|\mathbf{H}_{\text{eff}}\right)}_{I_0} + \underbrace{I\left(X_1; \mathbf{Y}_{\mathrm{MF}}|\mathbf{H}_{\text{eff}}\right)}_{I_{1\mathrm{PIA}}}. \tag{4.6}$$

The total mutual information for the system with joint ML detection can be numerically approximated using the expression (4.7), taking into account the modulation orders M_0 and M_1 of both codewords:

$$I\left(X_0, X_1; \mathbf{Y}|\mathbf{H}_{\text{eff}}, M_0, M_1\right) = \log(M_0 M_1) - \frac{1}{M_0 M_1 N_n} \times \left(\sum_{\mathbf{x} \in \mathcal{Q}^{M_0, M_1}} \sum_{n \in \mathcal{N}} \log \sum_{\mathbf{x}' \in \mathcal{Q}^{M_0, M_1}} \exp\left[\frac{-\|\mathbf{y} - \mathbf{H}_{\text{eff}}\mathbf{x}'\|^2 + \|\mathbf{n}\|^2}{N_0} \right] \right), \tag{4.7}$$

where the noise realization is drawn from a set $n \in \mathcal{N}$ with the maximum number of channel realizations $|\mathcal{N}| = N_n$.

The first codeword demonstrates identical levels of the mutual information for our PIA and SIC receivers, which can be approximated as follows:

$$I_0\left(X_0; \mathbf{Y}_{\text{MF}}|\mathbf{H}_{\text{eff}}, M_0, M_1\right) = \log M_0 - \frac{1}{M_0 M_1 N_n} \times \left(\sum_{x_0 \in \mathcal{Q}^{M_0}} \sum_{x_1 \in \mathcal{Q}^{M_1}} \sum_{n \in \mathcal{N}} \log \frac{\sum\limits_{x_0' \in \mathcal{Q}^{M_0}} \sum\limits_{x_1' \in \mathcal{Q}^{M_1}} \exp\left[-\frac{1}{N_0} \|\mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0' - \mathbf{h}_1 x_1'\|^2 \right]}{\sum\limits_{x_1'' \in \mathcal{Q}^{M_1}} \exp\left[-\frac{1}{N_0} \|\mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0 - \mathbf{h}_1 x_1''\|^2 \right]} \right). \tag{4.8}$$

However, it is not the case for the second codeword, since it enjoys interference-free detection with the SIC receiver and thus achieves higher mutual information than with the PIA detection at the same SNR level. The approximated values of mutual information of the second codeword in our SIC receiver can be computed by excluding the uncertainty from the denominator, as the receiver knows the exact symbol x_0:

$$I_1\left(X_1; \mathbf{Y}_{\text{MF}}|X_0, \mathbf{H}_{\text{eff}}, M_0, M_1\right) = \log M_1 - \frac{1}{M_0 M_1 N_n} \times \left(\sum_{x_0 \in \mathcal{Q}^{M_0}} \sum_{x_1 \in \mathcal{Q}^{M_1}} \sum_{n \in \mathcal{N}} \log \frac{\sum\limits_{x_1' \in \mathcal{Q}^{M_1}} \exp\left[-\frac{1}{N_0} \|\mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0 - \mathbf{h}_1 x_1'\|^2 \right]}{\exp\left[-\frac{1}{N_0} \|\mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0 - \mathbf{h}_1 x_1\|^2 \right]} \right). \tag{4.9}$$

To compute the mutual information values for the second codeword in the PIA receiver, we switch the symbols x_0 and x_1 in (4.8):

$$I_{1PIA}\left(X_1; \mathbf{Y}_{\text{MF}}|\mathbf{H}_{\text{eff}}, M_0, M_1\right) = \log M_1 - \frac{1}{M_0 M_1 N_n} \times \left(\sum_{x_0 \in \mathcal{Q}^{M_0}} \sum_{x_1 \in \mathcal{Q}^{M_1}} \sum_{n \in \mathcal{N}} \log \frac{\sum\limits_{x_0' \in \mathcal{Q}^{M_0}} \sum\limits_{x_1' \in \mathcal{Q}^{M_1}} \exp\left[-\frac{1}{N_0} \|\mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0' - \mathbf{h}_1 x_1'\|^2 \right]}{\sum\limits_{x_0'' \in \mathcal{Q}^{M_1}} \exp\left[-\frac{1}{N_0} \|\mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0'' - \mathbf{h}_1 x_1\|^2 \right]} \right). \tag{4.10}$$

To quantify the achievable performance and compare the theoretical potential of our PIA and SIC receivers, we analyze the total level of the mutual information for the MIMO system with two codewords in Rayleigh environment (Fig. 4.5).

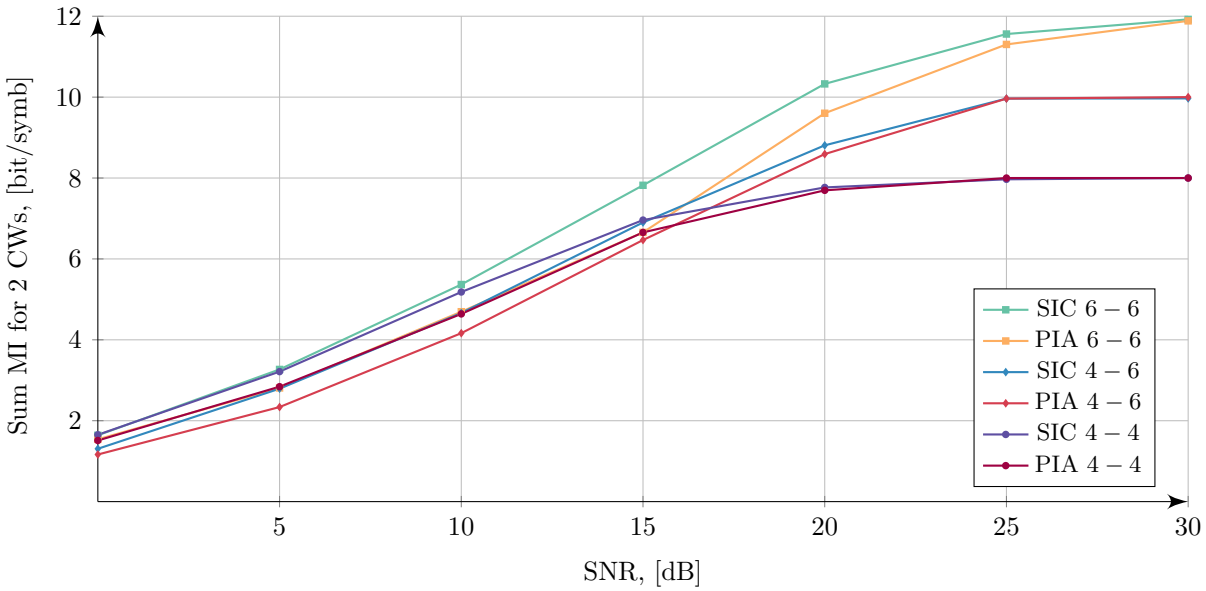


Figure 4.5: Mutual Information potential for our SIC and PIA receivers in Rayleigh fading for the codewords with the modulation orders $M_0, M_1 \in \{4, 6\}$. The performance potential of the SIC is richer than the one of the PIA receiver.

The total level of the mutual information for the system employing the SIC receiver is higher than if the PIA receiver was used, as it was expected in theory (4.5), (4.6).

The mutual information analysis provides the number of findings. The codewords mapped onto 16QAM constellations are preferable for low SNR regimes as this regime guarantees performance level very close to achievable by 64QAM-64QAM modulation (Fig. 4.5) while the computational time is significantly reduced, as we saw in previous section in Fig. 4.4. While the 64QAM-64QAM settings provide significant benefits at moderate and high SNR levels, the preferred use of 16QAM-64QAM modulation has a narrow region of application. This can be explained by the fact that the mutual information of the first codeword is a function of the constellation size of the second codeword, and thus tends to decrease as the modulation order of the second codeword increases.

4.5 Practical Results

We examine the empirical throughput, obtained as a result of the link-level simulations using OAI [OAI, 2017] downlink simulator. The simulator provides the flexibility to vary the LTE transmission parameters as well as to customize the propagation environment.

4.5.1 Simulation Parameters

The simulations were performed for Rayleigh and Rician flat fading channels. For the Rician channel, Angle-of-Arrival $\alpha = 45°$ radians and K-factor of 9.5 dB were chosen. Given 5MHz of LTE bandwidth (25 Resource Blocks), 3000 packets with 1 Physical Downlink Control Channel (PDCCH) symbol were transmitted over a wide range of noise variances. The transmission of two spatially multiplexed codewords was performed using a 2×2 antenna configuration. We chose $10 \leq \text{MCS}_0 \leq \text{MCS}_1 \leq 28$, Block Error Rate (BLER) of 10^{-2} and applied a perfect channel estimation (PCE) and a Least Square (LS) channel estimation at the UE. The simulation parameters are summarized in Table 4.1.

Table 4.1: Simulation Parameters

Antenna configuration	2×2
Transmission Scheme	TM4 Closed Loop Spatial Multiplexing
Number of Spatial Layers	2
Number of Resource Blocks	25
Number of PDCCH symbols	1
LTE Bandwidth	5 MHz
Subframe Duration	1 ms
Cyclic Prefix	Normal
Fading Environment	1 tap Rayleigh, 1 tap Rician with AoA $\alpha = 45°$ and K = 9.5 dB
Receiver Architecture	R-ML Parallel Interference Aware, R-ML Successive Interference Canceling
Channel Estimation	Perfect and Least Square
MCS	$10 \leq \text{MCS}_0 \leq \text{MCS}_1 \leq 28$
Target BLER	10^{-2}

4.5.2 Empirical throughput and MCS optimization

In a real LTE system, the throughput is limited by the rate R defined by the MCS [3GPP, 2016]. SIC receivers are sensitive to the choice of MCS: if the instantaneous channel does not support the rate of MCS_0, CW_0 is not decoded and the SIC procedure is thus not triggered. On the other hand, if the first stream is decoded, the second stream becomes interference-free and can potentially carry higher information rates. The PIA detection is less sensitive to the non-optimal MCS choice since the probabilities of the successful decoding of the codewords are dependent only on the channel conditions and the modulation order of the interferer, and do not hold a direct dependency between each other.

We aim to define the optimal value of $\text{MCS}_0^\star$ and $\text{MCS}_1^\star$ that maximize the throughput for our receivers in different SNR regimes by applying a brute force search method to the previously computed traces of the link-level simulations. The traces contain the per-stream and the total throughputs averaged across the channel realizations (assuming constant channel during 1 ms) for a wide range of SNR values for all MCS combinations.

The values of throughput are computed based on the BLER statistics from our simulator, acquired at each SNR point:

$$T_{\text{tot, sim}}(R_0, R_1, \text{SNR}) = T_0(R_0, R_1, \text{SNR}) + T_1(R_0, R_1, \text{SNR}). \tag{4.11}$$

The throughput values for the first codeword T_0 and the second codeword T_1 are obtained as

$$\begin{aligned} T_0(R_0, R_1, \text{SNR}) &= R_0(1 - BLER_0(R_0, R_1, \text{SNR})), \\ T_1(R_0, R_1, \text{SNR}) &= R_1(1 - BLER_1(R_0, R_1, \text{SNR})), \end{aligned} \tag{4.12}$$

where $BLER_0$ and $BLER_1$ are the corresponding Block Error Rates for the first and the second codewords. The BLER values depend on the SNR level, and on the rates of the desired and interfering codewords. The fact that the second codeword is attempted to be decoded only if the first one is decoded is incorporated in the Block Error Rate terms. Unlike with the PIA receiver, where the MCS choice on the streams is independent from each other, with the SIC receiver the decodability of the second stream is a function of BLER of the first stream, and we thus seek to define an optimal *combination* of the MCS.

For each SNR point, we select the values of $R_0^\star(\text{SNR})$ and $R_1^\star(\text{SNR})$ that provide the maximum throughput $T^\star_{\text{tot,sim}}(R_0^\star(\text{SNR}), R_1^\star(\text{SNR}), \text{SNR})$. This method is described in Algorithm 1:

Algorithm 1 Link-Level Simulations Based Brute Force Search of $R_0^\star, R_1^\star$

Input: SNR = 0 to $\text{SNR}_{\max}$, pregenerated traces.
Output: $R_0^\star, R_1^\star, T^\star_{\text{tot,sim}}(R_0^\star, R_1^\star, \text{SNR})$.

1: Going through all possible MCS combinations in pregenerated traces of LLS, for each SNR value find the pair of $\text{MCS}_0^\star(\text{SNR})$ and $\text{MCS}_1^\star(\text{SNR})$, that provides maximum throughput $T^\star_{\text{tot,sim}}\,(R_0^\star, R_1^\star, \text{SNR}) = \max\left[T_{\text{tot,sim}}(R_0, R_1, \text{SNR})\right]$ among all computed values $T_{\text{tot,sim}}(R_0, R_1, \text{SNR})$.
2: Obtain rates $R_0^\star$ and $R_1^\star$ corresponding to $\text{MCS}_0^\star(\text{SNR})$ and $\text{MCS}_1^\star(\text{SNR})$.

The empirically optimized throughput for our SIC receiver in Rayleigh channel is illustrated in Fig. 4.6.The solid lines represent the throughput $T_{\text{tot,sim}}(R_0, R_1, \text{SNR})$ obtained for all possible MCS combinations for CW_0 and CW_1, the red circles mark the optimized throughput $T^\star_{\text{tot,sim}}(R_0^\star, R_1^\star, \text{SNR})$, which is an envelope of the family of the solid lines. Consider the case, where both streams use $\text{MCS}_0^\star$ and $\text{MCS}_1^\star$ which are set to 28. Then, according to [3GPP, 2016], the transport block size for the for each stream is 18336 bits. This gives us 36.67 Mbit/s as the maximum throughput in our settings if both codewords are fully decoded without errors and retransmissions. However, zero-outage is not always possible for our link-level simulations due to the channel conditions, receiver imperfections, round-off errors, and fixed point implementation. These factors cause the degradation of empirical throughput $T^\star_{\text{tot,sim}}(R_0^\star, R_1^\star, \text{SNR})$ to 34 Mbit/s.

Fig. 4.7 illustrates both the throughput loss due to non-perfect CE and the benefits of the SIC receiver compared to PIA in the Rayleigh fading environment. The penalty for non-PCE does not exceed $1.5-2$ dB over the whole dynamic range of both receivers. The

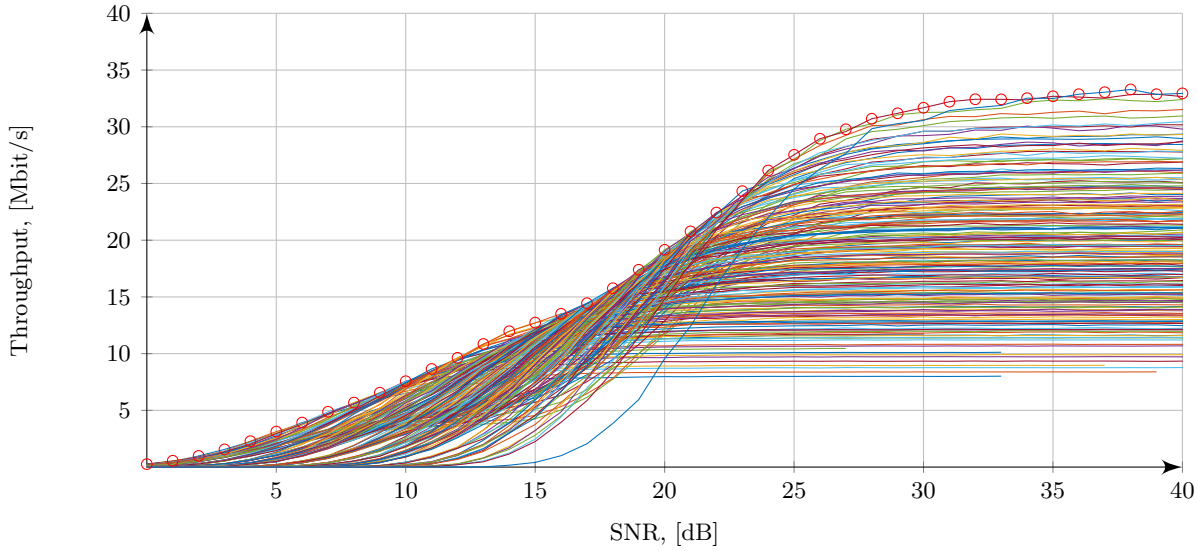


Figure 4.6: Optimized practically achieved throughput for our SIC receiver in Rayleigh channel with perfect channel estimation. The solid lines represent the empirical throughput $T_{\text{tot,sim}}$, obtained for all possible MCS combinations for both codewords, and the red circles mark the optimized throughput $T^{\star}_{\text{tot,sim}}$, which is an envelope of the family of the solid lines.

SIC receiver outperforms the PIA-based receiver in the moderate and high SNR regime and gains up to 1.8 Mbit/s. In the low SNR regime, the difference in the performance is hardly distinguishable.

A comparative analysis of the performance in the Rayleigh and in the Rician fading (Fig. 4.8) reveals that the gains of the SIC receiver depend on the presence of Line-of-Sight. For the Rician channel the gains of the SIC receiver are more significant and achieve 4 Mbit/s against 1.8 Mbit/s in Rayleigh fading. The obtained gains are expected to scale with bandwidth.

The levels of the optimized throughput are tightly linked to the corresponding values of the MCS. Fig. 4.9 illustrates the the optimized MCS values $\text{MCS}^{\star}_0$ and $\text{MCS}^{\star}_1$ for the SIC (left column of Fig. 4.9) and PIA (right column of Fig. 4.9) receivers in the Rician and Rayleigh fading. The PIA receiver uses almost identical values of the MCS for both streams. However, this does not hold for the SIC receiver, where the interference-free detection allows the receiver to support higher rate MCS on the second stream. In the Rician fading, the gap between the MCS values employed on the first and the second stream is higher.

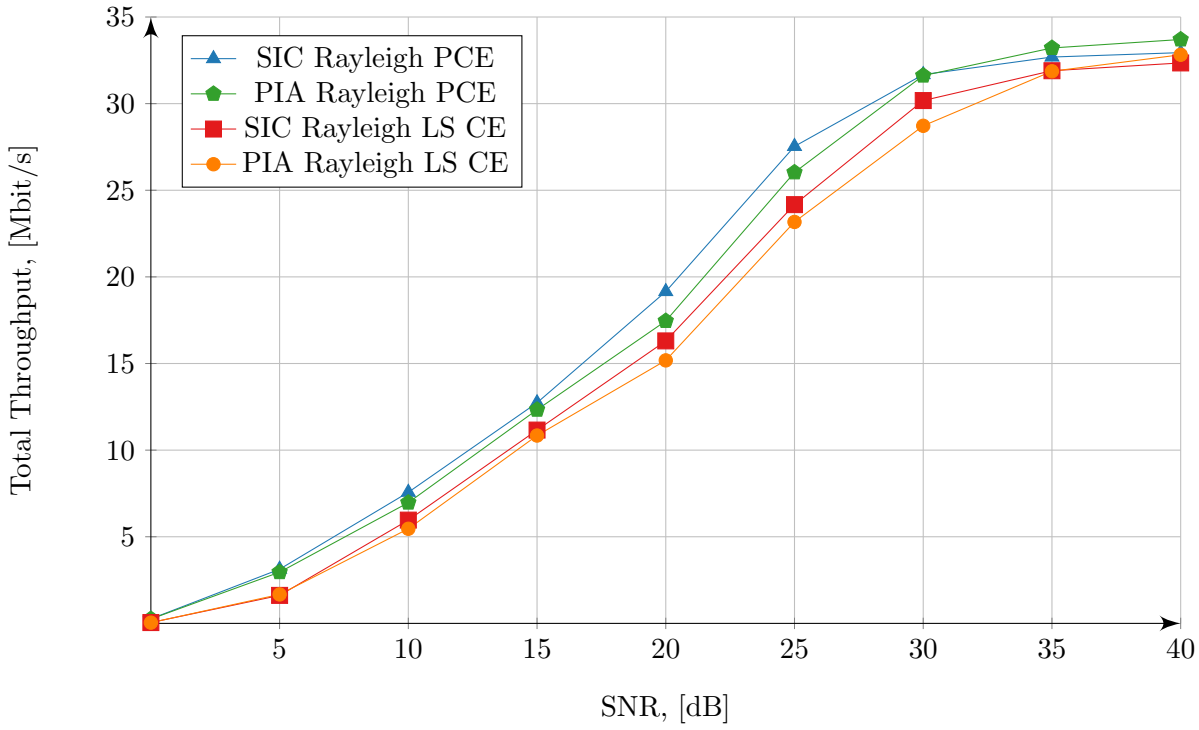


Figure 4.7: The optimized throughput $T^{\star}_{\text{tot,sim}}$ for the SIC and PIA receivers in Rayleigh fading with perfect channel estimation (PCE) and least square (LS) channel estimation. The penalty for non-PCE does not exceed $1.5-2$ dB over the whole dynamic range of both receivers. The SIC receiver outperforms the PIA-based receiver in the moderate and high SNR regime and gains up to 1.8 Mbit/s.

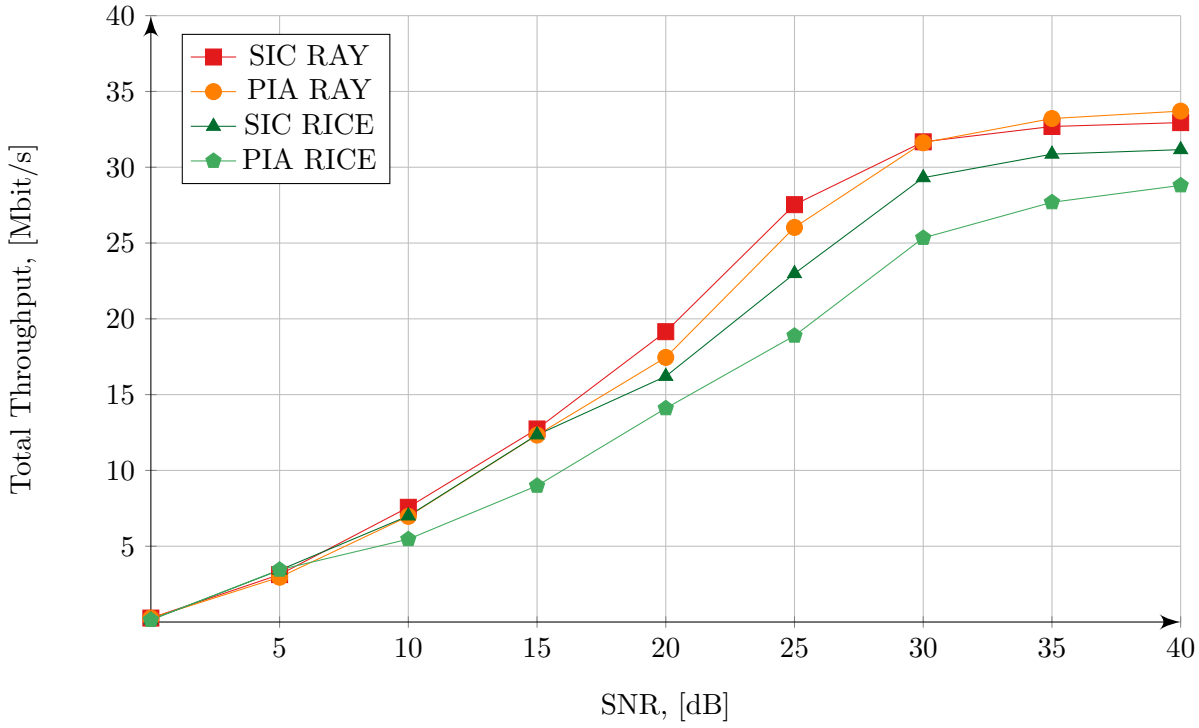


Figure 4.8: The optimized throughput $T^{\star}_{\text{tot,sim}}$ for the SIC and PIA receivers with perfect channel estimation in Rayleigh and Rician fading environments. The SIC receiver gains up to 4 Mbit/s in moderate and high SNR regime, while in the Rayleigh channel this gap is not so significant and achieves 1.8Mbit/s.

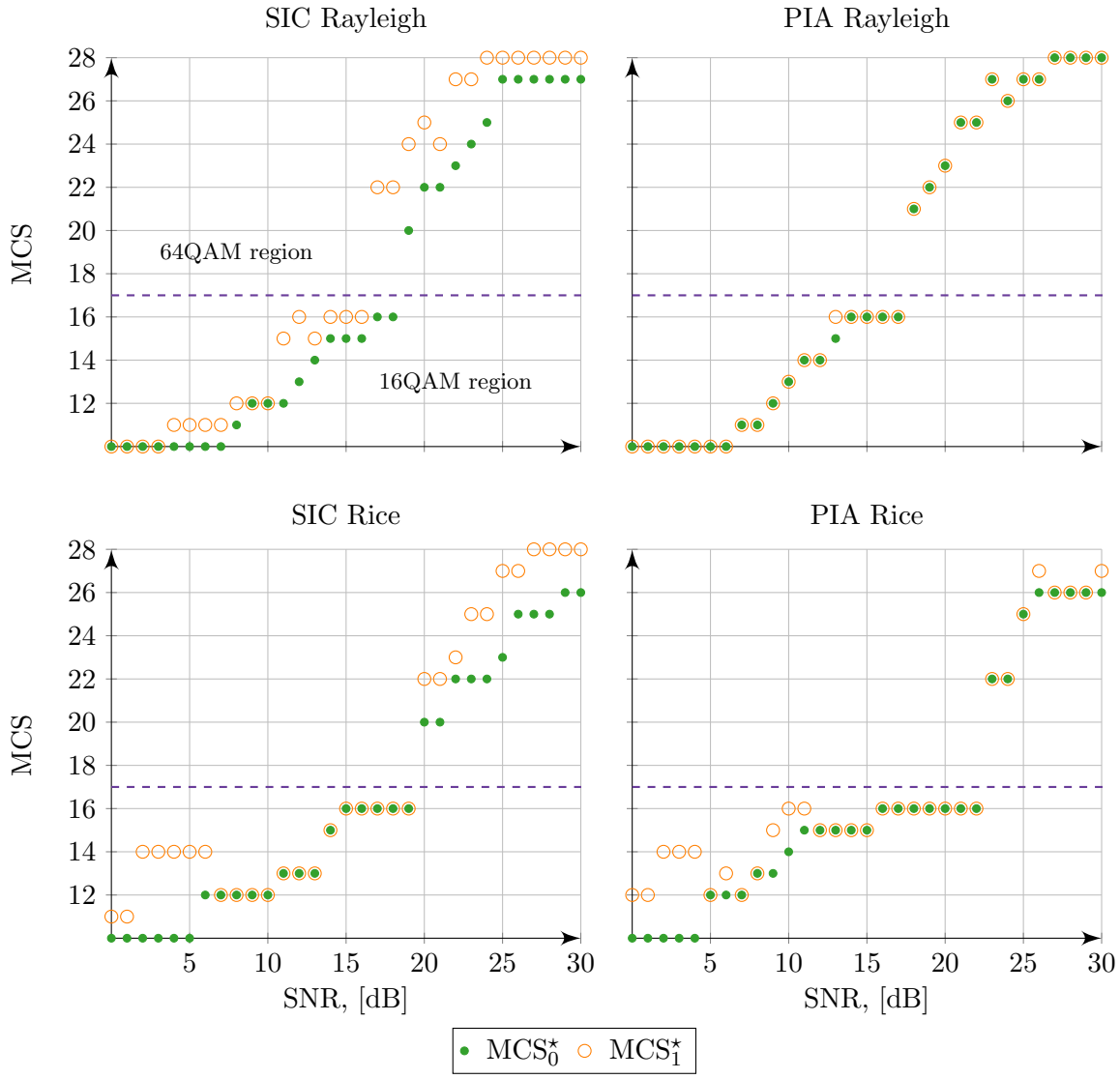


Figure 4.9: $\text{MCS}_0^\star$ and $\text{MCS}_1^\star$ evolution for SIC (left column) and PIA (right column) receivers with perfect channel estimation in Rayleigh and Rician fading environment. The interference-free detection allows the SIC receiver to support higher rate MCS on the second stream. The throughput of the SIC receiver is thus higher than the one of the PIA receiver.

Suggestions for the optimal modulation order that we deducted from the mutual information analysis in Section 4.4, are now confirmed by the empirical results shown in Fig. 4.9. For both fading environments, 16QAM-16QAM modulation is preferred in the low SNR regime of the Rayleigh channel. For the high SNR levels, the receiver confidently uses 64QAM modulation for both codewords, making use of the highest MCS 28 on both streams for the SNR around 30 dB, and 16QAM-64QAM combination is almost never applied.

4.6 Practice vs Theoretical Expectations

To quantify how much the results from our link-level simulations are affected by the receiver imperfections, round-off errors, and fixed point implementation, we perform a comparison between the empirical throughput and the theoretical expectations delivered by the Mutual Information analysis under the idealistic assumptions in Section 4.4.

To bring the theoretical expectations together with the empirical results, we apply the concept of *Mutual Information outage probability* [Wang and Giannakis, 2004b]. This allows to take into account the events where the capacity of the channel (in LTE systems with Bit Interleaved Coded Modulation represented by the mutual information) is lower than the rate of the MCS currently being used. In this analysis, we consider our SIC receiver, and the results for the PIA detection can be derived in the similar way.

The probabilities of outage for the first and the second codeword using our SIC receiver are defined respectively as:

$$P_{\text{out}_0}(R_0, \text{SNR}) = \Pr\left(I_0\left(\mathbf{H}_{\text{eff}}, \text{SNR}\right) < R_0\right). \tag{4.13}$$

In our SIC receiver, the probability of the outage event for the second codeword is conditioned on the the probability of the outage for the first codeword:

$$P'_{\text{out}_1}(R_0, R_1, \text{SNR}) = P_{\text{out}_0}(R_0, \text{SNR}) P_{\text{out}_1}(R_1, \text{SNR}), \tag{4.14}$$

where the term $P_{\text{out}_1}(R_1, \text{SNR})$ is the outage probability of the second codeword if it were independent from the successful decoding of the first codeword, and can be computed as follows:

$$P_{\text{out}_1}(R_1, \text{SNR}) = \Pr\left(I_1\left(\mathbf{H}_{\text{eff}}, \text{SNR}\right) < R_1\right). \tag{4.15}$$

The terms $I_0\left(\mathbf{H}_{\text{eff}}, \text{SNR}\right)$ and $I_1\left(\mathbf{H}_{\text{eff}}, \text{SNR}\right)$ represent the instantaneous mutual information for the first and the second codeword and can be computed using expressions (4.8) and (4.9):

$$\begin{aligned} I_0\left(\mathbf{H}_{\text{eff}}, \text{SNR}\right) &= I\left(X_0; \mathbf{Y}_{\text{MF}} | \mathbf{H}_{\text{eff}}, M_0, M_1\right), \\ I_1\left(\mathbf{H}_{\text{eff}}, \text{SNR}\right) &= I\left(X_1; \mathbf{Y}_{\text{MF}} | X_0, \mathbf{H}_{\text{eff}}, M_0, M_1\right). \end{aligned}$$

Let $T_{\text{tot}}(R_0, R_1, \text{SNR})$ be the total throughput of the system, then

$$\begin{aligned} T_{\text{tot}}(R_0, R_1, \text{SNR}) = {} & R_0(1 - P_{\text{out}_0}(R_0, \text{SNR})) \\ & + R_1(1 - P_{\text{out}_0}(R_0, \text{SNR}))(1 - P_{\text{out}_1}(R_1, \text{SNR})). \end{aligned} \tag{4.16}$$

We have provided the description of the analytical tools which are required for the comparison of the theoretical potential, discussed in Section 4.4, and our empirical results obtained via link-level simulations in Section 4.5.2. The empirical long-term throughput for the SIC receiver employing the optimal combination of the $\text{MCS}_0^\star$ and $\text{MCS}_1^\star$ is generated from the computed in advance traces of link-level simulations. The method is briefly described in Algorithm 1 and was previously used in Section 4.5.2.

To compare the empirical throughput with the theoretical expectations, we develop Hybrid and Upper-bound methods utilizing Mutual Information outage probability. For

each value of R_0, R_1, and SNR we compute and store the corresponding probabilities $P_{\text{out}_0}(R_0, \text{SNR})$ and $P'_{\text{out}_1}(R_1, \text{SNR})$. The Hybrid method (Algorithm 2) aims to identify the theoretical throughput that can be expected with the optimal $\text{MCS}_0^\star$ and $\text{MCS}_1^\star$, chosen in Section 4.5.2:

$$T_{\text{tot,hybr}}(SNR) = T_{\text{tot}}(R_0^\star(\text{SNR}), R_1^\star(\text{SNR}), \text{SNR}), \tag{4.17}$$

where the optimal rates $R_0^\star$ and $R_1^\star$ corresponding to the $\text{MCS}_0^\star$ and $\text{MCS}_1^\star$ maximize the *empirical* system throughput $T_{\text{tot,sim}}(R_0, R_1, \text{SNR})$:

$$R_0^\star(\text{SNR}), R_1^\star(\text{SNR}) = \underset{R_0,R_1}{\text{argmax}}\, T_{\text{tot,sim}}(R_0, R_1, \text{SNR}). \tag{4.18}$$

To provide an upper-bound for an empirical throughput, we develop Algorithm 3 where throughput (4.16) is computed for each possible MCS combination (not only optimal) and for each SNR value the maximum value of $T_{\text{tot,up-b}}$ is identified:

$$T_{\text{tot,up-b}}(\text{SNR}) = T_{\text{tot}}(R_{0,\text{MI}}^\star(\text{SNR}), R_{1,\text{MI}}^\star(\text{SNR}), \text{SNR}), \tag{4.19}$$

where the optimal rates $R_{0,\text{MI}}^\star$ and $R_{1,\text{MI}}^\star$ corresponding to the $\text{MCS}_0^\star$ and $\text{MCS}_1^\star$ maximize the *total* system throughput $T_{\text{tot}}(R_0, R_1, \text{SNR})$:

$$R_{0,\text{MI}}^\star(\text{SNR}), R_{1,\text{MI}}^\star(\text{SNR}) = \underset{R_0,R_1}{\text{argmax}}\, T_{\text{tot}}(R_0, R_1, \text{SNR}). \tag{4.20}$$

Algorithm 2 Hybrid throughput $T_{\text{tot, hybr}}$

Input: SNR = 0 to $\text{SNR}_{\max}$, $R_0^\star$, $R_1^\star$, $P_{\text{out}_0}(R_0, \text{SNR})$, $P_{\text{out}_1}(R_1, \text{SNR})$.
Output: $T_{\text{tot,hybr}}^\star(R_0^\star, R_1^\star, \text{SNR})$.

1: Given $R_0^\star, R_1^\star$ from Algorithm 1, find $P_{\text{out}_0}(R_0^\star, \text{SNR})$ and $P_{\text{out}_1}(R_1^\star, \text{SNR})$.
2: To get $T_{\text{tot,hybr}}^\star(R_0^\star, R_1^\star, \text{SNR})$, substitute $P_{\text{out}_0}(R_0^\star, \text{SNR})$, $P_{\text{out}_1}(R_1^\star, \text{SNR})$, $R_0^\star$ and $R_1^\star$ into (4.16).

Algorithm 3 Upper-Bound throughput $T_{\text{tot, up-b}}$

Input: SNR = 0 to $\text{SNR}_{\max}$, $P_{\text{out}_0}(R_0, \text{SNR})$, $P_{\text{out}_1}(R_1, \text{SNR})$.
Output: $T_{\text{tot,up-b}}(\text{SNR})$.

1: Find $R_{0,\text{MI}}^\star(\text{SNR}), R_{1,\text{MI}}^\star(\text{SNR}) = \underset{R_0,R_1}{\text{argmax}} T_{\text{tot}}(R_0, R_1, \text{SNR})$.
2: Substitute $R_{0,\text{MI}}^\star(\text{SNR})$, $R_{1,\text{MI}}^\star(\text{SNR})$, $P_{\text{out}_0}(R_{0,\text{MI}}^\star(\text{SNR}), \text{SNR})$ and $P_{\text{out}_1}(R_{1,\text{MI}}^\star(\text{SNR}), \text{SNR})$ into (4.16) to get $T_{\text{tot,up-b}}(\text{SNR}) = T_{\text{tot}}(R_{0,\text{MI}}^\star(\text{SNR}), R_{1,\text{MI}}^\star(\text{SNR}), \text{SNR})$..

Fig. 4.10 illustrates the gap between the empirical results of our SIC receiver, throughput obtained via our Hybrid method, and upper-bound throughput. The throughput values $T_{\text{tot,sim}}^\star$ and $T_{\text{tot,hybr}}^\star$ are very close when the UE is in the low SNR regime and uses 16QAM-16QAM or 16QAM-64QAM modulations. However, as soon as the receiver goes

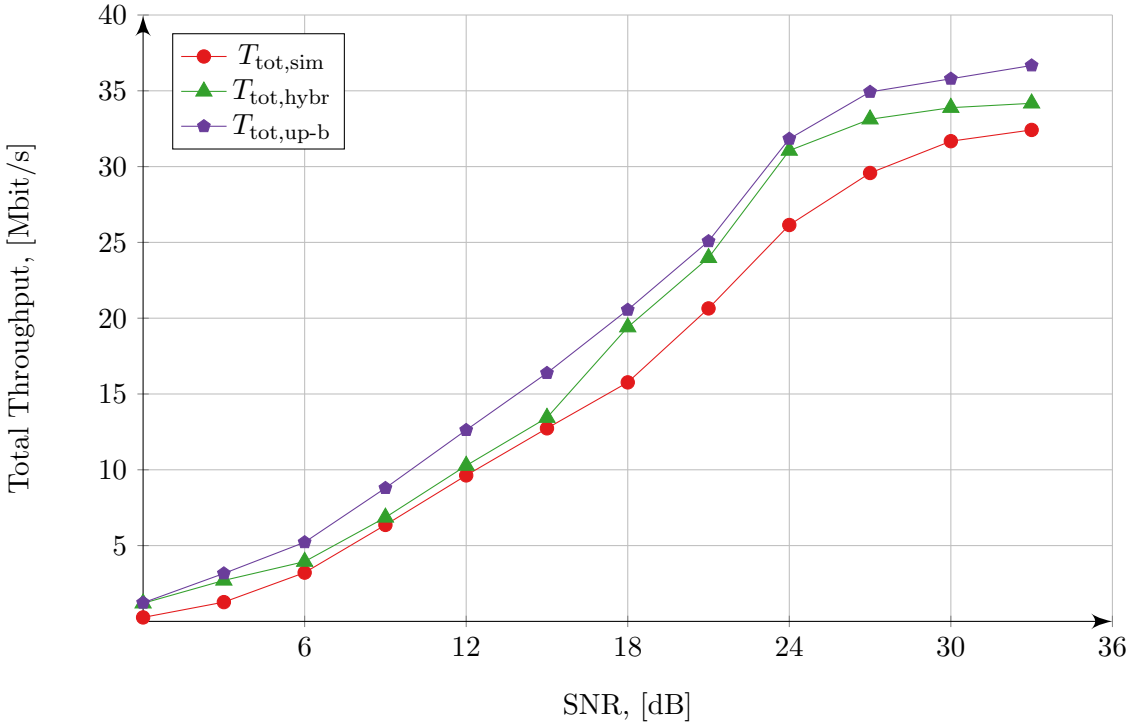


Figure 4.10: Empirical, Hybrid and Upper-Bound throughput for our SIC receiver in the Rayleigh channel. The throughput values $T^{\star}_{\text{tot,sim}}$ and $T^{\star}_{\text{tot,hybr}}$ are very close when the UE is in the low SNR regime where 16QAM-16QAM or 16QAM-64QAM is applied. However, the gap increases in the 64QAM-64QAM region.

to the high SNR regime, where 64QAM is used for both codewords, the gap between actual and predicted levels throughput increases. This can be explained by the fact that we model Mutual Information for the infinite codeblock length, while the codeblock length is in fact limited by the transport block size in a real LTE system. Perhaps, some results from [Polyanskiy et al., 2010] could provide more insights.

4.7 Conclusion

In this chapter, we have presented a comparative study of the Reduced Complexity Parallel Interference Aware receiver and Successive Interference Aware Canceling receiver based on our implementation in the OAI simulator.

Based on link-level simulations, the SIC receiver outperforms the PIA receiver by 4 Mbit/s in Rician flat fading and 1.8 Mbit/s in Rayleigh flat fading given 5 MHz LTE bandwidth. The gains are expected to scale with the bandwidth. Intuitively it could be expected that in the lower SNR regime the PIA receiver outperforms the SIC receiver, since both codewords undergo decoding attempts. However, both empirical throughput and theoretical potential results demonstrate that the SIC receiver actually always outperforms the PIA detection.

The quantification of the computational effort demonstrates that for high modulation

orders our SIC receiver is 25% more time efficient than the PIA receiver, thanks to replacing the time consuming IA LLR metric of the second codeword with the light-weight SIC block. For the moderate modulation orders, the computational effort is approximately at the same level.

The empirical performance is close to the theoretical expectations based on the mutual information and probability of outage analysis in the low and moderate SNR regimes. However, the gap increases at high SNR levels, which can be explained by the fact that we model Mutual Information for the infinite codeblock length, while the codeblock length is in fact limited by the transport block size in a real LTE system.

All the empirical results presented in this chapter are obtained from the experiments in our downlink simulator. As a next step, our receivers will be integrated into the OAI emulator — `oaisim` — as a part of OAI SoftUE. It will be then possible to compare the obtained levels of throughput with those reached with full LTE protocol stack. Furthermore, even though detection mechanisms in chipsets of commercial UEs are not disclosed, it would be still interesting to compare the throughput levels in our settings with the throughput obtained in the same scenario (i.e. two layer transmission) by a commercial UE in real time.

Chapter 5

HARQ Protocol for the R-ML PIA and SIC Receivers

Long Term Evolution (LTE) has been designed to allow maximum flexibility in exploiting the benefits of Multiple-Input-Multiple-Output (MIMO) channels. The so-called Transmission Modes (TM) range from transmit diversity over beamforming to spatial multiplexing. At the same time, Hybrid Automatic Repeat Request (HARQ) retransmission protocols were developed to reduce transmission errors over fading channels by using multiple retransmissions [Caire and Tuninetti, 2001; Lin et al., 1984].

The signal may experience deep fades, as it travels through the propagation medium. A deep fading environment may have a harmful effect on the signal, complicate the detection, and even lead to the loss of the codeword. In this situation, the UE may request from the base station to retransmit the distorted package.

The HARQ protocol incorporates the mechanism of signaling the successful/erroneous detection from the UE to the base station. The signaling is performed by sending the acknowledging message ACK or the non-acknowledging message NACK. Based on the delivered information, the base station makes a decision if the codeword needs to be retransmitted. If the retransmission is scheduled, the UE combines the knowledge of the codeword from the previous transmission with the new data, and has thus higher chances to successfully decode the codeword.

Single-User MIMO systems with two spatially multiplexed codewords require ACK and NACK messages to be reported on per-codeword basis [3GPP, 2015a]. Hence, the detection method at the UE inevitably influences the HARQ protocol handling. In this chapter, we provide a detailed look at HARQ processing in Single-User MIMO systems with our Parallel Interference-Aware (PIA) and Successive Interference Aware Canceling (SIC) receivers.

The HARQ retransmissions are handled on both sides of the MIMO system: at the base station and at the receiver. The majority of the literature is dedicated to the HARQ processing at the UE, and the 3GPP standard describes the processing at the base station. However, the practical aspects of the HARQ in the Single-User MIMO systems are not fully defined, especially for the case where the UE successfully decodes only one of the codewords.

Contributions We first implement Incremental Redundancy (IR) HARQ support for the Single-User MIMO system with the PIA and the SIC receiver architectures in our downlink simulator. We provide a detailed description of the protocol and discuss the blocks involved in the retransmission process. The base station processing does not depend on the receiver architecture, and is thus common for the MIMO systems with our PIA and SIC receivers. The UE processing is receiver specific and is discussed for each of the receivers. Special attention is dedicated to the multi-round SIC procedure, available for our SIC receiver: once the first codeword is decoded, it is possible to reconstruct the interference-free second codeword from *all* the preceding rounds and to perform Log-Likelihood Ratio (LLR) combining.

Secondly, we quantify the performance of our receivers with HARQ protocol in terms of throughput and reliability.

Finally, we test different retransmission schemes in the scenarios with one successfully decoded codeword and the other one is in error.

Contents

5.1 State of the Art

A fundamental analysis of the HARQ protocols in Gaussian collision channels was performed by [Caire and Tuninetti, 2001], who derived the closed-form expression for the throughput metrics. Their work was extended to MIMO V-BLAST systems by [Dekorsy, 2005] based on the conditional cut-off rate of MIMO transmission.

One of the features of the LTE technology is the ability of the eNodeB to adjust the modulation and coding scheme (MCS) to the instantaneous channel conditions, based on the Channel State Information (CSI). The benefits of adaptive HARQ with IR when MCS values change over the retransmission rounds were investigated by [Szczecinski et al., 2010]. His simulations show that the adaptive HARQ brings significant gains compared to non-adaptive schemes for high SNR in channels with an outdated CSI at the transmitter. The interesting results of the adaptive HARQ was proposed by [Villa et al., 2012], where the authors achieve performance close to the ergodic capacity with only one bit of the

feedback. However, the 3GPP LTE standards [3GPP, 2015a, 2016] define downlink HARQ protocol as non-adaptive and asynchronous.

Advanced MIMO receiver architecture takes cross-layer interference into account. This may put specific constrains on the design of the retransmission scheme [Toumpakaris et al., 2008]. Interference-aware successive decoding with symbol-level combining scheme was proposed by [Kwon et al., 2013]. The detailed analysis of the optimal combining schemes for MIMO systems with HARQ was offered by [Jang et al., 2007].

The majority of the state-of-the-art literature focuses on techniques to handle HARQ protocol in MIMO systems at the UE side. At the same time, there are very few sources describing the retransmission options of the eNodeB. In June 2016, we performed a drive test campaign in Sophia Antipolis, France, during which the full message flow between the eNodeB and the UE was captured. This allowed us to study the HARQ implementation in a practical MIMO LTE system with Ericsson eNodeB, configured in Cyclic Delay Diversity (TM 3) using Downlink Control Information (DCI) format 2A with two antenna ports. When the UE awaits to receive two transport blocks per subframe, they are associated with the same HARQ process. If only one codeword is in error, the eNodeB performs a retransmission of only the erroneous codeword, deactivating the codeword that was successfully decoded. However, there is no precoding information field for the two antenna port configuration in the DCI format 2A. The transmit diversity (Alamouti precoding) is the only option available for the eNodeB for a retransmission of a single codeword in TM 3. This inspired us to study retransmission strategies for TM 4 based on our OpenAirInterface [OAI, 2017] downlink simulator. TM 4 uses the DCI format 2 with an active precoding information field for two antenna port transmission. For the single CW retransmission, the eNodeB uses Temporary Precoder Matrix Indicator (TPMI) in the DCI format 2 to signal the retransmission scheme.

5.2 System Model

We consider a 2×2 MIMO system, where the connection between the base station and the UE is established in TM 4 with DCI format 2. The HARQ-IR protocol is configured with at maximum four retransmission rounds. The base station sends TB_0 and TB_1 mapped onto the spatially multiplexed CW_0 and CW_1 at the initial transmission round $r = 0$, as illustrated in Fig. 5.1.

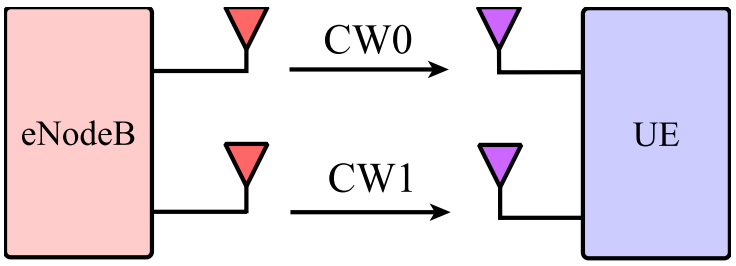


Figure 5.1: Diagram of the scenario for TM 4. Two transport blocks TB_0 and TB_1 are mapped onto the spatially multiplexed codewords CW_0 and CW_1 at the initial transmission round $r = 0$.

The received signal vector $\mathbf{y}_k \in \mathbf{C}^{2\times 1}$ for the k-th Resource Element, as seen by the UE, is given by

$$r = 0 : \mathbf{y}_k = \mathbf{H}_k \mathbf{P}_k \mathbf{x}_k + \mathbf{n}_k, \quad k = 1, 2..., K, \tag{5.1}$$

where $\mathbf{x}_k$ is the transmitted vector of two complex symbols x_0 and x_1 with variance of σ_0^2 and σ_1^2. The transmitted vector belongs to the $\mathcal{Q}^{M_0,M_1}$ alphabet, such that $\mathcal{Q}^{M_0,M_1} := \mathcal{Q}^{M_0} \times \mathcal{Q}^{M_1}$ is the Cartesian product of two modulation alphabets $\mathcal{Q}^{M_0}$ and $\mathcal{Q}^{M_1}$, and $M_0, M_1 \in \{2, 4, 6\}$ are the modulation orders of the QAM constellations. The vector $\mathbf{n}_k$ is Zero Mean Circularly Symmetric Complex Gaussian (ZMCSCG) white noise of double-sided power spectral density $N_0/2$ at two receive antennas of UE. The matrix $\mathbf{H}_k$ is a 2×2 channel matrix built with respect to one of the channel models, described in Section 3.2.1, and $\mathbf{P}_k$ is the precoding matrix employed by the eNodeB at the k-th Resource Element.

For the sake of simplicity, we drop the Resource Element index and replace the multiplication of $\mathbf{H}$ and $\mathbf{P}$ with the effective channel $\mathbf{H}_{\text{eff}}$:

$$\mathbf{y} = \mathbf{H}_{\text{eff}}\mathbf{x} + \mathbf{n}, \quad \mathbf{H}_{\text{eff}} = [\mathbf{h}_{\text{eff0}}\ \mathbf{h}_{\text{eff1}}]. \tag{5.2}$$

In the next section, we discuss the protocol implementation at both the base station and the UE following the described scenario.

5.3 Retransmission Protocols

In Single-User MIMO systems with two spatially multiplexed codewords, the HARQ protocol may follow one of the retransmission scenarios, illustrated in Fig. 5.2. The scenarios are defined by the outcome of the decoding procedures at the UE.

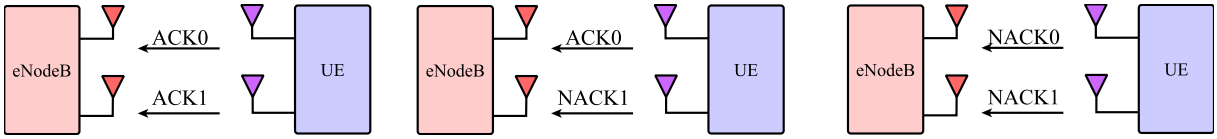


Figure 5.2: Possible retransmission scenarios for Single User MIMO system: the base station receives two acknowledgments and no retransmission is requested; one ACK and one NACK is received, and only one codeword is retransmitted; the base station receives a double NACK, and two spatially multiplexed codewords are prepared for a retransmission.

The receiver architecture plays an important role in the way retransmissions are handled at the UE. With our SIC receiver, the decodability of the second codeword is tightly linked to successful decoding of the first codeword, since no attempt to decode the second codeword is made if the first codeword is not decoded successfully. Thus if the first codeword is not decoded, the SIC receiver sends two non-acknowledgment messages NACK_0 and NACK_1 for codewords. This is not the case for our PIA receiver, where both streams are decoded independently and a failure to decode the first codeword does not necessary lead to the erroneous detection of the second codeword. Note that the index of acknowledgment and non-acknowledgment messages ACK_{TB} and NACK_{TB} indicates the transport block, and not the codeword.

During the initial transmission, or when the base station receives two non-acknowledgment messages NACK_0 and NACK_1, the eNodeB transmits two transport blocks TB_0 and TB_1 that are mapped onto codewords CW_0 and CW_1:

$$\text{TB}_0 \longmapsto \text{CW}_0, \quad \text{TB}_1 \longmapsto \text{CW}_1.$$

The setting when only one of the codewords is decoded (which corresponds to the second scenario in Fig. 5.2) attracted our attention since the standard does not explicitly specify how the base station should perform a retransmission in this case. In this case, the retransmission of the two codewords would lead to a non-efficient use of the spectral resources. The base station thus disables the corresponding transport block for the next retransmission rounds, and retransmits only the erroneous codeword. The deactivation is performed through the DCI: the MCS for the deactivated transport block is set to zero, while the corresponding redundancy version is set to one. According to the standard, if only one transport block is active, it should be mapped onto the first codeword CW_0 [3GPP, 2015b]. If the data conveyed by the first transport block TB_0 is successfully decoded, the remaining second transport block TB_1 is mapped onto the first codeword CW_0:

$$\text{TB}_1 \longmapsto \text{CW}_0, \quad \text{TB}_0 \text{ is disabled.}$$

Similarly, in the reverse situation, when the second transport block TB_1 is successfully decoded while the first transport block TB_0 is in error, TB_0 remains mapped onto the first codeword CW_0:

$$\text{TB}_0 \longmapsto \text{CW}_0, \quad \text{TB}_1 \text{ is disabled.}$$

After one of the codewords is disabled, the base station keeps DCI format 2 and should choose a precoding vector, which is then signaled to the UE via TPMI bit field in DCI. The options for precoding include Alamouti precoding and single layer precoding. When the single layer precoding is used, the base station may select the precoder from the available precoding codebook or use the one suggested in the latest CSI report on PUSCH.

Table 5.1: TPMI bit field interpretation for a dual CW transmission

Bit field	TPMI interpretation
0	$\frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$
1	$\frac{1}{2}\begin{bmatrix}1 & 1\\ j & -j\end{bmatrix}$
2	last PMI on PUSCH

As we have mentioned before, the base station transmits the information about the precoder matrix indicator via TPMI bit field in DCI. The TPMI bit field interpretation depends on the number of active codewords, and is different for single codeword and dual codeword transmissions. The interpretations for both cases are given in Table 5.1 and Table 5.2 following [3GPP, 2015b].

In the next sections, we discuss the design of the IR-HARQ retransmission protocol for our SIC and PIA receivers. We first provide the general algorithm of the HARQ protocol

Table 5.2: TPMI bit field interpretation for a single CW retransmission

Bit field	TPMI interpretation
0	Alamouti
1	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$
2	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$
3	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\j\end{bmatrix}$
4	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-j\end{bmatrix}$
5	1st column of the last PMI on PUSCH
6	2nd column of the last PMI on PUSCH

for Single-User MIMO system. Then, we look at the base station HARQ processing and the UE HARQ processing in details.

5.3.1 HARQ protocol for PIA Receiver

The main checkpoints of the retransmission protocol are the ACK and NACK messages, which are sent by the UE on per-codeword basis and indicate the result of the decoding procedure. When the base station receives two acknowledging messages ACK_0 and ACK_1, the retransmissions are terminated and the new pair of codewords is prepared for a transmission. However, in poor channel conditions, a few retransmission rounds might be required to successfully decode both codewords. To illustrate the protocol handling, we introduce the retransmission algorithm for TM4 PIA detection (Algorithm 4) which describes the overall methodology. We then detail several procedures corresponding to single and dual codeword signal processing at the base station (Algorithm 7, Algorithm 5), and discuss the signal processing with HARQ at the UE (Algorithm 8, Algorithm 6).

When the base station receives two non-acknowledgment messages NACK_0 and NACK_1, it means that the UE has failed to decode both codewords, and has requested a new redundancy version of both of them. In this case, the base station prepares a new retransmission via triggering the **Dual Codeword eNodeB Proc** procedure (Algorithm 5). This procedure performs rate matching with the new redundancy version using the Cyclic Buffer that was constructed during the encoding process before the initial transmission. The rate matching is followed by the classic signal processing chain with the modulation, layer mapping and precoding.

In our scenario, we assume that the dual codeword transmission always uses CSI-based precoding, that is sent by the UE on PUSCH. The single codeword transmission is studied with a few precoding options.

Following the dual codeword signal processing at the base station, the UE also launches the corresponding **Dual Codeword PIA Proc** procedure (Algorithm 6). Given the received signal vector $\mathbf{y}$, channel estimates, and TPMI (available from the DCI), the UE

Algorithm 4 Retransmission algorithm for TM4 PIA detection

$r \leftarrow 0$, TPMI $\leftarrow 2$, $r_{\max} \leftarrow 4$.
TB_0_flag $\leftarrow$ *enabled*, TB_1_flag $\leftarrow$ *enabled*, $\text{TB}_0 \longmapsto \text{CW}_0$, $\text{TB}_1 \longmapsto \text{CW}_1$
$\text{ACK}_0 \leftarrow false$, $\text{ACK}_1 \leftarrow false$
while $(r \leq (r_{\max} - 1))$ **do**
    if ACK_0 == false && ACK_1 == false **then**
        [eNodeB]: call **Dual Codeword eNodeB Proc**.
        [UE]: measurements, feedback reporting on PUSCH.
        [UE]: call **procedure Dual Codeword UE Proc**.
    else if ACK_0 == true && ACK_1 == false **then**
        [eNodeB]: call **Single Codeword eNodeB Proc**.
        [UE]: measurements, feedback reporting on PUSCH.
        [UE]: call **procedure Single Codeword UE Proc**.
    else if ACK_0 == false && ACK_1 == true **then**
        [eNodeB]: call **Single Codeword eNodeB Proc**.
        [UE]: measurements, feedback reporting on PUSCH.
        [UE]: call **procedure Single Codeword UE Proc**.
    else if ACK_0 == true && ACK_1 == true **then**
        break ▷ Move to the next package
    end if
    $r \leftarrow r + 1$
end while

applies the Matched filter and demodulates the received signal:

$$\mathbf{y}_{\text{MF}} = \mathbf{H}_{\text{eff}}^{H}\mathbf{y}.$$

The demodulation is followed by the per-codeword LLR computation using the interference-aware LLR metrics `M0_M1_llr` and `M1_M0_llr` initially proposed by [Ghaffar and Knopp, 2010a]. The calculated LLR values are then fed to two independent turbo decoders. Depending on the results of the decoding process, the UE generates acknowledgment and non-acknowledgment messages that are then sent to the base station.

If one of the codewords is decoded while another one is in error, the base station receives one acknowledgment and one non-acknowledgment message. It would be a waste

Algorithm 5 Dual Codeword eNodeB processing

procedure Dual Codeword eNodeB Proc(r)
    if $r == 0$ **then**
        perform encoding of two TBs.
        perform rate matching, modulation, layer mapping and precoding
        transmit two spatially multiplexed CWs.
    else if $0 < r \leq (r_{\max} - 1)$ **then**
        perform rate matching with new RVs using the same Cyclic Buffer.
        perform modulation, layer mapping and precoding
        transmit two spatially multiplexed CWs.
    else
        break ▷ Move to the next package
    end if
end procedure

Algorithm 6 Dual Codeword PIA processing

```
1: procedure Dual Codeword PIA Proc(y, channel estimates, TPMI)
2:     demodulate both codewords by applying precoder matrix corresponding to TPMI.
3:     compute LLR_0 and LLR_1, combine with LLRs from previous round if available.
4:     forward updated LLR_0 and LLR_1 to 2 turbo-decoders.
5:     if TB_0 is decoded && TB_1 is decoded then
6:         ACK_0 ← true, ACK_1 ← true
7:         send ACK_0 and ACK_1 message
8:         break
9:     else if TB_0 is decoded && TB_1 is not decoded then
10:        ACK_0 ← true, ACK_1 ← false
11:        send ACK_0 and NACK_1
12:    else if TB_0 is not decoded && TB_1 is decoded then
13:        ACK_0 ← false, ACK_1 ← true
14:        send NACK_0 and ACK_1
15:    else if TB_0 is not decoded && TB_1 is not decoded then
16:        ACK_0 ← false, ACK_1 ← false
17:        send NACK_0 and NACK_1
18:    end if
19: end procedure
```

of the resources to transmit two codewords again. Therefore, the base station deactivates the decoded transport block and switches to single codeword transmission by triggering the **Single Codeword eNodeB Proc** procedure (Algorithm 7).

As we discussed in the beginning of this section, if only one transport block is transmitted, it should be mapped onto the first codeword. After making the corresponding changes in the DCI, the UE preforms rate matching for the transport block that is still active using a new redundancy version, modulation and precoding. In the last CSI report, the base station received the Precoder Matrix Indicator that corresponds to the dual layer transmission (Table 5.1), and thus cannot use it directly. The eNodeB may impose the precoding according to its own choice from the Table 5.2, such as Alamouti precoding of beamforming, or use the first or the second column of the precoding matrix **P** suggested by the UE in the CSI report. The eNodeB precoder choice is encoded in the DCI.

To let the UE know that the decoded codeword is deactivated, the base station sets the corresponding MCS value to zero, and the redundancy version to one. After decoding the DCI, the UE understands that only one codeword was transmitted, and triggers the **Single Codeword UE Proc** procedure (Algorithm 8). After interpreting the TPMI bit field according to the Table 5.2 and selecting a corresponding precoder, the UE performs the Matched Filter based demodulation of the active codeword. After Maximum Ratio Combining, the LLR can be computed using interference-free LLR metrics `M0_llr` or `M1_llr`, depending on which transport block is disabled. The computed values of LLRs are then combined with the values of LLRs from the previous round and are fed to the turbo-decoder. Based on the results of the decoding procedure, the UE generates acknowledgment or non-acknowledgment messages and sends them to the base station.

The HARQ protocol handling with the SIC receiver has common steps with the strategy applied by our PIA receiver, such as codeword processing at the base station. How-

Algorithm 7 Single Codeword eNodeB processing

procedure Single Codeword eNodeB Proc(r, ACK_0, ACK_1)
  if $0 < r \leq (r_{\max} - 1)$ **then**
    if $\text{ACK}_0 == true$ **then**
      [eNodeB]: TB_0_flag $\leftarrow$ *disabled*, TB_1_flag $\leftarrow$ *enabled*
      [eNodeB]: $\text{MCS}_0 = 0$, $\text{RV}_0 = 1$
      $\text{TB}_1 \longmapsto \text{CW}_0$.
      perform rate matching with new RVs using the same Cyclic Buffer.
      perform modulation and precoding
      transmit CW_0.
    else
      [eNodeB]: TB_0_flag $\leftarrow$ *enabled*, TB_1_flag $\leftarrow$ *disabled*
      [eNodeB]: $\text{MCS}_1 = 0$, $\text{RV}_1 = 1$
      $\text{TB}_0 \longmapsto \text{CW}_0$.
      perform rate matching with new RVs using the same Cyclic Buffer.
      perform modulation and precoding
      transmit CW_0.
    end if
  else
    break ▷ Move to the next package
  end if
end procedure

Algorithm 8 Single Codeword UE processing

procedure Single Codeword PIA Proc(**y**, channel estimates, TPMI, TB_0_flag, TB_1_flag)
  if TB_0_flag == enabled **then**
    demodulate CW_0 by applying precoder vector corresponding to TPMI.
    compute LLR_0 and combine with LLRs from previous rounds.
    forward combined LLR_0 to turbo-decoder.
    if TB_0 is decoded **then**
      $\text{ACK}_0 \leftarrow true$
      send ACK_0
      break
    else if TB_0 is not decoded **then**
      $\text{ACK}_0 \leftarrow false$
      send NACK_0
    end if
  else if TB_1_flag == enabled **then**
    demodulate CW_0 applying precoder vector corresponding to TPMI.
    compute LLR_1 and combine with LLRs from previous rounds.
    forward combined LLR_1 to turbo-decoder.
    if TB_1 is decoded **then**
      $\text{ACK}_1 \leftarrow true$
      send ACK_1
      break
    else if TB_1 is not decoded **then**
      $\text{ACK}_0 \leftarrow false$
      send NACK_1
    end if
  end if
end procedure

ever, since the decodability of the second codeword is tightly coupled with the successful decoding of the first codeword, the SIC receiver treats retransmission differently. The originality comes from the so-called multi-round SIC procedure, where once the first codeword is decoded, the second codeword can be reconstructed from all the previous rounds and the LLR values for the second codeword are immediately combined. In the next section, we will detail this methodology.

5.3.2 HARQ protocol for SIC Receiver

The main difference between the HARQ treatment with SIC and PIA receivers comes from the tight link between the decodability of the second codeword and successful decoding of the first codeword. It means that the base station is configured for the dual codeword transmission till the moment when the first codeword is decoded. Similarly to the analysis of the PIA receiver, we first present the overall strategy (Algorithm 9), and then detail the dual and single codeword handling at the base station and a the UE.

Algorithm 9 Retransmission algorithm for TM4 SIC detection

1: $r \leftarrow 0$, TPMI $\leftarrow 2$, $r_{\max} \leftarrow 4$.
2: TB_0_flag $\leftarrow$ *enabled*, TB_1_flag $\leftarrow$ *enabled*, $\mathrm{TB}_0 \longmapsto \mathrm{CW}_0$, $\mathrm{TB}_1 \longmapsto \mathrm{CW}_1$
3: $\mathrm{ACK}_0 \leftarrow false$, $\mathrm{ACK}_1 \leftarrow false$
4: **while** $(r \leq (r_{\max} - 1))$ **do**
5: **if** ACK_0 == false && ACK_1 == false **then**
6: [eNodeB]: call **Dual Codeword eNodeB Proc**.
7: [UE]: measurements, feedback reporting on PUSCH.
8: [UE]: call **procedure Dual Codeword SIC Proc**.
9: **else if** ACK_0 == true && ACK_1 == false **then**
10: [eNodeB]: call **Single Codeword eNodeB Proc**(r, ACK_0, ACK_1).
11: [UE]: measurements, feedback reporting on PUSCH.
12: [UE]: call **procedure Single Codeword UE Proc**.
13: **else if** ACK_0 == true && ACK_1 == true **then**
14: **break** ▷ Move to the next package
15: **end if**
16: $r \leftarrow r + 1$
17: **end while**

As in the HARQ strategy with PIA receiver, for the SIC receiver strategy (Algorithm 9) the key triggers for the retransmissions are the acknowledgment and the non-acknowledgment ACK and NACK messages received by the base station. If the base station receives a double non-acknowledgment, the next retransmission is processed identically to the case with the PIA receiver.

Assume that the first codeword CW_0 is decoded on the round $r = r_{\mathrm{dec}}$. For rounds $r \leq r_{\mathrm{dec}}$, the base station applies the **Dual Codeword eNodeB Proc** procedure (Algorithm 6) by sending a new RV of spatially multiplexed CW_0 and CW_1 every round. During this time, the transport blocks TB_0 and TB_1 are mapped onto the corresponding codewords CW_0 and CW_1:

$$r \leq r_{\mathrm{dec}} : \mathrm{TB}_0 \longmapsto \mathrm{CW}_0, \mathrm{TB}_1 \longmapsto \mathrm{CW}_1.$$

Upon reception of the downlink data, the UE decodes the DCI and calls the **Dual Codeword SIC Proc** procedure (Algorithm 10). This procedure reflects the SIC architecture, where the decodability of the second codeword depends on the successful decoding of the first codeword. Similarly to the PIA receiver, the dual codeword processing in our SIC receiver begins with the Matched Filter based demodulation of the received signal $\mathbf{y}$:

$$\mathbf{y}_{\mathrm{MF}} = \mathbf{H}_{\mathrm{eff}}^{H}\mathbf{y}.$$

Algorithm 10 Dual Codeword SIC processing

```
1: procedure Dual Codeword SIC Proc(y, channel estimates, TPMI)
2:     demodulate both CWs by applying the precoder matrix corresponding to TPMI.
3:     compute LLR_0 and combine with the LLR values from the previous round, if available.
4:     forward the combined LLR_0 the turbo-decoder.
5:     if TB_0 is decoded then
6:         call procedure Multi-Round SIC Proc
7:         if TB_1 is decoded then
8:             ACK_0 ← true, ACK_1 ← true
9:             send ACK_0 and ACK_1
10:            break
11:        else
12:            ACK_0 ← true, ACK_1 ← false
13:            send ACK_0 and NACK_1
14:        end if
15:    else
16:        ACK_0 ← false, ACK_1 ← false
17:        send NACK_0 and NACK_1
18:    end if
19: end procedure
```

Contrary to the PIA detection, where demodulation is followed by the LLR computation for both codewords, the SIC receiver computes in the beginning the LLR values LLR_0 only for the first codeword. If this is not the first round of transmission, the values LLR_0 are updated with the corresponding values from the previous rounds. The LLR values are then sent to the turbo-decoder. If the first codeword is successfully decoded, the UE triggers the **Multi-Round SIC Proc** procedure, detailed in Algorithm 11. It is important that the receiver keeps the channel coefficient and compensated received signal $\mathbf{y}_{\mathrm{MF}}$ so that the second codeword can be reconstructed in all the preceding rounds.

Since the UE now knows the transmitted symbol x_0, the channel coefficients and the compensated received signal $y_{\mathrm{MF}_1 r}$, stored in a special buffer during the preceding retransmissions, can be used to reconstruct the interference-free second codeword on each retransmission round $r = 0, 1, \ldots, r_{\mathrm{dec}}$. The decoded bit sequence of the first codeword is re-encoded and passed through the rate matching with the appropriately picked redundancy version. Then, the bits are mapped onto the modulated symbols and multiplied with the corresponding channel coefficients ρ_r^*, which can be computed as:

$$\rho_r^* = h_{\mathrm{eff}01}^* h_{\mathrm{eff}00} + h_{\mathrm{eff}11}^* h_{\mathrm{eff}10}. \tag{5.3}$$

Algorithm 11 Multi-round SIC Procedure

procedure SIC(r_{dec}, MF outputs from previous rounds)
  $r_{\text{sic}} \leftarrow 0$
  while $0 \leq r_{\text{sic}} \leq r_{\text{dec}}$ **do**
    Re-encode the decoded bit sequence of the CW_0.
    Perform rate matching with $\text{RV}_0 = r_{\text{dec}}$.
    Perform mapping onto modulation symbols x_0.
    Multiply x_0 with ρ_r^*.
    Pass $\rho_r^* x_0$ to the subtracting unit to obtain interference-free CW_1.
    Compute $\text{LLR}_{1r_{\text{sic}}}$ and combine with LLR_1 from the previous rounds.
    Pass the updated LLR_1 to the turbo-decoder.
    if CW_1 is decoded **then**
      $\text{ACK}_0 \leftarrow true$, $\text{ACK}_1 \leftarrow true$.
      send ACK_0 and ACK_1.
      break
    else
      $r_{\text{sic}} \leftarrow r_{\text{sic}} + 1$
    end if
  end while
  $\text{ACK}_0 \leftarrow true$, $\text{ACK}_1 \leftarrow false$.
  send ACK_0 and NACK_1.
end procedure

The subtracting unit cancels $\rho_r^* x_0$ from the stored compensated received signal $y_{\text{MF}_1 r}$ (5.4) on the second antenna

$$y_{\text{MF}_1 r} = \rho_r^* x_0 + \|\mathbf{h}_{\text{eff1}}\|^2 x_1 + n'_{1r}, \tag{5.4}$$

and we obtain the interference-free signal $\widetilde{y}_{\text{MF}_1 r}$ for each round $r = 0, 1, \ldots, r_{\text{dec}}$:

$$\widetilde{y}_{\text{MF}_1 r} = \|\mathbf{h}_{\text{eff1}}\|^2 x_1 + n'_{1r}, \tag{5.5}$$

where the noise term $n'_1 = \mathbf{h}_{\text{eff1}}^H n_1$. Then, the corresponding multi-round interference-free LLR_{1r} metrics are calculated. The UE makes an attempt to decode TB_1 at each round after LLR combining. This is done to avoid crossing the succeeding rounds of the multi-round SIC procedure if the LLR values are already reliable enough. If the decoding is successful, the UE sends the acknowledgment messages ACK_0 and ACK_1 to the base station and awaits the reception of the next package. Otherwise, if the codeword CW_1 is still in error after r_{dec} decoding attempts, the ACK_0 and NACK_1 messages are sent.

If the second codeword is not decoded during the multi-round SIC procedure and the limit of retransmission rounds has not been reached yet, the eNodeB retransmits a new redundancy version $\text{RV}_1 = (r_{\text{dec}} + 1)$ of TB_1 using single codeword transmission following the **Single Codeword eNodeB Proc** procedure (Algorithm 7). Similarly, the UE triggers the **Single Codeword UE Proc** procedure (Algorithm 8), which is identical for our PIA and SIC receivers.

The precoding scheme for the single codeword transmission may impact the system performance in terms of throughput and reliability. In the next section we detail the precoding options that we apply in our simulation scenarios.

5.3.3 Precoding Schemes for Single Codeword Retransmission

After disabling one of the transport blocks, the eNodeB keeps the DCI format 2, which has an active TPMI field for the transmission with two antenna ports. However, the last TPMI was set for two layer transmission (Table 5.1), and thus must be reconfigured. The possible TPMI options, including the Alamouti precoding and the single layer precoding, are listed in Table 5.2 [3GPP, 2015b]. In this work, we focus on three options: TPMI $\in \{0, 5, 6\}$, where TPMI5 corresponds to the first column of the precoding matrix reported by the UE, and TPMI6 is interpreted as the second column. Our choice of the TPMI is driven by the fact that the Alamouti precoding (TPMI0) is designed to increase the reliability in low SNR regime, and TPMI5 and TPMI6 are expected to improve system performance at high SNR, as they match the estimated channel.

Apart from the different TPMI options, we also consider two possibilities: if there is an actual CSI for each retransmission round r, or if the last CSI was received on r_{dec} and thus is outdated for the r rounds, with $r_{\text{dec}} < r \leq (r_{\max} - 1)$, when a single codeword is retransmitted. Our precoder selection strategy maximizes the SNR level for the first codeword. Therefore, when the transport block is mapped to CW_0, the TPMI5 is expected to perform better than the TPMI6. This creates an interesting comparison between the TPMI5 with actual CSI (assuming only PMI feedback, Channel Quality Indicator is not taken into account) and the TPMI6 with outdated CSI.

TPMI0 The eNodeB configures the retransmission in Alamouti mode. In this case, the received signal for rounds $r_{\text{dec}} < r \leq (r_{\max} - 1)$ can be computed by:

$$\mathbf{y} = \frac{1}{\sqrt{2}}\mathbf{X}\mathbf{h} + \mathbf{n}, \quad \text{where } \mathbf{X} = \begin{bmatrix} x_0 & x_1 \\ -x_1^* & x_0^* \end{bmatrix}. \tag{5.6}$$

The detailed derivations on the Alamouti precoding for MIMO systems are given in Section 3.2.2.2.

TPMI5 and TPMI6 The base station configures the single layer retransmission based on the latest PMI report on PUSCH, using the precoder $\mathbf{p}$ from the first (TPMI5) or the second (TPMI6) column multiplied by $\sqrt{2}$ of latest reported PMI. In this case, the received signal for the r rounds, with $r_{\text{dec}} < r \leq (r_{\max} - 1)$, is given by:

$$\mathbf{y} = \mathbf{H}\mathbf{p}x + \mathbf{n}. \tag{5.7}$$

We have shown that the main difference in the HARQ protocol handling between two our receiver consists in the multi-round SIC procedure. To evaluate the retransmission gains and test the precoding options for single codeword retransmission, we perform numerical simulations.

5.4 Numerical Results and Discussion

In the previous section, we have discussed the implementation of the HARQ protocol for our PIA and SIC receivers. We now move forward to the quantification of the gains coming

from the multiple retransmissions via link level simulations performed in our downlink simulator. The empirical results are evaluated based on the reliability and maximum system throughput criteria.

5.4.1 Simulation parameters

For our empirical link level simulations, we used the following settings: 5 MHz of LTE bandwidth with 25 Resource Blocks were scheduled to a single user with two receive antennas. The propagation environment was simulated using the 8-tap Rayleigh fading channel model with exponential Power Delay Profile (PDP) and the Extended Pedestrian A (EPA) channel, whose PDP is defined by [3GPP, 2015d] model with zero Doppler frequency. The EPA channel was generated with the low and medium correlation matrix $\mathbf{R}$. The low correlation matrix $\mathbf{R}_{\text{low}}$ is defined as an identity matrix $\mathbf{I}_4$ of size 4×4, while the medium correlation matrix for 2×2 MIMO system can be obtained as

$$\mathbf{R}_{\text{medium}} = \begin{bmatrix} 1 & 0.9 & 0.3 & 0.27 \\ 0.9 & 1 & 0.27 & 0.3 \\ 0.3 & 0.27 & 1 & 0.9 \\ 0.27 & 0.3 & 0.9 & 1 \end{bmatrix}. \tag{5.8}$$

The base station sends 3000 packets with one Physical Downlink Control Channel symbol over a wide range of noise variances. Every retransmission round is drawn from a new channel realization. Since the TM 4 is designed for high data rates transmission, we chose the values of MCS_0 and MCS_1, such that $10 \leq \text{MCS}_0 \leq \text{MCS}_1 \leq 28$. The single codeword retransmission is simulated with three precoding options: TPMI0 (Alamouti precoding), TPMI5 (the first column of the precoder matrix suggested in CSI report) with actual CSI, and TPMI6 (the second column of the precoder matrix suggested in CSI report) with outdated CSI.

The parameters, used in the simulation scenarios in this Chapter are summarized in Table 5.3.

5.4.2 MCS Optimization and Throughput Analysis for the Multiple Rounds

To quantify the performance gains from the retransmission rounds, we investigate the empirical multi-round throughput achieved if the optimal combination of the MCS on both streams is used. A similar analysis was performed in Section 4.5.2 for a single round transmission, where our SIC receiver outperforms the PIA receiver by 1.8 Mbit/s in the Rayleigh flat fading channel, and by 4 Mbit/s in Rician flat fading channel, given 5 MHz of LTE bandwidth. We now extend the comparison to multiple HARQ rounds and frequency-selective channel models, such as the 8-tap Rayleigh channel and the EPA channel. In particular, we investigate whether the multiple rounds bring throughput gains and if these gains depend on the TPMI employed during the single codeword transmission. Following the LTE standards, we focus on non-adaptive HARQ protocol.

The total system throughput of the MIMO system with two transport blocks is composed from the throughput contribution of the first transport block TB_0 and the throughput contribution of the second transport block TB_1. The throughput of the MIMO system

Table 5.3: Simulation Parameters

Antenna configuration	2×2
Transmission Scheme	TM4 Closed Loop Spatial Multiplexing
Number of Spatial Layers	2
Number of Resource Blocks	25
Number of PDCCH symbols	1
LTE Bandwidth	5 MHz
Subframe Duration	1 ms
Cyclic Prefix	Normal
Fading Environment	8 tap Rayleigh with exponential PDP, EPA channel [3GPP, 2015d] with medium and low correlation
Doppler Frequency	0 Hz
Receiver Architecture	R-ML Parallel Interference Aware, R-ML Successive Interference Canceling
Channel Estimation	Perfect
MCS	$10 \leq \mathrm{MCS}_0 \leq \mathrm{MCS}_1 \leq 28$
Dual Layer TPMI	TPMI2
Single Layer TPMI	TPMI0, TPMI5, TPMI6
Target BLER	10^{-2}

with our settings can be thus computed as following:

$$T_{\text{tot, sim}}(R_0, R_1, \text{SNR}) = \sum_{r=0}^{r_{\max}-1} (T_{0r,\text{sim}}(R_0, R_1, \text{SNR}) + T_{1r,\text{sim}}(R_0, R_1, \text{SNR})), \tag{5.9}$$

with:

$$\begin{aligned} T_{0r,\text{sim}}(R_0, R_1, \text{SNR}) &= \frac{1}{r+1} R_0(1 - BLER_{0r}(R_0, R_1, \text{SNR})), \\ T_{1r,\text{sim}}(R_0, R_1, \text{SNR}) &= \frac{1}{r+1} R_1(1 - BLER_{1r}(R_0, R_1, \text{SNR})), \end{aligned} \tag{5.10}$$

where $r_{\max}$ is the maximum number of HARQ rounds that were performed. The term $T_{\text{TB}r}$ represents the throughput values for the transport block TB on the round r, R_0 and R_1 are the rates corresponding to MCS_0 and MCS_1, and $BLER_{0r}$ and $BLER_{1r}$ are the corresponding Block Error Rates. The fact that CW_1 is attempted for a decoding only if CW_0 is decoded is taken into account in BLER. The spectral efficiency declines with every retransmission round due to reuse of space-frequency resources by the same codeword. Therefore, the maximum contribution to the averaged long-term throughput is added during the first round.

In Section 4.5.2, we described a brute force approach that was used to identify the optimal combinations of the MCS. We now extend this method to the multiple HARQ rounds.

$$T^{\star}_{\text{tot,sim}}(\text{SNR}) = T_{\text{tot,sim}}(R^{\star}_0(\text{SNR}), R^{\star}_1(\text{SNR}), \text{SNR}), \tag{5.11}$$

where the optimal rates $R_0^\star$ and $R_1^\star$ corresponding to the $\mathrm{MCS}_0^\star$ and $\mathrm{MCS}_1^\star$ maximize the *empirical* system throughput $T_{\mathrm{tot,sim}}(R_0, R_1, \mathrm{SNR})$:

$$R_0^\star(\mathrm{SNR}), R_1^\star(\mathrm{SNR}) = \underset{R_0,R_1}{\operatorname{argmax}} T_{\mathrm{tot,sim}}(R_0, R_1, \mathrm{SNR}). \tag{5.12}$$

Fig. 5.3 illustrates the total system throughput $T^\star_{\mathrm{tot,\,sim}}$ after four HARQ rounds and the throughput $T_{00,\mathrm{sim}} + T_{10,\mathrm{sim}}$ after the first round achieved with $\mathrm{MCS}_0^\star$ and $\mathrm{MCS}_1^\star$ for SIC and PIA detection. Our SIC receiver outperforms the PIA receiver on the first round in all the channel models with gains varying from $2-4$ Mbit/s in high SNR regime to $5-7$ Mbit/s in low SNR regime. Multiple retransmission rounds reduce this gap, but the SIC receiver still performs better at high SNR (up to $2-3$ Mbit/s). The throughput contribution of the retransmission rounds becomes more significant with the increase of the channel correlation. However, the TPMI during a single codeword retransmission does not have a noticeable impact on the throughput. There is a slight preference to TPMI5 and TPMI6 in actual and outdated CSI scenarios over Alamouti coding in the frequency-selective Rayleigh channel, while in EPA channels there is no visible difference.

5.4.3 Reliability Analysis

In this section, we focus on the contribution of the multiple HARQ rounds to the communicational reliability of the MIMO system with our SIC receiver. The MCS optimization methodology applied in the previous section fits well to provide an initial idea about the throughput achievable with multiple HARQ rounds. However, it reflects poorly on the TPMI influence on the system performance in cases where a single codeword retransmission is used since it averages multiple channel realizations. For the following reliability analysis, we detach from the optimized $\mathrm{MCS}_0^\star$ and $\mathrm{MCS}_1^\star$ and consider a few particular MCS. To cover all the modulation orders, we select three meaningful MCS combinations: $(\mathrm{MCS}_0, \mathrm{MCS}_1) \in \{(12, 16), (16, 22), (20, 26)\}$.

The multiple retransmissions significantly contribute to the communicational reliability of the MIMO system in the lower SNR regime. Fig. 5.4 illustrates the BLER of the first transport block for four HARQ rounds.

Since the first stream is treated by the SIC and PIA receivers in the same way, both receivers are expected to show identical performance for the first transport block. At the BLER level of 10^{-1} (the Channel Quality Indicator target for LTE systems), the $(12, 16)$ MCS combination receives 10 dB and 13.5 dB improvements in the Rayleigh and the EPA channels respectively. The SNR gain increases for the higher values of MCS and reaches up to 13 dB in the Rayleigh channel 28 dB in the EPAM channel. However, these benefits come at price of a spectral efficiency degradation with every retransmission round since the space-frequency resources are reused by the same informational stream. Overall, the reliability is significantly suppressed with the increase of channel correlation.

The performance of the second codeword is tightly linked to the receiver architecture. We illustrate the analysis of the second stream with examples from the EPAM channel, as it shows a significant throughput contribution ($\approx$ 5Mbit/s) of the retransmission rounds (Fig. 5.3).

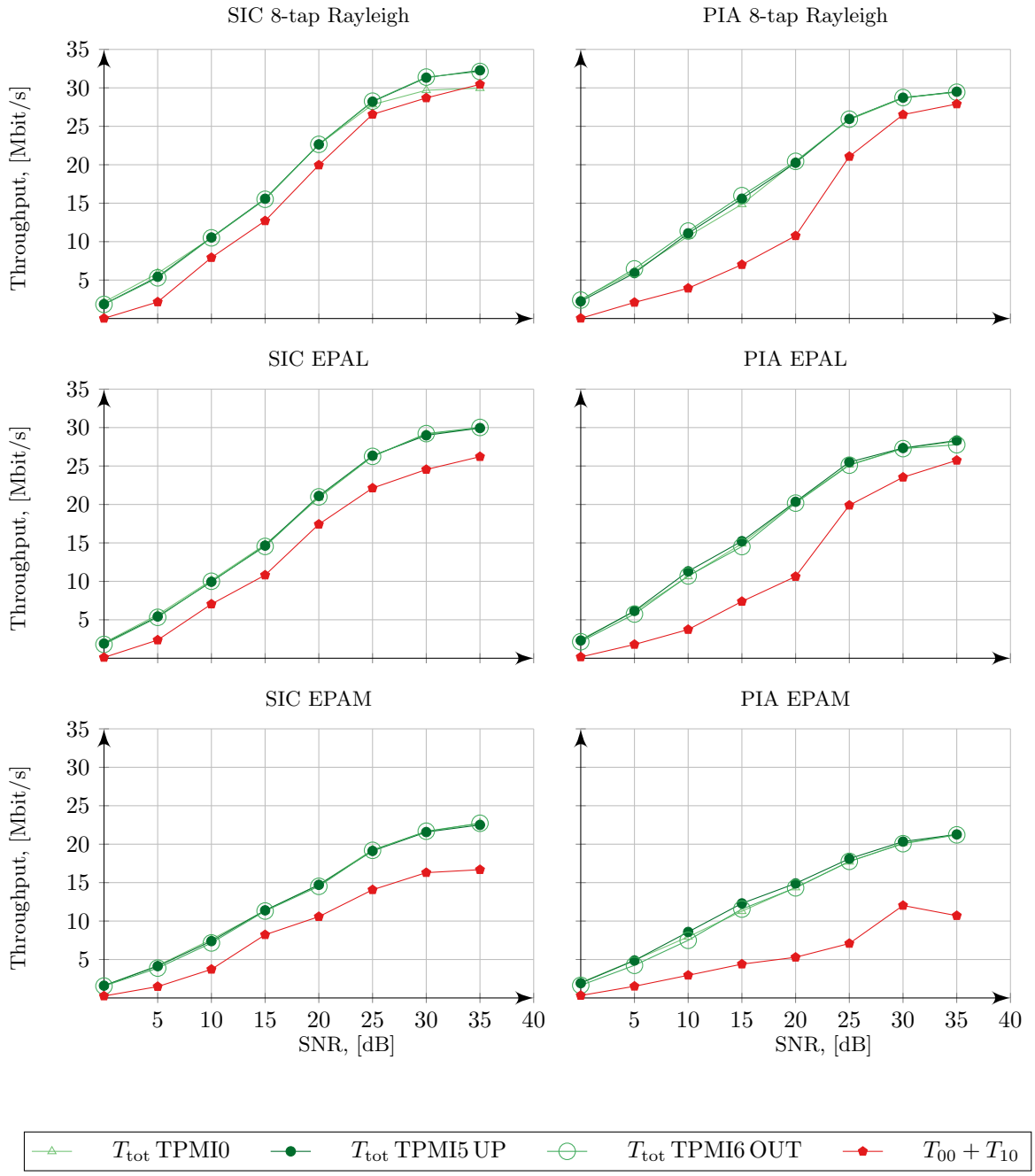


Figure 5.3: Comparison of SIC and PIA receivers' total throughput T_{tot} after 4 HARQ rounds and sum throughput $T_{00} + T_{10}$ for two codewords after the first round $r = 0$ applying *optimized* $\text{MCS}_0^\star$ and $\text{MCS}_1^\star$. We consider TPMI$\in \{0, 5, 6\}$ during the retransmissions of a single codeword. TPMI5 is studied in an updated CSI scenario, while TPMI6 is considered in an outdated CSI environment. The throughput $T_{00} + T_{10}$ after the first round is independent from the TPMI used during the retransmissions of a single codeword. The throughput plots for different TPMI values (depicted in shades of green) overlap each other, meaning that the performance does not significantly deviate with respect to the precoder used for the single codeword transmission.

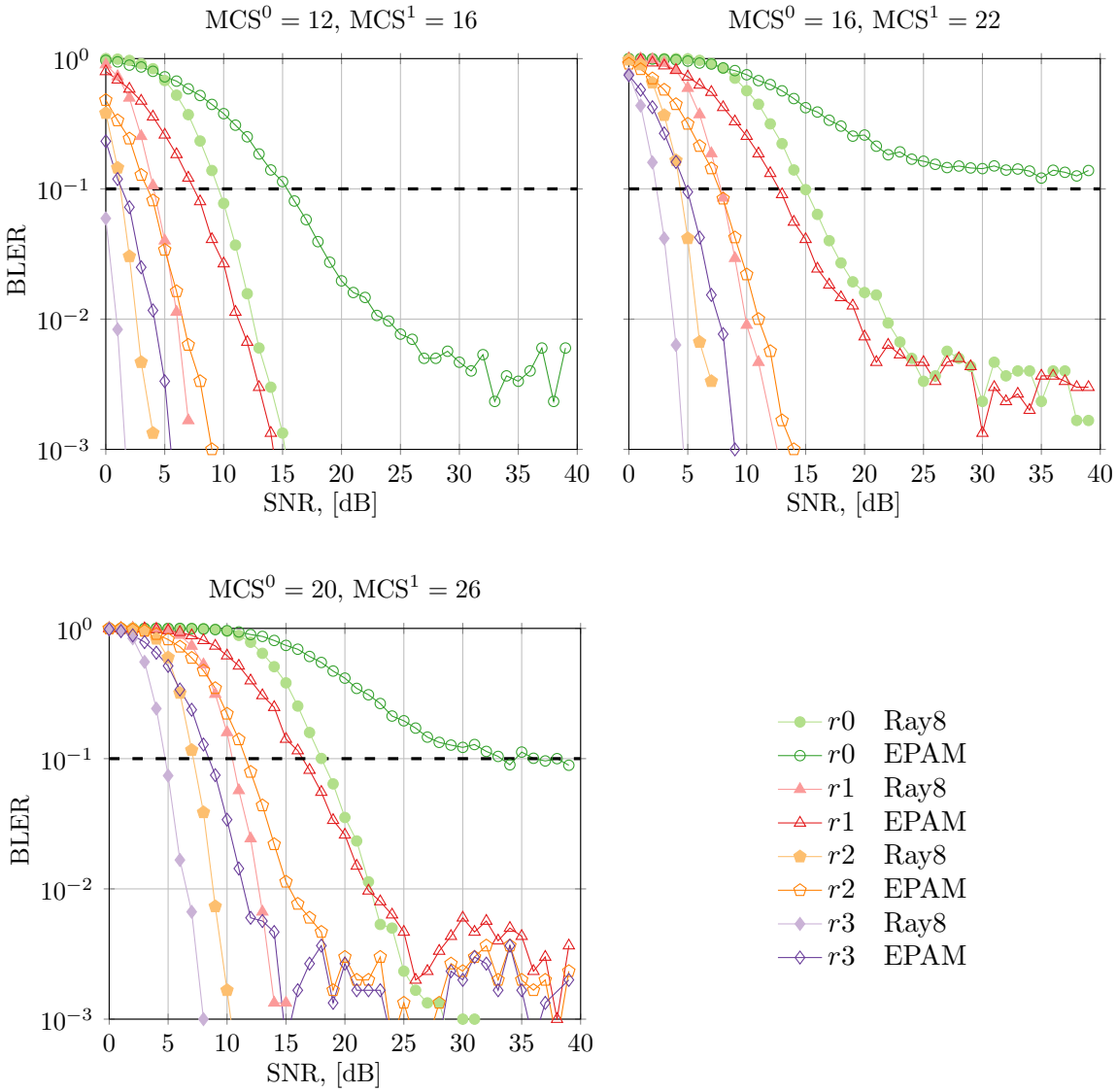


Figure 5.4: BLER of TB_0 in 8-tap Rayleigh and EPAM channels for 4 HARQ rounds (identical for SIC and PIA receiver) for $(\mathrm{MCS}_0, \mathrm{MCS}_1) \in \{(12, 16), (16, 22), (20, 26)\}$. The contribution of the every next round becomes more significant with the increase of the MCS values.

After the first codeword is decoded on round r_{dec}, the SIC receiver launches the multi-round SIC procedure. The decoded codeword is reconstructed with the corresponding redundancy version for the current and preceding r rounds, such that $0 \leq r \leq r_{\mathrm{dec}}$, and then subtracted from the Matched Filter outputs on every retransmission round. After subtraction, the LLR values are computed for the now interference-free second codeword with LLR values from preceding rounds. Following that, the decoding attempt is performed.

The number of the frames belonging to the second codeword that are passed to the

multi-round SIC procedure and that are decoded via it on round r is presented in Fig. 5.5. In the low SNR regime, a negligible amount of attempts was done at the initial transmission round due to the high BLER of the first codeword, shown in Fig. 5.4. A significant amount of decoding attempts was performed in the next rounds since the BLER of the first codeword is remarkably lower after retransmissions, even at low SNR. However, the majority of attempts failed, meaning that the gain from the LLR combining through multi-round SIC procedure is not sufficiently high to improve the performance in the low SNR regime. At moderate SNR level, about 50% of the attempts are decoded through the SIC procedure, while in high SNR level this value reaches 100%. We can conclude that our SIC receiver implementation benefits from multi-round LLR combining at moderate SNR and the majority of combining gain is coming from the second round, while gains are not that impressive in the low SNR regime.

At each retransmission round ($r > 0$), the total amount of retransmissions ret_r^{tot} is composed from the fraction of single codeword retransmissions ret_r^{single} and the fraction of two codewords retransmission ret_r^{multipl} as follows:

$$ret_r^{\text{single}} = ret_r^{\text{tot}} - ret_r^{\text{multipl}}. \tag{5.13}$$

Fig. 5.6 illustrates the amount of requests of single codeword retransmissions ret_r^{single} on round r sent by our SIC and PIA receivers in EPAM channel. In the low SNR regime the amount of single codeword retransmissions is almost identical, while in moderate and high SNR regimes the SIC receiver sends about 50% less requests of single codeword retransmissions ret_r^{single} thanks to the benefits of LLR combining via multi-round SIC procedure. This supports the idea that the main benefits of the SIC receiver are achieved during the first two transmission rounds.

It was shown that our SIC and PIA receivers demonstrate identical levels of reliability of communication for the first transport block. However, this is not the case for the second transport blocks since it receives a different treatment in our receivers. At the initial transmission round, our SIC receiver significantly outperforms PIA detection due to the SIC procedure. For example, it achieves 10% of errors at SNR level 10 dB lower than for PIA receiver for $(12, 16)$ MCS combination (Fig. 5.7). After the first round, the SNR level required to achieve less than 10% of errors is slightly lower for the SIC receiver compared to PIA detection. The consequent retransmission rounds bring significant benefits to the PIA detection, while for the SIC receiver gain between the third and the fourth retransmission round is not remarkable. The TPMI influence is more valuable than for the throughput evaluation: both receivers show preference for the Alamouti precoding. This means that the retransmission scheme that was employed by the Ericsson eNodeB for TM 3 during our drive test campaign is the optimal retransmission scheme for TM 4.

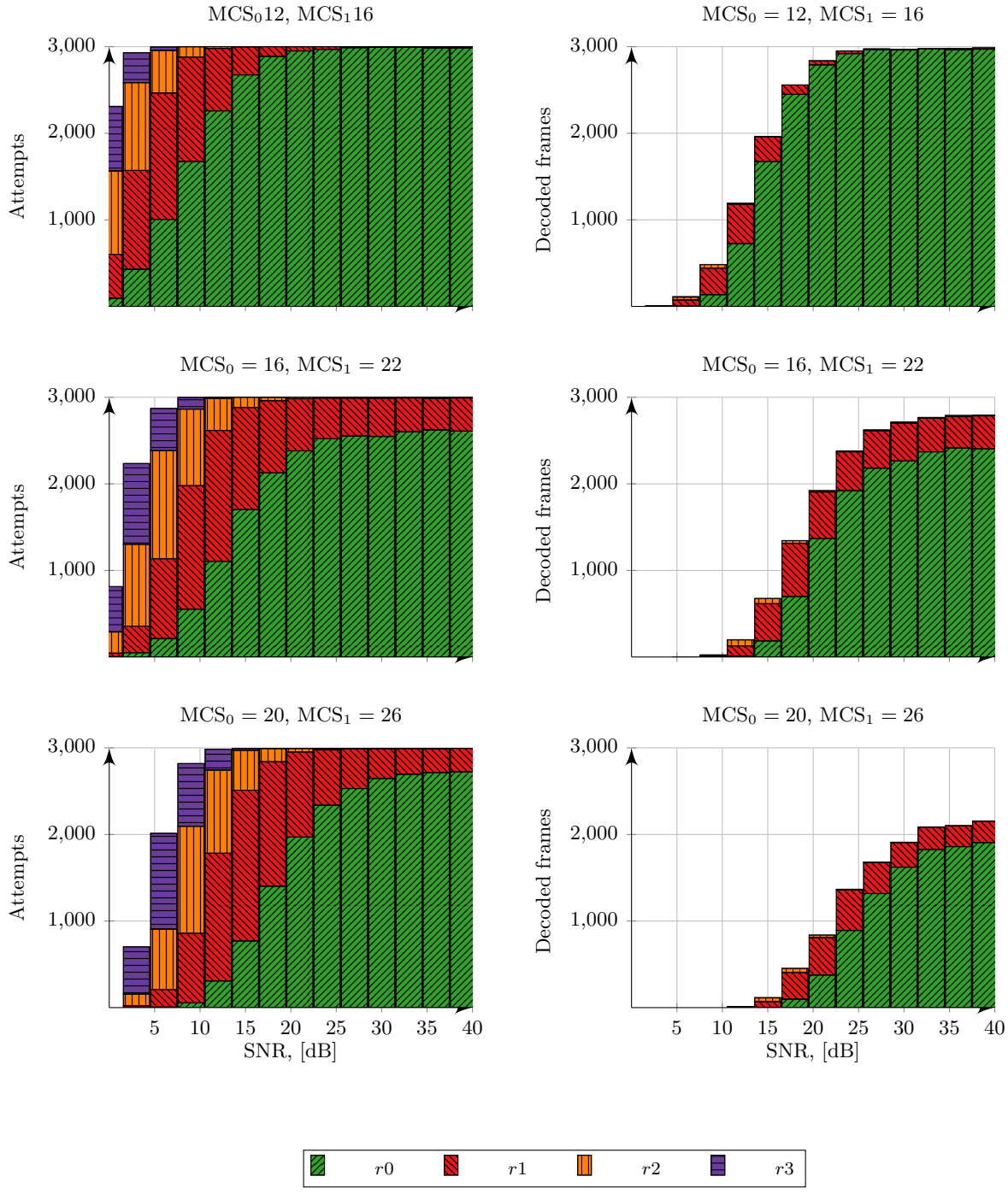


Figure 5.5: Amount of frames belonging to the second codeword that were attempted to be decoded (left column) and successfully decoded (right column) through multi-round SIC procedure on round r in EPAM channel. To cover all the modulation orders, we select three meaningful MCS combinations: $(\mathrm{MCS}_0, \mathrm{MCS}_1) \in \{(12, 16), (16, 22), (20, 26)\}$.

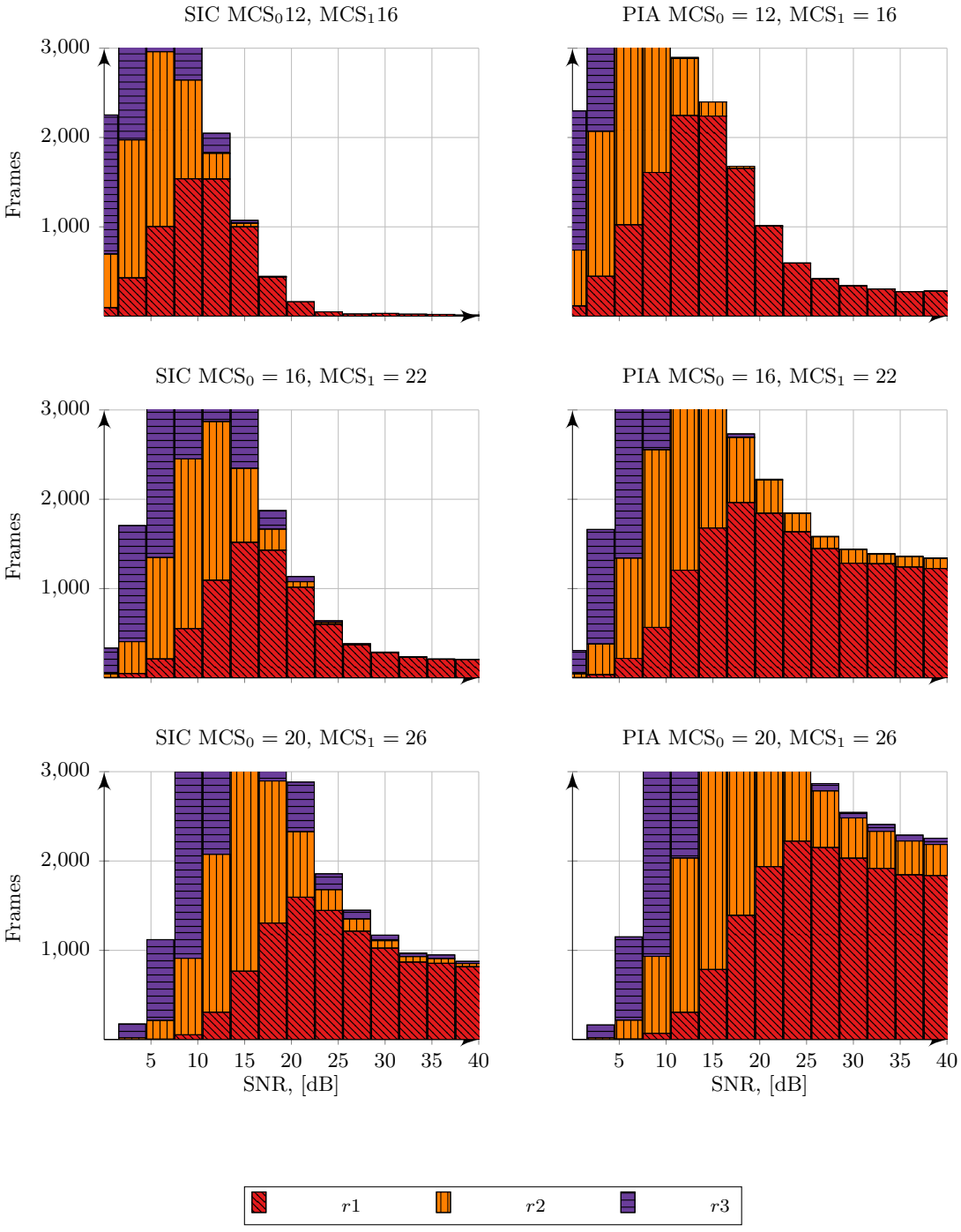


Figure 5.6: Amount of single codeword retransmissions ret_r^{single} on round r in EPAM channel for our SIC (left column) and PIA (right column) receivers for the MCS combinations, where $(\text{MCS}_0, \text{MCS}_1) \in \{(12, 16), (16, 22), (20, 26)\}$.

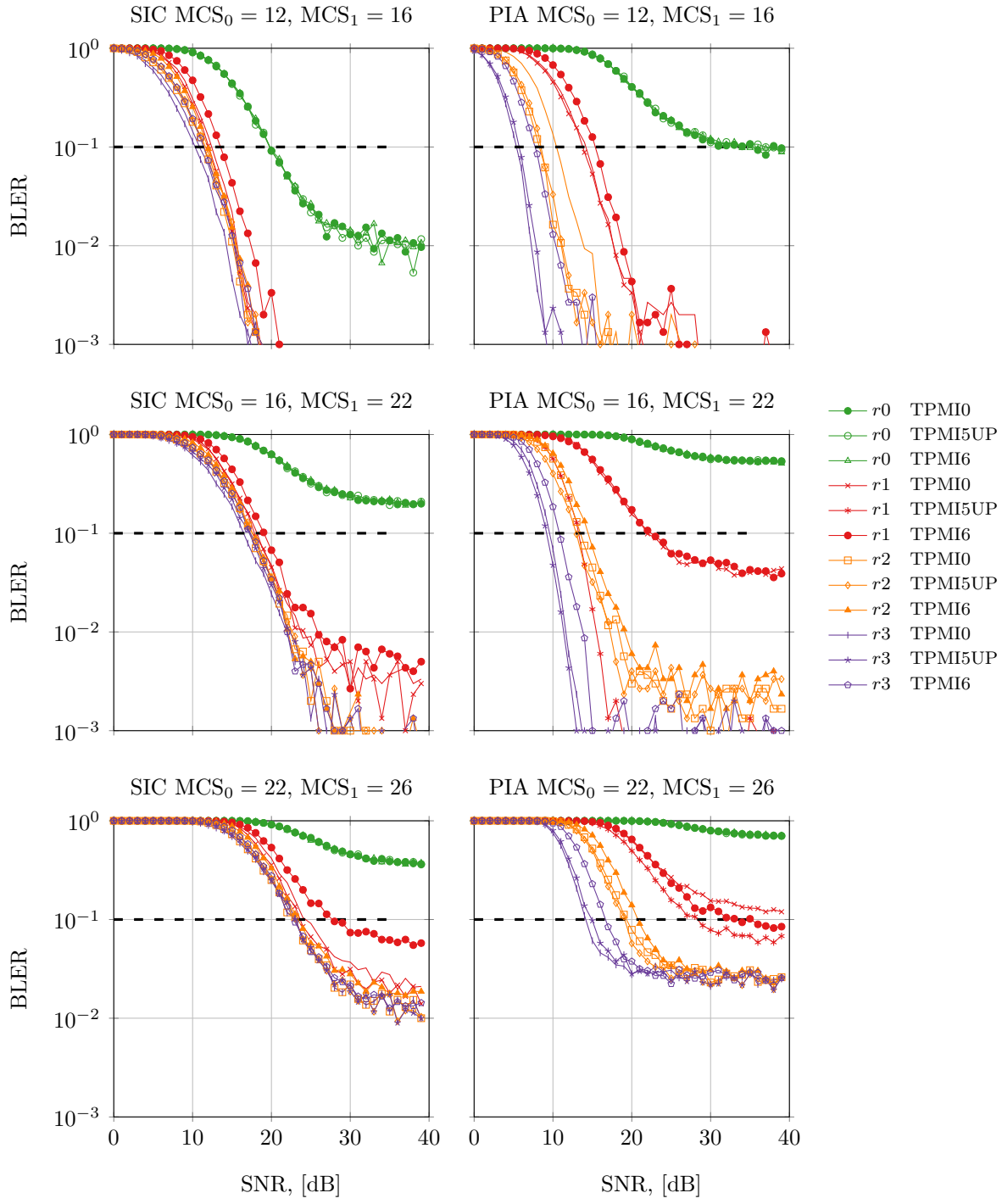


Figure 5.7: BLER of TB_1 in EPAM channel for 4 HARQ rounds for SIC (left column) and PIA (right column) receiver for TPMI=$\{0, 5, 6\}$. TPMI5 is studied in updated CSI scenario, while TPMI6 is considered in an outdated CSI environment; $(MCS_0, MCS_1) \in \{(12, 16), (16, 22), (20, 26)\}$.

5.4.4 A Note on Computational Effort

In Section 4.3.3, we have shown that our SIC receiver is 25% more time efficient than our PIA receiver given only one transmission round. For multiple rounds the SIC receiver sends fewer requests for single codeword retransmissions in high SNR regime, meaning that the overall processing time is further reduced. On the other hand, the SIC detection clearly requires higher memory consumption since we need to store the Matched Filter outputs and channel coefficients for each retransmission to be able to reconstruct the signal from preceding rounds.

5.5 Conclusion

Inspired by the drive test measurement campaign, we have presented an HARQ protocol methodology for our Reduced Complexity Parallel Interference Aware receiver and Successive Interference Aware Canceling receiver. The protocol was then implemented in our OAI downlink simulator LTE TM 4 setup. We quantified the throughput gains after multiple retransmission rounds in frequency selective environment and performed an analysis of the reliability of the downlink data traffic after multiple HARQ rounds.

The SIC receiver achieves higher throughput in all the fading scenarios. However, despite the LLR combining which is accessed through the multi-round SIC procedure, its gains are delivered by the first two rounds of transmissions, while the PIA receiver clearly benefits from all four retransmission rounds. For both receivers, the first transport block achieves huge performance improvements in the low SNR regime thanks to HARQ (up to 10 dB reduction to achieve BLER levels at 10^{-1}), since the price of spectral efficiency degradation with every retransmission round, as the space-frequency resources are reused by the same data stream.

The Ericsson eNodeB that was used to perform the drive tests was configured in TM 3 and applied Alamouti precoding as the only option allowed by the DCI format 2A for the retransmissions of the single codeword. Our analysis for the DCI format 2 showed that Alamouti precoding is favorable for the retransmissions of the single codeword in TM 4 from the reliability point of view, but there is no noticeable preference to any of the retransmission schemes from the throughput point of view.

Chapter 6

PHY Abstraction for the R-ML PIA and SIC Receivers

Physical layer abstraction provides an access to high speed link-level computations by quantifying the system performance through efficient techniques that rely on the transmission parameters and the channel conditions. The transmission parameters may include a transmission mode, modulation and coding schemes, and a redundancy version. The performance is evaluated in terms of *Block Error Rate* (BLER).

The simulation time is reduced by replacing the expensive procedures in the transmitter/receiver chain with mapping blocks capable of predicting the link level performance. In traditional LTE simulators, the full set of the physical layer procedures such as coding, modulation, convolution, demodulation and decoding occupies more than 80% of simulation time [Bilel et al., 2011], and is not affordable in terms of consumed time and CPU; the use of PHY abstraction may reduce computational time by a factor of 100 [Latif et al., 2013].

The second application of the physical layer abstraction is Link Adaptation: taking into account channel conditions, the UE calculates the Rank Indicator, the Precoder Matrix Indicator and the Channel Quality Indicator which maximize the instantaneous system throughput [Schwarz and Rupp, 2011].

The two widely studied techniques on physical layer abstraction are the Exponential Effective Signal-to-Noise Ratio Mapping (EESM) [Tuomaala and Wang, 2005] and the Mutual Information Effective Signal-to-Noise Ratio Mapping (MIESM) [Lei and Jin-bao, 2010; Sayana et al., 2007] methodologies. The EESM strategy appears light-weight and straight forward to implement, but the accuracy is not always satisfying. In contrast, the MIESM approach is more challenging to implement but promises higher accuracy.

To be eligible for real-time devises, the abstraction algorithms should satisfy certain requirements, such as low memory consumption, time efficiency, and feasible accuracy. Furthermore, it should be generalized for most common propagation scenarios and transmission settings, since the UE is not always capable of identifying the distribution of the channel coefficients. To switch dynamically between the calibration coefficients tuned for a particular value of the Modulation and Coding Scheme (MCS) in a particular channel model is also not always possible. In the state-of-art solutions, the researchers often em-

phasize on the accuracy, while the practical constrains do not receive sufficient attention.

Contributions In this chapter, we investigate the abstraction methodologies for Physical Downlink Shared Channel in Single-User Multiple-Input-Multiple-Output (MIMO) Long Term Evolution (LTE) systems using our Reduced Complexity Maximum Likelihood (R-ML) Parallel Interference Aware and Successive Interference Aware Canceling receivers, studied in Chapter 4. We develop and validate the MIESM methodology based on look-up tables (LUT) with levels of Mutual Information, initially proposed for multi-user MIMO by [Latif et al., 2012]. An analysis of look-up table quantization analysis is performed, and abstraction results are compared with equivalent results obtained from direct precise computations of the mutual information levels. We detail the advantages and drawbacks of the MEISM look-up table methodology, as well as the feasibility of its deployment for real-time transmissions. Finally, we apply a light-weight EESM methodology and approximate the per-stream performance of our receivers with the Minimum Mean Square Error (MMSE) receiver and the receiver with perfect interference canceling, and provide the analysis of the calibration coefficients.

Contents

6.1 State of the Art

Physical later abstraction is a powerful tool for quantifying system performance and can be used in link level simulations as well as in real time measurements. Furthermore, it serves as a part of the UE and eNodeB emulators in radio network performance prediction devices.

The two most studied PHY abstraction approaches are the EESM [Tuomaala and Wang, 2005], which is widely used for linear receivers, and the MIESM [IEEE, 2009; Lei and Jin-bao, 2010; Sayana et al., 2007], which reflects the nature of the ML-family of receivers. Both methods estimate the post-processed per-subcarrier Link Quality Metric (LQM) that can be represented by the SINR or mutual information level. The MIESM approach is proven to outperform EESM [Hanzaz and Schotten, 2013], but is lacking in terms of computation complexity [Galiotto et al., 2015]. Moreover, the EESM approach is

a weak choice in the presence of non-Gaussian interference [Latif et al., 2012; Stancanelli et al., 2011], since interference is — in this case — absorbed into Gaussian noise.

While post-processed SINR computations for the linear receivers are straightforward, it remains a formidable challenge for the non-linear ones, where joint detection is performed over all spatial layers. The post-processed SINR for a ML receiver can be estimated via bounding with upper- and lower-limits provided by the SINR of respectively Interference-free (IF) and MMSE receivers and applying calibration coefficients [Moon et al., 2012]. This approach was then extended to the R-ML IA case in [Lee et al., 2014], where the adjusting coefficients also depend on the interference strength. Another technique to estimate post-processed SINR for ML receiver through polynomial approximation was presented in [IEEE, 2009], but is not well adapted to MIMO and does not consider the interference-aware case. PHY abstraction for MIMO ML-receivers was developed in [Ramesh et al., 2009] based on QR and QL factorization of the channel matrix. The authors compute an upper and lower bound for the performance of each of the streams and then use the average value in order to characterize the achievable performance. To avoid the time consuming online mapping between channel gains and mutual information, [Latif et al., 2012] proposed to store channel statistics and the corresponding mutual information levels in a look-up table for a Multi-User MIMO system with R-ML IA receiver.

6.2 Scenario Description

We consider Transmission Mode 4 (TM 4) scenario with Closed Loop Spatial Multiplexing (CLSM) transmission. The base station is equipped with $n_{\text{tx}} = 2$ antennas and transmits two spatially multiplexed codewords to the UE with $n_{\text{rx}} = 2$ receive antennas. The codewords CW_0 and CW_1 belong to MCS_0 and MCS_1, with rates R_0 and R_1. We refer to the lower-rate R_0 codeword as CW_0, and CW_1 is always provided with an equal or higher rate R_1.

The received signal vector $\mathbf{y}_k$ for the k-th Resource Element observed by the UE is given by

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{P}_k\mathbf{x}_k + \mathbf{n}_k, \quad k = 1, 2, ..., K,$$

where $\mathbf{x}_k$ is the transmitted vector of two complex symbols x_0 and x_1 with variance of σ_0^2 and σ_1^2. The transmitted vector belongs to the $\mathcal{Q}^{M_0,M_1}$ alphabet, such that $\mathcal{Q}^{M_0,M_1} := \mathcal{Q}^{M_0} \times \mathcal{Q}^{M_1}$ is the Cartesian product of two modulation alphabets $\mathcal{Q}^{M_0}$ and $\mathcal{Q}^{M_1}$, and $M_0, M_1 \in \{2, 4, 6\}$ are the modulation orders of the QAM constellations. The vector $\mathbf{n}_k$ is Zero Mean Circularly Symmetric Complex Gaussian (ZMCSCG) white noise of double-sided power spectral density $N_0/2$ at two receive antennas of UE. The matrix $\mathbf{H}_k$ is a 2×2 channel matrix built with respect to one of the channel models, described in Section 3.2.1, and $\mathbf{P}_k$ is the precoding matrix employed by the eNodeB at the k-th Resource Element.

For the sake of simplicity, we drop the Resource Element index and replace the multiplication of $\mathbf{H}$ and $\mathbf{P}$ with the effective channel $\mathbf{H}_{\text{eff}}$:

$$\mathbf{y} = \mathbf{H}_{\text{eff}}\mathbf{x} + \mathbf{n}, \mathbf{H}_{\text{eff}} = [\mathbf{h}_{\text{eff0}}\ \mathbf{h}_{\text{eff1}}]. \tag{6.1}$$

Single-User MIMO multi-stream interference-aware detection falls into three groups: joint Maximum Likelihood detection, Parallel Interference Aware (PIA) detection and

Successive Interference Aware Canceling (SIC) detection. Due to high complexity of the conventional ML detection, we focus on PIA and SIC receivers. The receiver design was described in details in Chapter 4.

The signal processing starts with a linear Matched Filter (MF) operation, common for the both PIA and SIC receivers. The received signal $\mathbf{y}$ (6.1) is transformed into

$$\mathbf{y}_{\mathrm{MF}} = \mathbf{H}_{\mathrm{eff}}^{H}\mathbf{y}. \tag{6.2}$$

The effective compensated channels and the received vector after Matched Filtering can be obtained as follows:

$$\mathbf{y}_{\mathrm{MF}} = \boldsymbol{\alpha}x_0 + \boldsymbol{\gamma}x_1 + \begin{bmatrix} \mathbf{h}_{0\mathrm{eff}}^{H} \\ \mathbf{h}_{1\mathrm{eff}}^{H} \end{bmatrix} \mathbf{n}, \tag{6.3}$$

where the terms $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ stand for the vectors of the channel coefficients grouped in the following way:

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{0\mathrm{eff}}^{H}\mathbf{h}_{0\mathrm{eff}} \\ \mathbf{h}_{1\mathrm{eff}}^{H}\mathbf{h}_{0\mathrm{eff}} \end{bmatrix}, \tag{6.4}$$

$$\boldsymbol{\gamma} = \begin{bmatrix} \gamma_0 \\ \gamma_1 \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{0\mathrm{eff}}^{H}\mathbf{h}_{1\mathrm{eff}} \\ \mathbf{h}_{1\mathrm{eff}}^{H}\mathbf{h}_{1\mathrm{eff}} \end{bmatrix}, \tag{6.5}$$

For our SIC receiver, the term α_1 becomes zero due to subtraction of previously decoded and then reconstructed signal term $\mathbf{h}_{1\mathrm{eff}}^{H}\mathbf{h}_{0\mathrm{eff}}x_0$.

To understand the potential performance benefits that the R-ML SIC architecture might possess, we perform a comparative information theoretical analysis of our SIC and PIA receives.

6.3 An Information Theoretic Analysis of ML PIA and SIC Detection

A detailed information theoretic analysis of the SIC and PIA detection was presented in Section 4.4. We now briefly recall the fundamental insights that are required for physical layer abstraction.

Mutual Information between the transmitted and the received signals is a powerful instrument to quantify the performance of the MIMO system and is an informative measure of the quality of the radio link. The concept of mutual information can also be used to compare the performance of different detection mechanisms and receiver architectures.

Based on the chain rule, mutual information of joint ML detection equals the mutual information of the SIC receiver, while the PIA detection is characterized by lower levels of mutual information:

$$I_{\mathrm{ML}} = I_{\mathrm{SIC}} = I_{\mathrm{SIC}_0} + I_{\mathrm{SIC}_1} = \underbrace{I\left(X_0; \mathbf{Y}_{\mathrm{MF}} | \mathbf{H}_{\mathrm{eff}}\right)}_{\text{MI for the first CW}} + \underbrace{I\left(X_1; \mathbf{Y}_{\mathrm{MF}} | X_0, \mathbf{H}_{\mathrm{eff}}\right)}_{\text{MI for the second CW using SIC}}, \tag{6.6}$$

$$I_{\mathrm{ML}} \geq I_{\mathrm{PIA}} = I_{\mathrm{PIA}_0} + I_{\mathrm{PIA}_1} = \underbrace{I\left(X_0; \mathbf{Y}_{\mathrm{MF}} | \mathbf{H}_{\mathrm{eff}}\right)}_{\text{MI for the first CW}} + \underbrace{I\left(X_1; \mathbf{Y}_{\mathrm{MF}} | \mathbf{H}_{\mathrm{eff}}\right)}_{\text{MI for the second CW using PIA}} . \tag{6.7}$$

The performance of the interference-free (IF) receiver (or the receiver with perfect interference canceling) is an upper-bound for the performance of the the maximum-likelihood detection, and the MMSE receiver serves as a low-bound [Moon et al., 2012]. We extend this comparison to the SIC and PIA detection:

$$I_{\mathrm{MMSE}} \leq I_{\mathrm{PIA}} \leq I_{\mathrm{SIC}} = I_{\mathrm{ML}} = I_{\mathrm{IF}}.$$

Unfortunately, there is no closed-form expression to evaluate the mutual information level for discrete alphabets. Instead, it can be approximated using Monte-Carlo simulations over a large number of noise realizations. The simple expressions for the capacity for the IF and MMSE receivers are given by

$$C_{\mathrm{MMSE}} = \sum_{l=1}^{L} \log_2(1 + \mathrm{SINR}_{\mathrm{MMSE}_l}),$$
$$C_{\mathrm{IF}} = \sum_{l=1}^{L} \log_2(1 + \mathrm{SINR}_{\mathrm{IF}_l}),$$

where per-stream l corresponding SINR values can be computed as follows

$$\mathrm{SINR}_{\mathrm{MMSE}_l} = \frac{1}{[(\mathbf{I}_{n_{tx}} + \frac{1}{N_0 n_{\mathrm{tx}}} \mathbf{H}_{\mathrm{eff}}^H \mathbf{H}_{\mathrm{eff}})^{-1}]_{ll}} - 1, \tag{6.8}$$

$$\mathrm{SINR}_{\mathrm{IF}_l} = \frac{1}{N_0 n_{\mathrm{tx}}} \|\mathbf{h}_{\mathrm{eff,l}}\|^2. \tag{6.9}$$

We have shown that the SIC receiver performs closely to the optimum joint ML detection, and its performance is located inside the bounds of the IF receiver and the MMSE receiver. This setting-up phase will help us to build the required blocks for the MIESM and EESM abstraction techniques.

6.4 PHY Abstraction Methodologies

The physical layer abstraction aims to avoid expensive link-level computations. Instead, given some knowledge about the radio link (for example channel gains), it allows to efficiently estimate the system performance indicators, such as the reliability or the throughput by consulting the look-up tables with the performance statistics. We define the parameter that characterizes the radio link (SNR or mutual information) as the *Link Quality Metric* (LQM).

The subcarriers of the OFDM systems in frequency-selective fading experience different channel gains. However, it is not practically feasible to consider performance on per-subcarrier basis. Rather, performance should be averaged over a few resource blocks. Furthermore, the key transmission parameters for the link adaptation in the LTE systems,

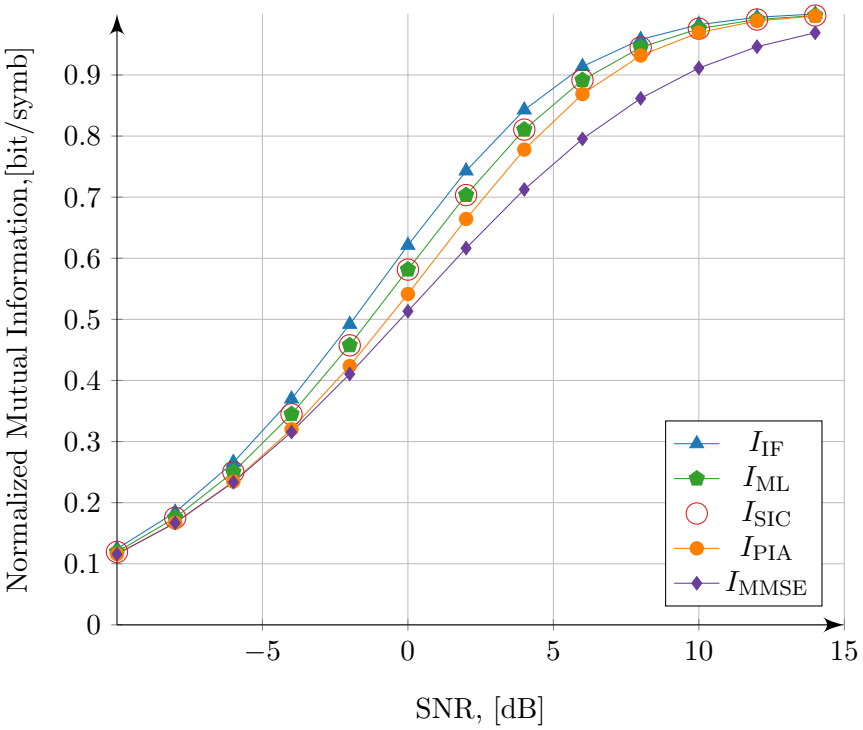


Figure 6.1: Normalized Mutual Information of the MIMO system with two codewords in Rayleigh channel for IF, joint ML, SIC, PIA and MMSE detection. The performances of the IF receiver and the MMSE receiver provide respectively the upper- and lower-bounds to the performance of the ML-based detection. The performance of our SIC receiver closely matches the ML joint detection, while our PIA receiver shows a slightly weaker performance.

such as the CQI, the PMI and the RI, have certain granularity and may vary from subband to subband or remain constant for the whole bandwidth dedicated to a particular user. This evokes the necessity for some sort of averaging of the per-subcarrier link quality metrics. Simple arithmetic averaging would not be able to capture the effects of the bit interleaving or how the channel gains vary inside the group of the radio elements, and thus does not provide an adequate performance [Nanda and Rege, 1998]. To circumvent this issue, the *mapping function* $\mathcal{I}_{\text{map}}$ was introduced. The purpose of the mapping function is at first to compress the multiple per-subcarrier link quality metrics LQM_m to a single effective value of "*information measure*"[Latif et al., 2012] or *compressed* link quality metric $\text{LQM}_{\text{compr}}$, and then to find a corresponding effective value SNR_{eff} using the *reverse mapping* $\mathcal{I}^{-1}_{\text{map}}$. This mapping process is the heart and the most challenging point of the physical layer abstraction, since it greatly influences the accuracy of the performance estimation. The EESM mapping and the MIESM mapping were proposed as potential techniques to capture the effect of frequency selectivity and have been widely studied in literature.

In fact, both MIESM and EESM abstraction methodologies can be described by the general Algorithm 12. although the link quality metrics and mapping functions applied in these two methods are different. The per-subcarrier link quality metric for EESM

Algorithm 12 General PDSCH Abstraction Algorithm

Input: MCS_0, MCS_1, $\mathbf{H}_{\text{eff}}$
Output: SNR_{eff}.

1: **for** each stream $0 \leq l \leq L$ **do**
2: **for** each subcarrier $k = 1, ...K$ **do**
3: compute an individual value of the $\text{LQM}_{l,k}$.
4: **end for**
5: Compress multiple $\text{LQM}_{l,k}$ values to a single value $\text{LQM}_{\text{compr},l}$ using the mapping function $\mathcal{I}_{\text{map}}$.
6: Map $\text{LQM}_{\text{compr}}$ to the effective $\text{SINR}_{\text{compr},l}$ using the reversed mapping function $\mathcal{I}_{\text{map}}^{-1}$.
7: Find the estimated $\text{BLER}_{\text{eff},l}$ corresponding to $\text{SINR}_{\text{eff},l}$ from the precomputed MCS-dependent AWGN curves.
8: Calibrate results using the adjustment MCS-dependent coefficients $\beta_{0,l}$ and $\beta_{1,l}$.
9: **end for**

mapping is represented by the instantaneous SNR_k, and the mapping function can be computed following the Chernoff union bound:

$$\mathcal{I}(\text{SNR}_k) = 1 - \exp(-\text{SNR}_k)$$

In the MIESM concept, the per-subcarrier link quality metric is represented by the instantaneous mutual information I_k. This technique is more complex than EESM since there is no closed form expression to compute the mutual information of a constrained alphabet, and the numerical approximation for high modulation orders via Monte-Carlo simulation would take a few hours for each SNR value. Various approaches have been proposed to circumvent this problem and to provide the efficient estimation of the mutual information, as it was discussed in Section 6.1.

We sought to extend the recently proposed for multi-user MIMO with interference aware receiver promising MIESM abstraction scheme, based on precomputed look-up tables, [Latif et al., 2012], and to adapt this approach to our Single-User MIMO SIC and PIA receivers. In the second part of this chapter, we will investigate an EESM approach as well.

6.5 Mutual Information Effective SNR Mapping

6.5.1 MIESM Methodology

The MIESM algorithm can be deducted from the general Algorithm 12 by straightforwardly replacing the link quality metric term with the mutual information, as shown in Algorithm 13.

The mapping process is the heart and the most challenging point of the abstraction methodology. Mapping can be done via the direct computations of the instantaneous per-subcarrier mutual information $I_{l,k}$, which would be very precise but is time and CPU consuming as it must be averaged across a large number of noise $\mathbf{z}$ realizations. The

Algorithm 13 MIESM PDSCH Abstraction Algorithm for PIA and SIC receivers

Input: MCS_0, MCS_1, $\mathbf{H}_{\text{eff}}$
Output: SNR_{eff}.

1: **for** each stream $0 \leq l \leq L$ **do**
2: **for** each subcarrier $k = 1, ...K$ **do**
3: compute an individual value of the $I_{l,k}$ using the mapping function $\mathcal{I}_{\text{map}}$.
4: **end for**
5: Compress multiple $I_{l,k}$ values into the single value $I_{\text{compr},l}$.
6: Map $I_{\text{compr},l}$ to the effective $\text{SNR}_{\text{eff},l}$ using the reversed mapping function $\mathcal{I}_{\text{map}}^{-1}$.
7: Find the estimated $\text{BLER}_{\text{eff},l}$ corresponding to $\text{SINR}_{\text{eff},l}$ from precomputed MCS-dependent AWGN curves.
8: Calibrate results using the adjustment MCS-dependent coefficients $\beta_{0,l}$ and $\beta_{1,l}$.
9: **end for**

mutual information levels for the first codeword of our SIC and PIA receiver are close and can be numerically approximated by (6.10). Applying the analysis performed in Section 6.3, the mutual information level for the SIC receiver can be obtained via (6.11). As the second stream of our PIA receiver is abstracted identically to the first stream, an expression for mutual information of the second codeword can be derived by a simple exchange of the symbols x_0 and x_1 in (6.10), resulting in (6.12).

$$I_{\text{SIC}_0} = I_{\text{PIA}_0} = I_0\left(X_0; \mathbf{Y}_{\text{MF}} | \mathbf{H}_{\text{eff}}, M_0, M_1, N_0\right) = \log M_0 - \frac{1}{M_0 M_1 N_n} \times \left(\sum_{x_0 \in \mathcal{Q}^{M_0}} \sum_{x_1 \in \mathcal{Q}^{M_1}} \sum_{n \in \mathcal{N}} \log \frac{\sum\limits_{x_0' \in \mathcal{Q}^{M_0}} \sum\limits_{x_1' \in \mathcal{Q}^{M_1}} \exp\left[-\frac{1}{N_0} \| \mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0' - \mathbf{h}_1 x_1' \|^2 \right]}{\sum\limits_{x_1'' \in \mathcal{Q}^{M_1}} \exp\left[-\frac{1}{N_0} \| \mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0 - \mathbf{h}_1 x_1'' \|^2 \right]} \right), \tag{6.10}$$

$$I_{\text{SIC}_1} = I_1\left(X_1; \mathbf{Y}_{\text{MF}} | X_0, \mathbf{H}_{\text{eff}}, M_0, M_1, N_0\right) = \log M_1 - \frac{1}{M_0 M_1 N_n} \times \left(\sum_{x_0 \in \mathcal{Q}^{M_0}} \sum_{x_1 \in \mathcal{Q}^{M_1}} \sum_{n \in \mathcal{N}} \log \frac{\sum\limits_{x_1' \in \mathcal{Q}^{M_1}} \exp\left[-\frac{1}{N_0} \| \mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0 - \mathbf{h}_1 x_1' \|^2 \right]}{\exp\left[-\frac{1}{N_0} \| \mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0 - \mathbf{h}_1 x_1 \|^2 \right]} \right), \tag{6.11}$$

$$I_{\text{PIA}_1} = I_1\left(X_1; \mathbf{Y}_{\text{MF}} | \mathbf{H}_{\text{eff}}, M_0, M_1, N_0\right) = \log M_1 - \frac{1}{M_0 M_1 N_n} \times \left(\sum_{x_0 \in \mathcal{Q}^{M_0}} \sum_{x_1 \in \mathcal{Q}^{M_1}} \sum_{n \in \mathcal{N}} \log \frac{\sum\limits_{x_0' \in \mathcal{Q}^{M_0}} \sum\limits_{x_1' \in \mathcal{Q}^{M_1}} \exp\left[-\frac{1}{N_0} \| \mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0' - \mathbf{h}_1 x_1' \|^2 \right]}{\sum\limits_{x_0'' \in \mathcal{Q}^{M_1}} \exp\left[-\frac{1}{N_0} \| \mathbf{y}_{\text{MF}} - \mathbf{h}_0 x_0'' - \mathbf{h}_1 x_1 \|^2 \right]} \right), \tag{6.12}$$

where the noise realization is drawn from a set $n \in \mathcal{N}$ with the maximum number of channel realizations $|\mathcal{N}| = N_n$.

In our approach, we propose to utilize precomputed look-up tables: channel statistics $||\boldsymbol{\alpha}_k||$, $||\boldsymbol{\gamma}_k||$, N_0 and corresponding $I_{l,k}$, can be obtained using (6.10), (6.11), (6.12) via Monte-Carlo simulations for a large number of channel realizations and noise variances, and then stored as multi-dimensional matrices:

$$
\begin{aligned}
I_{\mathrm{LUT}_{\mathrm{SIC}_0,k}}(||\boldsymbol{\alpha}_k||, ||\boldsymbol{\gamma}_k||, M_0, M_1, N_0) &= I\left(X_0; \mathbf{Y}_{\mathrm{MF}} \middle| \boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k, M_0, M_1, N_0\right), \\
I_{\mathrm{LUT}_{\mathrm{PIA}_0,k}}(||\boldsymbol{\alpha}_k||, ||\boldsymbol{\gamma}_k||, M_0, M_1, N_0) &= I\left(X_0; \mathbf{Y}_{\mathrm{MF}} \middle| \boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k, M_0, M_1, N_0\right), \\
I_{\mathrm{LUT}_{\mathrm{SIC}_1,k}}(||\boldsymbol{\alpha}_k||, ||\boldsymbol{\gamma}_k||, M_0, M_1, N_0) &= I\left(X_1; \mathbf{Y}_{\mathrm{MF}} \middle| X_0, \boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k, M_0, M_1, N_0\right), \\
I_{\mathrm{LUT}_{\mathrm{PIA}_1,k}}(||\boldsymbol{\alpha}_k||, ||\boldsymbol{\gamma}_k||, M_0, M_1, N_0) &= I\left(X_1; \mathbf{Y}_{\mathrm{MF}} \middle| \boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k, M_0, M_1, N_0\right).
\end{aligned}
$$

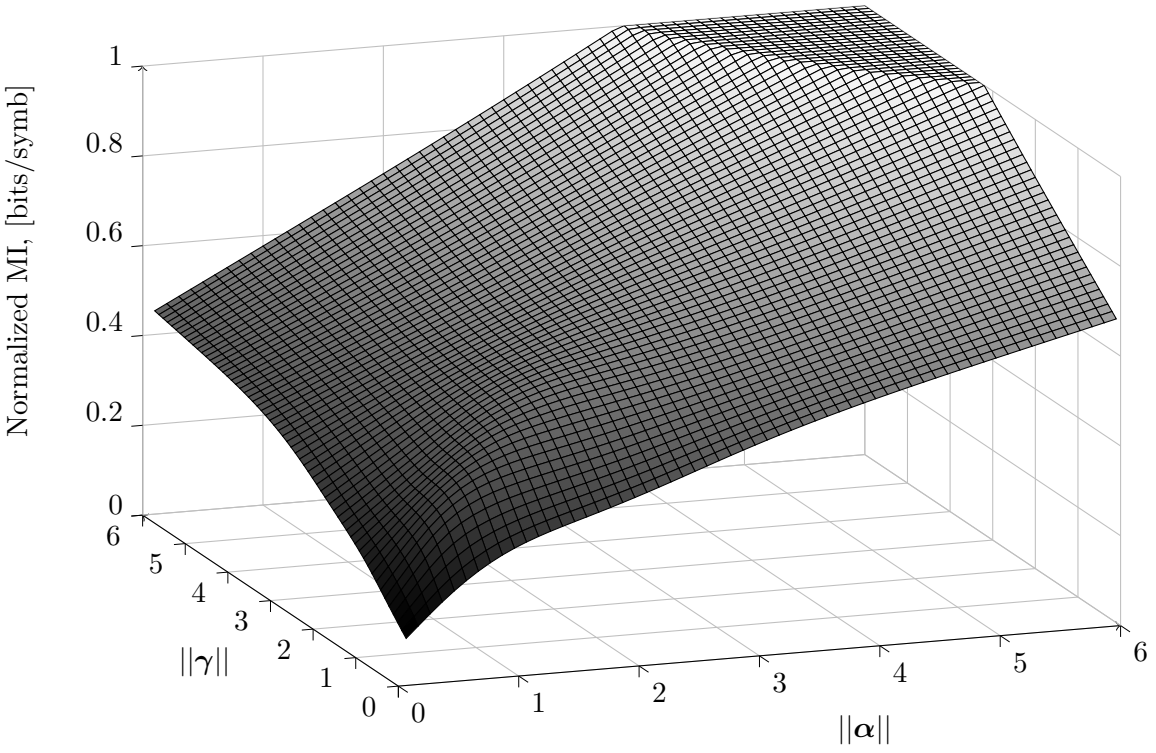


Figure 6.2: Normalized Mutual Information of the first stream $I_{\mathrm{LUT}_{l,k}} = \mathcal{F}(\boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k, N_0)$ taken at SNR slice 2 dB, when both codewords are mapped onto 16QAM. The horizontal x and y axes correspond to the norms of the channel columns $||\boldsymbol{\alpha}_k||$ and $||\boldsymbol{\gamma}_k||$.

The look-up table building strategy based on $||\boldsymbol{\alpha}_k||$, $||\boldsymbol{\gamma}_k||$ is motivated by the implementation complexity. We construct a gridded surface $I_{\mathrm{LUT}_{l,k}} = \mathcal{F}(||\boldsymbol{\alpha}_k||, ||\boldsymbol{\gamma}_k||, N_0)$ (Fig. 6.2) from the scattered set of the per-subcarrier mutual information values $I_{l,k}$. The surface is built with the help of data griding, interpolation and surface fitting tools,

which is not trivial when $\boldsymbol{\alpha}_k$ and $\boldsymbol{\gamma}_k$ are two-dimensional complex vectors. It could be possible to separate phase and amplitude for each of the channel coefficients, but this would lead to an unfeasibly high dimensionality of the look-up table.

The weak point of our look-up tables with $||\boldsymbol{\alpha}_k||$, $||\boldsymbol{\gamma}_k||$ is the loss of the channel phase information. To evaluate accuracy degradation in terms of Minimum Square Error (MSE), the results of the abstraction are compared with the direct computation method using (6.10), (6.11), (6.12) for a few modulation and coding schemes.

For systems with Bit Interleaved Coded Modulation, mutual information levels are limited by the modulation order M of a 2^M QAM constellation and saturate at different SNR or SINR values for different values M. Thus, a library of look-up tables is needed to take into account different combinations of modulation orders for both codewords and both receives.

For each stream l, the per-subcarrier values of $I_{\mathrm{LUT}_{l,k}}$ obtained from the look-up tables with the statistics for the stream and the receiver of interest are averaged, resulting in the compressed value $I_{\mathrm{compr},l}$:

$$I_{\mathrm{compr},l} = \frac{\sum_{k=1}^{K} I_{\mathrm{LUT}_{l,k}}}{\beta_{0,l} K}, \tag{6.13}$$

where $\beta_{0,l}$ is the first adjustment factor to compensate for modulation and coding rate.

To obtain the effective SNR, we seek a direct relation between $I_{\mathrm{compr},l}$ and $\mathrm{SINR}_{\mathrm{eff},l}$. We assume that there exists an equivalent one-tap SISO channel with a signal model (6.14) and averaged mutual information $I_{\mathrm{compr},l}$ (6.13):

$$\widetilde{y}_k = \widetilde{h}_k \widetilde{x}_k + \widetilde{n}_k, \tag{6.14}$$

where $\widetilde{x}_k \in \mathcal{Q}^M$ is a received complex symbol with variance of $\widetilde{\sigma}^2$, $\mathcal{Q}^M$ is a modulation alphabet of order $M \in \{2, 4, 6\}$, and the term $\widetilde{h}_k$ stands for the flat fading Rayleigh SISO channel with i.i.d. ZMCSCG random variables with a variance of 0.5 per dimension and $\widetilde{n}_k$ is ZMCSCG noise of double-sided power spectral density $N_0/2$. Then a one-to-one mapping between $I_{\mathrm{compr},l}$ and $\mathrm{SINR}_{\mathrm{eff},l}$ can be obtained via linear interpolation of (6.15) for a known value of $I_{\mathrm{compr},l}$. The mutual information for the single tap Rayleigh channel is computed offline using for all modulation alphabets and stored in the look-up tables.

$$I\left(X;Y|SNR,M\right) = \log M - \frac{1}{MN_hN_n}\left(\sum_{x\in\mathcal{Q}^M}\sum_{c}^{N_h}\sum_{z}^{N_n}\log\frac{\sum_{x\in\mathcal{Q}^M}\exp(-SNR\ |y-hx'|^2)}{\exp(-|n|^2)}\right), \tag{6.15}$$

where $SNR = \beta_{1,l}\mathrm{SINR}_{\mathrm{eff},l}$ and $\beta_{1,l}$ is the second adjustment factor to compensate for modulation and coding rate.

Assuming that effective SNR in a fading channel results in the same BLER as it would result in an AWGN channel, we have

$$BLER_l(\mathbf{H}_{\mathrm{eff}}, N_0, \mathrm{MCS}_l) \approxeq BLER_{\mathrm{AWGN}}(\beta_{1,l}\mathrm{SINR}_{\mathrm{eff},l}, \mathrm{MCS}_l). \tag{6.16}$$

It is important that the AWGN curves are precomputed for the full range of MCS and correspond to TM 4 since a various amount of Resource Elements is used for different transmission modes, resulting in a rate-dependent shift of the AWGN curves.

The calibration of the adjustment factors $\beta_{0,l}$ and $\beta_{1,l}$ is an important step of the physical layer abstraction validation. The detailed analysis of the calibration approaches has been done in [Brueninghaus et al., 2005]; the best fitting results are shown using logarithmic scale. The MSE between $\mathrm{SINR}_{\mathrm{eff},l}$ and $\mathrm{SINR}_{\mathrm{AWGN},l}$ is an adequate criteria for the training:

$$\beta_{\mathrm{opt}_0,l}, \beta_{\mathrm{opt}_1,l} = \underset{\beta_{0,l},\beta_{1,l}}{\arg\min} \frac{1}{N_H N_n} \left[\sum_{c}^{N_H} \sum_{z}^{N_n} |\mathrm{SINR}_{\mathrm{eff},l}(\beta_{0,l}, \beta_{1,l}, \mathrm{MCS}_l) - \mathrm{SINR}_{\mathrm{AWGN},l}(\mathrm{MCS}_l)|^2 \right]. \tag{6.17}$$

6.5.2 Results of the MIESM Approach

The traces used in the abstraction training were obtained via our OpenAirInterface downlink simulator `dlsim` for an LTE bandwidth of 5 MHz. The important part of the simulations was the careful generation of the input link level data for AWGN and frequency-selective channels. For the AWGN simulation, the channel was generated with the help of the spatial correlation matrix which nullifies cross-layer interference, and 10000 packets were transmitted. For the frequency-selective simulation, the 8-tap Rayleigh fading channel with i.i.d. entries and a delay spread of 0.8 microseconds was chosen, and 1000 packets were transmitted over 200 channel realizations for the wide range of noise variances targeting BLER of 10^{-2}. The AWGN and the frequency selective channel simulations were both carried out with the perfect channel estimation at the receiver. These traces were then used to obtain the calibration coefficients that are stored for each MCS and can be used for any random channel realization with an accuracy provided in Table 6.1, Table 6.2, Table 6.3.

Table 6.1: Calibration results for MIESM LUT abstraction of PIA receiver in 8-tap Rayleigh channel

MCS_0	MCS_1	$\mathrm{MSE}_{\mathrm{LUT},0}$	$\mathrm{MSE}_{\mathrm{LUT},1}$	$\beta_{\mathrm{LUTopt}_0,0}$	$\beta_{\mathrm{LUTopt}_0,1}$	$\beta_{\mathrm{LUTopt}_1,0}$	$\beta_{\mathrm{LUTopt}_1,1}$
2	2	0.0334	0.0544	8.5085	20.8125	0.2814	0.2485
4	4	0.0508	0.0571	2.9794	10.6384	0.2433	0.1608
4	12	0.0369	0.6021	11.2565	0.7811	0.1177	0.6284
10	10	0.3781	0.5707	0.6762	0.6772	0.4860	0.4315
12	12	0.8729	0.6986	0.8121	1.0228	0.6179	0.5991

The results of the MIESM methodologies are demonstrated in Fig. 6.3 and Fig. 6.4 for our PIA and SIC receiver respectively. The closer the Rayleigh curves (plotted in orange and green for the first and second codewords respectively) are to the corresponding AWGN curve (shown in red and blue), the more accurate is the abstraction. The abstracted performance for the high modulation orders does not match the AWGN curves as closely as for the lower modulations. This results from the loss of phase information during the look-up tables construction stage, which becomes significantly more critical with the increase

Table 6.2: Calibration results for MIESM LUT abstraction of SIC receiver in 8-tap Rayleigh channel

MCS_0	MCS_1	$MSE_{LUT,0}$	$MSE_{LUT,1}$	$\beta_{LUTopt_0,0}$	$\beta_{LUTopt_0,1}$	$\beta_{LUTopt_1,0}$	$\beta_{LUTopt_1,1}$
4	4	0.0533	0.0926	2.6906	18.2130	0.2634	0.1176
10	10	0.3791	0.8706	0.7201	0.8827	0.4698	0.4274
12	12	0.7480	1.1566	0.9950	0.9311	0.5356	0.5275

Table 6.3: Calibration results for the direct MIESM mapping abstraction methodology for PIA receiver in 8-tap Rayleigh channel

MCS_0	MCS_1	$MSE_{direct,0}$	$MSE_{direct,1}$	$\beta_{directopt_0,0}$	$\beta_{directopt_0,1}$	$\beta_{directopt_1,0}$	$\beta_{directopt_1,1}$
2	2	0.0364	0.0409	8.2438	4.9157	0.2300	0.2731
4	4	0.0217	0.0300	2.4298	2.3646	0.1839	0.1870
12	12	0.0904	0.1297	0.6829	0.7420	0.5465	0.5692

in the modulation. To support this reasoning, the calibration precision of the look-up table based method for QPSK constellation is verified with the direct MIESM mapping at the stage of obtaining the link quality metric per subcarrier. From a comparative analysis of Table 6.1, Table 6.3, a good match between these two methods is observed for QPSK constellation. However, this is not the case for higher modulation orders. Nevertheless, in the light of link adaptation the introduced accuracy satisfies the granularity of the transmission parameters, such as CQI.

In theory, the abstracted performance for the first codewords of our SIC receiver and PIA receiver should be similar. In practice, this is not always the case and the accuracy might vary even though the same look-up table was used, since $I_{SIC_0} = I_{PIA_0}$ (6.10). This is linked to the fact that our receivers are implemented in the fixed point and rounding errors are unavoidable. The results are expected to improve if the amount of the channel realizations used as input to the abstraction validation is increased, as well as if the look-up tables are built with higher precision.

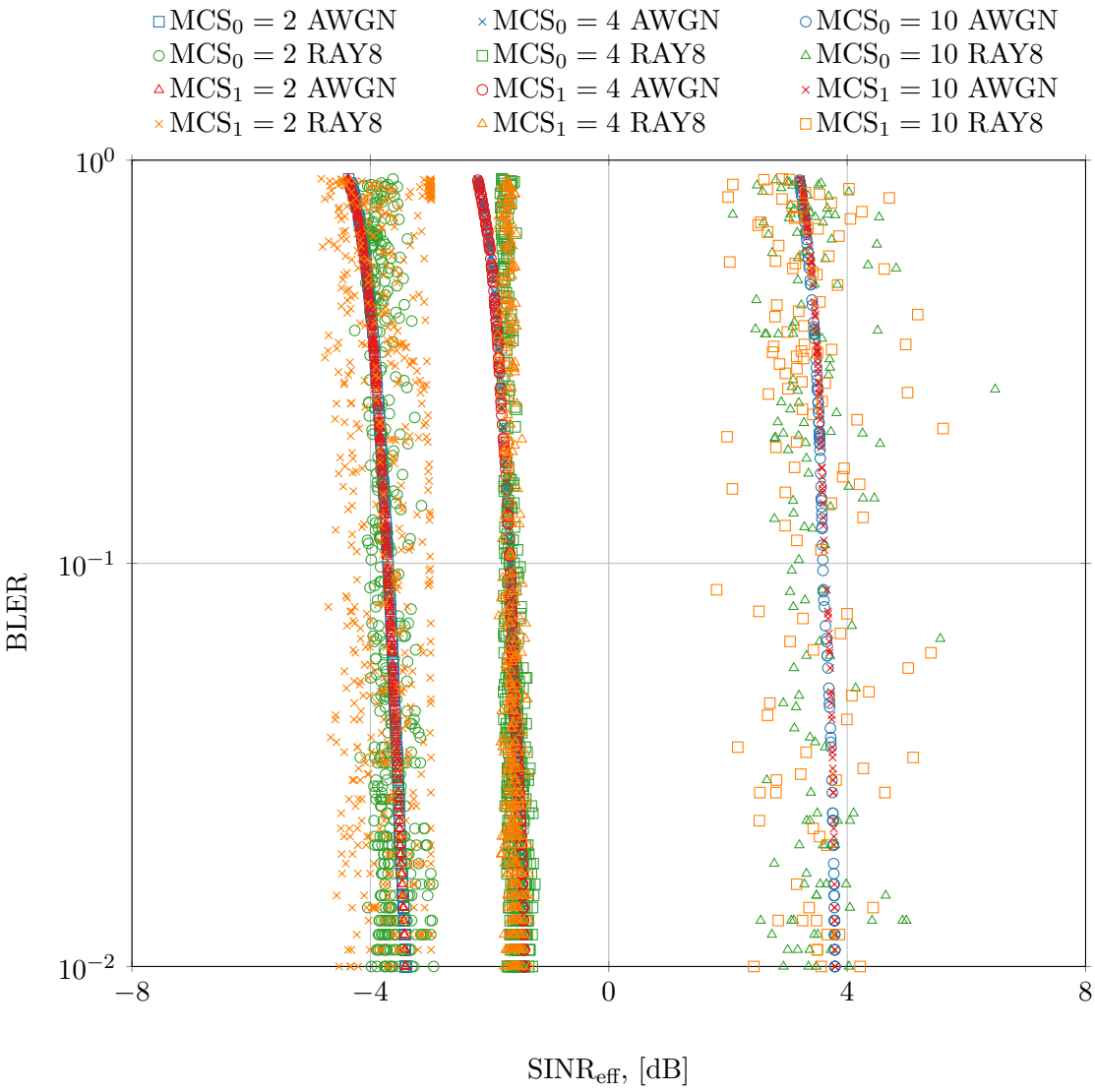


Figure 6.3: MIESM validation for our PIA receiver employing $(\mathrm{MCS}_0, \mathrm{MCS}_1) \in \{(2,2),(4,4),(10,10)\}$ in 8-tap Rayleigh channel. The calibration factors are taken into account. The closer the curves for the Rayleigh channel are to the corresponding AWGN values, the more accurate is the abstraction. As it can be seen, the accuracy degrades with the increase of the MCS values. This results from the loss of the phase information, which becomes significantly more critical with increasing transmission rate.

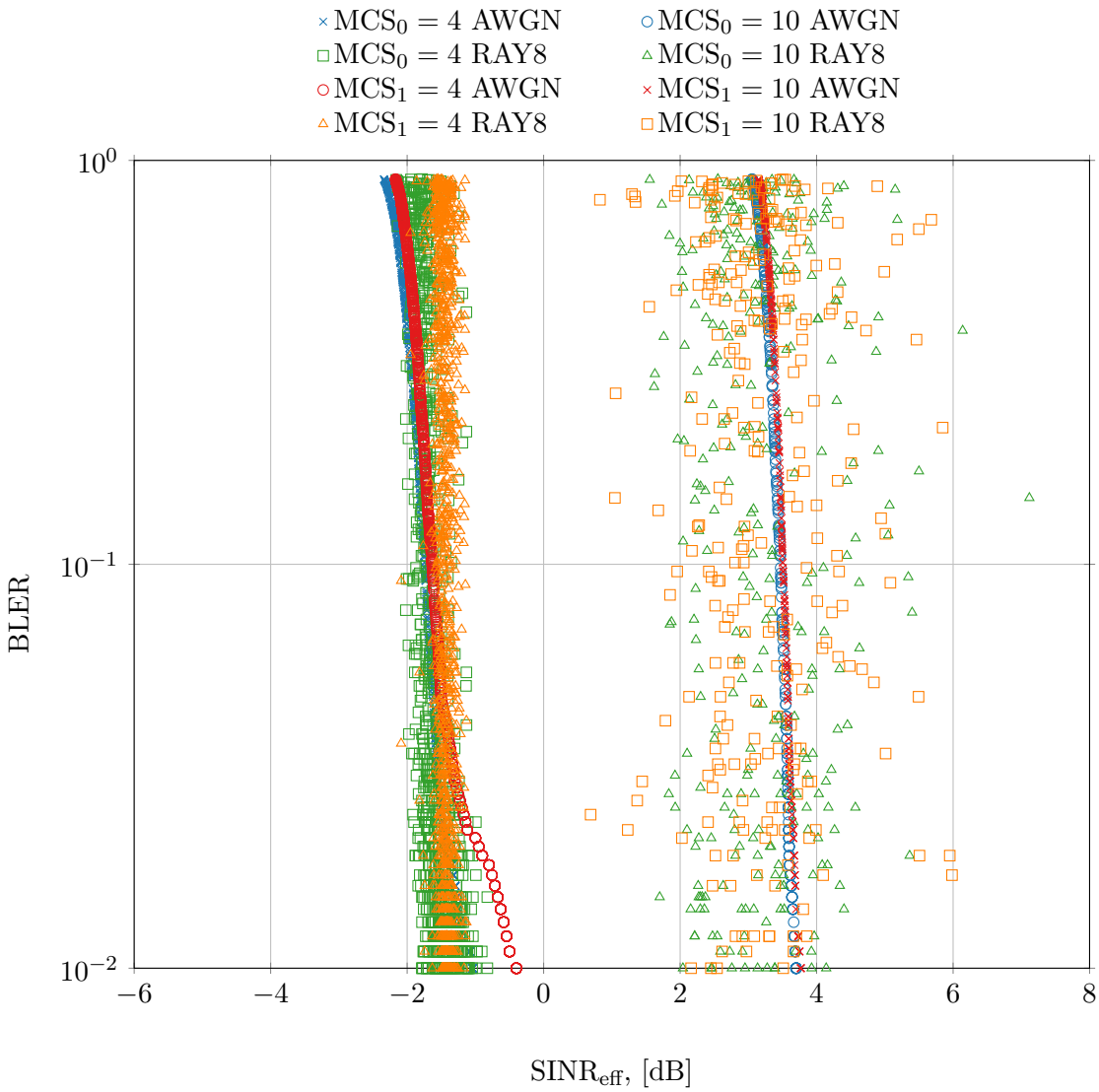


Figure 6.4: MIESM validation for our SIC receiver employing $(\text{MCS}_0, \text{MCS}_1) \in \{(4,4),(10,10)\}$ in 8-tap Rayleigh channel. The calibration factors are taken into account. Similar to the MIESM abstraction for the PIA receiver, the abstracted performance of our SIC receiver for the high modulation orders does not match the AWGN curves as well as for the lower modulations.

6.6 Exponential Effective SNR mapping

6.6.1 EESM Methodology

The state of the art literature dedicated to the MIESM abstraction has delivered a great contribution to the accuracy of the performance estimation, allowing for the perfect matches of the abstracted effective performance with the AWGN curves [Kim et al., 2010; Lei and Jin-bao, 2010; Olmos et al., 2010]. However, it is often associated with heavy computations which require a lot of offline calibration or look-up tables of unrealistic sizes [Latif et al., 2012], or exceed time limitations of a real time transmission. As seen from our experiments in Section 6.5, the MIESM approach greatly depends on the receiver implementation, fixed-point or float-point architecture, look-up table granularity and channel models, as the calibration coefficients might significantly vary from one fading environment to another [Chen et al., 2011a].

It is not exactly certain which techniques are used in the chipsets of real-time devices, but those most certainly are some very low-complexity techniques that yet guarantee satisfying levels of estimation. The EESM technique is proven to be a light-weight solution, although it loses in accuracy to the MIESM mapping [Brueninghaus et al., 2005]. Furthermore, it is often emphasized that the EESM mapping is not the optimum solution for the ML-like receivers since it is challenging to estimate the post-processed SINR_k in this case [Kim et al., 2010; Ramesh et al., 2009]. We do not claim that the EESM abstraction ideally reflects the maximum-likelihood or interference-aware nature of the receiver. Instead, we show that even the simple EESM architecture provides sufficient accuracy to be deployed in real-time systems and to ensure the efficient performance prediction. The EESM algorithm (Algorithm 14) follows the same main steps as MIESM algorithm (Algorithm 13), but requires a simple mapping function to compress the post-detection per-subcarrier SINR, and to map the compressed value to the effective SINR.

Algorithm 14 EESM PDSCH Abstraction Algorithm for PIA and SIC receivers

Input: MCS_0, MCS_1, $\mathbf{H}_{\text{eff}}$
Output: SNR_{eff}.

1: **for** each stream $0 \leq l \leq 1$ **do**
2: **for** each subcarrier $k = 1, ...K$ **do**
3: compute an individual value of the $\text{SINR}_{l,k}$ using the function $\mathcal{S}$.
4: **end for**
5: Compress multiple $\text{SINR}_{l,k}$ values to a single value $\text{SINR}_{\text{compr},l}$ using mapping function $\mathcal{I}_{\text{map}}$.
6: Map $\text{SINR}_{\text{compr},l}$ to the effective $\text{SINR}_{\text{eff},l}$ using the reversed mapping function $\mathcal{I}^{-1}_{\text{map}}$.
7: Find the estimated $\text{BLER}_{\text{eff},l}$ corresponding to $\text{SINR}_{\text{eff},l}$ from precomputed MCS-dependent AWGN curves.
8: Calibrate the results using the adjustment MCS-dependent coefficients $\beta_{0,l}$ and $\beta_{1,l}$.
9: **end for**

The post-detection SINR per-subcarrier is a function of the channel gains and noise power:

$$\mathrm{SINR}_k = \mathcal{S}(\mathbf{H}_k, N_0), \tag{6.18}$$

where the function $\mathcal{S}[.]$ is conditioned on the signal processing applied at the transmitter and the receiver [Bjerke et al., 2005]. We approximate the interference-aware streams of our receivers by the MMSE receiver ((6.19)), and the interference-free stream by the receiver with perfect interference canceling ((6.20)):

$$\mathrm{SINR}_{l,k} = \frac{1}{[(\mathbf{I}_{n_{tx}} + \frac{1}{N_0 n_{\mathrm{tx}}}\mathbf{H}_{\mathrm{eff},k}^H \mathbf{H}_{\mathrm{eff},k})^{-1}]_{ll}} - 1, \tag{6.19}$$

$$\mathrm{SINR}_{\mathrm{IF}_{\mathrm{sic1}}} = \frac{1}{N_0 n_{\mathrm{tx}}} \|\mathbf{h}_{\mathrm{eff1},k}\|^2. \tag{6.20}$$

Following the EESM mapping, the effective *per layer* $\mathrm{SNR}_{\mathrm{eff},l}$ is then obtained by compressing the per-subcarrier $\mathrm{SINR}_{l,k}$ via the mapping function $\mathcal{I}_{\mathrm{map}}$:

$$\mathrm{SINR}_{\mathrm{eff},l}(\beta_1, \beta_2) = \beta_1 \mathcal{I}_{\mathrm{map}}^{-1}\left[\frac{1}{K}\sum_{k=1}^{K} \mathcal{I}_{\mathrm{map}}\left(\frac{\mathrm{SINR}_{l,k}}{\beta_2}\right)\right],$$

where β_1, β_2 are the rate-dependent adjustment parameters, and $\mathcal{I}_{\mathrm{map}}$ is a mapping function that is calculated via Chernoff union bound for error probabilities and is straightforward to implement:

$$\mathcal{I}_{\mathrm{map}}(\mathrm{SINR}_{l,k}) = 1 - \exp(-\mathrm{SINR}_{l,k}).$$

In the low or zero mobility scenarios, we can assume that the effective SINR in the fading channel causes the same level of Block Error Rate as in the AWGN channel [Olmos et al., 2010], which can be expressed as follows:

$$\mathrm{BLER}_l(\mathbf{H}_{\mathrm{eff}}, N_0, \mathrm{MCS}_l) \approxeq \mathrm{BLER}_{\mathrm{AWGN}}(\beta_{1,l}\mathrm{SINR}_{\mathrm{eff},l}, \mathrm{MCS}_l).$$

The calibration of the adjustment coefficients is done based on MSE criterion:

$$\beta_{\mathrm{opt}_0,l}, \beta_{\mathrm{opt}_1,l} = \underset{\beta_{0,l},\beta_{1,l}}{\arg\min} \frac{1}{N_H N_n}\left[\sum_{c}^{N_H}\sum_{z}^{N_n} |\mathrm{SINR}_{\mathrm{eff},l}(\beta_{0,l}, \beta_{1,l}, \mathrm{MCS}_l) - \mathrm{SINR}_{\mathrm{AWGN},l}(\mathrm{MCS}_l)|^2\right].$$

6.6.2 Results of the EESM Approach

The abstracted performance of the PIA receiver with the approximation by the MMSE receiver shows satisfying levels of accuracy (Table 6.4, where the index after a comma stands for a layer number): the MSE error does not exceed 0.5 dB, which perfectly fits the granularity of the CQI ($0.7 - 1$ dB). Similar levels of accuracy are guaranteed for the first stream of the SIC receiver in Table 6.5. The MSE error does not show a significant growth with the increase of the modulation level. At the same time, the lower-bound approximation by the IF receiver is not as optimistic for the second codeword of the SIC receiver, as the $\mathrm{MSE}_{\mathrm{EESM},1}$ values are higher than it was expected and result in

approximately 1.5 dB (Fig. 6.6, Table 6.5). The reason for this is the dependency of the decodability of the second codeword on the successful decoding of the first codeword. There exist situations where the instantaneous values of SINR for the second codeword are strong enough to potentially allow successful decoding, but since the first codeword was not decoded, the second codeword is also lost.

Table 6.4: Calibration results for EESM abstraction of PIA receiver in 8-tap Rayleigh channel

MCS_0	MCS_1	$MSE_{EESM,0}$	$MSE_{EESM,1}$	$\beta_{EESM_0,0}$	$\beta_{EESM_0,1}$	$\beta_{EESM_1,0}$	$\beta_{EESM_1,1}$
4	4	0.0741	0.2597	1.5845	1.7147	1.6947	1.6950
10	10	0.0490	0.1402	4.2456	3.1394	4.7579	3.3168
12	12	0.0744	0.1399	4.5814	5.1916	4.7984	5.3779

Table 6.5: Calibration results for EESM abstraction of SIC receiver in 8-tap Rayleigh channel

MCS_0	MCS_1	$MSE_{EESM,0}$	$MSE_{EESM,1}$	$\beta_{EESM_0,0}$	$\beta_{EESM_0,1}$	$\beta_{EESM_1,0}$	$\beta_{EESM_1,1}$
4	4	0.3537	0.7674	1.0977	0.8812	0.4205	0.9527
10	10	0.0719	0.9561	3.6321	1.8303	3.8368	3.3085
12	12	0.1986	1.8764	6.4572	1.7754	7.0483	2.0731

Table 6.6: Comparison of the MIESM and EESM accuracy of the abstracted performance of the PIA receiver in 8-tap Rayleigh channel

MCS_0	MCS_1	$MSE_{LUT,0}$	$MSE_{EESM,0}$	$MSE_{LUT,1}$	$MSE_{EESM,1}$
4	4	0.0508	0.0741	0.0571	0.2597
10	10	0.3781	0.0490	0.5707	0.1402
12	12	0.8729	0.0744	0.6986	0.1399

Table 6.7: Comparison of the MIESM and EESM accuracy of the abstracted performance of the SIC receiver in 8-tap Rayleigh channel

MCS_0	MCS_1	$MSE_{LUT,0}$	$MSE_{EESM,0}$	$MSE_{LUT,1}$	$MSE_{EESM,1}$
4	4	0.0533	0.3537	0.0926	0.7674
10	10	0.3791	0.0719	0.8706	0.9561
12	12	0.7480	0.1986	1.1566	1.8764

A comparative analysis of the accuracy, introduced by the EESM and MIESM mapping methods, is presented in Table 6.6 and Table 6.7. For the QPSK constellations, the MIESM approach outperforms the EESM methodology for both streams of our PIA and SIC receivers. For the codewords decoded in the presence of interference and mapped onto constellations with high cardinality, the EESM mapping surprisingly outperforms our heavy MIESM approach. Furthermore, the MSE values of the EESM mapping remain at low level with the increase of the MCS index, compared to the MIESM (as one can notice that $MSE_{EESM,0} \ll MSE_{LUT,0}$). This makes EESM more favorable for practical

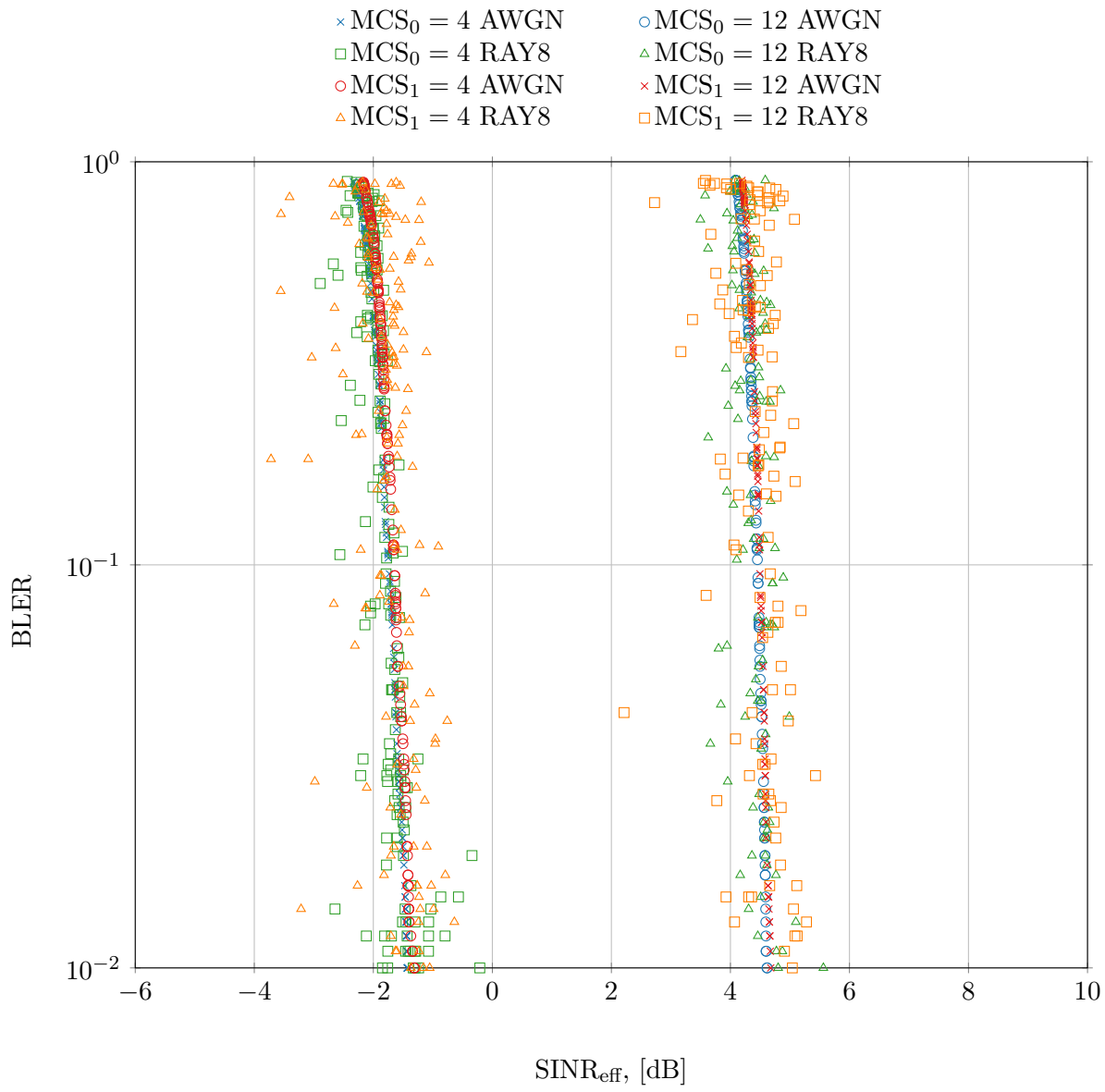


Figure 6.5: EESM validation for our PIA receiver with $(\text{MCS}_0, \text{MCS}_1) \in \{(4,4),(12,12)\}$ in 8-tap Rayleigh channel. The calibration factors are taken into account. In contrast with the MIESM approach, the MSE does not grow with the increase of the transmission rate, and remains at approximately the same level.

.

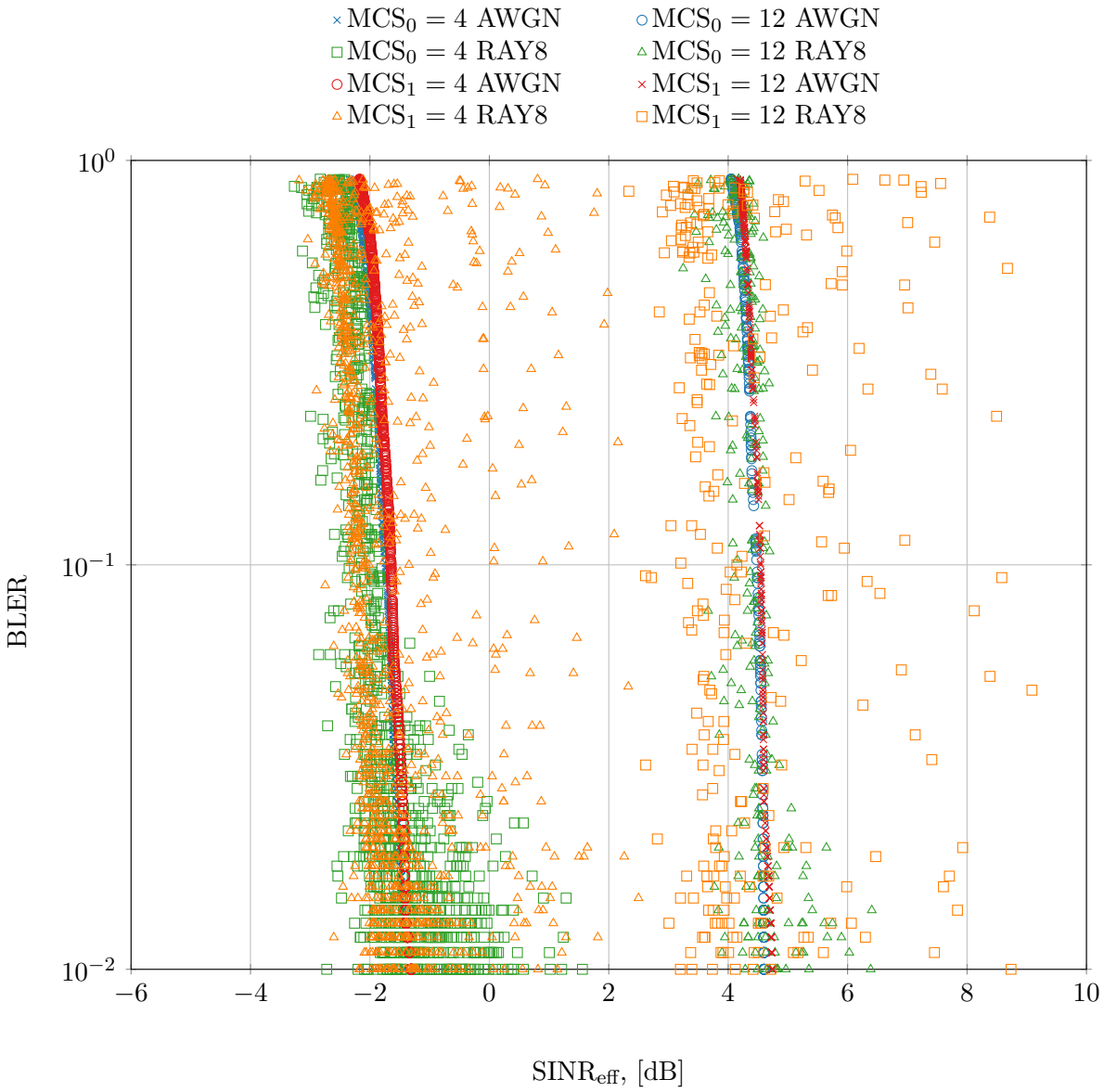


Figure 6.6: EESM validation for our SIC receiver with $(\text{MCS}_0, \text{MCS}_1) \in \{(4,4),(12,12)\}$ in 8-tap Rayleigh channel. The calibration factors are taken into account. The accuracy of abstracted performance of the interference-free stream is reduced competed to the performance of the first stream.

.

use. However, the EESM abstraction for interference-free detection of the SIC receiver is not as reliable, as for the first stream, and provides around 1 dB of accuracy.

The optimized MCS-dependent adjusting coefficients $\beta_{\mathrm{LUTopt}_0,0}$ and $\beta_{\mathrm{LUTopt}_1,0}$ for the first codeword, and $\beta_{\mathrm{LUTopt}_0,1}$ and $\beta_{\mathrm{LUTopt}_1,1}$ for the second codeword are not always known. During the CQI computation, the UE possesses a knowledge of the current MCS, which can then be used to estimate an effective SINR and throughput for the current MCS in given instantaneous channel conditions. Ideally, the UE has to calculate the throughput prediction for the full set of 28 MCS values (or 15 CQI positions). This process takes a significant amount of time that the UE may not have. We now verify if the proposed abstraction methodology is adequate when no calibration is applied ($\beta_{\mathrm{LUTopt}_0,0} = \beta_{\mathrm{LUTopt}_1,0} = \beta_{\mathrm{LUTopt}_0,1} = \beta_{\mathrm{LUTopt}_1,1}$). The MIESM approach is vulnerable to the absence of calibration and the performance prediction would be unreliable (Table 6.9, Table 6.11). Fortunately, the EESM mapping is less vulnerable to the absence of a calibration and results in maximum loss of 2 CQI positions, which could be considered practically feasible (Table 6.10, Table 6.10).

Table 6.8: The accuracy of the EESM abstraction of PIA receiver in 8-tap Rayleigh channel without calibration

MCS_0	MCS_1	$MSE_{EESM,0,calibr}$	$MSE_{EESM,1,calibr}$	$MSE_{EESM,0,\ nocalibr}$	$MSE_{EESM,1,nocalibr}$
4	4	0.0741	0.2597	0.0909	0.3119
10	10	0.0490	0.1402	0.5711	1.4187
12	12	0.8729	0.6986	0.6749	1.6768

Table 6.9: The accuracy of the MIESM LUT abstraction of PIA receiver in 8-tap Rayleigh channel without calibration

MCS_0	MCS_1	$MSE_{LUT,0,calibr}$	$MSE_{LUT,1,calibr}$	$MSE_{LUT,0,\ nocalibr}$	$MSE_{LUT,1,nocalibr}$
2	2	0.0576	0.0819	1.0012	10.5408
4	4	0.0508	0.0571	0.3905	9.2519
10	10	0.3781	0.5707	1.0657	12.7933
12	12	0.8729	0.6986	1.9643	13.4768

Table 6.10: The accuracy of the EESM abstraction of SIC receiver in 8-tap Rayleigh channel without calibration

MCS_0	MCS_1	$MSE_{EESM,0,calibr}$	$MSE_{EESM,1,calibr}$	$MSE_{EESM,0,\ nocalibr}$	$MSE_{EESM,1,nocalibr}$
4	4	0.3537	1.2539	0.3652	0.7674
10	10	0.0719	1.6563	0.5711	3.4223
12	12	0.1986	1.9709	1.7553	2.1868

Table 6.11: The accuracy of the MIESM LUT abstraction of SIC receiver in 8-tap Rayleigh channel without calibration

MCS_0	MCS_1	$MSE_{LUT,0,calibr}$	$MSE_{LUT,1,calibr}$	$MSE_{LUT,0,\ nocalibr}$	$MSE_{LUT,1,nocalibr}$
4	4	0.0533	0.0926	0.8006	8.1446
10	10	0.3791	0.8706	1.3145	10.7345
12	12	0.7480	1.1566	1.7768	10.6323

6.7 Discussion and Conclusion

In this chapter, we have studied two physical layer abstraction methodologies for our Reduced Complexity Maximum Likelihood PIA and SIC receivers: the interference aware look-up table based MIESM approach and the EESM approach based on the approximation with the MIMO MMSE receiver and the IF receiver. Both methodologies were evaluated through the link-level simulations carried out by the OAI downlink link-level simulator.

The MIESM approach was built upon the interference-aware strategy [Latif et al., 2012], where the channel statistics is stored in the form of look-up tables, and the link quality metrics is computed via accessing the tables with the direct and reverse mapping functions. This strategy provides satisfying levels of accuracy, but has a number of weak points. This methodology greatly depends on the receiver implementation, fixed-point or float-point architecture, look-up table granularity and channel models since the calibration coefficients might significantly vary from one fading environment to another [Chen et al., 2011a].

Moreover, even in our current setup the size of the look-up tables exceeds the practically feasible limits, and cannot be loaded into the memory of the UE at the same time for all the possible constellations. Furthermore, the practical LTE system system is capable of switching between the two transport blocks and single transport blocks transmission. In our setting, it would require further loading the look up tables for other transmission scenarios. Unfortunately, taking into account the time that is consumed to load the look-up table into memory of the device, it would be impossible to access them dynamically. All the tables would thus need to be loaded at the very beginning of the transmission, which would take a few hundred megabytes of memory, making it an unfeasible solution for real-time systems. Nevertheless, it is possible to use the look up tables in the simulators to speed-up the link-level computations.

To circumvent the drawbacks of the MIESM method, we evaluated a straightforward and light-weight EESM abstraction based on the approximation of the ML PIA and SIC receivers with the MIMO MMSE and IF detection. The abstracted performance is only slightly lower than the one of MIESM method for the low modulation order. In contrast with the MIESM approach, whose accuracy degrades with the increase of the modulation order, the MSE of the EESM approach remains at the low level even for high modulation orders.

The majority of the state-of-art abstraction methods are based on the calibration coefficients, which depend on the employed MCS and fading environment. However, in

practical Link Adaptation scenarios it is not always possible to obtain the correct values of the calibration coefficients. We have thus evaluated the if the proposed abstraction methodology is adequate without performing a calibration. Unlike MIESM mapping, the EESM approach provides practically acceptable levels of accuracy in this settings.

Chapter 7

Link Adaptation for the R-ML PIA and SIC Receivers

The concept of adaptive feedback communication for wireless systems was originally proposed in 1968 for noiseless and delay-free feedback channels [Hayes, 1968]. The idea behind an adaptive transmission consists in adjusting some key transmission parameters to guarantee the optimum key performance indicators (such as throughput, reliability, energy efficiency) depending on the variations of the wireless medium over time, frequency, and space. Non-adaptive systems are designed to support the worst-case conditions, which restricts them from reaching their full potential. At the same time, an adaptive coded modulation with turbo codes brings the performance in fading channel close to the Shannon capacity limit in AWGN channel [Goldsmith and Chua, 1998]. Fundamental theoretical studies have demonstrated that low complexity rate adaptation techniques perform close to the joint rate and power adaptation [Alouini and Goldsmith, 1999].

In modern wireless networks, the adaptive transmission techniques and the corresponding protocols are regrouped under the term of *Link Adaptation*. In Long Term Evolution (LTE) systems, the link adaptation can be performed in open or closed loop cycles. During an open loop cycle, the UE is not directly involved in the adaptation and the eNodeB determines the transmission parameters based on the ACK/NACK rates statistics [Gilbert et al., 1996]. However, the statistics must be collected during a certain period of time, which makes this strategy relatively slow [Tao and Czylwik, 2011].

Closed loop adaptation cycles are based on the transmission parameters suggested by the UE as a result of the calculations performed with the estimated channel coefficients. These parameters form the Channel State Information (CSI), and, depending on the configuration, the base station may or may not take them into account.

The CSI includes three informational items that are sent to the base station via one of the uplink channels: the Channel Quality Indicator (CQI), the Rank Indicator (RI) and the Precoder Matrix Indicator (PMI). During a single antenna transmission, only the CQI is required, while the PMI and the RI are necessary for optimization of the transmission via multiple antennas.

The problem of real-time feedback estimation is one of the challenging points in receiver design. Although the outdated CSI degrades the system performance compared to an ideal

or predicted CSI, it still provides significantly better performance than a blind Round-Robin scheduler assuming that the outdated channel is not completely independent from the current channel state [Goldenbaum et al., 2011].

The calculation of the CSI is not a trivial task, and various metrics and criteria have been proposed to provide an adequate degree of accuracy with the affordable time and complexity restrictions [Love and Heath, 2003; Schwarz et al., 2010]. However, the practical implementation for the integration into real chipsets remains up to the vendor.

Contributions In this chapter, we study a low-complexity Link Adaptation strategy compatible with our Reduced Complexity Parallel Interference Aware receiver and Successive Interference Aware Canceling receiver. In LTE Single-User systems with spatially multiplexed schemes, the CSI computations strictly depend on the number of the transmitted transport blocks. Indeed, if both codewords are active the UE operates with the precoder matrices, rather than vectors, and the CQI reporting is done for each transport block individually. Based on this, we follow a natural path to first decide on the RI, and then to perform a joint estimation of the PMI and the CQI. Our investigations demonstrate that the design of our receivers allows for the successful detection of both codewords, even if the channel is poorly-conditioned. Furthermore, our sub-optimal light-weight precoder calculation based on evaluation of the correlation coefficients of the channel matrix is numerically proven to perform closely to the optimal Maximum Mutual Information (MI) based criteria. Finally, we propose a light weight CQI adaptation based on abstraction techniques developed in the previous chapter and study joint CQI and PMI adaptation. We demonstrate that our SIC receiver always reports higher SNR on the second stream, than our PIA receiver given the the same radio conditions.

Contents

7.1 State of the Art

In this section, we provide an overview of existing methods for CSI computation. We then motivate our approach and provide the details on the implementation.

When the concept of link adaptation concept was initially introduced to the practical implementations of the wireless standards, [Catreux et al., 2002] discussed four main adaptation techniques: the Mean SINR [Nanda et al., 2000], the Multiple Statistics of the Received SINR, the Error Statistics Informations [Gilbert et al., 1996], and the Combination of the SINR and the Error Statistics Information. Researches have identified the main implementation issues: the determination of the channel adaptation thresholds, the computational cost and the signaling overhead for the per-subcarrier adaptation. The problem was complexified by the subband adaptation, which caused degradation in frequency-selective environments, and non-perfect feedback channel. Even though a few research efforts have been made since then to circumvent these challenges, the complexity-performance trade-off is still an open question.

7.1.1 Rank Indicator

Single-User Closed Loop Spatial Multiplexing transmission is designed for the high SINR regime, where significant capacity gains can be accessed through a simultaneous transmission via multiple spatial layers. In practice, the performance might be degraded by the channel correlation, noise distortion, multipath propagation, fading and interference, and the multiplexing gain might thus be not accessible. Instead, the system might benefit from a transmit diversity scheme to ensure reliable communication. The mechanism that evaluates the channel conditions and regulates the transition from one transmission scheme to another is called *rank adaptation*.

Various criteria have been proposed to define the threshold between these two transmission schemes and were compared by [Bai et al., 2010b]. Some criteria are based solely on the analysis of the channel characteristics: Singular Value Decomposition (SVD) Based Criteria and Maximum Capacity Criteria, while others such as Maximum MMSE post-processed Capacity Criterion and Maximum Mutual Information based selection also take into account receiver architecture. In the SVD method, the decision is based on the comparison of the singular values of the channel matrix and the SNR at the receiver. The maximum channel capacity method quantifies the capacity in the current channel conditions for both the spatial multiplexing and the transmit diversity and selects the transmission scheme that provides maximum capacity gains. The Maximum Mutual Information criteria is demonstrated to be the optimal since it provides the highest spectral efficiency gain. However, this criterion is expensive to implement for non-linear receivers.

Joint optimization of the RI and the PMI was proposed by [Schwarz and Rupp, 2011]. However, their method is built on mutual information estimation, and thus is challenging to implement on chipsets.

We thus seek a practical solution that allows to adequately adjust a transmission scheme to instantaneous channel conditions.

7.1.2 Precoding Matrix Indicator

The main purpose of the precoding block in LTE is to increase the system throughput and to reduce inter-layer interference by transforming the channel matrix based on the knowledge of the CSI. In an ideal system, the transmitter and the receiver both possess

perfect knowledge of the channel. In practice, the knowledge of the radio link conditions may be outdated (short coherence time) or fed back through an erroneous uplink channel.

In LTE systems, precoding can be codebooks-based and non-codebook based. In the first case, the precoding codebooks are standardized and shared between the base station and the UE [3GPP, 2015c]. In the latter case, so-called Demodulation Reference Signals are added before precoding and provide information about joint influence of the channel and precoding. For Transmission Mode 4, only a codebook based precoding is defined, and the applied precoder index is signaled between the eNodeB and the UE. The codebooks are standardized [3GPP, 2015c] and defined for different transmission settings, such as transmission rank, antenna configuration, etc. For our 2×2 system using Closed Loop Spatial Multiplexing (CLSM) transmission, the eNodeB has two possible choices for the precoding matrix $\mathbf{P}$:

$$\mathbf{P} \in \left\{ \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \right\}. \tag{7.1}$$

Based on the channel measurement and estimation, the UE calculates the optimal $\mathbf{P}$ and signals the corresponding index to the eNodeB. However, the calculation algorithms are not standardized and depend on the vendor's implementation. A few solutions have been proposed for precoder calculations, such as the maximum mutual information criterion [Schwarz et al., 2010] and the maximum SNR criterion [Ghaffar and Knopp, 2010b]. The maximum MI-based criterion is optimal, but is difficult to implement: it involves a time-consuming computation of Mutual Information for every subband or requires precomputed Look-Up Tables (LUT) with the values of Mutual Information corresponding to various channel conditions. The receiver architecture should also be taken into account as it may impact the total Mutual Information of the system (Section 2.1.2). On the other hand, in case of Successive Interference Canceling design, the intuitive and easy to implement solution is to minimize the SNR of the first stream since the successful decoding of the second codeword depends on the decodability of the first codeword [Ghaffar and Knopp, 2010b].

7.1.3 Channel Quality Indicator

The third essential component of CSI is the CQI which allows to adjust the transmission data rate to the instantaneous channel conditions and to avoid exceeding the outage capacity of the channel. By definition, the CQI represents a combination of the modulation and the coding rate that can be supported by an instantaneous channel with a certain level of probability of error which should not exceed 10%. The LTE standard defines 15 CQI values with a known spectral efficiency.

The CQI values, which the UE reports to the base station, are correlated with the channel characteristics and the receiver architecture since advanced receivers possess higher chances of successful detection of the signal in poor channel conditions. The link quality metric that unifies these influencing factors is the *post-processed* SINR, which is conditioned on the receiver architecture. If a channel is frequency flat, the post-processed SINR is uniformly distributed among Resource Elements, and thus a simple SINR averaging is sufficient for the CQI estimation. However, for practical scenarios, the *effective*

SINR concept is still required to reflect the non-uniform distribution of the post-processed SINR on the Resource Elements, caused by the multipath propagation and fading. The Effective Exponential SNR Mapping (EESM) and the Mutual Information Effective SNR Mapping (MIESM) methodologies, reviewed in Chapter 6, are widely studied for CQI estimation [Donthi and Mehta, 2011; Schwarz and Rupp, 2011]. We observed that they often require calibration coefficients, which depend on the channel model, the receiver architecture (and implementation), and the MCS value. This may limit their use for real-time transmissions.

Manufacturers of chipsets for the mobile stations most certainly apply low complexity techniques, that are, unfortunately, undisclosed. To our knowledge, only a few available scientific contributions consider the CQI estimation without calibration factors. [Chen et al., 2011b] observed that the channel quality degrade with the increase of the multipath delay spread, and proposed a joint mapping based on SINR and maximum multipath delay spread cut-off points. We aim to build the low-complexity CQI estimation methodology which allows to maximize throughput in real-time transmission and provides an acceptable level of accuracy without calibration.

In LTE, the base station schedules the same MCS to all the resource elements, dedicated to one user. However, due to the non-uniform distribution of the post-processed SINR, the radio conditions for some Resource Elements might be characterized by higher or lower values of CQI than other. To capture this phenomenon, the LTE standard [3GPP, 2016] divided the LTE bandwidth into fractions — *subbands* — and defined *subband CQI* and *wideband CQI* reporting. The number of available subbands and their size depend on the LTE bandwidth.

The wideband CQI provides a single CQI value for the whole bandwidth, while the subband CQI provides information about radio conditions for selected K-best subbands. This approach is valid not only for the CQI reporting, but is also extended to PMI feedback. For Single-User Closed Loop Spatial Multiplexing the standard defines multiple *feedback reporting modes*, which consist of combinations of CQI and PMI reporting.

7.2 Rank Adaptation

7.2.1 Methodology

Depending on the number of active spatial layers, the transmission mode 4 may apply a CLSM transmission scheme if the RI is two, or fall back to Alamouti precoding if the reported rank value is one. While designing our rank estimation methodology, we sought a low-complexity solution that provides an adequate performance and is fast enough to be used in a real system. Our first approach for the rank estimation was to reuse the Maximum Capacity Criteria proposed by [Bai et al., 2010b] and to compare the non-constrained preprocessed channel capacity for CLSM and Alamouti schemes. The UE then would report the RI that corresponds to the transmission scheme that guarantees the maximum capacity in instantaneous channel conditions.

Recall that the ergodic capacity of the baseline 2×2 MIMO system using closed loop

spatial multiplexing can be computed using the following formula:

$$C_{\text{CLSM}} = \log_2 \det\left(\mathbf{I} + \gamma \mathbf{H}\mathbf{H}^H\right), \tag{7.2}$$

where $\mathbf{I}$ is 2×2 identity matrix, $\gamma = E_s/N_0$ is a pre-transmitted SNR value (we assume unit signal energy $E_s = 1$), and $\mathbf{H}$ is a channel matrix. Similarly, the capacity achieved by the Alamouti precoding in the identical 2×2 MIMO system can be computed as

$$C_{\text{AL}} = \log_2\left(1 + \gamma \left\|\mathbf{H}\right\|_{\text{F}}^2\right), \tag{7.3}$$

where $\|.\|_{\text{F}}$ is Frobenius norm.

However, this method did not yield to a valid estimation in our settings, since the channel capacity with Alamouti precoding C_{AL} is always less or equal than the channel capacity of the closed loop spatial multiplexing C_{CLSM}. The analytical proof of this fact is provided in the next lines.

Assume that $C_{\text{AL}} \leq C_{\text{CLSM}}$. Since the logarithm function is a monotonically increasing, the comparison of the functions C_{AL} and C_{CLSM} reduces to the comparison of the arguments $\left(1 + \gamma \left\|\mathbf{H}\right\|_{\text{F}}^2\right)$ and $\det\left(\mathbf{I} + \gamma \mathbf{H}\mathbf{H}^H\right)$:

$$C_{\text{AL}} \leq C_{\text{CLSM}} \Leftrightarrow \left(1 + \gamma \left\|\mathbf{H}\right\|_{\text{F}}^2\right) \leq \det\left(\mathbf{I} + \gamma \mathbf{H}\mathbf{H}^H\right). \tag{7.4}$$

We first develop the expression for the closed loop spatial multiplexing C_{CLSM}. Consider $\mathbf{H}\mathbf{H}^H$:

$$\mathbf{H}\mathbf{H}^H = \begin{bmatrix} h_{00} & h_{01} \\ h_{10} & h_{11} \end{bmatrix} \begin{bmatrix} h_{00}^* & h_{10}^* \\ h_{01}^* & h_{11}^* \end{bmatrix} = \begin{bmatrix} |h_{00}|^2 + |h_{01}|^2 & h_{00}h_{10}^* + h_{01}h_{11}^* \\ h_{10}h_{00}^* + h_{11}h_{01}^* & |h_{10}|^2 + |h_{11}|^2 \end{bmatrix}.$$

With this, we have

$$\det\left(\mathbf{I} + \gamma \mathbf{H}\mathbf{H}^H\right) = \det \begin{bmatrix} 1 + \gamma\left(|h_{00}|^2 + |h_{01}|^2\right) & \gamma\left(h_{00}h_{10}^* + h_{01}h_{11}^*\right) \\ \gamma\left(h_{10}h_{00}^* + h_{11}h_{01}^*\right) & 1 + \gamma\left(|h_{10}|^2 + |h_{11}|^2\right) \end{bmatrix}. \tag{7.5}$$

Using the fact that the determinant of a 2×2 matrix equals the difference between the products of the diagonal elements: $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$, (7.5) can be further developed as:

$$\begin{aligned} \det\left(\mathbf{I} + \gamma \mathbf{H}\mathbf{H}^H\right) = 1 &+ \gamma^2\left(|h_{00}|^2|h_{10}|^2 + |h_{01}|^2|h_{11}|^2\right) \\ &+ \gamma^2\left(|h_{00}|^2|h_{11}|^2 + |h_{01}|^2|h_{10}|^2\right) \\ &+ \gamma\left(|h_{00}|^2 + |h_{01}|^2 + |h_{10}|^2 + |h_{11}|^2\right) \\ &- \gamma^2\left(|h_{00}|^2|h_{10}|^2 + |h_{01}|^2|h_{11}|^2\right). \end{aligned} \tag{7.6}$$

Applying the definition of the Frobenius norm, we obtain

$$\left(1 + \gamma \left\|\mathbf{H}\right\|_{\text{F}}^2\right) = 1 + \gamma\left(|h_{00}|^2 + |h_{01}|^2 + |h_{10}|^2 + |h_{11}|^2\right). \tag{7.7}$$

Taking into account (7.6) and (7.9), we can rewrite (7.4) as

$$\begin{aligned} 1+\gamma\left(|h_{00}|^2+|h_{01}|^2+|h_{10}|^2+|h_{11}|^2\right) \leq 1+\gamma\left(|h_{00}|^2+|h_{01}|^2+|h_{10}|^2+|h_{11}|^2\right) \\ +\gamma^2\left(|h_{00}|^2|h_{11}|^2+|h_{01}|^2|h_{10}|^2\right), \\ 0 \leq \gamma^2\left(|h_{00}|^2|h_{11}|^2+|h_{01}|^2|h_{10}|^2\right), \end{aligned}$$

which is valid for any value of SNR γ and the channel coefficients h_{00}, h_{01}, h_{10}, and h_{11}. Thus, we have proved that the Maximum Capacity Criteria is not a valid methodology for our 2×2 scenario.

The performance of the spatial multiplexing transmission schemes is sensitive to the transmit and receive correlation [Bölcskei et al., 2002]. An objective measure of the channel correlation is the conditional number $\mathcal{K}(\mathbf{H})$. Channel coefficients are estimated based on the interpolation of the sequence of reference symbols with certain periodicity. Therefore, the UE receives constant updates on the channel gains and can provide an estimation of the conditional number. In linear algebra, the condition number $\mathcal{K}(\mathbf{H})$ of the matrix $\mathbf{H}$ is defined as the ratio between the maximal and minimal singular values $\sigma_{max}(\mathbf{H})$ and $\sigma_{min}(\mathbf{H})$ of the matrix:

$$\mathcal{K}(\mathbf{H}) = \frac{\sigma_{\max}(\mathbf{H})}{\sigma_{\min}(\mathbf{H})}. \tag{7.8}$$

The ideal condition number is $\mathcal{K}(\mathbf{H}) = 0$ dB, which means that the singular values are of equal power. In practice, the channel is considered as well-conditioned if $\mathcal{K}(\mathbf{H})$ does not exceed 10 dB [Agilent, 2009]. While the computation of the condition number for high dimensional matrices is tricky and time consuming, it is straightforward for our 2×2 matrices: the computations are simplified to the product of the n-norms of the channel matrix $\mathbf{H}$ and inverse channel matrix $\mathbf{H}^{-1}$:

$$\mathcal{K}(\mathbf{H}) = \|\mathbf{H}\|_{\mathrm{F}} \left\|\mathbf{H}^{-1}\right\|_{\mathrm{F}}. \tag{7.9}$$

The computations are done using streaming SIMD instructions [Skillicorn, 1990], which allow to execute four multiplications simultaneously. The conditional number is estimated on a per-subcarrier basis, and if the majority of the values are below a certain threshold $\mathcal{T}$, the RI is set to two and the spatial multiplexing transmission scheme is applied. Otherwise the UE reports the RI one and the base station transmits using Alamouti precoding scheme.

7.2.2 Numerical Results and Discussion

Our link-level simulations were carried out with LTE bandwidth of 5 MHz and 8-tap Rayleigh fading channel and Extended Pedestrian A (EPA) channel 3GPP [2015d] with zero Doppler frequency. For the EPA channel modeling, we applied a moderate and high correlation matrix, referred to as to EPAM and EPAH respectively. The eNodeB sends 3000 packets with 1 Physical Downlink Control Channel symbol over a wide range of noise variances. To consider different modulation orders, we chose the following pairs of

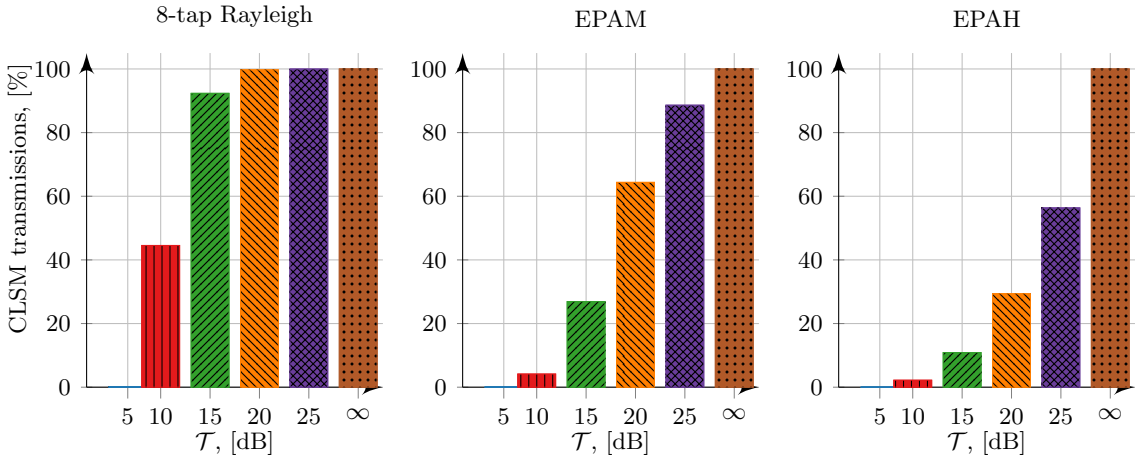


Figure 7.1: Percentage of spatially multiplexed transmissions for different levels of threshold $\mathcal{T}$ in the 8-tap Rayleigh, the EPAM and the EPAH fading channls.

the MCS: $(\mathrm{MCS}_0, \mathrm{MCS}_1) \in \{(4,4),(12,12),(22,26)\}$. For each fading environment, we consider the set of thresholds $\mathcal{T} \in \{5, 10, 15, 20, 25, \infty\}$ dB.

In the 8-tap Rayleigh channel, 90% of the transmissions belong to spatial multiplexing, if the condition number threshold $\mathcal{T}$ is set to 15 dB, and 100% if $\mathcal{T}$ is higher than 20 dB, as illustrated in Fig. 7.1. With the increase of correlation level, the amount of simultaneous transmissions of the two transport blocks given the same values of threshold $\mathcal{T}$ decreases: for $\mathcal{T} = 25$ dB there are only 90% of transmissions in the EPAM channel, while this value achieves only 57% in the EPAH channel. This implies that we might expect, in theory, that rank adaptation brings significant benefits in throughput in highly correlated channels.

Fig. 7.2 and Fig. 7.3 illustrate the empirical throughput achieved by our PIA and SIC for different values of threshold $\mathcal{T}$. For both constellations, the receivers show the highest throughput when the eNodeB always sends two transport blocks simultaneously. Therefore, our receivers successfully manage the detection of both codewords even in ill-conditioned channels.

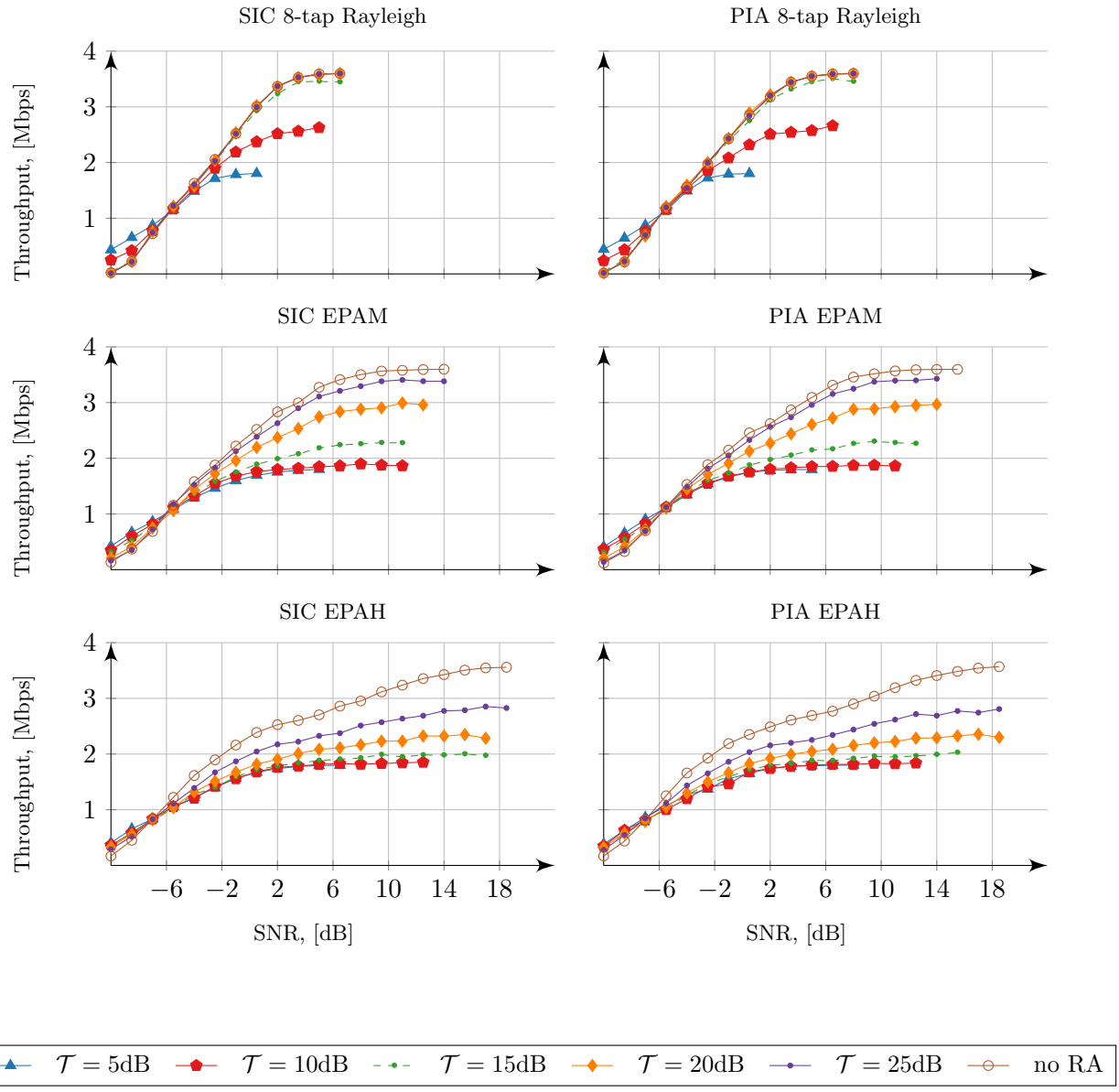


Figure 7.2: Sum system throughput for the PIA and SIC receivers with $\text{MCS}_0 = 4$ and $\text{MCS}_1 = 4$ in the 8-tap Rayleigh, the EPAM and the EPAH channels. For SNR values higher than -4 dB, the highest throughput is achieved if the condition number threshold $\mathcal{T}$ is set to infinity.

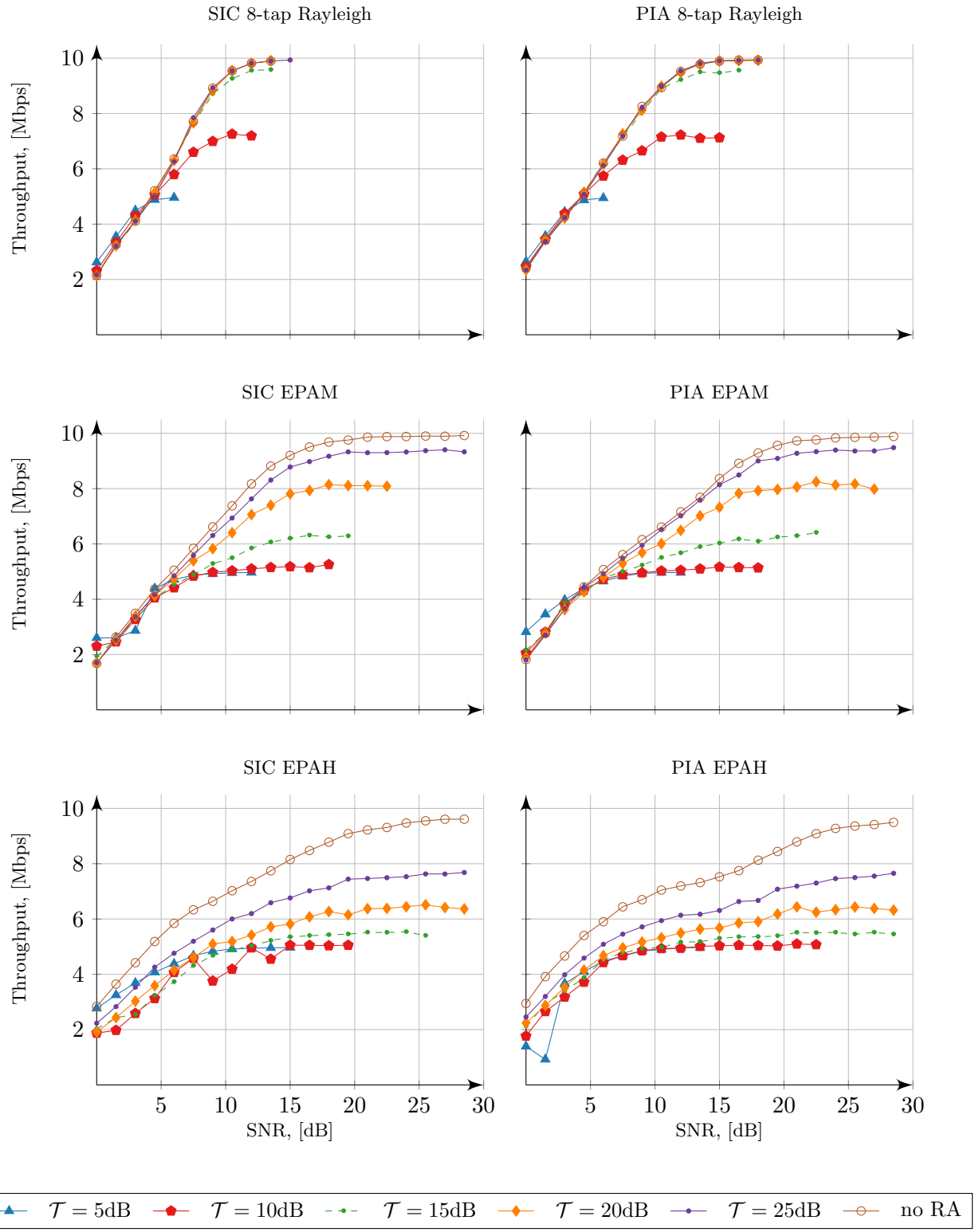


Figure 7.3: Sum system throughput for the PIA and SIC receivers after one round of the transmissions in the 8-tap Rayleigh, the EPAM and the EPAH channels. The codewords belong to 16QAM constellation: $\text{MCS}_0 = 12$ and $\text{MCS}_1 = 12$. The throughput values are the highest in the setting where the base station always sends two transport blocks regardless the channel.

7.3 Precoder Estimation

The maximum mutual information based criterion is optimal but is difficult to implement. The mutual information levels depend on the receiver architecture and should be computed based on (4.5), (4.6) for all precoding options. The UE then selects the precoder matrix that provides the highest level of mutual information. On the other hand, in case of Successive Interference Canceling design, the intuitive and easy to implement solution is to maximize the SNR of the first stream since the successful decoding of the second codeword depends on the decodability of the first codeword [Ghaffar and Knopp, 2010b]. The UE computes two ratios between the SNR values of the first stream and the second stream:

$$\left(\frac{\|\mathbf{h}_0 + \mathbf{h}_1\|^2}{\|\mathbf{h}_0 - \mathbf{h}_1\|^2}, \frac{\|\mathbf{h}_0 + j\mathbf{h}_1\|^2}{\|\mathbf{h}_0 - j\mathbf{h}_1\|^2} \right). \tag{7.10}$$

If the first ratio is bigger than the second one, then the UE selects the precoder matrix with real values, otherwise with complex values. This computation can be simplified to the evaluation of the real and imaginary parts of the correlation coefficient $\rho_{10} = \mathbf{h}_1^H \, \mathbf{h}_0$, as we show below.

We aim to determine the regions on the complex plane $\mathbb{C}$, where

$$\frac{\|\mathbf{h}_0 + \mathbf{h}_1\|^2}{\|\mathbf{h}_0 - \mathbf{h}_1\|^2} \geq \frac{\|\mathbf{h}_0 + j\mathbf{h}_1\|^2}{\|\mathbf{h}_0 - j\mathbf{h}_1\|^2}. \tag{7.11}$$

Consider two complex vectors a and b. Then, the norm of the sum of $\mathbf{a}$ and $\mathbf{b}$ can be developed in the following way:

$$\|\mathbf{a} + \mathbf{b}\|^2 = (\mathbf{a} + \mathbf{b})^H (\mathbf{a} + \mathbf{b}) = \|\mathbf{a}\|^2 + \mathbf{a}^H \mathbf{b} + \mathbf{b}^H \mathbf{a} + \|\mathbf{b}\|^2 .$$

Similarly, the numerators and denominators of (7.11) can be rewritten as

$$\begin{aligned} \|\mathbf{h}_0 + \mathbf{h}_1\|^2 &= \|\mathbf{h}_0\|^2 + \rho^* + \rho + \|\mathbf{h}_1\|^2 , \\ \|\mathbf{h}_0 - \mathbf{h}_1\|^2 &= \|\mathbf{h}_0\|^2 - \rho^* - \rho + \|\mathbf{h}_1\|^2 , \\ \|\mathbf{h}_0 + j\mathbf{h}_1\|^2 &= \|\mathbf{h}_0\|^2 + j\rho^* - j\rho + \|\mathbf{h}_1\|^2 , \\ \|\mathbf{h}_0 - j\mathbf{h}_1\|^2 &= \|\mathbf{h}_0\|^2 - j\rho^* + j\rho + \|\mathbf{h}_1\|^2 . \end{aligned} \tag{7.12}$$

Multiplying both sides of (7.11) with the denominators $\|\mathbf{h}_0 - \mathbf{h}_1\|^2$ and $\|\mathbf{h}_0 - j\mathbf{h}_1\|^2$, we obtain

$$\|\mathbf{h}_0 + \mathbf{h}_1\|^2 \|\mathbf{h}_0 - j\mathbf{h}_1\|^2 \geq \|\mathbf{h}_0 + j\mathbf{h}_1\|^2 \|\mathbf{h}_0 - \mathbf{h}_1\|^2. \tag{7.13}$$

Applying (7.12) to (7.13), we can further simplify to:

$$\|\mathbf{h}_0 + \mathbf{h}_1\|^2 \left(\rho^* + \rho + j\rho - j\rho^*\right) \geq 0, \tag{7.14}$$

where $\|\mathbf{h}_0 + \mathbf{h}_1\|^2$ is always non-negative, and thus

$$\left(\rho^* + \rho + j\rho - j\rho^*\right) \geq 0. \tag{7.15}$$

Decomposing ρ into real and imaginary parts, we have

$$\rho = \Re(\rho) + j\Im(\rho), \quad \rho^* = \Re(\rho) - j\Im(\rho).$$

Now (7.15) can be further developed as

$$j(\Re(\rho) + j\Im(\rho) - \Re(\rho) + j\Im(\rho)) + \Re(\rho) - j\Im(\rho) + \Re(\rho) + j\Im(\rho) \geq 0 \iff \Re(\rho) \geq \Im(\rho).$$

Thus, Inequality (7.11) holds for any ρ such that $\Re(\rho) \geq \Im(\rho)$. Similarly, the values of ρ such that $\Re(\rho) < \Im(\rho)$ satisfy inequality

$$\frac{\|\mathbf{h}_0 + \mathbf{h}_1\|^2}{\|\mathbf{h}_0 - \mathbf{h}_1\|^2} < \frac{\|\mathbf{h}_0 + j\mathbf{h}_1\|^2}{\|\mathbf{h}_0 - j\mathbf{h}_1\|^2}.$$

We can now define the criterion of the precoder matrix selection based on the correlation coefficient ρ between the columns $\mathbf{h}_0$ and $\mathbf{h}_1$ of the channel matrix $\mathbf{H}$:

$$\mathbf{P} = \begin{cases} \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \text{ for } \Re(\rho_{10}) \geq \Im(\rho_{10}); \\ \\ \frac{1}{2}\begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}, \text{ for } \Re(\rho_{10}) < \Im(\rho_{10}). \end{cases} \tag{7.16}$$

Since our solution does not guarantee the maximum throughput, we perform a numerical analysis of the potential performance loss due to the SNR-based criterion compared to the mutual information based PMI selection. The levels of the mutual information are obtained via Monte-Carlo simulations of (4.5), (4.6) for a wide range of SNR values.

As it can be seen in Fig. 7.4, the mutual information based criterion outperforms the suboptimal SNR-based computation only when the first codeword is mapped on 16QAM constellation, and the second codeword belongs to 64QAM. However, this gap vanishes when both codewords belong to the same constellation and almost complete match is observed. The SNR-based criterion is thus light-weight solution for the PMI computation, and offers performance levels that are close to the optimal maximum mutual information criterion.

7.4 CQI Adaptation

7.4.1 Methodology

In the previous sections, we have proposed the methodologies for the RI and the PMI estimation which are designed to optimize the number of spatial layers and reduce inter-layer interference. We now seek a practical CQI estimation methodology that allows to adjust transmission data rates to instantaneous channel conditions.

As a preliminary step for our CQI estimation methodology, we build a look-up table for effective SINR to CQI mapping. The standard describes the CQI values that posses a certain spectral efficiency and approximate LTE code rate (Table 7.1).

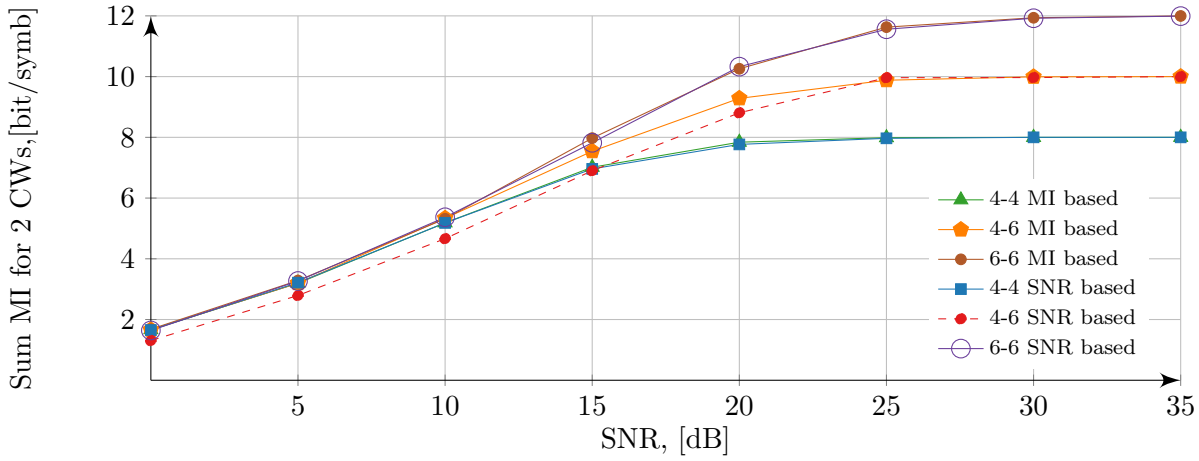


Figure 7.4: Potential mutual information levels for two codewords using our R-ML SIC receiver with MI-based and SNR-based precoder selection for the CWs with the modulation orders $M_0, M_1 \in \{4, 6\}$. The comparison is done for the frequency flat Rayleigh channel. The SINR-based criterion leads to the same performance levels as the MI-based criterion, when the codewords are mapped onto the same constellation.

Table 7.1: LTE CQI Table [3GPP, 2016]

CQI	Modulation	code rate ×1024	efficiency	Code LTE
0		out of range		
1	QPSK	78	0.1523	0.076172
2	QPSK	120	0.2344	0.11720
3	QPSK	193	0.3770	0.18850
4	QPSK	308	0.6016	0.30080
5	QPSK	449	0.8770	0.43850
6	QPSK	602	1.1758	0.58790
7	16QAM	378	1.4766	0.36914
8	16QAM	490	1.9141	0.47852
9	16QAM	616	2.4063	0.60157
10	64QAM	466	2.7305	0.45508
11	64QAM	567	3.3223	0.53717
12	64QAM	666	3.9023	0.65038
13	64QAM	772	4.5234	0.75390
14	64QAM	873	5.1152	0.85253
15	64QAM	948	5.5547	0.92578

However, the rate matching for the downlink data transmissions is based on 28 available values of MCS. The base station thus needs to translate the reported CQI values into MCS.

Each MCS possesses a certain coding rate, which varies depending on multiple factors, such as a number of allocated Physical Resource Blocks, number of PDCCH symbols, etc. Using our downlink simulator, we obtained the performance of Single-User MIMO CLSM transmissions in AWGN channel for the full set of MCS values. For each MCS value we then defined a corresponding SINR value such that the Block Error Rate (BLER) does not exceed 10% (Fig. 7.5). These SINR values are identical for our PIA and SIC receivers,

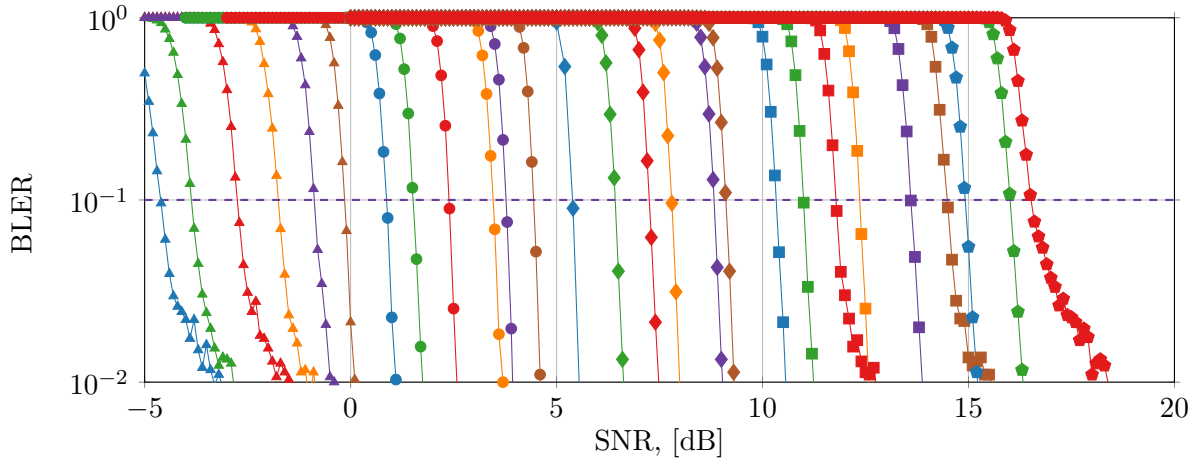


Figure 7.5: Block Error Rates curves in AWGN channel. The curves correspond to MCS1-28, plotted from left to right in ascending order.

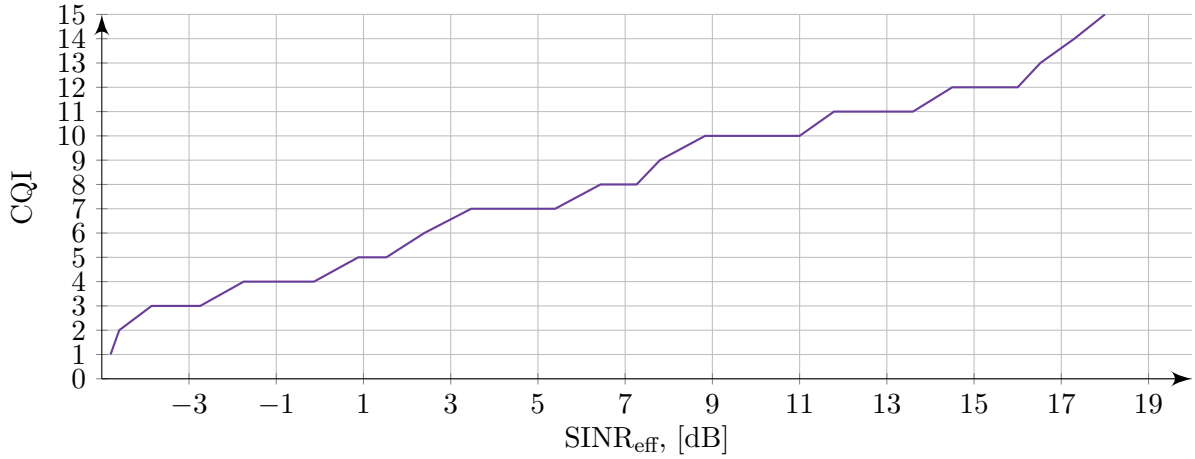


Figure 7.6: CQI-to-SINR mapping for 2×2 LTE Single-User MIMO system with 25 Physical Resource Blocks and 1 PDCCH symbol based on Table 7.2. Different mapping might be required for another setting.

since two codewords do not suffer from interference in the AWGN channel.

The selected SINR values correspond to the *effective* SINR, that causes the same level of Block Error Rate in frequency selective channel as in the AWGN channel. During the simulations, we consider 5 MHz of LTE bandwidth (25 Physical Resource Blocks) and one PDCCH symbol. To perform the mapping, the coding rates of MCS and approximate LTE code rate of CQI are compared, and the corresponding pairs are created. This data is further summarized in Table 7.2 and illustrated in Fig. 7.6; one CQI value may have multiple corresponding MCS values. Depending on the implemented algorithm, available resources, and feedback error margin the base station selects one of the suitable CQI values. In our scenario, the base station always chooses the highest possible MCS.

The obtained look up table with SINR-CQI-MCS mapping can now be used for the CQI estimation. Our methodology for the PIA and SIC receivers is built upon the results

Table 7.2: MCS to CQI mapping

MCS	Modulation	SNR_{eff}, [dB]	code rate	CQI
0	QPSK	−4.8	0.094444	1
1	QPSK	−4.6083	0.125556	2
2	QPSK	−3.8572	0.15222	3
3	QPSK	−2.7431	0.196667	3
4	QPSK	−1.7460	0.250000	4
5	QPSK	−0.8757	0.307778	4
6	QPSK	−0.1344	0.36111	4
7	QPSK	0.8804	0.432222	5
8	QPSK	1.5240	0.485556	5
9	QPSK	2.3938	0.556667	6
10	16QAM	3.4707	0.278333	7
11	16QAM	3.7824	0.305000	7
12	16QAM	4.4562	0.345000	7
13	16QAM	5.3956	0.398333	7
14	16QAM	6.4353	0.448333	8
15	16QAM	7.2630	0.501667	8
16	16QAM	7.7969	0.537222	9
17	64QAM	8.8338	0.358148	10
18	64QAM	9.1140	0.370000	10
19	64QAM	10.3429	0.423333	10
20	64QAM	11	0.458889	10
21	64QAM	11.7885	0.494444	11
22	64QAM	12.3711	0.530000	11
23	64QAM	13.6	0.582222	11
24	64QAM	14.5	0.626667	12
25	64QAM	15	0.65333	12
26	64QAM	16	0.706667	12
27	64QAM	16.52	0.733333	13
28	64QAM	17.3	0.848889	14

of the light weight EESM abstraction without calibration factors, which we have developed and validated in Chapter 6. The performance of the interference-aware streams is approximated with the per-stream performance of the MMSE receiver, while the performance of the interference-free second stream of our SIC receiver is approximated by the receiver with perfect interference canceling. Our methodology is based on the UE measurements, such as noise variance N_0 and channel estimates $\mathbf{H}$, and does not require any knowledge of currently applied MCS.

The joint wideband CQI/wideband PMI estimation (Algorithm 15) starts with the selection of a single preferred precoding matrix $\mathbf{P}$ for the whole bandwidth. This is done by computation and accumulation of the correlation coefficient ρ_{10} for each Resource Element. The accumulated value is then evaluated based on (9.28) and a single PMI value is selected. After that, we compute the per-stream SINR value for each Resource Element taking the PMI into account. For our PIA receiver and for the first stream of the

SIC receiver, the SINR is computed following the approximation by the MMSE receiver:

$$\mathrm{SINR}_{\mathrm{MMSE}\,l,k} = \frac{1}{[(\mathbf{I}_{n_{tx}} + \frac{1}{N_0 n_{\mathrm{tx}}}\mathbf{H}_{\mathrm{eff},k}^{H}\mathbf{H}_{\mathrm{eff},k})^{-1}]_{ll}} - 1. \tag{7.17}$$

The SINR values for the second stream of the SIC receiver are computed using the following expression:

$$\mathrm{SINR}_{\mathrm{IF}_1} = \frac{1}{N_0 n_{\mathrm{tx}}}\|\mathbf{h}_{\mathrm{eff}1,k}\|^2. \tag{7.18}$$

The SINR values per Resource Element are then compressed into the effective value $\mathrm{SINR}_{\mathrm{eff}_l}$ using the EESM mapping functions:

$$\mathrm{SINR}_{\mathrm{eff}_l} = -\ln\sum_{k=1}^{K}\exp(-\mathrm{SINR}_{l,k}). \tag{7.19}$$

The per-stream CQI value is now available from Table 7.2.

If the feedback transmission mode is set to wideband CQI/subband PMI reporting (Algorithm 16), the full transmission bandwidth is divided into seven subbands, and the precoder matrix is estimated for each of them individually. The SINR values for each Resource Element are then computed taking the per-subband PMI into account.

In this work, we consider the following scenarios: wideband CQI and wideband PMI (corresponds to Mode 1-1 in [3GPP, 2016]); wideband CQI and subband PMI (Mode 1-2). During the wideband CQI and wideband PMI transmission, one single precoding matrix is estimated for the whole set of subbands, and the CQI is calculated taking it into account. If the wideband CQI and the subband PMI feedback is configured, the preferred precoder matrix is estimated for each subband, and a single CQI value is calculated using a per-subband precoding. We do not consider the subband CQI transmission, as it is required for the effective scheduling of multiple users: the base station allocates the particular subband to the user who reported the highest CQI; it is thus not of particular interest in our setup, where the full bandwidth is dedicated to the single user.

7.4.2 Numerical Results

To verify our CQI estimation methodology and compare different feedback modes, we perform numerical simulations. In our settings, we consider a 2×2 MIMO system where the base station always transmits two codewords on two spatially multiplexed layers. We generate 1000 channel realizations for 300 Resource elements (25 Physical Resource Blocks) and a set of pre-transmitted SNR values $\gamma = 1/N_0 \in \{-5, 0, \dots 40\}$ dB. The SINR to CQI mapping was performed based on Table 7.2.

First we compare the per-stream CQI values reported by our PIA and SIC receiver in two joint feedback modes: wideband CQI/wideband PMI, and wideband CQI/subband PMI (Fig. 7.7). Given the same channel conditions (pre-transmitted SNR γ and channel matrix), first streams of both PIA and SIC receiver report the same CQI, since they are treated equally. At the same time, the second stream of the SIC receiver reports CQI values that are 2-4 positions higher than on the second stream of the PIA receiver,

Algorithm 15 Joint Wideband CQI and Wideband PMI Adaptation for the PIA and the SIC Receivers

Input: $\mathbf{H}$, N_0.
Output: CQI_l.

for each Resource Element $k = 1, ...K$ **do**
  compute and accumulate $\rho_{10} = \mathbf{h}_{1,k}^H \, \mathbf{h}_{0,k} + \rho_{10}$.
end for
Evaluate ρ_{10} and select preferred PMI and precoding matrix $\mathbf{P}$.
for each Resource Element $k = 1, ...K$ **do**
  Obtain $\mathbf{H}_{\text{eff}\,k} = \mathbf{H}_k \mathbf{P}$.
end for
for each stream $0 \leq l \leq 1$ **do**
  for each Resource Element $k = 1, ...K$ **do**
    compute an individual value of the $\text{SINR}_{l,k} = \mathcal{S}(\mathbf{H}_{\text{eff}k}, N_0, \text{rec. arch})$.
  end for
  Obtain the effective $\text{SINR}_{\text{eff},l}$ using EESM mapping functions.
  Find the corresponding CQI_l value from Table 7.2.
end for

Algorithm 16 Joint Wideband CQI and Subband PMI Adaptation for the PIA and the SIC Receivers

Input: $\mathbf{H}$, N_0.
Output: CQI_l.

for each subband $s = 1, ...S$ **do**
  for each Resource Element $k = 1, ...K_s$ **do**
    compute and accumulate $\rho_{10,s} = \mathbf{h}_{1,s,k}^H \, \mathbf{h}_{0,s,k} + \rho_{10,s}$.
  end for
  Evaluate $\rho_{10,s}$ and select preferred PMI and precoding matrix $\mathbf{P}_s$.
end for
for each subband $s = 1, ...S$ **do**
  for each Resource Element $k = 1, ...K$ **do**
    Obtain $\mathbf{H}_{\text{eff},s,k} = \mathbf{H}_{s,k} \mathbf{P}_s$.
  end for
end for
Concatenate per-subband effective channel matrices into one vector $\mathbf{H}_{\text{eff}}$
for each stream $0 \leq l \leq 1$ **do**
  for each Resource Element $k = 1, ...K$ **do**
    compute an individual value of the $\text{SINR}_{l,k} = \mathcal{S}(\mathbf{H}_{\text{eff}k}, N_0, \text{rec. arch})$.
  end for
  Obtain the effective $\text{SINR}_{\text{eff},l}$ using EESM mapping functions.
  Find the corresponding CQI_l value from Table 7.2.
end for

since it is approximated by the interference-free receiver. However, the per-stream CQI values for both feedback modes are identical. The CQI granularity is non-uniform and

varies between 0.5 to 2 dB (Fig. 7.6), which makes capturing the precoder reporting harder.

The obtained CQI values are mapped onto the corresponding MCS (Fig. 7.8, left) values using Table 7.2. The MCS values for the SIC receiver predictably reach high levels. However, these levels are lower the ones obtained by our MCS optimization strategy in Chapter 4 at every particular value of the pre-transmitted SNR, which leads to throughput degradation (Fig. 7.8, right). This is a result of a few factors. First, our MCS optimization in Chapter 4 was performed in a joint manner for both codewords, which took into account the influence of the modulation order and rate of the interfering codeword. Second, our EESM methodology without calibration factors brings some estimation errors, which become significant with the increase of the MCS values.

7.5 Conclusion

In this chapter, we have provided a methodology to compute CSI feedback for our Reduced Complexity Parallel Interference Aware receiver and Successive Interference Aware Canceling receiver. We propose to first estimate the RI, and then perform joint estimation of the PMI and the CQI.

The RI estimation based on condition number of channel matrix revealed that our receivers achieve high throughput values from spatial multiplexing transmissions even in channels with high correlation. This finding remains valid for the whole set of available MCS values.

The proposed light weight PMI adaptation, which maximizes the SINR on the first stream, is based on the evaluation of the correlation coefficient of the estimated channel matrix. This method is straightforward to implement, and provides mutual information levels close to the optimum maximum Mutual Information criterion.

Our CQI methodology is based on approximation of our PIA and SIC receivers with MMSE and Interference-free receivers. The empirical results revealed that for the SIC receiver the reported CQI values and the level of predicted throughput are higher than for the PIA receiver given the same channel conditions. In Chapter 4 we have proposed an MCS optimization strategy as an analog of a Slow Rate Adaptation technique. Our methodology, presented in this chapter, predicts lower levels of throughput than the MCS optimization strategy since it does not exploit the knowledge of the interference. Another reason for the throughput degradation is fusion of the CQI granularity and non-negligible errors in the effective SINR estimation. However, the loss due to CQI errors can usually be compensated by HARQ retransmissions [Yu et al., 2010]. Our CQI estimation methodology should thus be studied in the scenarios with HARQ protocol support. Another potential strategy to verify would be to avoid taking AWGN curves into consideration and instead apply the results of the MCS optimization as a reference.

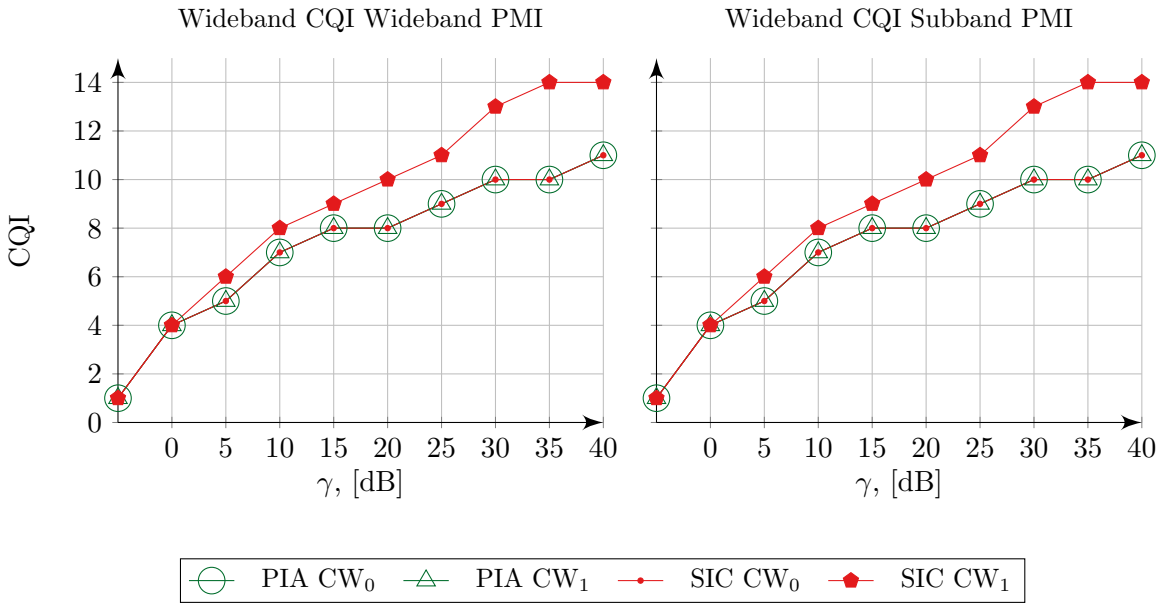


Figure 7.7: Per stream CQI values for joint wideband CQI/wideband PMI reporting (left) and wideband CQI/subband PMI reporting (right) for our PIA and SIC receivers. Since the CQI granularity varies between 0.5 to 2 dB, the influence of subband PMI is too small to be captured.

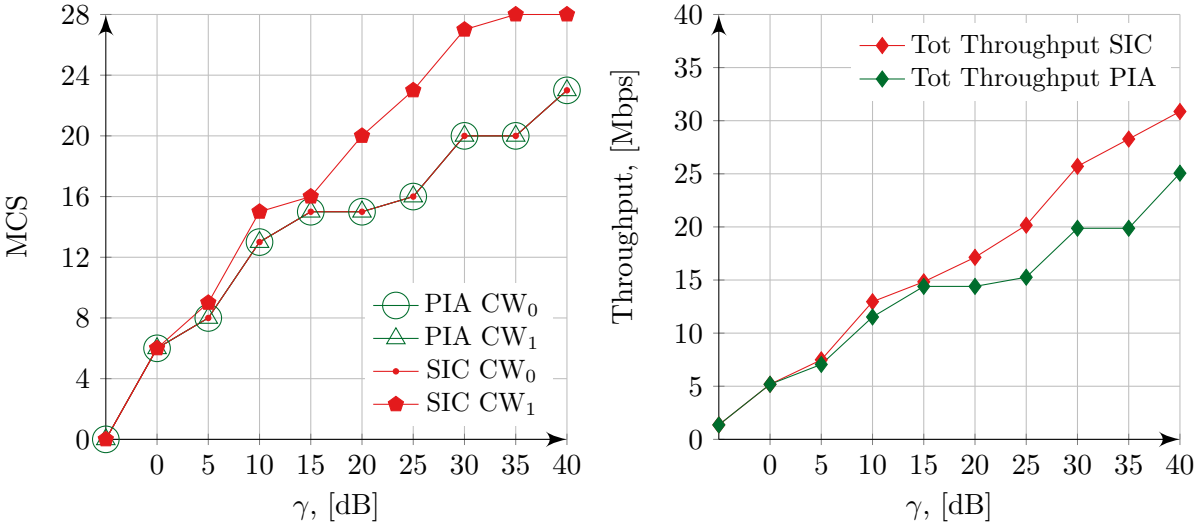


Figure 7.8: Per-stream MCS values (left), corresponding to obtained CQI values, and the achieved total throughput for our PIA and SIC receivers.

Chapter 8

Conclusions and Future Research

8.1 Conclusions

In this thesis, we have studied low complexity alternatives to joint ML detection in Single-User MIMO systems. Precisely, we have implemented and made compatible with the practical LTE MIMO systems two reduced complexity receiver architectures: the ML Parallel Interference Aware (PIA) detection and the ML Successive Interference Aware Canceling (SIC) detection.

We initially expected that the SIC receiver would outperform the PIA detection in high and medium SNR regimes, thanks to the interference-free detection of the second codeword. At the same time, it also seemed logical that our PIA receiver would enjoy some benefits in the low SNR region since the probabilities of successful decoding of both codewords are decoupled from each other. In contrast, the decoding attempt for the second codeword in the SIC receiver could be potentially blocked by the erroneous detection of the first codeword. However, both the information theoretical analysis and the empirical throughput results, developed in Chapter 4, have demonstrated that the SIC receiver is never in fact outperformed by the PIA detection and is a feasible solution in all SNR regimes. The obtained optimal MCS levels confirmed that the second stream of the SIC detection is capable of carrying higher data rates than the first stream.

The outcomes of the computational effort quantification are also positive for our SIC receiver, as we have shown in Chapter 4. Precisely, the single-thread implementation of our SIC receiver is 25% more computationally efficient for high modulation orders than the PIA receiver. This gain was enabled thanks to replacing the *interference aware* LLR metric for the second codeword with the SIC block and *interference-free* LLR metric.

To make receivers compatible with practical LTE systems, we had to revisit essential LTE protocols such as the Incremental Redundancy HARQ and the Link Adaptation. The HARQ protocol was discussed in Chapter 5 from both the base station processing and the UE processing points of view. Special care was given to multi-round SIC procedures, available at the SIC receiver.

To improve the performance levels claimed in Chapter 4, we then sought a light weight Link Adaptation solution (Chapter 7). Our PMI estimation methodology is based on maximizing the SNR level of the first stream. This approach demonstrated performance levels close to the results obtained with the optimum maximum Mutual Information (MI)

based criterion. The Rank estimation based on a condition number threshold revealed that contrary to intuition, both our receivers achieve the highest throughput if the base station always schedules two codewords. To perform the CQI estimation we contemplated on a low complexity abstraction methodology that would provide a fast estimate of the effective SNR and then map it onto CQI. The initial solution in Chapter 6 — MIESM — was based on look up tables and provided acceptable accuracy, if calibrated. Unfortunately, the size of the required tables made it unfeasible for real-time systems. Nevertheless, it is possible to use these look up tables in simulators to speed-up link-level computations. To circumvent the drawbacks of the MIESM method, we evaluated a straightforward EESM abstraction. The second solution showed satisfying results both when calibration coefficients are available or absent.

8.2 Future Perspective

In this thesis, we have integrated our receivers into MIMO systems based on LTE standard. The general concept of PIA and SIC detection can be applied to 5G with some modifications.

Although the standardization phase of New Radio networks (3GPP Release 14 and Release 15) is still ongoing and some features are not fully defined yet, it is already known that one of the baseline MIMO schemes will be Single-User and Multi-User spatial multiplexing. According to 3GPP Study Item on New Radio Physical Layer [3GPP, 2017], if the number of spatial layers is limited to four, only one codeword is used for PDSCH transmission, while 5 to 8-layer transmission allows spatially multiplexed transmission of two independent codewords. The mapping principle follows the LTE mapping [3GPP, 2015c]. These mapping scenarios will cause the interference between the symbols belonging to the same codeword, but mapped onto different spatial layers. This interference must be treated properly as the Gaussian assumption would significantly degrade the performance.

One potential solution would be to extend the existing dual-stream LLR metrics [Ghaffar and Knopp, 2010a] to take into account four interfering symbols. However, this is an extremely challenging task, which will most certainly not pass the computational effectiveness test described in Chapter 4. A more perspective solution are advanced receiver architectures that reduce the 4×4 channel to 2×2 interfering blocks. One of these architectures is a practically feasible Block QR decomposition [Thomas et al., 2014]. Once one of the symbols is decoded, the SIC procedure can be launched. However, if the SIC detection is applied, the probability of error propagation is not negligible. In the scenarios considered in the scope of this thesis, error propagation was circumvented by the Cyclic Redundancy Checks (CRC). However, since the eNodeB performs layer mapping after the CRC bits are added, this trick might not be available in 5G scenarios. Instead, the probability of error propagation can be minimized by applying intelligent ordering techniques. The LLR-based ordering criteria, proposed by [Kim and Kim, 2006] is unfeasible due to extremely high complexity since it would require to estimate LLR levels for each stream, and start detection with the stream possessing the highest level. Instead, low complexity ordering based on the norm of the columns of the channel matrix and post-

processed SINR[Cho et al., 2010] is perhaps worth investigating for scenarios with only one codeword. These criteria normally are not valid for the baseline scenario of this thesis, where two codewords with different rates are spatially multiplexed, and the higher SINR levels do not necessary guarantee higher probability of successful detection. However, the scenarios with a single codeword require further investigation.

Another direction of the research is physical layer abstraction for the messages with a short codeword length. This scenario falls into Ultra Reliable Low Latency Communication (URLLC) use case of the 5G technology, also characterized by a bursty traffic nature and a non-Gaussian interference distribution. The existing abstraction techniques might not be applicable (the MIESM is currently based on infinite codeblock length expressions for MI), and the future abstraction methodologies could potentially be built upon the results developed Verdu [2012], Polyanskiy et al. [2010], Polyanskiy et al. [2011]. In addition, a short codelength means the size of the control information (metadata) is comparable with the size of the user data, and the BLER needs to be estimated jointly for both control and data channels.

Chapter 9

Résumé Français

Nous présentons dans cette thèse une évaluation pragmatique d'alternatives faible complexité à la détection conjointe optimale du maximum de vraisemblance (Maximum Likelihood – ML) dans les sytèmes mono-utilisateur (Single-User – SU) entrées multiples, sorties multiples (Multiple-Input Multiple-Output – MIMO) avec transmissions spatialement multiplexées. La faible complexité est obtenue grâce à l'utilisation de métriques dites "interférence-aware" ML proposées par Ghaffar et Knopp. Nous concentrons notre étude sur deux architectures de récepteurs faible complexité : parallèle interférence-aware (Parallel Interference Aware – PIA) et annulation successive d'interférences (Successive Interference Canceling – SIC). Pour être déployable dans des modems réels, l'architecture d'un récepteur doit remplir certains critères de rapidité de temps de calcul et être évaluée dans des simulateurs ou émulateurs conformes aux normes. Nos simulations et expériences sont réalisées dans le simulateur downlink OpenAirInterface (OAI) que nous avons grandement développé au cours de ces travaux.

Dans la première partie de cette thèse, nous présentons une étude comparative complète de nos récepteurs PIA et SIC, en mettant l'accent sur leurs implémentations, les régimes des rapports signal/bruit (Signal-to-noise Ratio – SNR) et les schémas modulation/codage (Modulation and Coding Scheme – MCS) optimaux, et sur l'effort computationnel. Une analyse théorique des performances potentielles, soutenue par un large éventail de résultats pratiques, démontre que notre récepteur SIC surpasse le récepteur PIA en termes de débit et d'effort computationnel. Cette amélioration du temps de calcul est obtenue en remplaçant l'utilisation coûteuse de métriques interférence-aware décisions douces (log likelihood ratio – LLR) d'un des flux spatiaux par l'utilisation plus affiace d'un bloc SIC avec métriques sans interférence. Ces résultats sont particulièrement importants car la consommation de temps pour le traitement d'un signal est l'un des indicateurs de performance les plus significatifs et est un critère de déploiement dans les modems réels.

Pour rendre nos récepteurs compatibles avec de véritable systèmes LTE MIMO, nous avons revisité, dans une seconde partie de cette thèse, les protocoles essentiels de Long Term Evolution (LTE) tels que le protocole de requête de répétition automatique hybride à redondance incrémentale (Hybrid Automatic Repeat Request – HARQ) et le protocole d'adaptation de lien. En titant profit de la structure particulière de notre récepteur SIC, nous montrons comment la procédure SIC multi-ronde peut être utilisée pour effectuer le traitement de la retransmission à la station mobile.

L'estimation des paramétres en temps réel pour les transmissions adaptatives est l'un aspect difficile des procédures des récepteurs. Les principaux facteurs limitants sont causés par signaling overhead, les seuils d'adaptations de canaux, et la présence de canaux feedback imparfaits. Bien que d'importants efforts aient été dédiés pour limiter l'influence de ces facteurs, le compromis optimal entre complexité et performance reste un problème ouvert. Nous proposons une estimation faible complexité de trois composants: l'indicateur de rang (Rank Indicator – RI), l'indicateur de matrice précodeur (Precoder Matrix Indicator – PMI), et l'indicateur de qualité de canal (Channel Quality Indicator – CQI). Les résultats obtenus en utilisant notre méthode d'estimation RI révèle que nos deux récepteurs atteignent un début maximal lorsque la station de base effectue une transmission spatialement multiplexée, même dans des cannaux mal conditionnés. Le fossé entre les niveaux de performance de notre estimation PMI et la stratégie optimale basée sur le critère d'information mutuelle maximale est négligeable.Notre estimation CQI, basée sur des techniques d'abstraction de EESM ne permet pas d'atteindre la performance optimale. Néanmoins, les erreurs CQI peuvent être réparées avec l'utilisation de retransmissions HARQ.

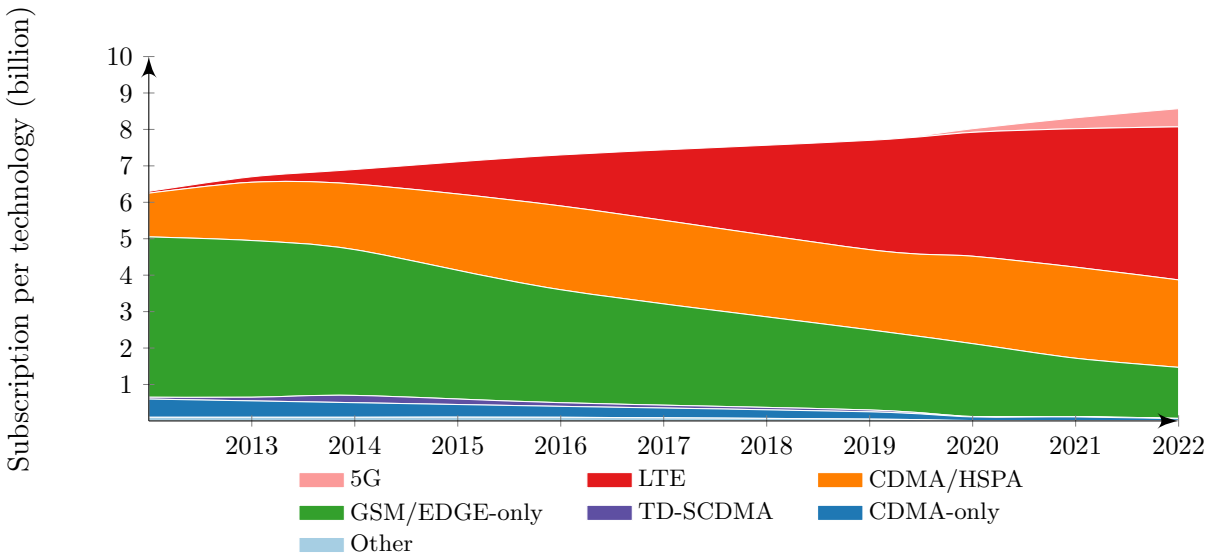


Figure 9.1: Abonnement mobile par technologie en milliards, selon [Ericsson, 2016].

9.1 Chapitre 1 · Introduction

Les communications sans fil font désormais partie intégrante de la vie des gens. En effet, l'utilisation de liaisons radio entre la station de base et la station mobile a supprimé la nécessité d'être à proximité géographique du terminal fixe pour accéder aux services en introduisant le concept de *mobilité*.

En 2009, le projet de partenariat de troisième génération (3GPP) a introduit la quatrième génération (4G) de technologie mobile avec une interface radio à commutation de paquets *Long Term Evolution* (LTE). La technologie LTE a gagné en popularité grâce à ses débits de pointe élevés et révolutionnaires, sa faible latence et sa qualité de service améliorée. [Ericsson, 2016] prévoit que la technologie LTE sera la plus utilisée d'ici 2019, avec 4,6 milliards de clients.

La spécificité de la détection MIMO mono-utilisateur est que, les deux mots de passe étant porteurs de données utiles, un récepteur pour être efficace doit appliquer des techniques avancées pour les récupérer avec la plus grande probabilité. La conception des récepteurs pour les systèmes MIMO peut être basée sur des filtres linéaires, tels que zero-forcing (ZF) ou l'erreur quadratique moyenne (Minimum Mean Square Error – MMSE), appelés récepteurs *linéaires*; ou appliquer des techniques non-linéaires telles que les métriques à maximum de vraisemblance (Maximum Likelihood – ML) ou l'annulation d'interférence successive (Successive Interference Canceling – SIC), les récepteurs de ce type sont dits *non-linéaires*.

Le sujet principal de cette thèse concerne les alternatives de faible complexité à la détection ML conjointe dans les systèmes MIMO mono-utilisateur. Précisément, nous avons étudié deux architectures de récepteurs à complexité réduite : la détection ML parallèle interférence-aware (Parallel Interference Aware Successive Interference Canceling – PIA) et la detection ML interférence-aware annulation d'interférence successive (Interference-

Aware – IA-SIC).

Pour rendre nos récepteurs compatibles avec les systèmes LTE MIMO utilisés pratiques, nous revisitons les protocoles LTE essentiels tels que le protocole hybrid automatic repeat request (HARQ) et l'adaptive modulation et coding (Link Adaptation – LA).

9.2 Chapitre 2 · Notion

Le processus d'extraction des données d'un signal reçu corrompu par des distorsions de bruit et des interférences est appelé *détection*. Le détecteur reçoit une observation du signal reçu et, suivant un certain critère, doit calculer une estimation du signal transmis. Le critère de détection joue un rôle de premier plan dans l'architecture du récepteur : il définit les comportements du système, la sensibilité aux perturbations et aux distorsions du bruit, ainsi que la complexité du calcul.

Considérons un modèle de signal MIMO en liaison descendante classique:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{9.1}$$

où $\mathbf{y}$ est le vecteur de signal reçu sur les antennes n_{rx} de le terminal mobile (User Equipment – UE), $\mathbf{x}$ est le vecteur de signal transmis depuis n_{tx} antennes de la station de base (eNodeB), $\mathbf{H}$ est une matrice de canal et $\mathbf{n}$ est un bruit Gaussien. Nous discutons de la configuration où les symboles transmis $\mathbf{x}$ sont mappés sur un alphabet connu $\mathcal{Q}^M$ de l'ordre de modulation $M \in \{2, 4, 6\}$ pour 4QAM, 16 QAM et 64 QAM respectivement.

L'idée derrière la *détection linéaire* est de transformer le canal MIMO en un ensemble de sous-canaux SISO parallèles en éliminant les interférences entre les couches. Cela peut être effectué en utilisant des filtres linéaires basés sur un pseudo-inverse de canal, tels que ZF ($\mathbf{W}_{\mathrm{ZF}}$) et MMSE ($\mathbf{W}_{\mathrm{MMSE}}$) [Onggosanusi et al., 2002]:

$$\mathbf{W}_{\mathrm{ZF}} = \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H, \tag{9.2}$$

$$\mathbf{W}_{\mathrm{MMSE}} = \left(\mathbf{H}^H\mathbf{H} + \sigma_{\mathrm{n}}^2\mathbf{I}\right)^{-1}\mathbf{H}^H, \tag{9.3}$$

où σ_{n}^2 est la variance du bruit. Les récepteurs linéaires se caractérisent par une complexité de calcul faible et un niveau de performance acceptable par rapport aux récepteurs optimaux.

Les récepteurs MIMO non-linéaires optimaux ou quasi-optimaux offrent un gain de performance élevé, mais nécessitent souvent une complexité de calcul prohibitive. Les principaux représentants des détecteurs non-linéaires sont des récepteurs basés sur la métrique de distance minimale, tels que les récepteurs ML optimaux, leurs approximations et leurs variations, et les récepteurs SIC.

Les détecteurs ML MIMO classiques effectuent une recherche exhaustive sur tous les candidats vectoriels possibles, et le *candidat le plus probable* se situe à la distance minimale du vecteur reçu:

$$\widehat{\mathbf{x}} = \underset{\mathbf{x}\in\mathcal{Q}^{\mathrm{M}}}{\arg\min} \left\|\mathbf{y} - \mathbf{H}\mathbf{x}\right\|^2. \tag{9.4}$$

L'architecture sensible aux interférences des récepteurs avancés apporte des gains significatifs aux performances du système. Si les symboles interférents proviennent d'un

alphabet fixe, comme c'est le cas des systèmes LTE, le récepteur peut deviner l'ordre de modulation du mot de code interférent (ou acquérir cette connaissance) et l'appliquer pendant le processus de détection [Ghaffar and Knopp, 2010a] .

Le récepteur SIC est un autre représentant de la famille des récepteurs non-linéaires. Chaque filtre de la banque réceptrice détecte un des flux de données, qui est ensuite itérativement soustrait du signal restant, ce qui résulte en un signal sans interférence à la sortie de l'étape finale [Wolniansky et al., 1998].

9.2.1 Méthodologies d'abstraction PHY

Abstraction de couche physique est un outil utile pour la prévision des performances du système. L'idée derrière ce concept est de cartographier les conditions de canal instantanées aux indicateurs de performance. Le canal instantané est généralement représenté par le SNR et les gains de canal, et les performances sont souvent mesurées en termes de *Taux d'erreur sur les blocs* (Block Error Rate – BLER).

La technique de mappage est le cœur de l'abstraction, car elle définit le niveau de précision avec lequel la performance est quantifiée. Dans les environnements sélectifs en fréquence, chaque sous-porteuse OFDM réalise différents gains de canal. Il serait trop long de considérer les estimations des indicateurs de performance (par exemple, le débit) sur une base par sous-porteuse. Pour contourner ce défi, le concept de *rapport signal sur bruit efficace* $\mathrm{SNR}_{\mathrm{eff}}$ a été introduit par [Nanda and Rege, 1998]. Le SNR effectif est un SNR équivalent dont le signal bénéficierait dans le canal AWGN et provoque le même BLER. Grâce à cette propriété, le processus de mappage est souvent appelé *fonction de compression*, où le vecteur de SNR_k par sous-porteuse k est condensé en une seule valeur $\mathrm{SNR}_{\mathrm{eff}}$. Le principe du mappage consiste à calculer la valeur de $\mathrm{SNR}_{\mathrm{eff}}$ de manière à minimiser l'erreur (au carrée) entre $\mathrm{SNR}_{\mathrm{eff}}$ et le l'équivalent $\mathrm{SNR}_{\mathrm{AWGN}}$ dans le canal AWGN généré pour le même schéma de modulation et de codage.

Les deux principales approches pour le mappage sont le *mappage exponentiel du rapport signal sur bruit efficace* (Exponential Effective SNR Mapping – EESM) [Kim et al., 2011; Sandanalakshmi et al., 2007; Stancanelli et al., 2011] et le *mappage du rapport signal sur bruit efficace des informations mutuelles* (Mutual Information Effective SNR Mapping – MIESM) [Olmos et al., 2010; Srinivasan et al., 2008]. L'approche MIESM est la plus difficile car elle implique des calculs d'Information Mutuelle par sous-porteuse.

9.3 Chapitre 3 · Modèle de signal et cadre de simulation

Le département Systèmes de communication d'EURECOM est le fondateur d'OpenAirInterface (OAI) [OAI, 2017] — une plate-forme de prototypage expérimental open source unique de 4G – 5G de logiciel/matériel qui offre la flexibilité d'implémentation de logiciels tout en assurant la conformité aux normes 3GPP. L'OAI offre la possibilité d'effectuer des expériences en simulation, en émulation et en mode temps réel. Plusieurs simulations unitaires pour la couche physique sont disponibles pour reproduire le comportement du transport et des canaux physiques dans les transmissions en liaison montante et descendante. L'un d'eux, le simulateur pour les canaux DLSCH et PDSCH, est utilisé comme

plate-forme expérimentale pour la conception du récepteur dans cette thèse. Nos modèles théoriques sont implémentés en langage de programmation `C` et les résultats des simulations au niveau du lien sont utilisés pour valider les méthodologies proposées.

9.3.1 Simulateur de liaison descendante

Notre simulateur de liaison descendante `dlsim` modélise les comportements de liaison du système LTE pendant la transmission de liaison descendante entre la station de base et l'UE. Il est logiquement divisé en trois blocs fonctionnels : les procédures de la station de base, la génération du milieu de propagation et la convolution avec le signal transmis, et les procédures UE.

Le traitement du signal de la couche physique suit strictement les normes 3GPP [3GPP, 2015b,c, 2016]. Dans certains modes de transmission, la rétroaction d'informations d'état des canaux (Channel State Information – CSI) est disponible.

9.3.2 Modèles analytiques

9.3.2.1 Modélisation des canaux

Les simulations sont effectuées à l'aide du bruit blanc Gaussien additif (Additive White Gaussian Noise – AWGN), de l'évanouissement à plat et de l'évanouissement sélectif des fréquences basé sur les modèles de canaux de Rayleigh, Rician et EPA [3GPP, 2015d].

9.3.2.2 Modélisation du signal

Mono-utilisateur MIMO est le sujet principal de cette thèse. Ce mode de transmission a été introduit dans la version LTE 8. Il est utilisé pour maximiser le débit d'un seul utilisateur, et est donc conçu pour l'environnement et les scénarios SINR élevés, où l'UE est proche de la station de base.

Considérons un système MIMO de taille 2×2 fonctionnant en mode de transmission 4 avec le système de transmission CLSM (Close Loop Spatial Multiplexing). La station de base envoie deux blocs de transport TB_0 et TB_1 mappés sur les mots de code spatialement multiplexé CW_0 et CW_1. Le vecteur de signal reçu $\mathbf{y}_k \in \mathbf{C}^{2\times 1}$ pour l'élément de ressource k-ième vu par l'UE est donné par :

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{P}_k \mathbf{x}_k + \mathbf{n}_k, \quad k = 1, 2..., K, \tag{9.5}$$

où $\mathbf{x}_k$ est le vecteur transmis de deux symboles complexes x_0 et x_1 avec variance de σ_0^2 et σ_1^2. Le vecteur transmis appartient à l'alphabet $\mathcal{Q}^{M_0,M_1}$, tel que $\mathcal{Q}^{M_0,M_1} := \mathcal{Q}^{M_0} \times \mathcal{Q}^{M_1}$ est le produit cartésien de deux alphabets de modulation $\mathcal{Q}^{M_0}$ et $\mathcal{Q}^{M_1}$, et $M_0, M_1 \in \{2, 4, 6\}$ sont les ordres de modulation des constellations QAM.

Le vecteur $\mathbf{n}_k$ est un bruit Gaussien de densité spectrale de puissance double face $N_0/2$ à deux antennes de réception d'UE. La matrice $\mathbf{H}_k$ est une matrice de 2×2 construite par rapport à l'un des modèles de canaux, décrits dans la section précédente, et $\mathbf{P}_k$ est la matrice de précodage utilisée par le eNodeB au k-ième élément ressource.

Outre le scénario décrit, le mode de transmission 4 peut être utilisé dans une configuration avec un seul mot de code actif. On a alors :

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{p}_k x_k + \mathbf{n}_k, \tag{9.6}$$

où $\mathbf{p}_k$ est un vecteur précodeur pour la transmission à couche unique, et x_k est un symbole de transmission du mot de code unique CW_0 tiré de l'alphabet $\mathcal{Q}^{M_0}$.

9.4 Chapitre 4 · Architecture des récepteurs R-ML

La détection interférence-aware multi-canal MIMO mono-utilisateur se divise en trois groupes : ML classique, parallèle interférence-aware R-ML (Parallel Interference Aware Successive Interference Canceling – PIA) et annulation successive d'interférences (Successive Interference Canceling – SIC). Dans ce chapitre, nous discutons de nos récepteurs PIA et SIC et quantifions l'effort de calcul pour le traitement du signal. L'analyse théorique des limites de performance est suivie d'une évaluation du débit empirique obtenu via des simulations au niveau des liaisons.

9.4.1 Mise en oeuvre des récepteurs

Nous considérons un scénario de multiplexage spatial TM4 en boucle fermée. Le modèle de signal est identique à (9.5). Par souci de simplicité, nous supprimons l'index élément ressource et remplaçons la multiplication de $\mathbf{H}$ et $\mathbf{P}$ par le canal effectif $\mathbf{H}_{\text{eff}}$:

$$\mathbf{y} = \mathbf{H}_{\text{eff}}\mathbf{x} + \mathbf{n}, \quad \mathbf{H}_{\text{eff}} = [\mathbf{h}_{\text{eff0}}\mathbf{h}_{\text{eff1}}]. \tag{9.7}$$

Les schémas de blocs illustrant l'architecture des récepteurs R-ML PIA et R-ML SIC sont respectivement représentés dans Fig. 9.2 et Fig. 9.3. Les récepteurs traitent le premier mot de code de la manière habituelle, mais traitent le second mot de code différemment. Nous décrivons les blocs de traitement des signaux communs et détaillons ensuite les procédures individuelles. Les deux récepteurs exploitent des métriques de LLR de faible complexité [Ghaffar and Knopp, 2010a].

9.4.1.1 Blocs IA R-ML communs

Après le filtrage correspondant (9.7), les récepteurs PIA et SIC fournissent un traitement identique au premier mot de code de débit inférieur en calculant les valeurs de LLR_0 en utilisant le bit IA `M0_M1_llr` [Ghaffar and Knopp, 2010a], qui traite le second mot de code comme une interférence:

$$\mathbf{y}_{\text{MF}} = \frac{1}{\sqrt{|h_{\text{av}}|^2}} \mathbf{H}_{\text{eff}}^H \mathbf{y}, \tag{9.8}$$

où $|h_{\text{av}}|^2$ est le niveau moyen du canal. Les valeurs LLR souples sont ensuite transmises à la procédure de débrouillage et par conséquent au décodeur turbo.

Si le premier mot de code est décodé correctement, le récepteur SIC déclenche la procédure SIC. Cependant, aucune tentative n'est faite pour décoder le deuxième mot

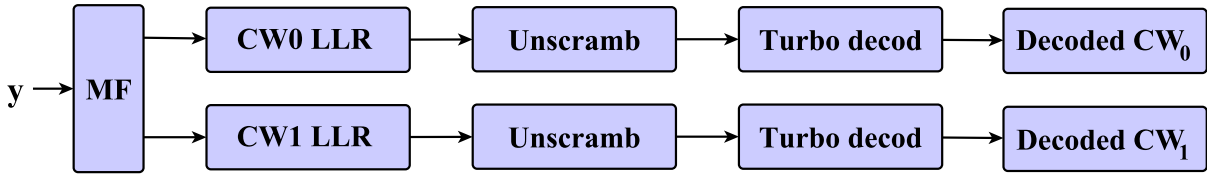


Figure 9.2: Schéma de bloc du récepteur R-ML PIA. Les deux mots de code sont traités de la même manière en utilisant des métriques de bits de faible complexité IA.

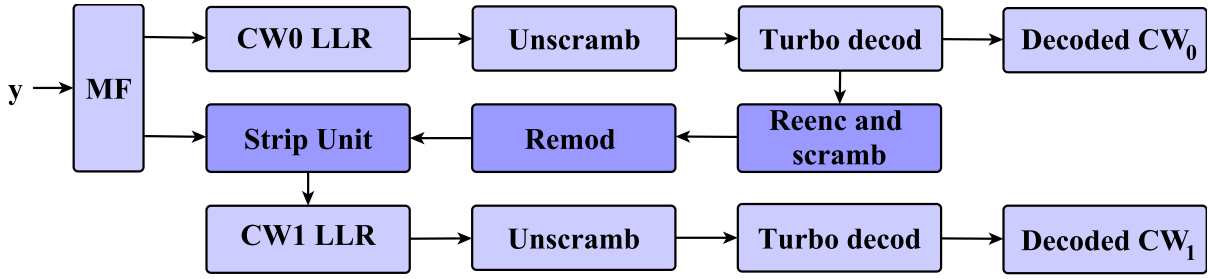


Figure 9.3: Schéma de bloc du récepteur R-ML SIC. Le bloc SIC effectue le recodage et le brouillage du signal brouilleur précédemment décodé, la remodulation, l'unité de soustraction, le calcul LLR en utilisant des métriques de LLR de faible complexité, le débrouillage et le décodage turbo.

de code si les premiers mots de code sont erronés. En revanche, pour le récepteur PIA (Fig. 9.2), les valeurs LLR pour le premier et le second mots de code sont calculées en parallèle à l'aide du fonctions `M0_M1_llr` et `M1_M0_llr` [Ghaffar and Knopp, 2010a], et sont transmises à deux turbo-décodeurs indépendants. La probabilité que le deuxième mot de code soit décodé avec succès ne dépend donc pas de la décodabilité du premier mot de code.

9.4.1.2 Procédure SIC

La procédure SIC commence par le recodage de la séquence de bits récemment décodée du premier mot de code et par leur mappage sur les symboles de modulation x_0 (voir Fig. 9.3). Le signal reçu compensé sur la deuxième antenne est donné par

$$y_{\mathrm{MF}_1} = \left(\rho^* x_0 + \|\mathbf{h}_{\mathrm{eff1}}\|^2 x_1\right) + n_1', \tag{9.9}$$

où le terme de bruit $n_1' = \mathbf{h}_{\mathrm{eff1}}^H n_1$ et le coefficient de corrélation $\rho^* = h_{\mathrm{eff01}}^* h_{\mathrm{eff00}} + h_{\mathrm{eff11}}^* h_{\mathrm{eff10}}$. Après multiplication par x_0, obtenu par le décodage réussi du premier mot de code, avec le coefficient de corrélation ρ^* et soustraction du résultat de y_{MF_1} dans l'unité de soustraction, x_1 bénéficie d'une détection sans interférence :

$$\widetilde{y}_{\mathrm{MF}_1} = \|\mathbf{h}_{\mathrm{eff1}}\|^2 x_1 + n_1'. \tag{9.10}$$

Les valeurs LLR pour le décodage du deuxième mot de code peuvent maintenant être calculées à l'aide de la métrique légère sans interférence à l'aide de la fonction `M1_llr`.

9.4.2 Effort de calcul

Pour être pratiquement réalisable pour les transmissions en temps réel, la conception du récepteur doit être efficace en termes de calcul. Dans un vrai modem LTE, le rapport ACK/NACK doit être généré dans la troisième sous-trame après la réception des données, ce qui donne une fenêtre de 2 ms pour traiter les données. En utilisant des mesures en temps réel, nous vérifions que nos récepteurs répondent à cette exigence. Les mesures sont effectuées en utilisant un thread d'une machine à 64 bit avec un processeur à 2.10 GHz et une mémoire de 8 GB.

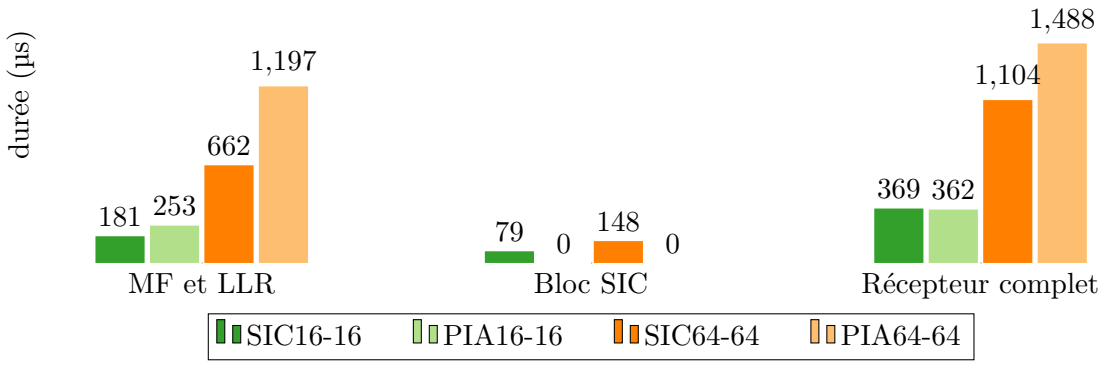


Figure 9.4: Temps consommé par nos récepteurs SIC et PIA pour traiter les données en liaison descendante en fonction des constellations sur lesquelles les mots de code sont mappés.

Les mesures prennent en compte les calculs MF et LLR, le débrouillage, le bloc SIC et le décodage. La durée de traitement des sous-trames est moyenné sur 10000 sous-trames. Notre récepteur SIC est 25% plus efficace pour les ordres de modulation élevés (Fig. 9.4). Nos deux récepteurs prennent moins de 1.5 ms pour traiter une sous-trame de liaison descendante dans une bande passante de 5 MHz, et peuvent donc être déployés dans des systèmes réels. Pour une bande passante supérieure, nous proposons d'utiliser le multi-threading.

9.4.3 Analyse des informations mutuelles

Pour les systèmes LTE pratiques où le vecteur transmis appartient à un alphabet discret fini, il n'y a pas d'expression de forme fermée pour l'information mutuelle (Mutual Information - MI). Au lieu de cela, elle peut être approximée numériquement via des simulations type Monte-Carlo. Suivant la règle de la chaîne MI, le MI total du système MIMO avec le décodage conjoint I_{ML} ML peut être décomposé en I_0 et I_1, sans compromettre les performances:

$$\underbrace{I\left(X_0, X_1; \mathbf{Y}|\mathbf{H}_{\mathrm{eff}}\right)}_{I_{\mathrm{SIC}}=I_{\mathrm{ML}}} = \underbrace{I\left(X_0; \mathbf{Y}_{\mathrm{MF}}|\mathbf{H}_{\mathrm{eff}}\right)}_{I_0} + \underbrace{I\left(X_1; \mathbf{Y}_{\mathrm{MF}}|X_0, \mathbf{H}_{\mathrm{eff}}\right)}_{I_1}, \tag{9.11}$$

où I_0 représente le MI conditionnel entre le vecteur de signal reçu $\mathbf{Y}_{\mathrm{MF}}$ et le symbole transmis X_0, étant donné la connaissance de la matrice de canaux estimée $\mathbf{H}_{\mathrm{eff}}$. Le second composant I_1 représente le MI conditionnel entre le vecteur de signal reçu $\mathbf{Y}_{\mathrm{MF}}$ et le

symbole transmis X_1, étant donné la connaissance non seulement de la matrice estimée du canal, mais aussi du symbole X_0. A la suite de cette note, le récepteur SIC ne subit aucune perte d'information par rapport aux limites de détection communes ML de [Foschini and Gans, 1998]. Cependant, ce n'est pas le cas pour le principe PIA, où le récepteur connaît l'ordre de modulation du premier mot codé mais ne possède pas d'informations précises sur X_0 :

$$\underbrace{I\left(X_0, X_1; \mathbf{Y}|\mathbf{H}_{\text{eff}}\right)}_{I_{\text{ML}}} \leq \underbrace{I\left(X_0; \mathbf{Y}_{\text{MF}}|\mathbf{H}_{\text{eff}}\right)}_{I_0} + \underbrace{I\left(X_1; \mathbf{Y}_{\text{MF}}|\mathbf{H}_{\text{eff}}\right)}_{I_{1\text{PIA}}}. \tag{9.12}$$

9.4.4 Résultats pratiques

Nous examinons le débit empirique obtenu à la suite des simulations au niveau des liaisons à l'aide du simulateur de liaison descendante OAI [OAI, 2017]. Le simulateur offre la flexibilité de modifier les paramètres de transmission LTE et de personnaliser l'environnement de propagation.

9.4.4.1 Débit empirique et optimisation MCS

Dans un système LTE réel, le débit est limité par le taux R défini par le MCS [3GPP, 2016]. Les valeurs de débit sont calculées à partir des statistiques BLER de notre simulateur, acquises à chaque point SNR:

$$T_{\text{tot, sim}}(R_0, R_1, \text{SNR}) = T_0(R_0, R_1, \text{SNR}) + T_1(R_0, R_1, \text{SNR}). \tag{9.13}$$

Les valeurs de débit pour le premier mot de code T_0 et le second mot de code T_1 sont obtenues comme suit:

$$\begin{aligned} T_0(R_0, R_1, \text{SNR}) &= R_0(1 - BLER_0(R_0, R_1, \text{SNR})), \\ T_1(R_0, R_1, \text{SNR}) &= R_1(1 - BLER_1(R_0, R_1, \text{SNR})), \end{aligned} \tag{9.14}$$

où $BLER_0$ et $BLER_1$ sont les taux d'erreur de bloc correspondants pour les premier et deuxième mots de code. Une analyse comparative des performances dans les évanouissements de Rayleigh et de Rician (Fig. 9.5) révèle que les gains du récepteur SIC dépendent de la présence de ligne de mire. Pour le canal rician, les gains du récepteur SIC sont plus importants et atteignent 4 Mbit/s contre 1.8 Mbit/s en évanouissement de Rayleigh. Les gains obtenus devraient évoluer avec la bande passante.

9.4.4.2 Pratique et attentes théoriques

Pour quantifier l'impact des imperfections du récepteur, des erreurs d'arrondi et de l'implémentation du point fixe sur nos simulations au niveau des liens, nous effectuons une comparaison entre le débit empirique et les attentes théoriques fournies par l'analyse MI sous les hypothèses idéalistes.

Fig. 9.6 illustre l'écart entre les résultats empiriques de notre réception SIC. Les valeurs de débit $T^{\star}_{\text{tot, sim}}$ et $T^{\star}_{\text{tot, up-b}}$ sont très proches lorsque l'UE est dans le régime de faible rapport signal-bruit et utilise des modulations 16QAM-16QAM ou 16QAM-64QAM. Cependant, dès que le récepteur passe au régime SNR élevé, où une 64QAM est utilisée pour

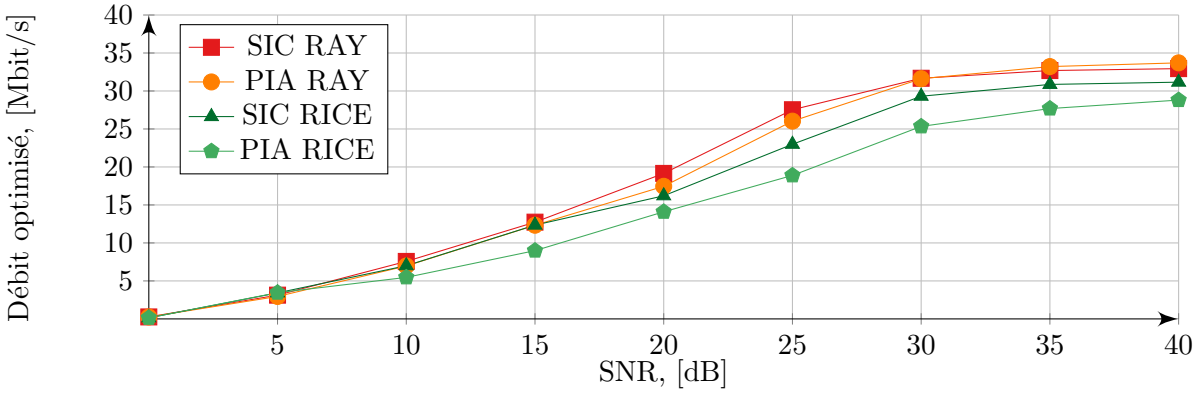


Figure 9.5: Le débit optimisé $T^{\star}_{\text{tot, sim}}$ pour les récepteurs SIC et PIA avec une estimation de canal parfaite dans les environnements à évanouissements Rayleigh et Rician. Le récepteur SIC gagne jusqu'à 4 Mbit/s en régime de SNR modéré et élevé, alors que dans le canal Rayleigh, cet écart n'est pas si important et permet d'atteindre 1.8Mbit/s.

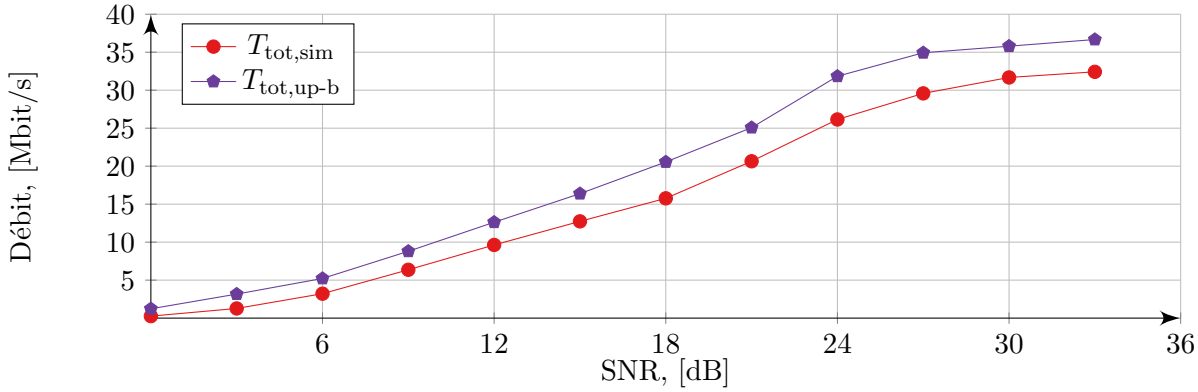


Figure 9.6: Débit empirique et limite supérieure pour notre récepteur SIC dans le canal Rayleigh. Les valeurs de débit $T^{\star}_{\text{tot, sim}}$ et $T^{\star}_{\text{tot, up-b}}$ sont très se ferme lorsque l'UE est dans le régime de SNR bas où 16QAM-16QAM ou 16QAM-64QAM est appliqué. Cependant, l'écart augmente dans la région 64QAM - 64QAM.

les deux mots de passe, l'écart entre le débit réel et le débit prévu augmente. Cela peut s'expliquer par le fait que nous modélisons MI pour la longueur de bloc de code infinie, alors que la longueur de bloc de code est en fait limitée par la taille du bloc de transport dans un système LTE réel. Peut-être que certains résultats de [Polyanskiy et al., 2010] pourraient fournir plus de renseignements.

9.4.5 Conclusion

Sur la base des simulations au niveau des liens obtenues à partir des expériences de notre simulateur de liaison descendante, le récepteur SIC surpasse le récepteur PIA de 4 Mbit/s en évanouissements russes et de 1.8 Mbit/s en évanouissement plat Rayleigh à 5 MHz LTE bande passante. Les gains devraient évoluer avec la bande passante.

La quantification de l'effort de calcul démontre que, pour les ordres de modulation élevés, notre récepteur SIC est 25% plus rapide que le récepteur PIA, grâce au remplacement de la métrique IA LLR du second mot de code par le bloc SIC léger. Pour les ordres de modulation modérés, l'effort de calcul est approximativement du même niveau.

Les principales contributions de ce chapitre sont reflétées dans la publication suivante:

- E. Lukashova, F. Kaltenberger, and R. Knopp, "Reduced Complexity ML Interference-Aware Parallel and SIC Receivers for LTE SU-MIMO," in 14th International Symposium on Wireless Communication Systems (ISWCS 2017), Bologna, Italy, Aug. 2017.

9.4.6 Chapitre 5 · HARQ Protocol pour les R-ML Récepteurs

Dans ce chapitre, nous implémentons d'abord la prise en charge protocole hybride à demande automatique de répétition (Hybrid Automatic Repeat Request – HARQ) pour le système MIMO mono-utilisateur avec les récepteurs PIA et SIC dans notre simulateur de liaison descendante. Le traitement du protocole au niveau de l'UE est spécifique au destinataire et fait l'objet d'une discussion pour chacun des récepteurs. Une attention particulière est consacrée à la procédure SIC à plusieurs tours, disponible pour notre récepteur SIC : une fois le premier mot de code décodé, il est possible de reconstruire le deuxième mot de code sans interférence à partir de *tous* les tours précédents et de combiner les rapports de vraisemblance logarithmique (Log-likelihood Ratio – LLR). Deuxièmement, nous quantifions les performances de nos récepteurs avec le protocole HARQ en termes de débit et de fiabilité. Enfin, nous testons différents schémas de retransmission dans le cas où un mot de code est décodé avec succès et un mot de code erroné.

9.4.7 Protocoles de retransmission

Dans les systèmes MIMO mono-utilisateur avec deux mots de code multiplexés spatialement, le protocole HARQ peut suivre l'un des scénarios de retransmission, illustrés dans Fig. 9.7.

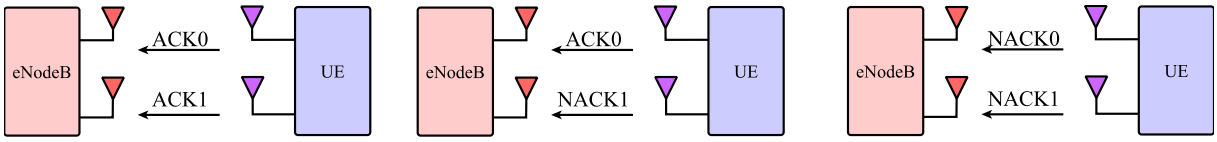


Figure 9.7: Scénarios de retransmission possibles pour le système MIMO mono-utilisateur : la station de base reçoit deux accusés de réception (Acknowledgment – ACK) et aucune retransmission n'est pas nécessaire; un ACK et un NACK sont reçus et un seul mot de code est retransmis; la station de base reçoit un double NACK et deux mots de code multiplexés spatialement sont retransmis.

Lors de la transmission initiale ou lorsque la station de base reçoit deux messages sans accusé de réception NACK_0 et NACK_1, l'eNodeB transmet deux blocs de transport TB_0 et TB_1 mappés sur des mots de code CW_0 et CW_1:

$$\mathrm{TB}_0 \longmapsto \mathrm{CW}_0, \quad \mathrm{TB}_1 \longmapsto \mathrm{CW}_1.$$

Le réglage lorsqu'un seul des mots de code est décodé (ce qui correspond au deuxième scénario de Fig. 9.7) a retenu notre attention car la norme ne spécifie pas explicitement comment la station de base doit effectuer une retransmission dans ce cas. La retransmission des deux mots de code conduirait à une utilisation non efficace des ressources spectrales. La station de base désactive ainsi le bloc de transport correspondant pour les prochains tours de retransmission et ne retransmet que le mot de code erroné.

9.4.7.1 Protocole HARQ pour le Récepteur PIA

À la réception des deux mots de code, l'UE lance la procédure **Traitement PIA à double mot de code** (Algorithm 17). Étant donné le vecteur de signal reçu $\mathbf{y}$, les estimations de canal et le TPMI (disponible auprès de la DCI), l'UE effectue la détection PIA en fonction des métriques LLR interférence-aware `M0_M1_llr` et `M1_M0_llr` proposé par [Ghaffar and Knopp, 2010a]. Les valeurs LLR calculées sont ensuite transmises à deux turbo décodeurs indépendants. En fonction des résultats du processus de décodage, l'UE génère des messages d'accusé de réception et de non-acquittement qui sont ensuite envoyés à la station de base.

Algorithm 17 Traitement PIA à double mot de code

procedure Traitement PIA à double mot de code($\mathbf{y}$, estimations de canal, TPMI)
    démoduler les deux mots de code en appliquant une matrice de précodeur correspondant à TPMI.
    calculer LLR_0 et LLR_1, combiner avec les LLR de l'étape précédente, si disponible.
    envoyer LLR_0 et LLR_1 à turbo-decodeurs.
    if TB_0 est décodé && TB_1 et décodé **then**
        $\text{ACK}_0 \leftarrow true$, $\text{ACK}_1 \leftarrow true$
        transmettre les messages ACK_0 et ACK_1
        break
    else if TB_0 est décodé && TB_1 n'est pas décodé **then**
        $\text{ACK}_0 \leftarrow true$, $\text{ACK}_1 \leftarrow false$
        transmettre ACK_0 et NACK_1
    else if TB_0 n'est pas décodé && TB_1 est décodé **then**
        $\text{ACK}_0 \leftarrow false$, $\text{ACK}_1 \leftarrow true$
        transmettre NACK_0 et ACK_1
    else if TB_0 n'est pas décodé && TB_1 n'est pas décodé **then**
        $\text{ACK}_0 \leftarrow false$, $\text{ACK}_1 \leftarrow false$
        transmettre NACK_0 et NACK_1
    end if
end procedure

Si l'un des mots de code est décodé alors qu'un autre est en erreur, la station de base reçoit un message d'accusé de réception et un message de non-acquittement. Elle désactive ensuite le bloc de transport décodé et bascule vers une transmission par bloc de transport mappée sur le premier mot de code conformément à la norme. Après avoir apporté les modifications correspondantes dans la DCI, l'UE effectue l'accomodation du taux pour le bloc de transport toujours actif en utilisant une nouvelle version de redondance, modulation et codage. L'eNodeB peut imposer le précodage, tel que le précodage Alamouti, ou utiliser la première ou la deuxième colonne de la matrice de précodage $\mathbf{P}$

suggérée par l'UE dans le rapport CSI. Le choix du précodeur eNodeB est codé dans le DCI.

Après le décodage de la DCI, l'UE comprend qu'un seul mot de code a été transmis et déclenche la procédure **Traitement PIA à mot de code unique** (Algorithm 18). Les valeurs calculées des LLR sont ensuite combinées aux valeurs des LLR du tour précédent et sont transmises au turbo-décodeur. Sur la base des résultats de la procédure de décodage, l'UE génère des messages d'accusé de réception ACK ou de non-acquittement NACK et les envoie à la station de base.

Algorithm 18 Traitement PIA à mot de code unique

procedure Traitement PIA à mot de code unique(**y**, estimations de canal, TPMI, TB_0_flag, TB_1_flag)
  if TB_0_flag == enabled **then**
    démoduler CW_0 en appliquant le vecteur de précodage correspondant à TPMI.
    calculer LLR_0 et combiner avec les LLR des tours précédents.
    transmettre LLR_0 au turbo-decodeur.
    if TB_0 est décodé **then**
      $ACK_0 \leftarrow true$
      transmettre ACK_0
      break
    else if TB_0 n'est pas décodé **then**
      $ACK_0 \leftarrow false$
      transmettre $NACK_0$
    end if
  else if TB_1_flag == enabled **then**
    démoduler CW_0 en appliquant le vecteur de précodeur correspondant à TPMI.
    calculer LLR_1 et combiner avec les LLR des tours précédents.
    transmettre LLR_1 au turbo-decodeur.
    if TB_1 est décodé **then**
      $ACK_1 \leftarrow true$
      transmettre ACK_1
      break
    else if TB_1 n'est pas décodé **then**
      $ACK_0 \leftarrow false$
      transmettre $NACK_1$
    end if
  end if
end procedure

9.4.7.2 Protocole HARQ pour le Récepteur SIC

La principale différence entre le traitement HARQ avec les récepteurs SIC et PIA provient du lien étroit qui existe entre la décodabilité du deuxième mot de code et le décodage réussi du premier mot de code. Cela signifie que la station de base est configurée pour la transmission en double mot codé jusqu'au moment où le premier mot codé est décodé.

À la réception des données de liaison descendante, l'UE décode la DCI et appelle la procédure **Traitement SIC à double mot de code** (Algorithm 19). Cette procédure

reflète l'architecture SIC, où la décodabilité du deuxième mot de code dépend du décodage réussi du premier mot de code.

Algorithm 19 Traitement SIC à double mot de code

procedure Traitement SIC à double mot de code(**y**, estimations de canal, TPMI)
    démoduler les deux mots de code en appliquant la matrice de précodage correspondant à TPMI.
    calculer LLR_0 et combiner avec les valeurs LLR du tour précédent, si disponibles.
    transmettre LLR_0 au turbo-decodeur.
    if TB_0 est décodé **then**
        call **procedure Traitement Multi-Tour SIC**
        if TB_1 est décodé **then**
            $\mathrm{ACK}_0 \leftarrow true$, $\mathrm{ACK}_1 \leftarrow true$
            transmettre ACK_0 et ACK_1
            break
        else
            $\mathrm{ACK}_0 \leftarrow true$, $\mathrm{ACK}_1 \leftarrow false$
            transmettre ACK_0 et NACK_1
        end if
    else
        $\mathrm{ACK}_0 \leftarrow false$, $\mathrm{ACK}_1 \leftarrow false$
        transmettre NACK_0 et NACK_1
    end if
end procedure

Algorithm 20 Traitement Multi-Tour SIC

procedure SIC(r_{dec}, MF outputs from previous rounds)
    $r_{\mathrm{sic}} \leftarrow 0$
    while $0 \leq r_{\mathrm{sic}} \leq r_{\mathrm{dec}}$ **do**
        Reencodez la séquence de bits décodée du CW_0.
        Effectuer une correspondance de taux avec $\mathrm{RV}_0 = r_{\mathrm{dec}}$.
        Effectuer un mappage sur les symboles de modulation x_0.
        Multiplier x_0 avec ρ_r^*.
        Donner $\rho_r^* x_0$ à l'unité soustractive pour obtenir sans interférence CW_1.
        Calculer $\mathrm{LLR}_{1r_{\mathrm{sic}}}$ et combiner avec LLR_1 des étapes précédentes.
        Donner LLR_1 au turbo-decodeur.
        if CW_1 est décodé **then**
            $\mathrm{ACK}_0 \leftarrow true$, $\mathrm{ACK}_1 \leftarrow true$.
            transmettre ACK_0 et ACK_1.
            break
        else
            $r_{\mathrm{sic}} \leftarrow r_{\mathrm{sic}} + 1$
        end if
    end while
    $\mathrm{ACK}_0 \leftarrow true$, $\mathrm{ACK}_1 \leftarrow false$.
    transmettre ACK_0 et NACK_1.
end procedure

Contrairement à la détection PIA, où la démodulation est suivie du calcul LLR pour les deux mots de code, le récepteur SIC calcule initialement les valeurs LLR_0 uniquement

pour le premier mot de code. S'il ne s'agit pas du premier tour de transmission, les valeurs LLR_0 sont mises à jour avec les valeurs correspondantes des tours précédents. Les valeurs LLR sont ensuite envoyées au turbo-décodeur. Si le premier mot de code est décodé avec succès (le tour $r = r_{\mathrm{dec}}$), l'UE déclenche la procédure **Traitement Multi-Tour SIC**, détaillée dans Algorithm 20. C'est important que le récepteur conserve le coefficient de canal et le signal reçu compensé afin que le deuxième mot de code puisse être reconstruit dans tous les tours précédents.

L'UE tente de décoder TB_1 à chaque tour après la combinaison LLR. Ceci est fait pour éviter de croiser les tours suivants de la procédure SIC à tours multiples si les valeurs de LLR sont déjà suffisamment fiables.

9.4.8 Résultats numériques et discussion

9.4.8.1 Paramètres de simulation

Pour nos simulations empiriques au niveau des liaisons, nous avons utilisé les paramètres suivants: une largeur de bande LTE de 5 MHz a été planifiée pour un seul utilisateur avec deux antennes de réception. L'environnement de propagation a été simulé à l'aide du modèle de canal à évanouissements de Rayleigh et du modèle de canal Extended Pedestrian A (EPA). La retransmission d'un mot codé unique est simulée avec trois options de précodage: TPMI0 (précodage Alamouti), TPMI5 (la première colonne de la matrice de précodage suggérée dans le rapport CSI) avec le CSI réel, et TPMI6 (la deuxième colonne) avec CSI obsolète.

9.4.8.2 Optimisation MCS et analyse du débit pour les tours multiples

Pour quantifier les gains de performance résultant des tours de retransmission, nous examinons si les tours multiples génèrent des gains de débit et si ces gains dépendent du TPMI utilisé lors de la transmission d'un mot de code unique. Fig. 9.8 illustre le débit total du système $T^{\star}_{\mathrm{tot,sim}}$ après quatre tours HARQ et le débit $T_{00,\mathrm{sim}} + T_{10,\mathrm{sim}}$ après le premier tour obtenu avec $\mathrm{MCS}^{\star}_0$ et $\mathrm{MCS}^{\star}_1$ pour SIC et détection de l'évaluation PIA dans le scénario avec modèle de canal EPA avec niveau de corrélation modéré (EPAM). Notre récepteur SIC surpasse le récepteur PIA au premier tour avec des gains allant de 2 à 4 Mbit/s en régime de SNR élevé à $5 - 7$ Mbit/s en régime de SNR faible. Les tours multiples de retransmission réduisent cet écart, mais le récepteur SIC fonctionne toujours mieux avec un SNR élevé (jusqu'à 2 ou 3 Mbit/s). Cependant, le TPMI lors de la retransmission d'un mot de code unique n'a pas d'impact significatif sur le débit.

9.4.8.3 Analyse de fiabilité

Les retransmissions multiples contribuent de manière significative à la fiabilité du système MIMO dans le régime SNR inférieur. Fig. 9.9 illustre le BLER du premier bloc de transport pour quatre tours HARQ.

Les deux récepteurs affichent des performances identiques pour le premier bloc de transport. Au niveau BLER de 10^{-1}, la combinaison $(12, 16)$ MCS reçoit des améliorations de 10 dB et de $13,5$ dB respectivement dans les canaux de Rayleigh et de l'EPA. Le gain

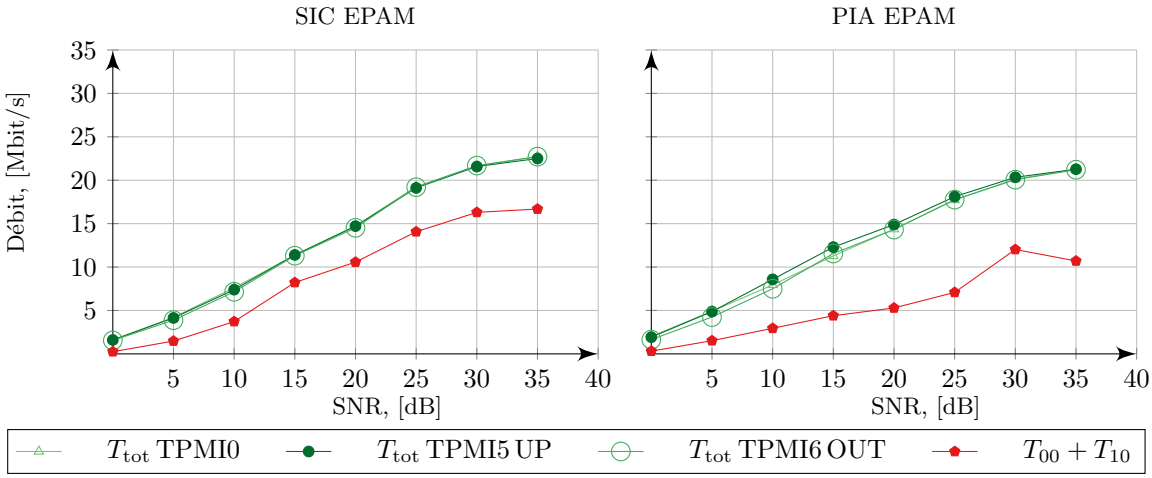


Figure 9.8: Comparaison du débit total des récepteurs SIC et PIA T_{tot} après 4 tours HARQ et du débit total $T_{00}+T_{10}$ pour deux mots de code après le premier tour $r = 0$ s'appliquant *optimisé* $\text{MCS}_0^\star$ et $\text{MCS}_1^\star$.

du SNR atteint jusqu'à 13 dB dans le canal de Rayleigh 28 dB dans le canal EPAM. Lors de la première transmission, notre récepteur SIC surpasse de loin la détection PIA grâce à la procédure SIC. Les tours consécutifs de retransmission apportent des avantages significatifs à la détection PIA, tandis que pour le récepteur SIC, le gain entre le troisième et le quatrième round de retransmission n'est pas remarquable (Fig. 9.10). Les deux récepteurs préfèrent le précodage Alamouti.

Conclusion Le récepteur SIC atteint un débit supérieur dans tous les scénarios d'évanouissement. Cependant, malgré la combinaison LLR à laquelle on accède via la procédure SIC à plusieurs tours, ses gains sont générés par les deux premiers tours de transmission, tandis que le récepteur PIA bénéficie clairement des quatre tours de retransmission. Notre analyse pour le format DCI 2 a montré que le précodage Alamouti était favorable aux retransmissions du mot de code unique dans TM 4 du point de vue de la fiabilité, mais il n'existait aucune préférence notable pour les schémas de retransmission du point de vue du débit. vue. Certaines parties de ce travail ont été publiées dans:

- E. Lukashova, F. Kaltenberger, and R. Knopp, "Single-User MIMO ML successive interference canceling receiver with HARQ-IR protocol," in 21st International ITG Workshop on Smart Antennas (WSA 2017), Berlin, Germany, March 2017.

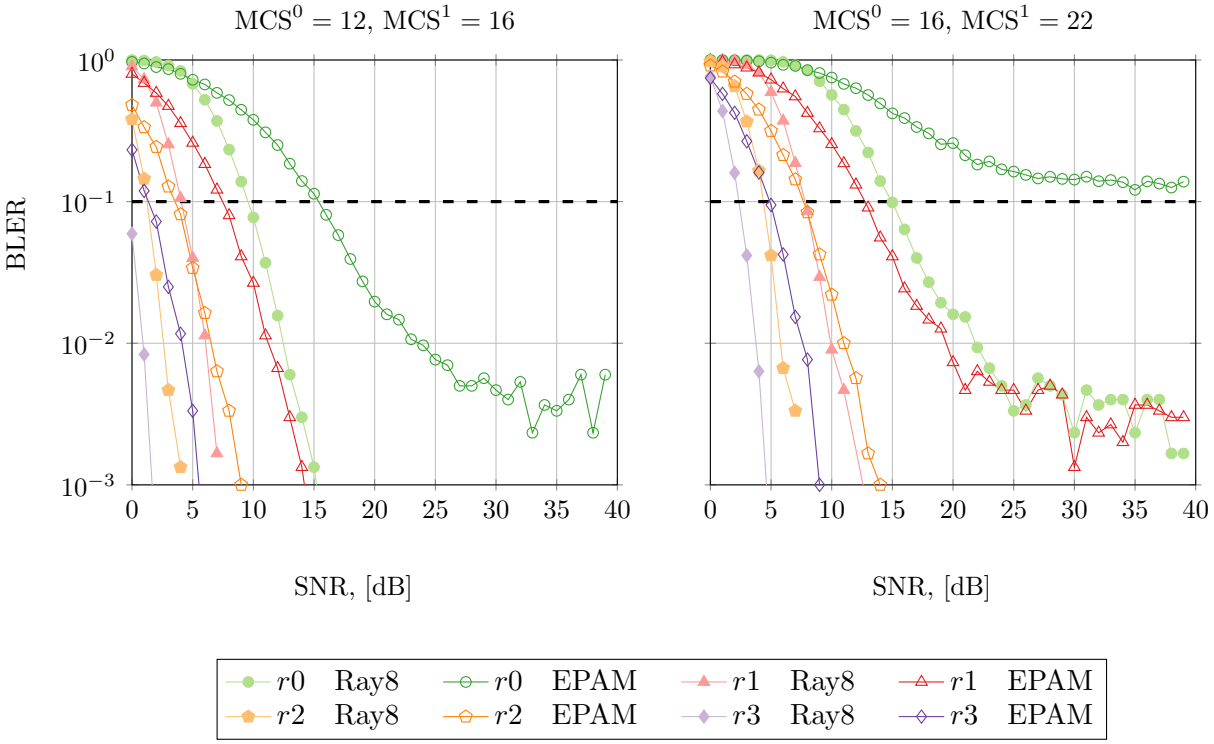


Figure 9.9: BLER de TB_0 à 8 -tap canaux Rayleigh et EPAM pour 4 tours HARQ pour $(MCS_0, MCS_1) \in \{(12, 16), (16, 22)\}$.

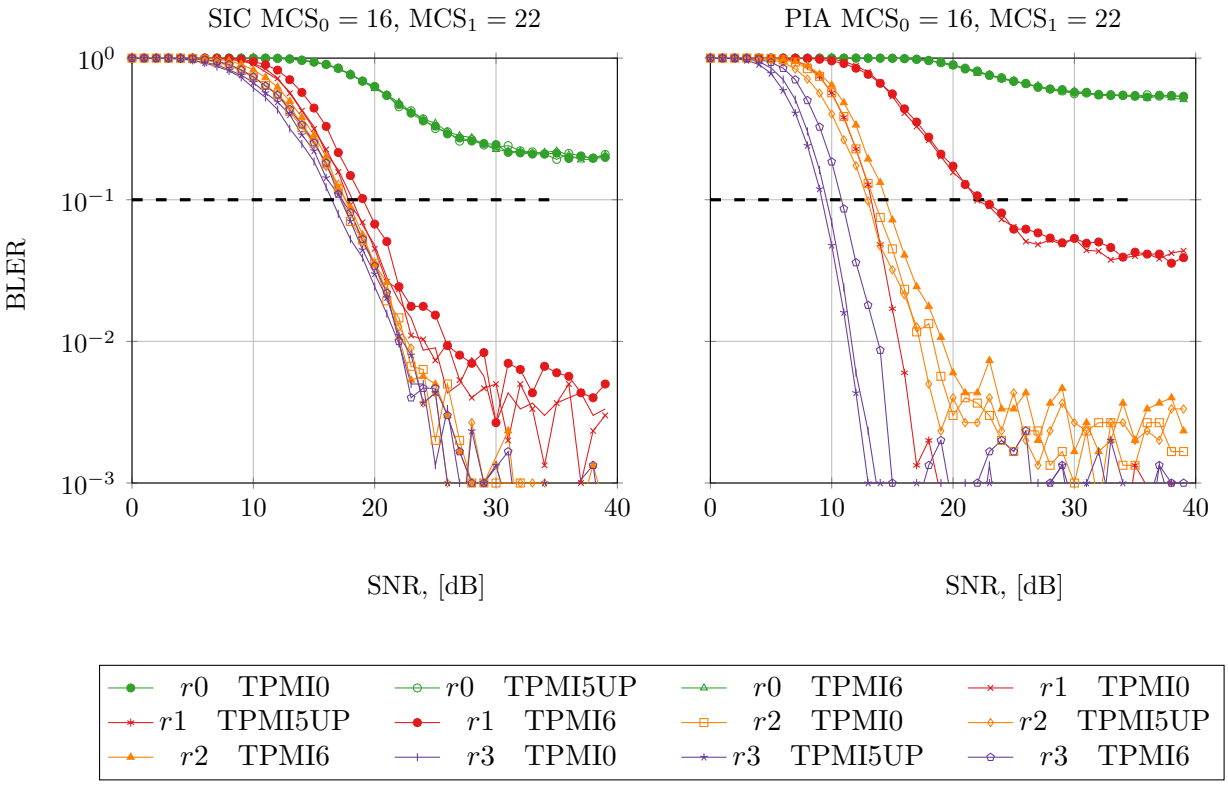


Figure 9.10: BLER de TB_1 dans le canal EPAM pour 4 tours HARQ pour nos récepteurs.

9.5 Chapitre 6 · Abstraction PHY pour les Récepteurs R-ML PIA et SIC

Dans ce chapitre, nous étudions les méthodologies d'abstraction pour les systèmes MIMO LTE mono-utilisateur utilisant nos récepteurs R-ML PIA et R-ML SIC. Nous détaillons les avantages et les inconvénients de notre méthodologie Mutuel d'information sur la Base de la Cartographie SINR Efficace (Mutual Information Effective SNR Mapping – MIESM) baseé dans les tables de consultation, ainsi que la faisabilité de son déploiement pour les transmissions en temps réel. Enfin, nous appliquons une méthodologie de la Cartographie Exponentielle Efficace SINR (Exponential Effective SINR Mapping – EESM) légère et approximons les performances par flux de nos récepteurs avec le récepteur l'erreur quadratique moyenne (Minimum Mean Square Error – MMSE) et le récepteur avec l'annulation parfaite des interférences (Interference-free – IF).

9.5.1 Description du Scénario

Nous considérons le scénario TM4 avec une transmission CLSM. L'eNodeB est équipé avec $n_{\text{tx}} = 2$ antennes et transmet deux mots de code à l'UE avec $n_{\text{rx}} = 2$ antennes de réception. Les mots de code CW_0 et CW_1 appartiennent à MCS_0 et MCS_1, avec les taux R_0 et R_1. Nous nous référons au mot de code plus faible R_0 comme CW_0, et CW_1 est toujours fourni avec un taux égal ou supérieur R_1.

Le vecteur de signal reçu $\mathbf{y}_k$ pour l'élément de ressource k observé par l'UE est donné par

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{P}_k\mathbf{x}_k + \mathbf{n}_k, \quad k = 1, 2, ..., K,$$

où $\mathbf{x}_k$ est le vecteur transmis de deux symboles complexes x_0 et x_1 avec une variance de σ_0^2 et σ_1^2. Le vecteur $\mathbf{n}_k$ est le bruit blanc gaussien, $\mathbf{H}_k$ est une matrice de canaux 2×2 et $\mathbf{P}_k$ est la matrice de précodage. Simplifiant davantage, on obtient :

$$\mathbf{y} = \mathbf{H}_{\text{eff}}\mathbf{x} + \mathbf{n}, \mathbf{H}_{\text{eff}} = [\mathbf{h}_{\text{eff0}}\mathbf{h}_{\text{eff1}}]. \tag{9.15}$$

Le traitement du signal commence par d'un filtre adapté (Matched Filter – MF) linéaire, commune aux récepteurs PIA et SIC. Le signal reçu $\mathbf{y}$ (9.15) est transformé en

$$\mathbf{y}_{\text{MF}} = \mathbf{H}_{\text{eff}}^H\mathbf{y} = \boldsymbol{\alpha}x_0 + \boldsymbol{\gamma}x_1 + \begin{bmatrix}\mathbf{h}_{0\text{eff}}^H & \mathbf{h}_{1\text{eff}}^H\end{bmatrix}^T\mathbf{n}, \tag{9.16}$$

où les termes $\boldsymbol{\alpha}$ et $\boldsymbol{\gamma}$ représentent les vecteurs des coefficients de canal groupés de la manière suivante:

$$\begin{aligned}\boldsymbol{\alpha} &= \begin{bmatrix}\alpha_0 & \alpha_1\end{bmatrix}^T = \begin{bmatrix}\mathbf{h}_{0\text{eff}}^H\mathbf{h}_{0\text{eff}} & \mathbf{h}_{1\text{eff}}^H\mathbf{h}_{0\text{eff}}\end{bmatrix}^T, \\ \boldsymbol{\gamma} &= \begin{bmatrix}\gamma_0 & \gamma_1\end{bmatrix}^T = \begin{bmatrix}\mathbf{h}_{0\text{eff}}^H\mathbf{h}_{1\text{eff}} & \mathbf{h}_{1\text{eff}}^H\mathbf{h}_{1\text{eff}}\end{bmatrix}^T.\end{aligned} \tag{9.17}$$

9.5.2 Mutuel d'information sur la Base de la Cartographie SINR Efficace

9.5.2.1 Méthodologie MIESM

Les étapes générales du MIESM pour les récepteurs PIA et SIC sont décrites dans Algorithm 21.

Algorithm 21 Algorithme d'abstraction MIESM PDSCH pour récepteurs PIA et SIC

Input: MCS_0, MCS_1, $\mathbf{H}_{\text{eff}}$
Output: SNR_{eff}.

1: **for** chaque flux $0 \leq l \leq L$ **do**
2: **for** chaque sous-porteuse $k = 1, ...K$ **do**
3: calculer une valeur individuelle de $I_{l,k}$ à l'aide de la fonction de mappage $\mathcal{I}_{\text{map}}$.
4: **end for**
5: Compresser plusieurs valeurs de $I_{l,k}$ en une valeur unique $I_{\text{compr},\,l}$.
6: Mapper $I_{\text{compr},\,l}$ à la valeur efficace $\text{SNR}_{\text{eff},\,l}$ en utilisant la fonction de mappage inversée $\mathcal{I}_{\text{map}}^{-1}$.
7: Trouver le $\text{BLER}_{\text{eff},l}$ estimé correspondant à $\text{SINR}_{\text{eff},l}$ précalculé dépendant de MCS courbes AWGN.
8: Calibrer les résultats en utilisant les coefficients d'ajustement MCS-dépendant$\beta_{0,l}$ et $\beta_{1,l}$.
9: **end for**

Le processus de cartographie est le cœur et le point le plus difficile de la méthodologie d'abstraction. Nous proposons d'utiliser des tables de consultation précalculées: statistiques de canal $||\boldsymbol{\alpha}_k||$, $||\boldsymbol{\gamma}_k||$, N_0 et $I_{l,k}$ correspondant peuvent être obtenus à l'aide d'expressions pour l'information mutuelle et de simulations Monte-Carlo pour un grand nombre de réalisations de canaux et de variances de bruit, puis stockées sous forme de matrices multidimensionnelles:

$$I_{\text{LUT}_{\text{SIC}_0,k}}(||\boldsymbol{\alpha}_k||, ||\boldsymbol{\gamma}_k||, M_0, M_1, N_0) = I\left(X_0; \mathbf{Y}_{\text{MF}} \middle| \boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k, M_0, M_1, N_0\right),$$

$$I_{\text{LUT}_{\text{PIA}_0,k}}(||\boldsymbol{\alpha}_k||, ||\boldsymbol{\gamma}_k||, M_0, M_1, N_0) = I\left(X_0; \mathbf{Y}_{\text{MF}} \middle| \boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k, M_0, M_1, N_0\right),$$

$$I_{\text{LUT}_{\text{SIC}_1,k}}(||\boldsymbol{\alpha}_k||, ||\boldsymbol{\gamma}_k||, M_0, M_1, N_0) = I\left(X_1; \mathbf{Y}_{\text{MF}} \middle| X_0, \boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k, M_0, M_1, N_0\right),$$

$$I_{\text{LUT}_{\text{PIA}_1,k}}(||\boldsymbol{\alpha}_k||, ||\boldsymbol{\gamma}_k||, M_0, M_1, N_0) = I\left(X_1; \mathbf{Y}_{\text{MF}} \middle| \boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k, M_0, M_1, N_0\right).$$

Pour chaque flux l, les valeurs par sous-porteuse de $I_{\text{LUT}_{l,k}}$ obtenues à partir des tables de consultation avec les statistiques du récepteur concerné sont moyennées, ce qui donne la valeur compressée $I_{\text{compr},\,l}$:

$$I_{\text{compr},l} = \frac{\sum_{k=1}^{K} I_{\text{LUT}_{l,k}}}{\beta_{0,l} K}, \tag{9.18}$$

où $\beta_{0,l}$ est le premier facteur d'ajustement pour compenser le taux de modulation et de codage.

Pour obtenir le SINR effectif, nous recherchons une relation directe entre $I_{\text{compr, }l}$ et $\text{SINR}_{\text{eff},l}$. Nous supposons qu'il existe un canal SISO équivalent à un modèle de signal (9.19) avec une tape et une information mutuelle moyenne $I_{\text{compr},l}$ (9.18):

$$\widetilde{y}_k = \widetilde{h}_k \widetilde{x}_k + \widetilde{n}_k, \tag{9.19}$$

où $\widetilde{x}_k$ est un symbole complexe reçu, $\widetilde{h}_k$ est un canal Rayleigh SISO d'évanouissement et $\widetilde{n}_k$ est un bruit gaussien blanc. Ensuite, une correspondance individuelle entre $I_{\text{compr, }l}$ et $\text{SINR}_{\text{eff},l}$ peut être obtenue par interpolation linéaire de (9.20) pour une valeur connue de $I_{\text{compr, }l}$. Les informations mutuelles pour le canal Rayleigh à prise unique sont calculées hors ligne pour tous les alphabets de modulation et stockées dans les tables de recherche.

$$I\left(X;Y|SNR,M\right) = \log M - \frac{1}{MN_hN_n}\Bigg(\sum_{x\in\mathcal{Q}^M}\sum_{c}^{N_h}\sum_{z}^{N_n}\log\frac{\sum_{x\in\mathcal{Q}^M}\exp(-SNR\ |y-hx'|^2)}{\exp(-|n|^2)}\Bigg), \tag{9.20}$$

où $SNR = \beta_{1,l}\text{SINR}_{\text{eff},l}$ et $\beta_{1,l}$ est le deuxième facteur d'ajustement pour compenser le taux de modulation et de codage.

En supposant que le SINR effectif dans un canal à évanouissements donne le même BLER que pour un canal AWGN, nous obtenons:

$$BLER_l(\mathbf{H}_{\text{eff}}, N_0, \text{MCS}_l) \approxeq BLER_{\text{AWGN}}(\beta_{1,l}\text{SINR}_{\text{eff},l}, \text{MCS}_l). \tag{9.21}$$

L'erreur carrée minimum entre $\text{SINR}_{\text{eff},l}$ et $\text{SINR}_{\text{AWGN},l}$ est un critère adéquat pour la formation:

$$\beta_{\text{opt}_0,l}, \beta_{\text{opt}_1,l} = \underset{\beta_{0,l},\beta_{1,l}}{\arg\min}\ \frac{1}{N_HN_n}\Bigg[\sum_{c}^{N_H}\sum_{z}^{N_n}|\text{SINR}_{\text{eff},l}(\beta_{0,l},\beta_{1,l},\text{MCS}_l) - \text{SINR}_{\text{AWGN},l}(\text{MCS}_l)|^2\Bigg], \tag{9.22}$$

où $\beta_{0,l}$ et $\beta_{1,l}$ sont les coefficients d'étalonnage.

9.5.2.2 Résultats de l'approche MIESM

Les traces utilisées lors de la formation d'abstraction ont été obtenues via notre simulateur de liaison descendante OAI `dlsim` pour une largeur de bande LTE de 5 MHz. Pour la simulation AWGN, le canal a été généré à l'aide de la matrice de corrélation spatiale, qui annule les interférences entre couches, et 10000 paquets ont été transmis. Pour la simulation sélective en fréquence, le canal d'atténuation de Rayleigh 8 tapes avec entrées indépendantes et identiquement distribuées et un écart de retard de 0.8 microsecondes ont été choisis, et 1000 paquets ont été transmis sur 200 réalisations de canal pour un large éventail de variances de bruit ciblant la BLER de 10^{-2}. Ces traces ont ensuite été utilisées pour obtenir les coefficients d'étalonnage stockés pour chaque MCS et peuvent être

Table 9.1: Résultats d'étalonnage pour l'abstraction MIESM LUT du récepteur PIA dans le canal Rayleigh à 8 tapes

MCS_0	MCS_1	$MSE_{LUT,0}$	$MSE_{LUT,1}$	$\beta_{LUTopt_0,0}$	$\beta_{LUTopt_0,1}$	$\beta_{LUTopt_1,0}$	$\beta_{LUTopt_1,1}$
4	4	0.0508	0.0571	2.9794	10.6384	0.2433	0.1608
10	10	0.3781	0.5707	0.6762	0.6772	0.4860	0.4315
12	12	0.8729	0.6986	0.8121	1.0228	0.6179	0.5991

Table 9.2: Résultats d'étalonnage pour l'abstraction MIESM LUT du récepteur SIC dans le canal Rayleigh à 8 tapes

MCS_0	MCS_1	$MSE_{LUT,0}$	$MSE_{LUT,1}$	$\beta_{LUTopt_0,0}$	$\beta_{LUTopt_0,1}$	$\beta_{LUTopt_1,0}$	$\beta_{LUTopt_1,1}$
4	4	0.0533	0.0926	2.6906	18.2130	0.2634	0.1176
10	10	0.3791	0.8706	0.7201	0.8827	0.4698	0.4274
12	12	0.7480	1.1566	0.9950	0.9311	0.5356	0.5275

utilisées pour toute réalisation de canal aléatoire avec la précision fournie dans Table 9.1, Table 9.2, Table 9.3.

Les résultats de la méthodologie MIESM basée sur LUT sont présentés dans Fig. 9.11. Les performances abstraites pour les ordres de modulation élevé ne correspondent pas aux courbes AWGN de manière aussi précise que pour les modulations inférieures. Ceci résulte de la perte d'informations de phase pendant la phase de construction des tables de consultation. La précision d'étalonnage de la méthode basée sur la table de conversion pour la constellation QPSK est vérifiée avec la cartographie MIESM directe. D'après une analyse comparative de Table 9.1, Table 9.3, une bonne correspondance est observée pour la constellation QPSK. Les résultats devraient s'améliorer si le nombre de réalisations de canaux utilisées en tant qu'entrée pour la validation d'abstraction est augmenté, ainsi que si les tables de consultation sont construites avec une précision plus élevée.

9.5.3 Exponentielle Efficace SINR

9.5.3.1 Méthodologie EESM

Il n'est pas connu précisement quelles techniques sont utilisées dans les jeux de puces de périphériques temps réel, mais il s'agit certainement de techniques de très faible complexité qui garantissent encore des niveaux satisfaisants d'estimation.

Nous montrons que même la simple architecture EESM fournit une précision suffisante

Table 9.3: Résultats d'étalonnage pour la méthodologie d'abstraction de cartographie MIESM directe pour le récepteur PIA dans un canal Rayleigh à 8 tapes

MCS_0	MCS_1	$MSE_{direct,0}$	$MSE_{direct,1}$	$\beta_{directopt_0,0}$	$\beta_{directopt_0,1}$	$\beta_{directopt_1,0}$	$\beta_{directopt_1,1}$
2	2	0.0364	0.0409	8.2438	4.9157	0.2300	0.2731
4	4	0.0217	0.0300	2.4298	2.3646	0.1839	0.1870
12	12	0.0904	0.1297	0.6829	0.7420	0.5465	0.5692

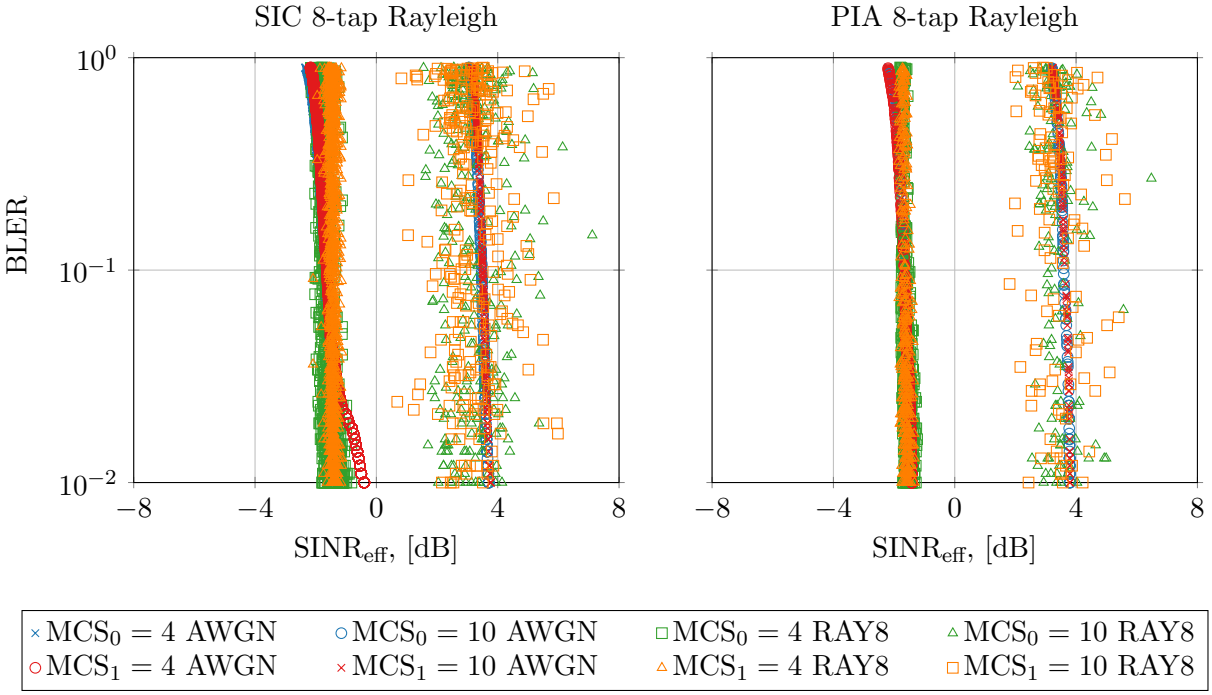


Figure 9.11: Validation MIESM pour nos récepteurs SIC et PIA utilisant $(\text{MCS}_0, \text{MCS}_1) \in \{(4,4),(10,10)\}$ dans le canal Rayleigh avec 8 tapes. Les facteurs d'étalonnage sont pris en compte.

pour être déployée dans des systèmes en temps réel et pour assurer une prévision de performance efficace. L'algorithme EESM suit les mêmes étapes principales que l'algorithme MIESM (Algorithm 21), mais nécessite une fonction de mappage simplifiée pour compresser le SINR de post-détection par sous-porteuse et pour mapper le valeur comprimée au SINR efficace.

Nous approximons les flux IA de nos récepteurs par le récepteur MMSE ((9.23)), ainsi que le flux sans interférences du récepteur avec l'annulation parfaite des interférences ((9.24)):

$$\text{SINR}_{l,k} = \frac{1}{[(\mathbf{I}_{n_{tx}} + \frac{1}{N_0 n_{\text{tx}}}\mathbf{H}_{\text{eff},k}^H \mathbf{H}_{\text{eff},k})^{-1}]_{ll}} - 1, \tag{9.23}$$

$$\text{SINR}_{\text{IF}_{\text{sic1}}} = \frac{1}{N_0 n_{\text{tx}}} \|\mathbf{h}_{\text{eff1},k}\|^2. \tag{9.24}$$

Après le mappage EESM, le *par couche* $\text{SNR}_{\text{eff},l}$ efficace est ensuite obtenu en compressant la sous-porteuse $\text{SINR}_{l,k}$ via la fonction de mappage $\mathcal{I}_{\text{map}}$:

$$\text{SINR}_{\text{eff},l}(\beta_1, \beta_2) = \beta_1 \mathcal{I}_{\text{map}}^{-1}\left[\frac{1}{K}\sum_{k=1}^{K} \mathcal{I}_{\text{map}}\left(\frac{\text{SINR}_{l,k}}{\beta_2}\right)\right],$$

où β_1, β_2 sont les paramètres d'ajustement dépendant du taux et $\mathcal{I}_{\text{map}}$ est une fonction de mappage calculée par l'union Chernoff liée à des probabilités d'erreur et facile à implémenter:

$$\mathcal{I}_{\text{map}}(\text{SINR}_{l,k}) = 1 - \exp(-\text{SINR}_{l,k}).$$

Table 9.4: Comparaison de la précision MIESM et EESM de la performance abstraite du récepteur PIA dans le canal Rayleigh à 8 tapes

MCS_0	MCS_1	$\mathrm{MSE}_{\mathrm{LUT},0}$	$\mathrm{MSE}_{\mathrm{EESM},0}$	$\mathrm{MSE}_{\mathrm{LUT},1}$	$\mathrm{MSE}_{\mathrm{EESM},1}$
4	4	0.0508	0.0741	0.0571	0.2597
10	10	0.3781	0.0490	0.5707	0.1402
12	12	0.8729	0.0744	0.6986	0.1399

Table 9.5: Comparaison de la précision MIESM et EESM de la performance abstraite du récepteur SIC dans un canal Rayleigh à 8 tapes

MCS_0	MCS_1	$\mathrm{MSE}_{\mathrm{LUT},0}$	$\mathrm{MSE}_{\mathrm{EESM},0}$	$\mathrm{MSE}_{\mathrm{LUT},1}$	$\mathrm{MSE}_{\mathrm{EESM},1}$
4	4	0.0533	0.3537	0.0926	0.7674
10	10	0.3791	0.0719	0.8706	0.9561
12	12	0.7480	0.1986	1.1566	1.8764

Dans les scénarios de mobilité faible ou nulle, nous pouvons supposer que le SINR effectif dans le canal à évanouissements provoque le même taux de taux d'erreur sur les blocs que dans le canal AWGN [Olmos et al., 2010], ce qui peut être exprimé comme suit:

$$\mathrm{BLER}_l(\mathbf{H}_{\mathrm{eff}}, N_0, \mathrm{MCS}_l) \approxeq \mathrm{BLER}_{\mathrm{AWGN}}(\beta_{1,l}\mathrm{SINR}_{\mathrm{eff},l}, \mathrm{MCS}_l).$$

L'étalonnage des coefficients d'ajustement est effectué sur la base du critère MSE:

$$\beta_{\mathrm{opt}_0,l}, \beta_{\mathrm{opt}_1,l} = \underset{\beta_{0,l},\beta_{1,l}}{\arg\min} \frac{1}{N_H N_n} \left[\sum_{c}^{N_H} \sum_{z}^{N_n} |\mathrm{SINR}_{\mathrm{eff},l}(\beta_{0,l}, \beta_{1,l}, \mathrm{MCS}_l) - \mathrm{SINR}_{\mathrm{AWGN},l}(\mathrm{MCS}_l)|^2 \right].$$

9.5.3.2 Résultats de l'approche EESM

L'analyse comparative de la précision, introduite par les méthodes de cartographie EESM et MIESM, est présentée dans Table 9.4 et Table 9.5. Pour les constellations QPSK, l'approche MIESM surpasse la méthodologie EESM pour les deux flux de nos récepteurs PIA et SIC. Pour les mots de code décodés en présence d'interférences et mappés sur des constellations à cardinalité élevée, la cartographie EESM surpasse de manière surprenante notre approche MIESM lourde. En outre, les valeurs MSE du mappage EESM restent à un niveau bas avec l'augmentation de l'indice MCS, par rapport à MIESM ($\mathrm{MSE}_{\mathrm{EESM},0} \ll \mathrm{MSE}_{\mathrm{LUT},0}$). Cela rend le EESM plus favorable pour une utilisation pratique. Cependant, l'abstraction EESM pour l'abstraction du récepteur SIC n'est pas aussi fiable que pour le premier flux et fournit une précision d'environ 1 dB.

9.5.4 Conclusion

L'approche MIESM offre des niveaux de précision satisfaisants, mais dépend beaucoup de l'implémentation du récepteur, de l'architecture à virgule fixe ou à virgule flottante, de la granularité de la table de consultation et des modèles de canal car les coefficients de

calibration peuvent varier de manière significative d'un environnement d'évanouissements à l'autre.

Afin de contourner les inconvénients de la méthode MIESM, nous avons évalué une abstraction EESM simple et légère basée sur l'approximation des récepteurs ML PIA et SIC avec les récepteurs triviale MIMO MMSE et IF. La performance abstraite n'est que légèrement inférieure à celle de la méthode MIESM pour l'ordre de faible modulation.

Les parties de ce chapitre consacrées à l'abstraction de MIESM pour notre récepteur PIA sont publiées dans:

- E. Lukashova, F. Kaltenberger, R. Knopp, and C. Bonnet, "PHY layer abstraction for SU-MIMO LTE system employing parallel Interference-Aware detection," in International Workshop on Link and System Level Simulations 2016 (IWSLS2'16), Vienna, Austria, June 2016.

9.6 Chapitre 7 · Adaptation de Liaison pour les Récepteurs IA PIA et SIC

Dans ce chapitre, nous étudions une stratégie d'adaptation de lien peu complexe compatible avec nos récepteurs R-ML PIA et R-ML SIC.

9.6.1 Adaptation du Rang

9.6.1.1 Méthodologie

Les performances des schémas de transmission à multiplexage spatial sont sensibles à la corrélation d'émission et de réception [Bölcskei et al., 2002]. Une mesure objective de la corrélation de canal est le conditionnement $\mathcal{K}(\mathbf{H})$. Le canal est considéré comme étant bien conditionné si $\mathcal{K}(\mathbf{H})$ ne dépasse pas 10 dB [Agilent, 2009]. Alors que le calcul du conditionnement pour les matrices de grandes dimensions est difficile et prend beaucoup de temps, il est facile pour nos matrices 2×2:

$$\mathcal{K}(\mathbf{H}) = \|\mathbf{H}\|_{\mathrm{F}} \left\|\mathbf{H}^{-1}\right\|_{\mathrm{F}}, \tag{9.25}$$

où $\mathbf{H}$ est la matrice de canaux.

Le nombre conditionnel est estimé par sous-porteuse. Si la majorité des valeurs sont inférieures à un certain seuil $\mathcal{T}$, Rang (Rank Indicatoor – RI) est assigné la valeur deux et le schéma de transmission à multiplexage spatial est appliqué. Sinon, l'UE signale le RI un et la station de base transmet en utilisant le schéma de précodage Alamouti.

9.6.1.2 Résultats numériques et discussion

Nos simulations au niveau des liaisons ont été effectuées avec 3000 paquets, une bande passante LTE de 5 MHz un canal EPA (Extended Pedestrian A) 3GPP [2015d] étendu avec une fréquence Doppler nulle et une matrice de corrélation modérée et élevée (EPAM et EPAH).

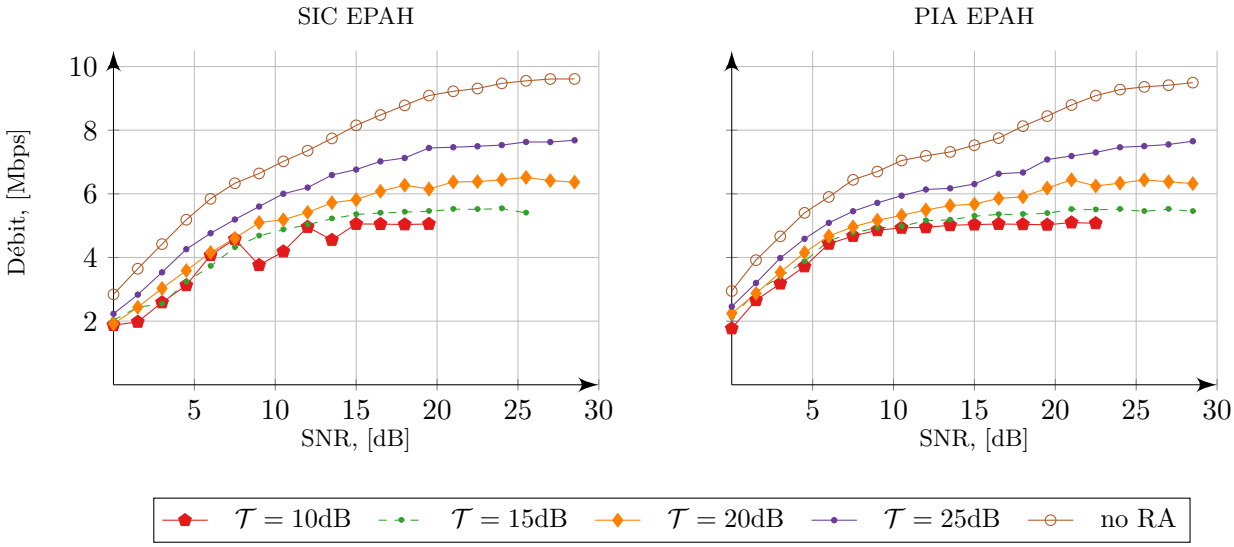


Figure 9.12: Somme totale du système pour les récepteurs PIA et SIC ($\text{MCS}_0 = 12$ et $\text{MCS}_1 = 12$) après un tour de transmissions du canal EPAH.

Les paires suivantes du MCS ont été choisies: $(\text{MCS}_0, \text{MCS}_1) \in \{(4,4),(12,12),(22,26)\}$. Pour chaque environnement d'évanouissements, nous considérons l'ensemble des seuils $\mathcal{T} \in \{5, 10, 15, 20, 25, \infty\}$ dB.

Nos expériences ont montré que dans le canal Rayleigh, 90% des transmissions appartiennent au multiplexage spatial, si le seuil du nombre de conditions $\mathcal{T}$ est défini sur 15 dB et 100% si $\mathcal{T}$ est supérieur à 20 dB. Avec l'augmentation du niveau de corrélation, le nombre de transmissions de deux blocs de transport pour un même seuil $\mathcal{T}$ diminue: pour $\mathcal{T} = 25$ dB, il n'y a que 90% de transmissions dans le canal EPAM, alors que cette valeur n'atteint que 57% dans le canal EPAH. Les récepteurs affichent le débit le plus élevé lorsque l'eNodeB envoie toujours deux blocs de transport simultanément (Fig. 9.12). Par conséquent, nos récepteurs gèrent avec succès la détection des deux mots de code, même dans des canaux mal conditionnés.

9.6.2 Estimation du Précodeur

Le critère maximum fondé sur l'information mutuelle est optimal mais difficile à mettre en œuvre. Les niveaux d'information mutuels dépendent de l'architecture du récepteur et doivent être calculés sur la base de (9.11), (9.12) pour toutes les options de précodage. En cas de conception d'annulation d'interférence successive, la solution intuitive et facile à mettre en œuvre consiste à maximiser le SNR du premier flux, car le décodage réussi du deuxième mot de code dépend de la décodabilité du premier mot de code [Ghaffar and Knopp, 2010b]. L'UE calcule deux rapports entre les valeurs SNR du premier flux et du second flux:

$$\left(\frac{\|\mathbf{h}_0 + \mathbf{h}_1\|^2}{\|\mathbf{h}_0 - \mathbf{h}_1\|^2}, \frac{\|\mathbf{h}_0 + j\mathbf{h}_1\|^2}{\|\mathbf{h}_0 - j\mathbf{h}_1\|^2} \right). \tag{9.26}$$

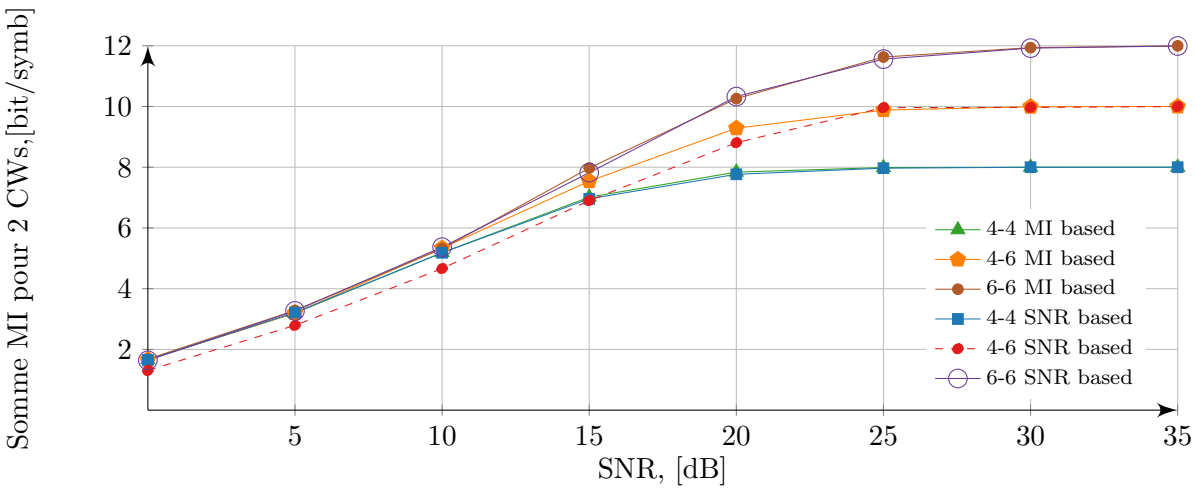


Figure 9.13: Niveaux d'information mutuels potentiels pour deux mots de code utilisant notre récepteur SIC R-ML avec sélection de précodeurs basée sur MI et sur SNR.

Si le premier rapport est supérieur au second, l'UE sélectionne la matrice de précodage avec des valeurs réelles, sinon avec des valeurs complexes. Ce calcul peut être simplifié à l'évaluation des parties réelle et imaginaire du coefficient de corrélation $\rho_{10} = \mathbf{h}_1^H \mathbf{h}_0$:

$$\frac{\|\mathbf{h}_0 + \mathbf{h}_1\|^2}{\|\mathbf{h}_0 - \mathbf{h}_1\|^2} < \frac{\|\mathbf{h}_0 + j\mathbf{h}_1\|^2}{\|\mathbf{h}_0 - j\mathbf{h}_1\|^2} \iff \Re(\rho) < \Im(\rho). \tag{9.27}$$

Nous pouvons maintenant définir le critère de sélection de la matrice de précodage sur la base du coefficient de corrélation ρ entre les colonnes $\mathbf{h}_0$ et $\mathbf{h}_1$ de la matrice de canal $\mathbf{H}$:

$$\mathbf{P} = \begin{cases} \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, & \text{for } \Re(\rho_{10}) \geq \Im(\rho_{10}); \\ \\ \frac{1}{2}\begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}, & \text{for } \Re(\rho_{10}) < \Im(\rho_{10}). \end{cases} \tag{9.28}$$

Comme on peut le voir dans Fig. 9.13, le critère basé sur le SNR est donc une solution légère pour le calcul de l'indicateur de la matrice de précodage (Precoder Matrix Indicator – PMI) et offre des niveaux de performance proches du critère optimal d'information mutuelle.

9.6.3 Adaptation du CQI

9.6.3.1 Méthodologie

Nous recherchons maintenant une méthodologie d'estimation d'un indicateur de qualité de canal reçu (Channel Quality Indicator – CQI) pratique permettant d'ajuster les débits de données de transmission aux conditions de canal instantanées.

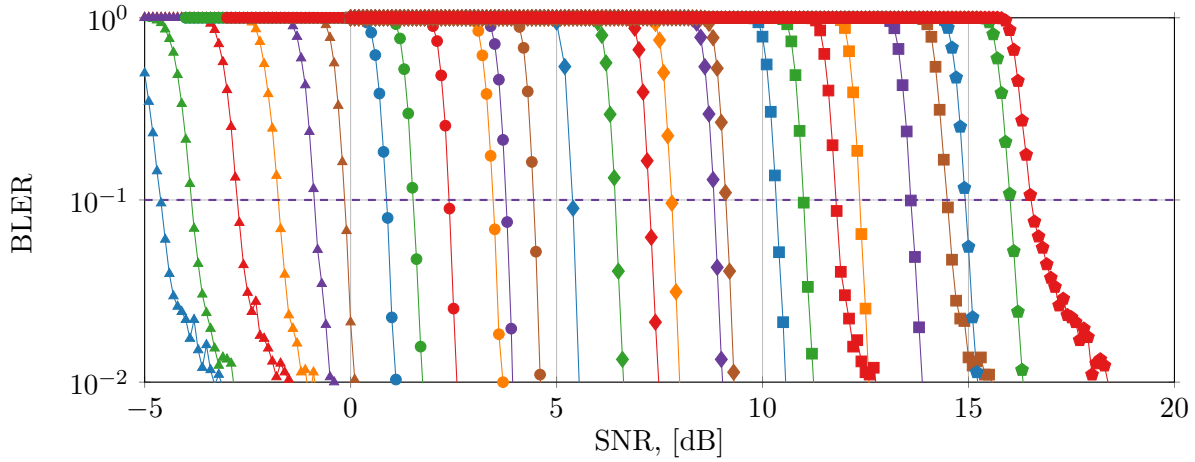


Figure 9.14: Courbes de BLER dans le canal AWGN. Les courbes correspondent aux MCS 1 - 28, tracés de gauche à droite par ordre croissant.

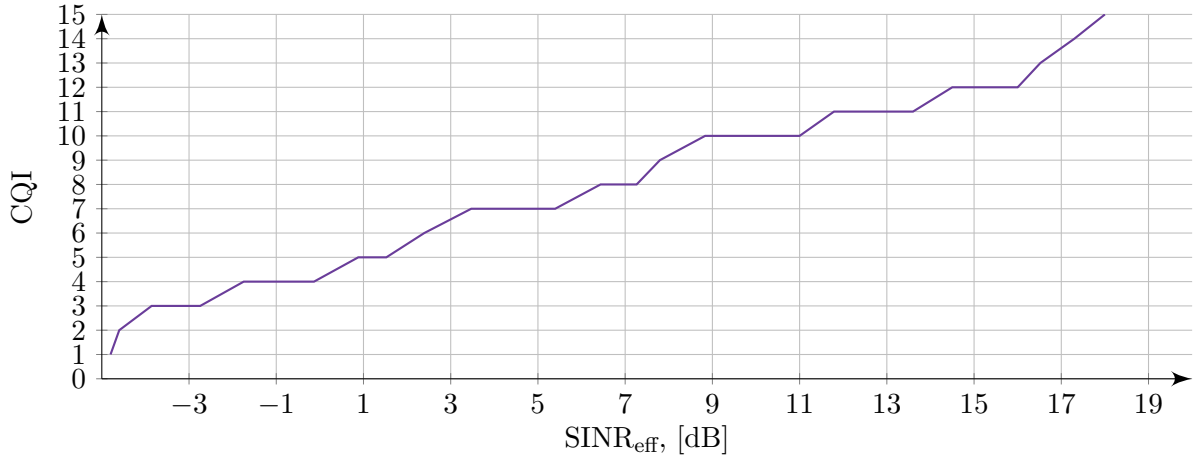


Figure 9.15: Mappage CQI vers SINR pour 2×2 Système MIMO mono-utilisateur LTE avec blocs de ressources physiques 25 et 1 symbole PDCCH. Un mappage différent peut être requis pour un autre paramètre.

La première étape de la méthodologie d'estimation CQI consiste à créer une table de consultation pour une cartographie efficace du rapport SINR à CQI. La correspondance de taux pour les transmissions de données en liaison descendante est basée sur 28 valeurs disponibles de MCS. La station de base doit donc traduire les valeurs CQI signalées en MCS. À l'aide de notre simulateur de liaison descendante, nous avons obtenu les performances des transmissions MIMO CLSM pour un utilisateur unique dans un canal AWGN pour l'ensemble des valeurs MCS. Pour chaque valeur MCS, nous avons ensuite défini une valeur SINR correspondante telle que le taux d'erreur sur les blocs (Block Error Rate – BLER) ne dépasse pas 10% (Fig. 9.14). Ces valeurs SINR sont identiques pour nos récepteurs PIA et SIC, car deux mots de code ne sont pas perturbés dans le canal AWGN.

Les valeurs SINR sélectionnées correspondent au SINR *efficace*, qui provoque le même

niveau de taux d'erreur de bloc dans le canal sélectif en fréquence que dans le canal AWGN. Au cours des simulations, nous supposons une largeur de bande LTE de 5 MHz (25 blocs de ressources physiques) et un symbole PDCCH. Pour effectuer le mappage, comparons les taux de codage de MCS et le taux de code LTE approximatif de CQI, puis créons les paires correspondantes (Fig. 9.15).

La table de consultation obtenue avec le mappage SINR-CQI-MCS peut maintenant être utilisée pour l'estimation CQI. Notre méthodologie pour les récepteurs PIA et SIC repose sur les résultats de notre captation légère EESM sans facteurs d'étalonnage: les performances des flux IA sont approximées avec les performances par flux du récepteur MMSE, tandis que les performances du système sans interférence le second flux de notre récepteur SIC est approximé par le récepteur avec une annulation parfaite des interférences.

Algorithm 22 Adaptation conjointe large bande CQI et PMI large bande pour les récepteurs PIA et SIC

Input: $\mathbf{H}$, N_0.
Output: CQI_l.

for chaque émlément $k = 1, ...K$ **do**
    calculer et accumuler $\rho_{10} = \mathbf{h}_{1,k}^H \mathbf{h}_{0,k} + \rho_{10}$.
end for
Evaluer ρ_{10} et sélectionner le PMI préféré et la matrix de précodage $\mathbf{P}$.
for chaque élément $k = 1, ...K$ **do**
    Calculer $\mathbf{H}_{\mathrm{eff}\,k} = \mathbf{H}_k \mathbf{P}$.
end for
for chaque flux $0 \leq l \leq L$ **do**
    for chaque élément $k = 1, ...K$ **do**
        calculer une valeur individuelle $\mathrm{SINR}_{l,k} = \mathcal{S}(\mathbf{H}_{\mathrm{eff}k}, N_0, \text{rec. arch})$.
    end for
    Calculer le $\mathrm{SINR}_{\mathrm{eff},l}$ efficace en utilisant les fonctions de mappage EESM.
    Trouver la valeur CQI_l correspondante avec Fig. 9.15.
end for

L'estimation conjointe large bande CQI/large bande PMI est décrite dans (Algorithm 22). Si le mode de transmission de retour est défini sur large bande CQI/sous-bande rapport PMI (Algorithm 23), la bande passante de transmission complète est divisée en sept sous-bandes et la matrice de précodage est estimée individuellement pour chacune d'elles. Les valeurs SINR pour chaque élément de ressource sont ensuite calculées en tenant compte du PMI par sous-bande.

9.6.3.2 Résultats numériques

Dans nos paramètres, nous considérons un système MIMO à 2×2 où la station de base transmet toujours deux mots de code sur deux couches multiplexées dans l'espace. Nous générons 1000 réalisations de canaux pour 300 éléments de ressources (des blocs de ressources physiques de 25) et un ensemble de valeurs de SNR pré-transmises $\gamma = 1/N_0 \in \{-5, 0, \ldots 40\}$ dB.

Algorithm 23 Adaptation conjointe large bande CQI et PMI sous-bande pour les récepteurs PIA et SIC

Input: $\mathbf{H}$, N_0.
Output: CQI_l.

for chaque sous-bande $s = 1, ...S$ **do**
  for chaque élément $k = 1, ...K_s$ **do**
    calculer et accumuler $\rho_{10,s} = \mathbf{h}_{1,s,k}^H \mathbf{h}_{0,s,k} + \rho_{10,s}$.
  end for
  Evaluer $\rho_{10,s}$ et sélectionner le PMI préféré et la matrix de précodage $\mathbf{P}_s$.
end for
for chaque sous-bande $s = 1, ...S$ **do**
  for chaque élément $k = 1, ...K$ **do**
    Calculer $\mathbf{H}_{\text{eff},s,k} = \mathbf{H}_{s,k}\mathbf{P}_s$.
  end for
end for
Concatener matrices de canaux efficaces par sous-bande en un seul vecteur $\mathbf{H}_{\text{eff}}$
for chaque flux $0 \leq l \leq L$ **do**
  for chaque élément $k = 1, ...K$ **do**
    calculer une valeur individuelle $\text{SINR}_{l,k} = \mathcal{S}(\mathbf{H}_{\text{eff}k}, N_0, \text{rec. arch})$.
  end for
  Calculer le $\text{SINR}_{\text{eff},l}$ efficace en utilisant les fonctions de mappage EESM.
  Trouver la valeur CQI_l correspondante avec Fig. 9.15.
end for

Étant donné des conditions de canal identiques (SNR pré-transmis γ et matrice de canal), les premiers flux des récepteurs PIA et SIC signalent le même CQI, car ils reçoivent le même traitement (Fig. 9.16) . Dans le même temps, le second flux du récepteur SIC indique des valeurs CQI supérieures de 2 à 4 au second flux du récepteur PIA, car il est approximé par le récepteur sans interférence. Cependant, les valeurs CQI par flux pour les deux modes de retour sont identiques. La granularité CQI est non uniforme et varie entre 0.5 et 2 dB (Fig. 9.15), ce qui rend plus difficile la capture du précodeur.

Les valeurs CQI obtenues sont mappées sur le MCS correspondant (Fig. 9.17). Les valeurs MCS du récepteur SIC atteignent de manière prévisible des niveaux élevés.

9.6.3.3 Conclusion

Dans ce chapitre, nous avons fourni une méthodologie permettant de calculer le retour d'information CSI pour nos récepteurs PIA R-ML et SIC R-ML. Nous proposons d'abord d'estimer le RI, puis d'effectuer une estimation conjointe du PMI et du CQI.

L'estimation RI basée sur le conditionnement de la matrice de canaux a révélé que nos récepteurs atteignent des valeurs de débit élevées grâce aux transmissions de multiplexage spatial, même dans les canaux à corrélation élevée. Ce résultat reste valable pour l'ensemble des valeurs MCS disponibles.

L'adaptation PMI légère proposée, qui maximise le SINR sur le premier flux, est basée sur l'évaluation du coefficient de corrélation de la matrice de canaux estimée. Cette

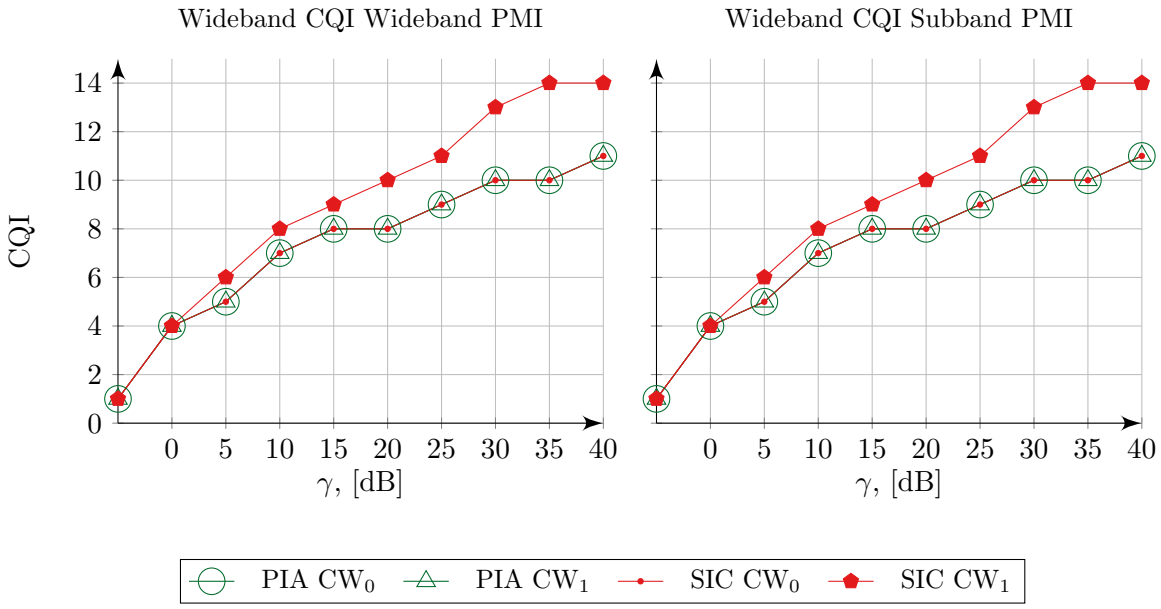


Figure 9.16: Valeurs CQI par flux pour les rapports communs PQ large bande / PMI large bande (à gauche) et les rapports PMI CQI large bande / à large bande (à droite) pour nos récepteurs PIA et SIC. Étant donné que la granularité CQI varie entre 0, 5 et 2 dB, l'influence du PMI de la sous-bande est trop faible pour être capturée.

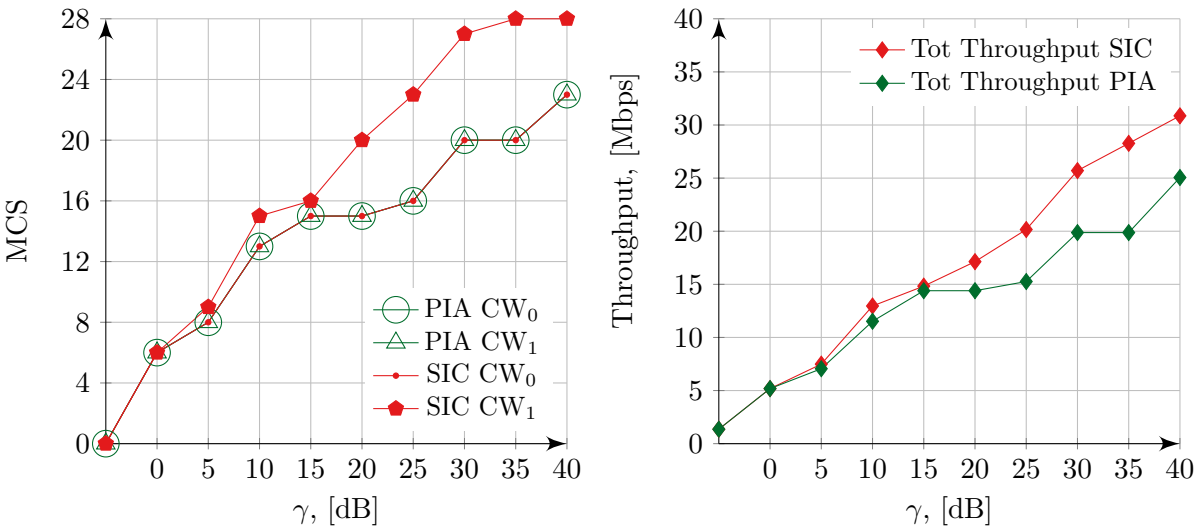


Figure 9.17: Valeurs MCS par flux (à gauche), correspondant aux valeurs CQI obtenues et au débit total atteint pour nos récepteurs PIA et SIC.

méthode est simple à mettre en œuvre et fournit des niveaux d'information mutuelle proches du critère optimal d'information mutuelle maximum.

Notre méthodologie CQI est basée sur l'approximation de nos récepteurs PIA et SIC avec des récepteurs MMSE et sans interférence. Les résultats empiriques ont révélé que, pour le récepteur SIC, les valeurs CQI signalées et le niveau de débit prédit sont supérieurs à ceux du récepteur PIA dans les mêmes conditions de canal.

9.7 Chapitre 8 · Conclusion

Dans cette thèse, nous avons étudié des alternatives de faible complexité à la détection conjointe ML dans les systèmes MIMO mono-utilisateur. Précisément, nous avons implémenté et rendu compatible avec les systèmes pratiques LTE MIMO deux architectures de récepteurs de complexité réduite: la détection ML parallèle interférence-aware (R-ML PIA) et la détection interférence-aware annulation successive d'interférences (R-ML IA SIC).

Nous nous attendions initialement à ce que le récepteur SIC surpasse la détection PIA dans les régimes de SNR élevé et moyen, grâce à la détection sans interférence du second mot de code. Dans le même temps, il semblait également logique que notre récepteur PIA bénéficierait de certains avantages dans la région à faible SNR lorsque le décodage du premier mot de code échoue, car le deuxième mot de code subit toujours la procédure de décodage. En revanche, la tentative de décodage du deuxième mot de code dans le récepteur SIC est bloquée par la détection erronée du premier mot de code. Cependant, une analyse théorique de l'information et des résultats de débit empiriques ont démontré que le récepteur SIC n'est jamais surperformé par la détection PIA et constitue une solution réalisable dans tous les régimes de SNR.

Les résultats de la quantification de l'effort de calcul sont également positifs pour notre récepteur SIC, il est environ 25% plus informatique efficace que le récepteur PIA pour les ordres de modulation élevée. Ce gain a été rendu possible en remplaçant la métrique LLR *interférence-aware* pour le deuxième mot de code par le bloc SIC et la métrique LLR de *sans interférences*.

Pour rendre les récepteurs compatibles avec les systèmes LTE pratiques, nous avons dû revoir les protocoles LTE essentiels tels que le système HARQ à redondance incrémentielle et l'adaptation de liaison. Un soin particulier a été apporté aux procédures SIC à plusieurs tours disponibles au récepteur SIC, ce qui permet d'obtenir des gains de débit importants lors des premier et deuxième tours de retransmission.

Pour améliorer les performances, nous avons ensuite recherché une solution d'adaptation de lien légère. Notre méthodologie d'estimation PMI est basée sur la maximisation du niveau de SNR du premier flux. Cette approche a démontré des performances proches du critère optimal optimal basé sur les informations mutuelles. L'estimation du classement basée sur un seuil de nombre de conditions a révélé que, contrairement à l'intuition, nos deux récepteurs atteignent le débit le plus élevé si la station de base transmet toujours deux mots de code. Pour effectuer l'estimation de le CQI, nous avons envisagé une méthodologie d'abstraction de faible complexité qui fournirait une estimation rapide du SNR eficace, puis le mapperait sur le CQI. La solution initiale — MIESM — était

basée sur des tables de consultation et fournissait une précision acceptable, si elle était étalonnée. Malheureusement, la taille des tables requises rend les systèmes en temps réel irréalisables. Néanmoins, il est possible d'utiliser ces tables de recherche dans des simulateurs pour accélérer les calculs au niveau des liens. Pour contourner les inconvénients de la méthode MIESM, nous avons évalué une simple abstraction EESM. La deuxième solution a donné des résultats satisfaisants lorsque les coefficients d'étalonnage sont disponibles ou absents.

Appendices

Appendix A

Measurement Campaign

During this thesis, we have made a few contributions to French FUI FAPIS project "4G in Vitro". The goal of this project was to bridge the gap between simulations and live experiments by developing a tool that emulates field conditions in the lab. Mobipass is an advanced test system aiming to reproduce the behavior and radio propagation conditions of the network with multiple connected devices and real-world traffic models. We have performed a drive test campaign in Sophia Antipolis, France, to evaluate how close the performance predicted by Mobipass is to reality.

The campaigns aimed to collect field measurements composed of GPS positions, Key Performance Indicators (KPI) such as uplink (UL) and downlink (DL) throughput, Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), CQI and store the traces in the way that the user is able to easily reproduce the results in a virtual mode.

The results are obtained during three rounds of the drive test campaign, performed with approximately 6 months gaps. The first measurement campaign was held in February 2015, the second round was performed in November 2015, and the third round took place in June 2016 and was assisted by the Com4Innov team. The first two rounds were not very informative due to the technical limitations (equipment, base station functioning, multiple interferes). However, the collected data allowed us to understand better the constrains and the requirements of the scenario. The third round brought a fair comparison between a real network performance and Mobipass.

All the driving tests followed a similar scenario: the UE, represented by a dongle or a mobile phone, was connected to the laptop with TEMS^{TM} Investigation software, which collected the KPIs and stored the message exchange between the UE and the base station in the traces with tracking the positions of the UE. The collected data was then analyzed and compared with the Mobipass prediction.

A.1 Radio Network

Measurement campaigns were held in Sophia-Antipolis, France using Com4Innov EUTRAN composed of France Telecom Orange antenna site (depicted in orange in Fig. A.1) equipped with an Ericsson eNodeB and its 3 antenna sectors with corresponding Physi-

cal Cell ID 6, 7 and 8 set at 2,6 GHz frequency. The neighboring cell site of Bouygues Telecom Cica (depicted in blue in Fig. A.1) with Physical Cell ID 3, 4 and 5 was turned off during measurements collection.

Figure A.1: Sophia Antipolis EUTRAN positioning

Table A.1: Sophia Antipolis antenna sites parameters

Site	Longit.	Lat.	Sec.	Azimut	Cell ID	Ph. Cell-ID	Freq.
Cica - Pylone Bouygues Telecom	7° 2' 44.746"	43° 37' 37.865"	1	110°	0	3	FDD B7 - 2,6 Ghz
Cica - Pylone Bouygues Telecom	7° 2' 44.746"	43° 37' 37.865"	2	190°	1	4	FDD B7 - 2,6 Ghz
Cica - Pylone Bouygues Telecom	7° 2' 44.746"	43° 37' 37.865"	3	280°	2	5	FDD B7 - 2,6 Ghz
France Telecom Orange - Pylone	7° 3' 4.705"	43° 37' 1.549"	1	30°	0	6	FDD B7 - 2,6 Ghz
France Telecom Orange - Pylone	7° 3' 4.705"	43° 37' 1.549"	2	97°	1	7	FDD B7 - 2,6 Ghz
France Telecom Orange - Pylone	7° 3' 4.705"	43° 37' 1.549"	3	285°	2	8	FDD B7 - 2,6 Ghz

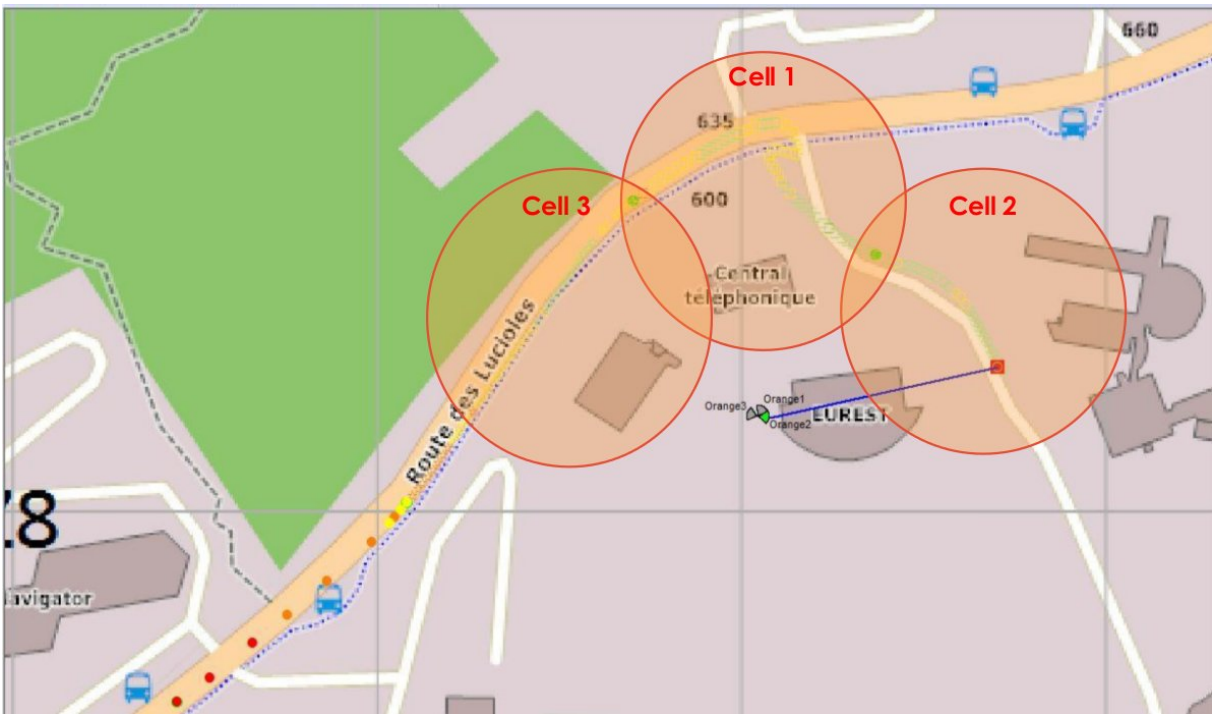


Figure A.2: Expected handover zones. Cell index corresponds to the index of the eNodeB sector.

The expected handover zones in the direct proximity of the France Telecom Orange - Pylone antenna site are presented in Fig. A.2. The eNodeB exploits Transmission Mode 2 (TM 2) and TM 3, what means that the full MIMO Downlink (DL) capacity, available in case of TM 4 (Closed-Loop Single User MIMO), is not accessed.

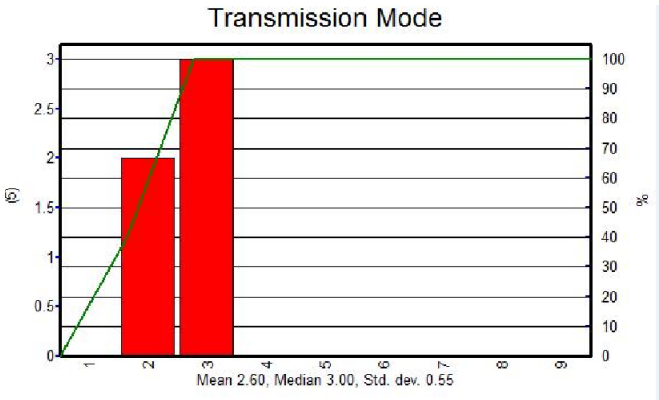


Figure A.3: DL Transmission Modes in the cells

A.2 First Measurements Campaign

The first measurement campaign was done using a dongle placed inside of the car, a GPS receiver loosely fixed on the roof of the car, and a laptop with the TEMS^{TM} Investigation software. The UL and DL traffic were simulated using the iperf tool with the Transmission Control Protocol (TCP).

The measurements were collected during one circle drive on the track presented in Fig. A.4. TEMS registered 10 handover events. Fig. A.5 shows the RSRP level that the UE receives from each of the cells on the track duration; vertical lines indicate HO event. This confirms RSRP-level criteria for HO in our case: the UE gets connected to the cell with the highest RSRP.

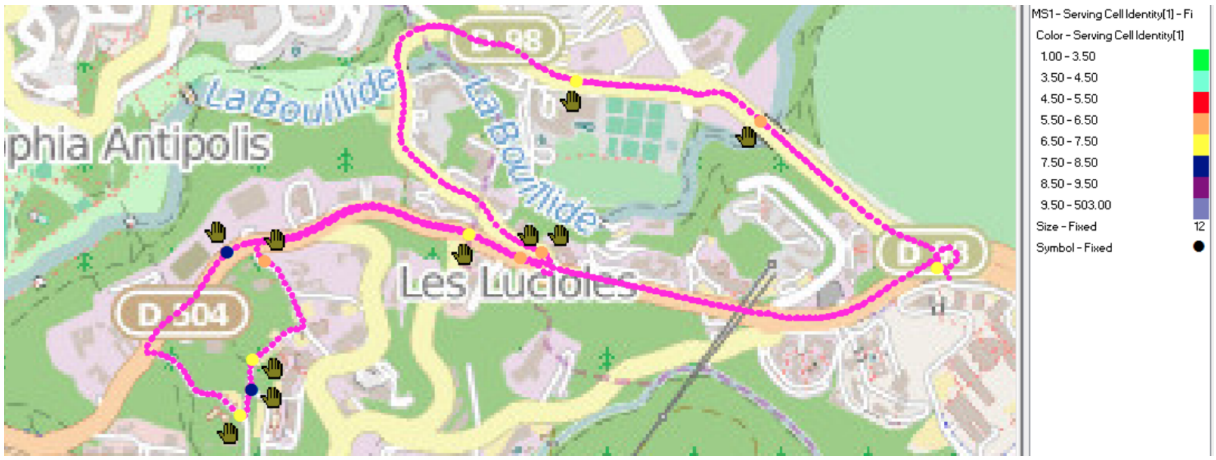


Figure A.4: Handover on a Driving Track Campaign 1 - February 2015

A.3 Second Measurements Campaign

The dongle used during the previews campaign belongs to UE family of Category 3, which limits its throughput potential. The equipment of the second measurement campaign was thus broaden with 2 antennas (Fig. A.6a), connected to the dongle, and smartphone Samsung Galaxy Note 4 Category 6 placed inside of the car (Fig. A.6). The routes were chosen such a way to avoid connection to Cell 8, where possible. Compared to the previous campaign, the route has been extended to the north.

The UL was simulated using iperf tool with the TCP and the DL was stimulated using File Transfer Protocol (FTP) server.

Dongle: DL was an FTP pulled from the server close to the eNodeB and to simulate UL, we connected to an iperf server close to the eNodeB and TCP protocol was used. The QoS was the default of the Ericsson eNodeB.

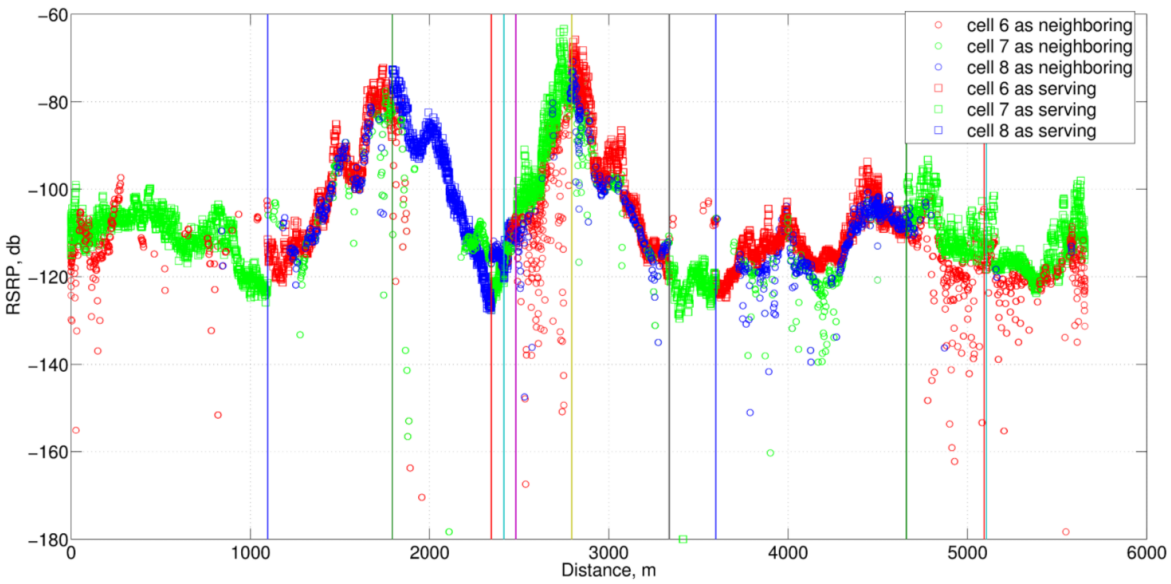
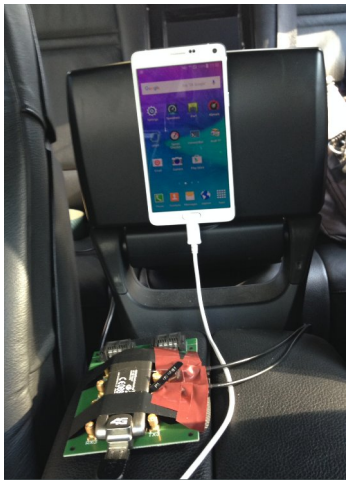


Figure A.5: RSRP of cells on a Driving Track Campaign 1 - February 2015

Smartphone: DL was taken as an FTP pulled from the server close to the eNodeB and UL was simulated using iperf application on the smartphone connected to an iperf server close to the eNodeB and TCP protocol was used. The QoS was a default one of the Ericsson eNodeB.

(a) External antennas and GPS receiver



(b) The UEs: dongle and smartphone

Figure A.6: Equipment November 2015

The following steps were performed:

1. Calibration of RSSI and RSRP with spectrum analyzer in cell ID 6 in line of sight conditions.

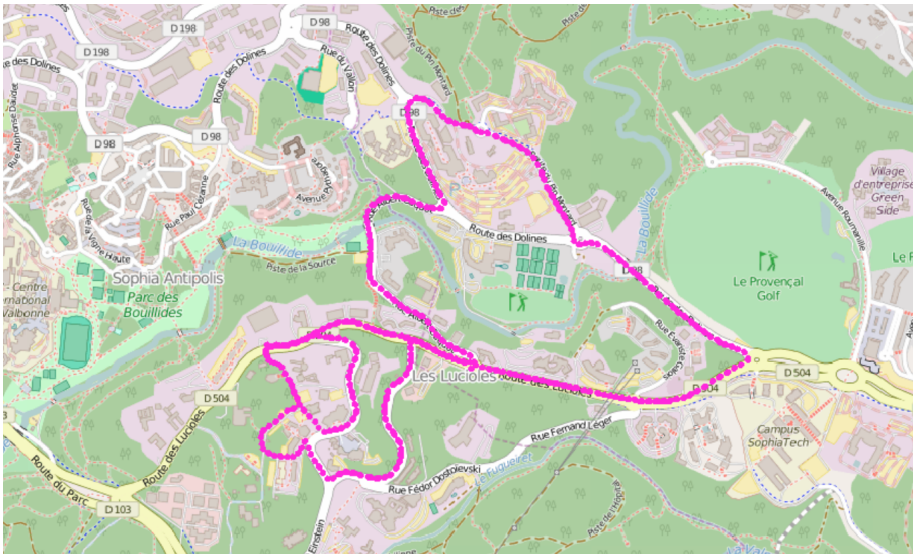


Figure A.7: The Driving Track for Campaign 2 - November 2015

2. Calibration of a smartphone.
3. Calibration of a dongle.
4. Drive a route 3 times with a smartphone.
5. Drive a route 3 times with a dongle.

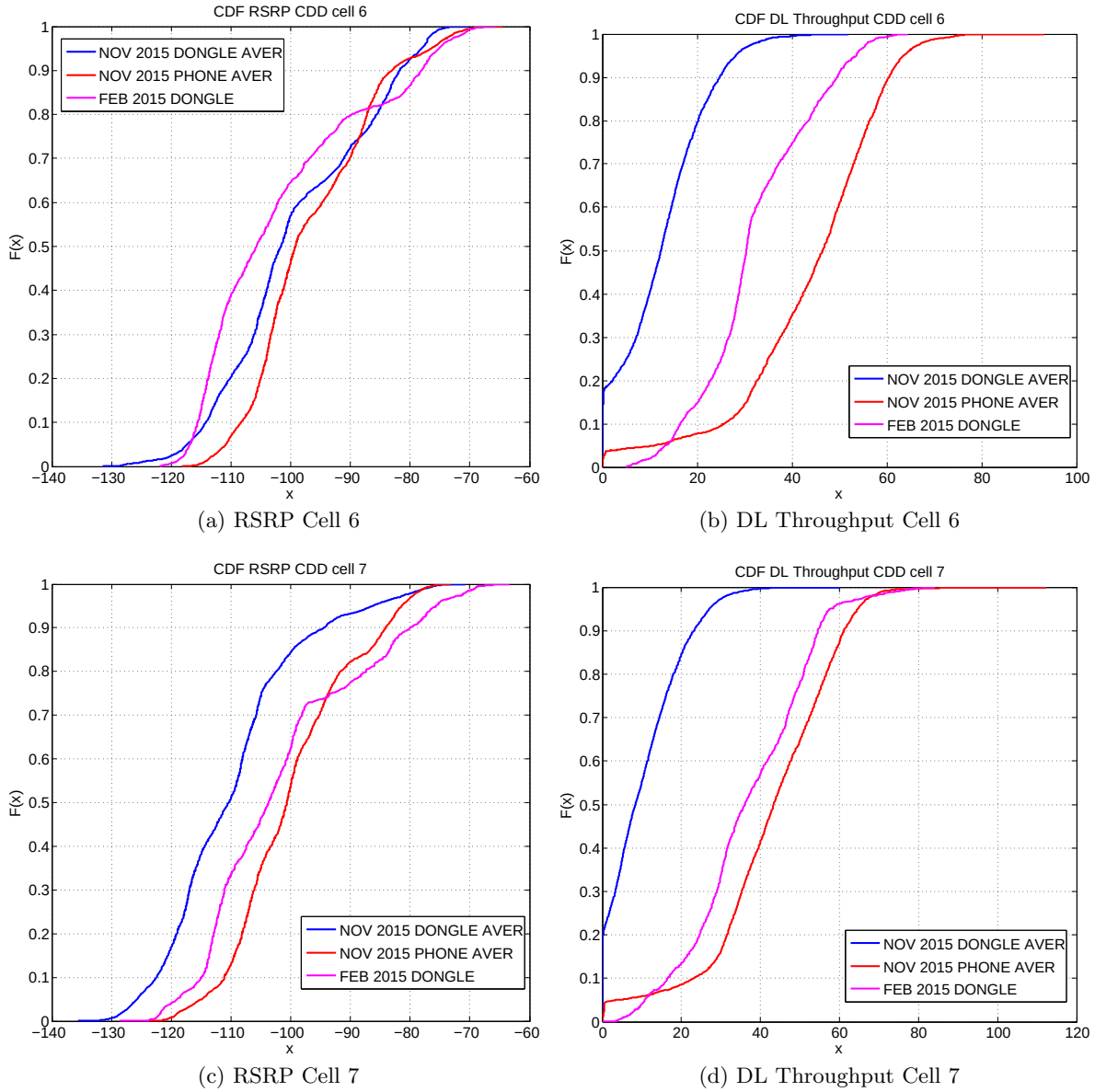
A.4 Comparative Analysis of Two Field Campaigns

In this section, we compare KPIs obtained from the two driving campaigns, and comment on possible causes in the differences between the dongle and the smartphone performances.

A.4.1 RSRP and DL Throughput Cell 6 vs Cell 7

The RSRP Cumulative Distribution Functions (CDF) from the dongle and smartphone in cell 6 and cell 7 are presented in (Fig. A.8a and Fig. A.8c). On average, the RSRP values in cell 6 for all three cases (dongle February 2015, dongle November 2015 and smartphone November 2015) are close; the path of cell 6 (pointing north) is on a higher ground and shows a better RSRP, while the path of cell 7 has been extended to the east in November 2015 in an area where the RSRP is lower. Comparing the RSRP with DL throughput, we can see that even if the RSRP is better than the February curves, the dongle throughput is much worse than in November 2015. On the other hand, the smartphone performed much better than dongle November 2015 with almost the same RSRP curve.
For the cell 7, the dongle in November 2015 has a much worst RSRP than the phone. It is not surprising that throughput is significantly worse. However, comparing the smartphone with the February measurements, we see that again the smartphone shows higher throughput than the dongle. Recall: the DL traffic in November 2015 was simulated using

FTP. It is important to notice that none of the UEs enjoys maximum throughput in any of cells; the highest value it reaches is 80 Mbps for a very short period time.



(a) RSRP Cell 6

(b) DL Throughput Cell 6

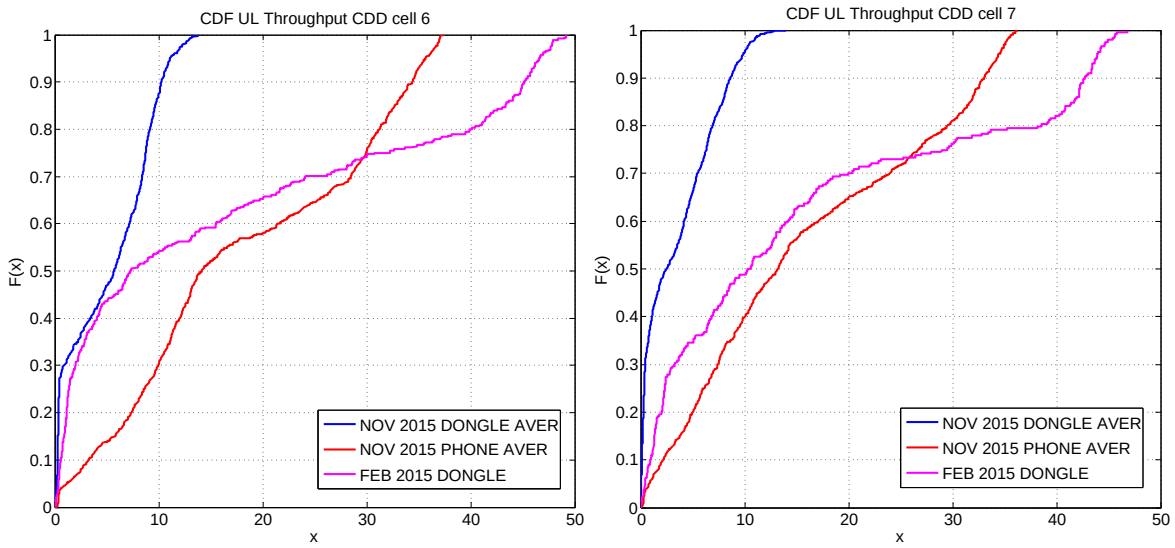
(c) RSRP Cell 7

(d) DL Throughput Cell 7

Figure A.8: DL Throughput and RSRP CDF for a smartphone and dongle in both campaigns in cell 6 and cell 7

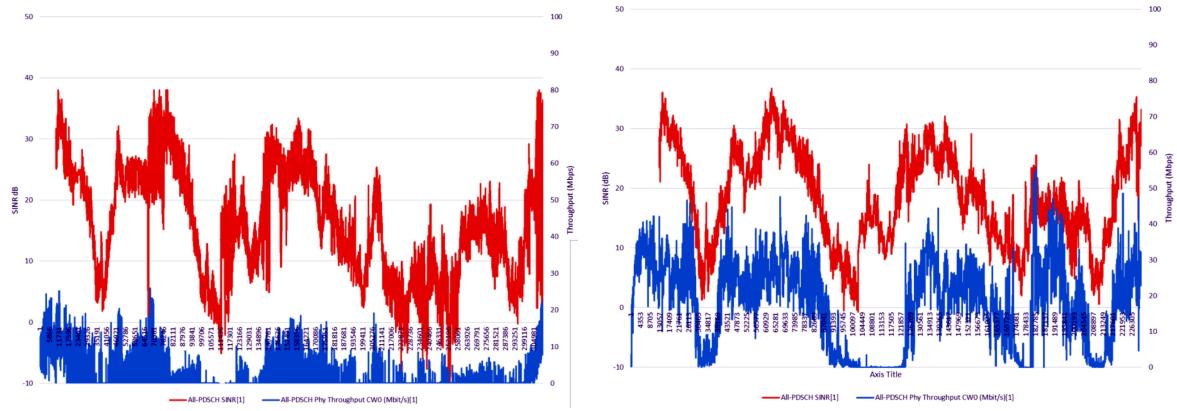
A.4.2 UL Throughput in cell 6 vs cell 7

Looking at the UL throughput values for cell 6 (Fig. A.9), we know that the RSRP in November is better than during the February drive; it is clear that the phone profits from this fact. Despite this fact, the dongle in February showed much higher values of UL



(a) Cell 6 (b) Cell 7

Figure A.9: UL Throughput CDF for a smartphone and dongle in both campaigns



(a) Dongle (b) Smartphone

Figure A.10: SINR vs Throughput

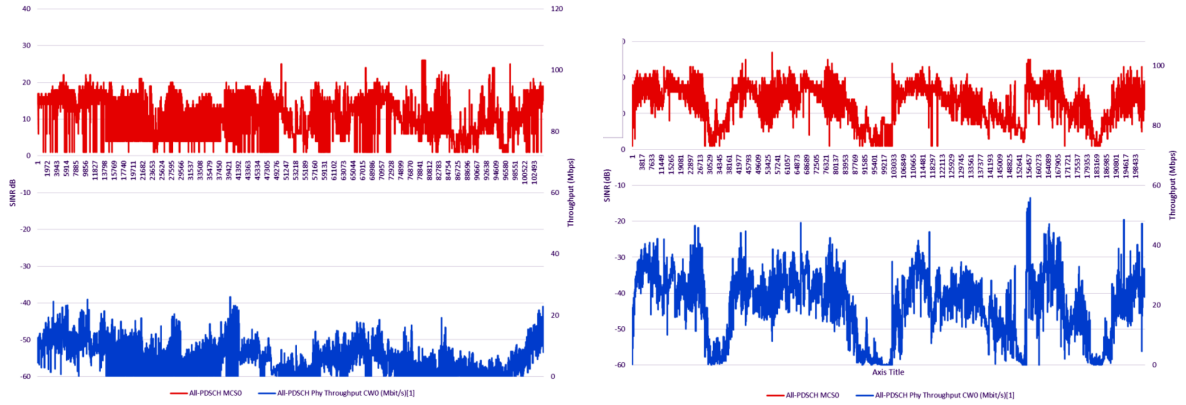
throughput, than in November. The explanation is not so trivial.

A.4.3 SNR vs Throughput

We focus on the performance of the dongle and the smartphone in November 2015. In LTE, throughput grows exponentially with SNR till some saturation point. SINR measurements from the dongle and smartphone are almost the same (Fig. A.10). However, throughput is very different: the phone registers higher throughput than the dongle given the same SNR.

A.4.4 MCS and DL Throughput

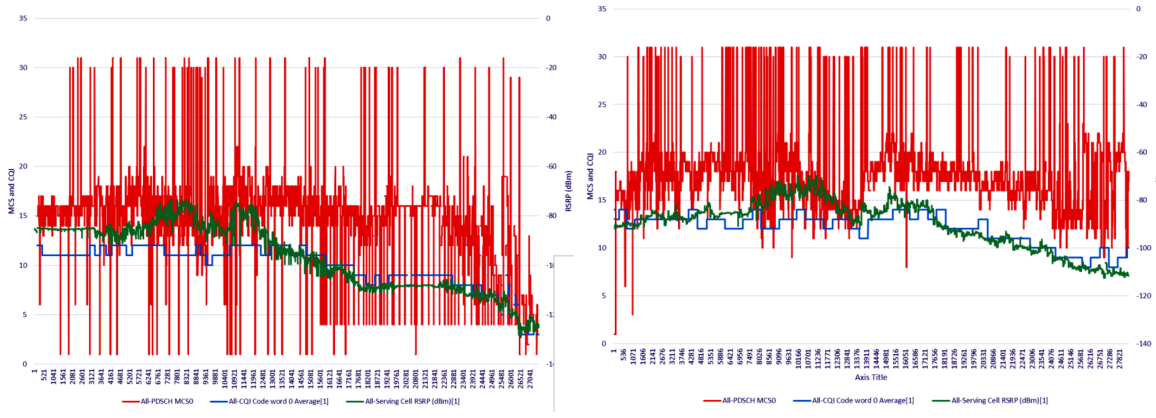
The smartphone shows fewer MCS switching events, and overall higher MCS values (Fig. A.11). In contrast, MCS values are clamped to a low value with the dongle even when the SINR is high, and this fact is reflected in the lower throughput values of the dongle.



(a) Dongle (b) Smartphone

Figure A.11: MCS and Throughput

A.4.5 MCS, CQI and HARQ



(a) Dongle (b) Smartphone

Figure A.12: RSRP, CQI and MCS

The MCS with values of 29, 30 and 31 represent the modulation index of a retransmission. We are zooming on the first 41 000 values of the drive test (Fig. A.12). The dongle has fewer retransmissions than the phone.

However, the dongle has multiple MCS values going to lower values like 1. The reason for the lower throughput might thus be a combination of low MCS and TCP protocol.

However, it is hard to guess why the modulation is going to lower values. The only difference was that the RSRP was lower for the dongle drive test.

A.5 Comparative Analysis of Field Campaigns and Mobipass Prediction

A.5.1 Mobipass Setup

For the lab measurements, Mentum Planet predictions for the pathloss for each point of the grid was used with further adding an EVA channel which simulates the slow and fast fading over the channel bandwidth. A UE of category 4 was considered with 20 MHz bandwidth dedicated to it. The UE is slave to the eNodeB so it is in TM 3. As for the rank, it is always 2, and we get up to 130 Mbps on a 20 MHz channel.

A.5.2 Results

In general, Mobipass forecasts much higher throughput level. It is hard to explain why when the UE experiences high SINR and a good channel, throughput is still limited to 60 Mbps, which is significantly far from a maximum value (Fig. A.13). In addition, it is

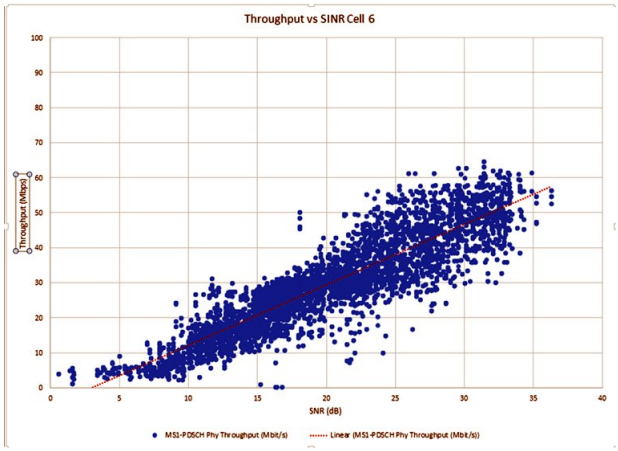
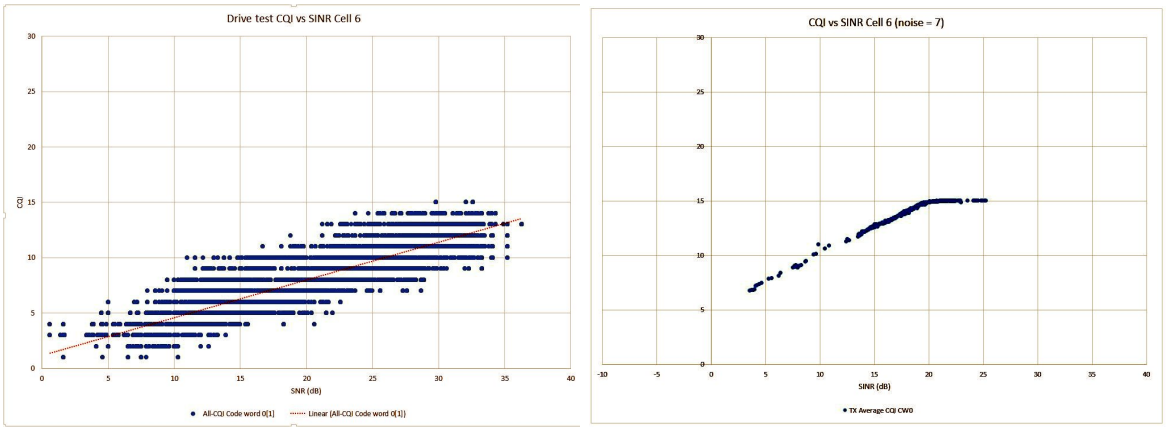


Figure A.13: DL Throughput vs SINR

not very clear, why the CQI values reported by the UEs have a wide spread. For example, for a SINR of 25 dB, the dongle can report a value between 7 to 14 (Fig. A.14). There could be a few reasons for this:

1. It is possible that CQI reported does not match the time when SINR is measured.
2. It is possible that the dongle does not report a high value of CQI consistently when it is over 25 dB for example. So the curve is actually shifted down and therefore not supplying the high throughput. Maybe the dongle is also reporting the CQI according to the BLER (which seems high in the drive test).
3. It could be due to the way small scale fading is taken into account in Mobipass.



(a) Field (b) Mobipass

Figure A.14: CQI vs SINR

4. The way UE computes SINR and CQI is unknown and and this might be different from Mobipass estimation.

A.6 Possible improvements for the next round of measurements

Even though the discussed driving campaigns did not provide a reliable data for a meaningful comparison, we have identified the weaknesses to be fixed in the next round of experiments:

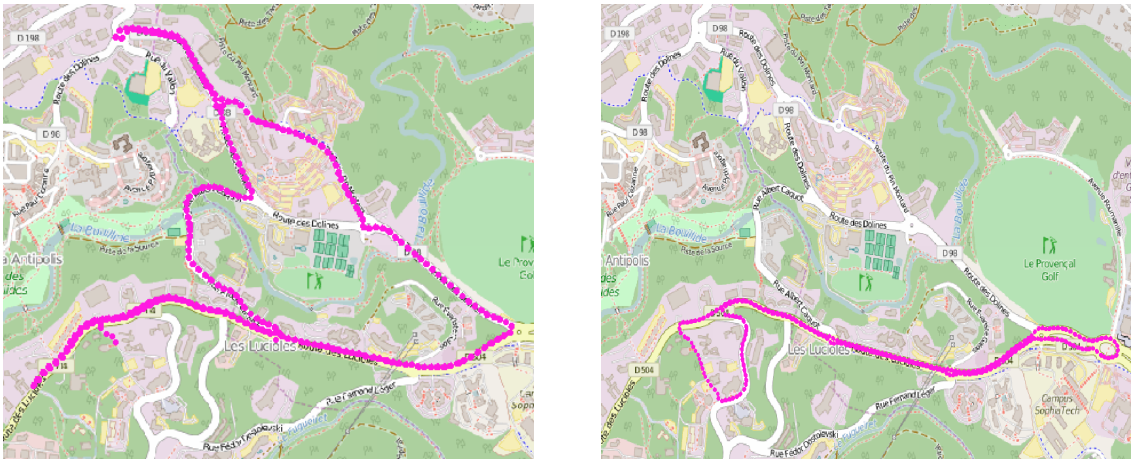
1. Use external antennas (placed on th roof of the car) for the smartphone (placed inside the car).
2. Switch off 2 out of 3 antenna sectors in order to get full throughput.
3. Use UDP for traffic generation. TCP is not the right way to test because of the way the TCP protocol adapts to the variation of the link. It tends to lower throughput. On contrary, UDP does not have a throughput regulating mechanism. It always sends the max that you tell it and at the other end, it calculates the packets that got through.

A.7 Third Measurements Campaign

The third measurement campaign was held in June 2016 in Sophia-Antipolis, France with the technical support from Com4innov team. There were three main lessons that were learned from the previous campaigns (Appendix A.6):

1. The maximum throughput can be achieved only if one sole sector of eNodeB is active.
2. The downlink and uplink must be generated by the means of iperf with UDP protocol.

3. The UE must be of category 4 or higher.



(a) Sector 1 Cell 6

(b) Sector 2 Cell 7

Figure A.15: Driving Tracks June 2016

This time, the experiments are thus done only with the smartphone. However, we failed to connect the external antennas and limited ourselves to the Smartphone Samsung Galaxy 4 of category 4 with Android version 4.3 statically fixed inside the car.

In this campaign, only one eNodeB sector was active in order the let the phone experience maximum throughput. Thus, the experiments for the Sector 1 and Sector 2 were performed following individual tracks (Fig. A.15a and Fig. A.15b).

Both UL and DL were simulated using iperf tool with UDP launching the following command on the client side

```
iperf -c (ip of the server) -u -b 100M -l 1000 -t 3600,
```

where b is bandwidth, l is datagram length and t is time period. The com4innov Virtual Machine (VM) was accessed through the different network using a dongle with Monaco Telecom sim card.

The following steps were performed:

1. Switch on Sector 1 and switch off Sector 2 and 3.
2. Calibration of RSSI and RSRP with spectrum analyzer in cell 6 in Line-of-Sight (LOS) conditions.
3. Drive a route for Sector 1 for downlink measurement collection.
4. Switch on Sector 2 and switch off Sector 1 and 3.
5. Calibration of RSSI and RSRP with spectrum analyzer in cell 7 in LOS conditions.
6. Drive a route for Sector 2 for downlink measurement collection.
7. Drive a route for Sector 2 for uplink measurement collection.
8. Drive a route for Sector 1 for uplink measurement collection.

A.8 Analysis of the Third Field Campaign

A.8.1 RSRP and DL Throughput

Maximum DL throughput for LTE Release 8, that can be supported by the phone of category 4, is 150 Mbps. Despite the fact that there is no significant difference in RSRP CDF compared to multiple previous trials, throughput, acquired during the last round of measurements, reached much higher values: 22% of time DL throughput was 100 Mbps and higher (Fig. A.16, Fig. A.17), while in the previous trials the maximum value was 80 Mbps (Fig. A.8). Moreover, it is valuable that when the UE is in LOS conditions, it achieves the maximum possible throughput. On the other hand, throughput CDF is not improved significantly, which can be explained by the huge fraction of track with very low connectivity or without any connectivity at all (Fig. A.19a, Fig. A.19b). These blind or semi-blind zones illustrate how the total coverage is decomposed into areas served by a particular sector with limited intersection between them. The intersections were designated by handover events observed during the first and second field campaigns (Fig. A.4).

A.8.2 UL Throughput

Similar to DL throughput in Appendix A.8.1, the phone registers a maximum UL throughput (50 Mbps) more often (Fig. A.18a, Fig. A.18b), than during the previous campaigns (Fig. A.9a, Fig. A.9b).

A.8.3 SNR vs Throughput

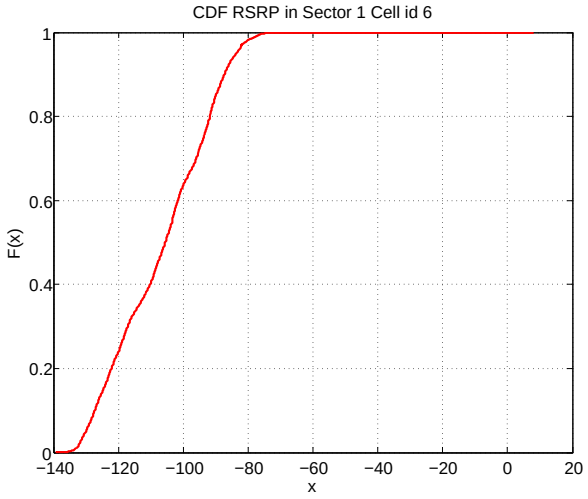
As depicted on Fig. A.20, Fig. A.21, SINR values close to 0 dB restrict throughput to almost zero, while the high throughput of 145 Mbps corresponds to the SINR value of 37 dB, compared to the previous campaign,where same high SINR level resulted in 70 Mbps (Fig. A.13). The results of the third campaign are significantly higher, which is explained by the fact that with the use of UDP, throughput adapts to the channel conditions much faster than when TCP is used.

A.8.4 MCS vs SINR

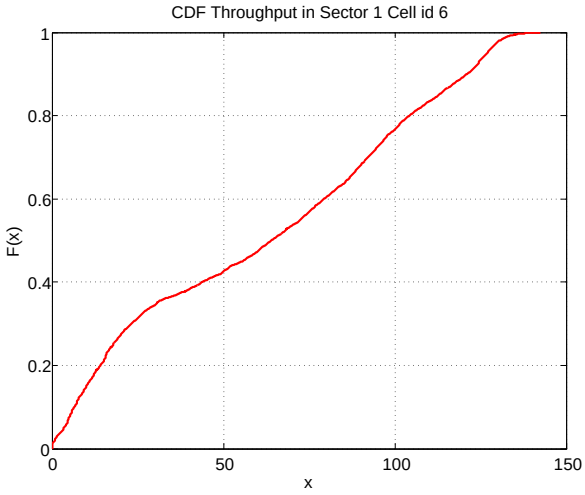
We observe frequent MCS fluctuations on Fig. A.23, which also supports the hypothesis that the reason for unsatisfying measurements during the previous campaigns was the TCP usage. The MCS distribution on Fig. A.22a shows that the majority of time MCS is higher than 20, while for the previous campaigns such high values were exceptional.

A.8.5 MCS, CQI and RSRP

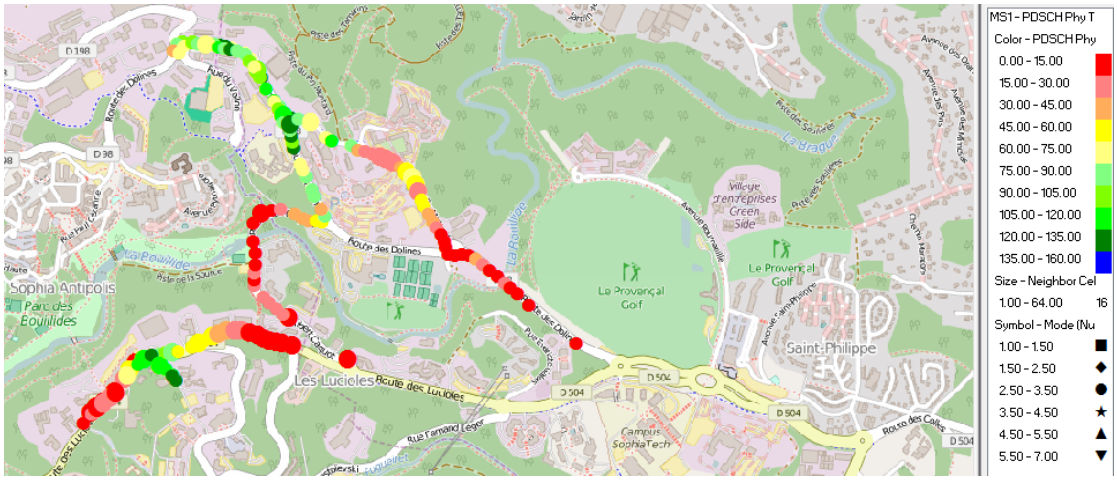
The RSRP and CQI conditions did not show a dramatical change, compared to the previous trials, however, the corresponding assigned MCS is higher (Fig. A.24), leading to higher throughput.



(a) RSRP Sector 1 Cell 6



(b) DL Throughput Sector 1 Cell 6

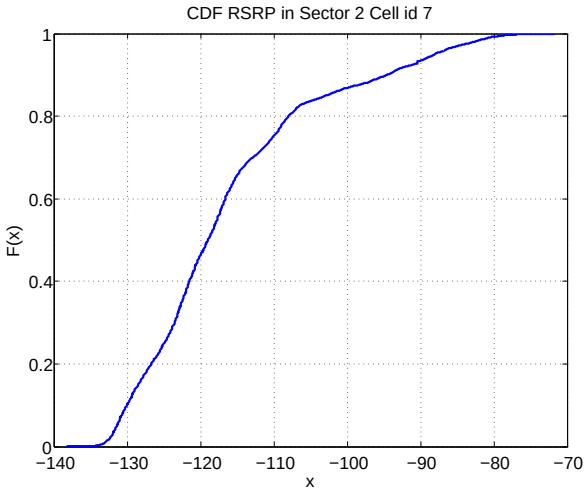


(c) DL Throughput Sector 1 Cell 6 on the track

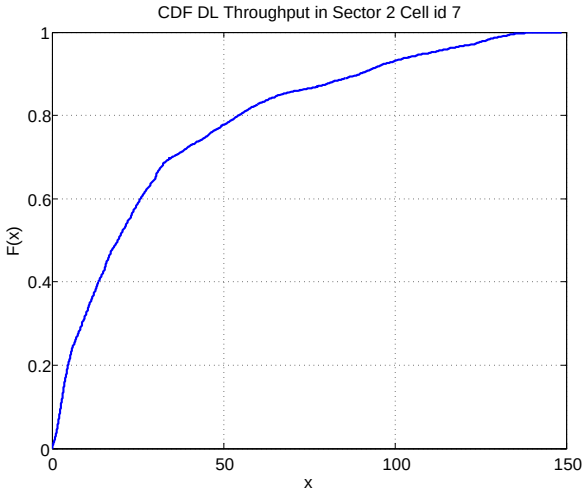
Figure A.16: RSRP and DL Throughput Sector 1 Cell 6

A.8.6 Transmission Mode, Transmission Scheme and Rank

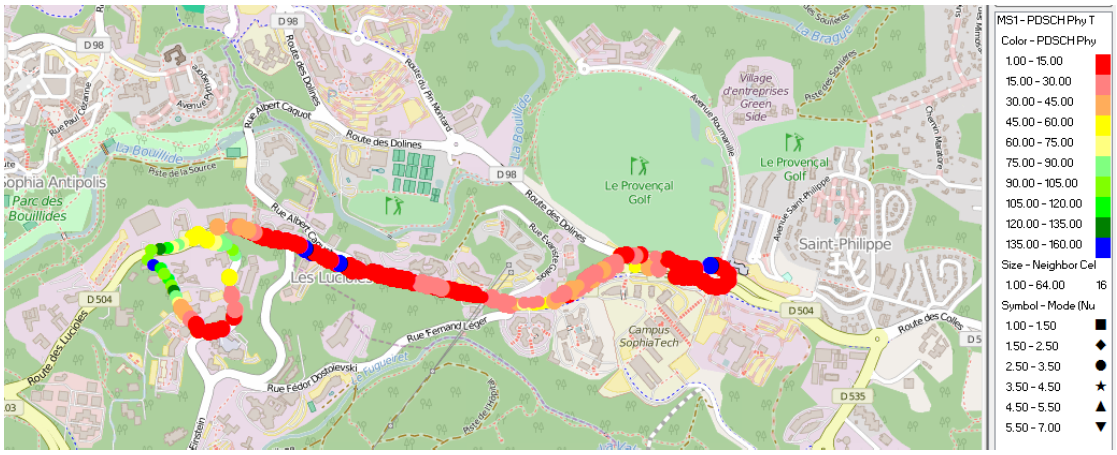
Multiple-Input Multiple-Output (MIMO) design is a key technique to increase throughput in wireless communication systems. Each particular MIMO configuration used in LTE system is called Transmission Mode (TM). Depending on Downlink Control Information (DCI) format being used, the TM follows one of the Transmission Schemes (TS), that are assigned based on scheduling politics and rank estimation [3GPP, 2016]. Rank of the channel matrix shows, what number of spacial layers the system can support under current channel conditions. For example, TM3 acts as Transmit Diversity (TS2), if Rank is 1, or Cyclic Delay Diversity (TS3), if Rank is 2, while TM2 is always configured in Transmit Diversity scheme.
As we see in Fig. A.28a and Fig. A.28b, the UE is configured in TM3 almost 75% of time, but the rank is 2 only during 60% of the time; this means that Transmit Diversity was



(a) RSRP Sector 2 Cell 7



(b) DL Throughput Sector 2 Cell 7



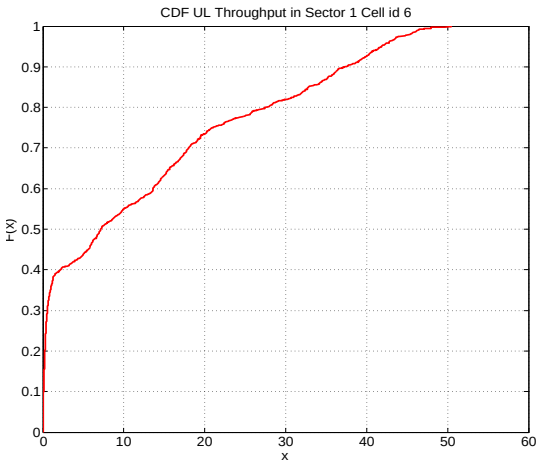
(c) DL Throughput Sector 1 Cell 6 on the track

Figure A.17: RSRP and DL Throughput Sector 2 Cell 7

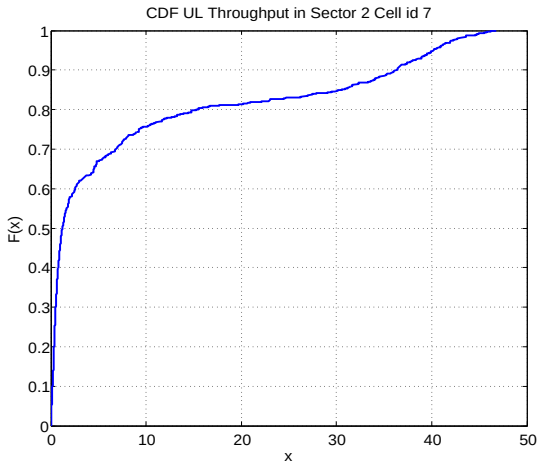
used on 40% of coverage, and another 60% were for Cyclic Delay Diversity (Fig. A.25, Fig. A.26, Fig. A.27).

A.8.7 Transmission Scheme,CQI and RSRP

At this point we want to check how the assigned transmission scheme depends on the channel conditions. CQI and RSRP parameters were chosen to characterize channel state. From Fig. A.29, it is clear that at the certain value of CQI and RSRP, the transmission scheme changes to the lower one in order to provide more reliable communications in poor channel conditions.

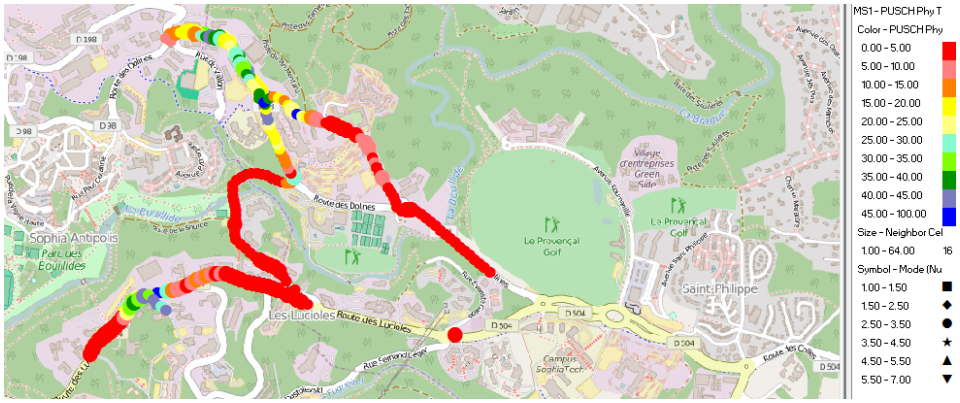


(a) CDF UL Throughput Sector 1 Cell 6



(b) CDF UL Throughput Sector 2 Cell 7

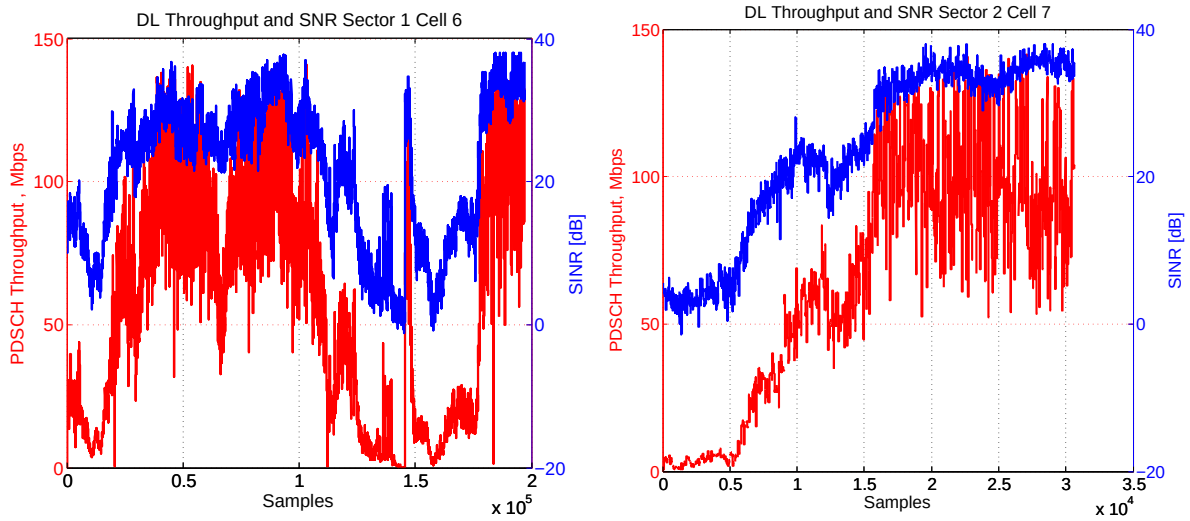
Figure A.18: CQI vs SINR



(a) Sector 1 Cell 6

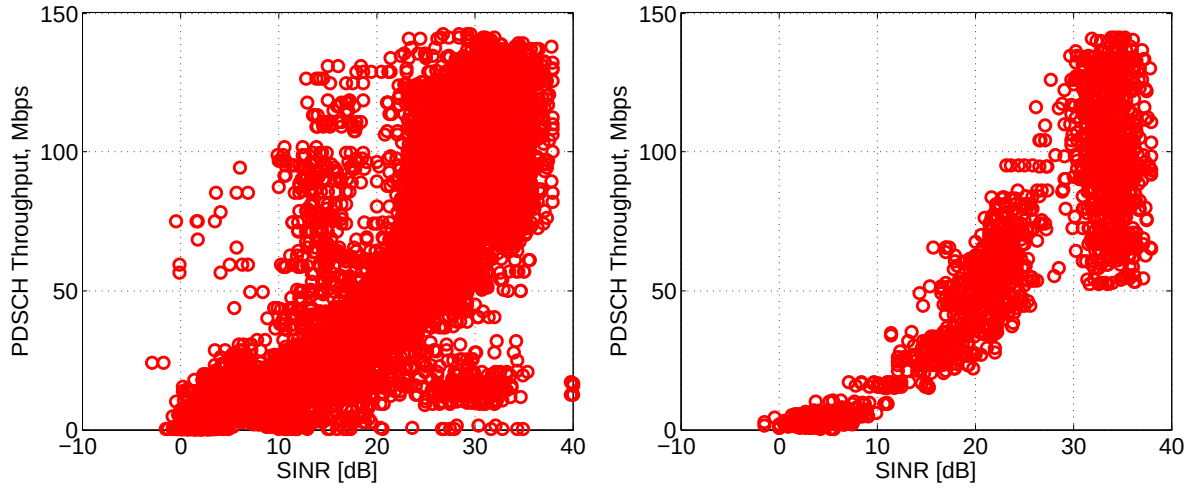
(b) Sector 2 Cell 7

Figure A.19: UL Throughput on the track



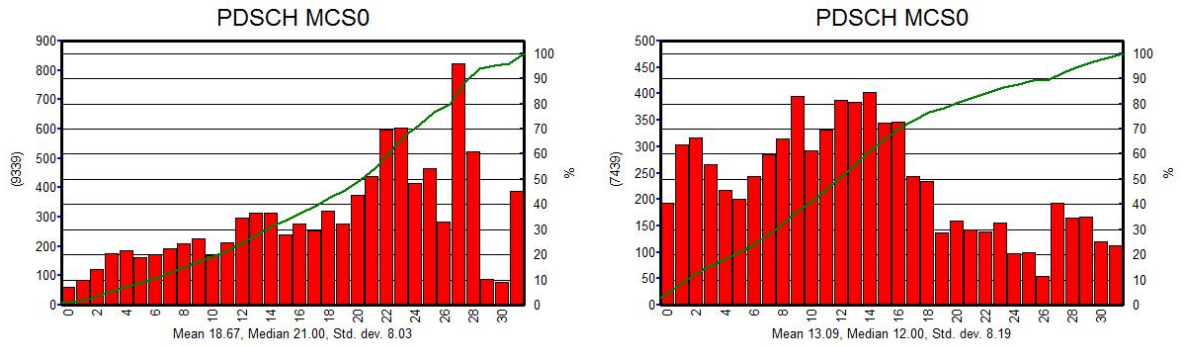
(a) Sector 1 Cell 6 (b) Sector 2 Cell 7

Figure A.20: Throughput Adaptation to channel conditions



(a) Sector 1 Cell 6 (b) Sector 2 Cell 7

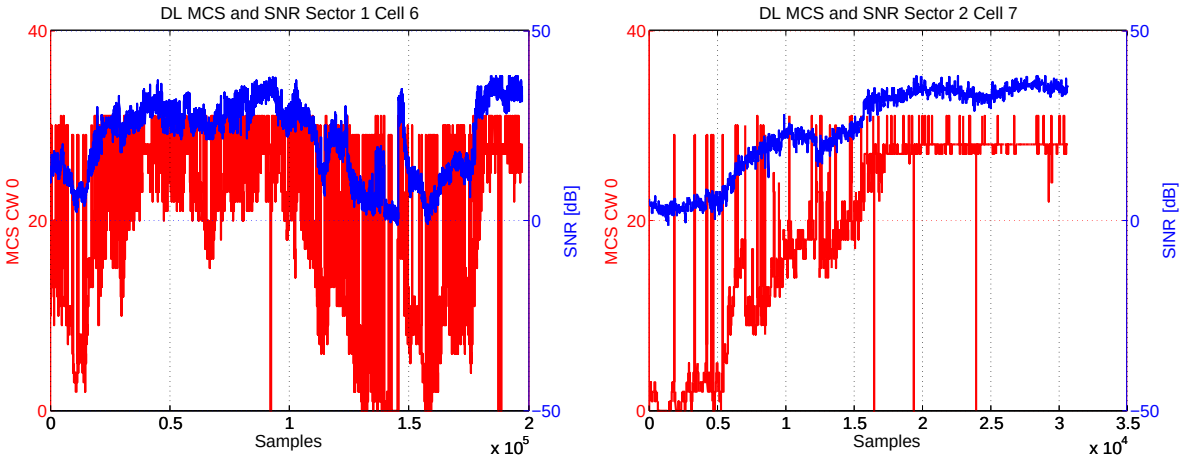
Figure A.21: SINR vs Throughput



(a) Sector 1 Cell 6

(b) Sector 2 Cell 7

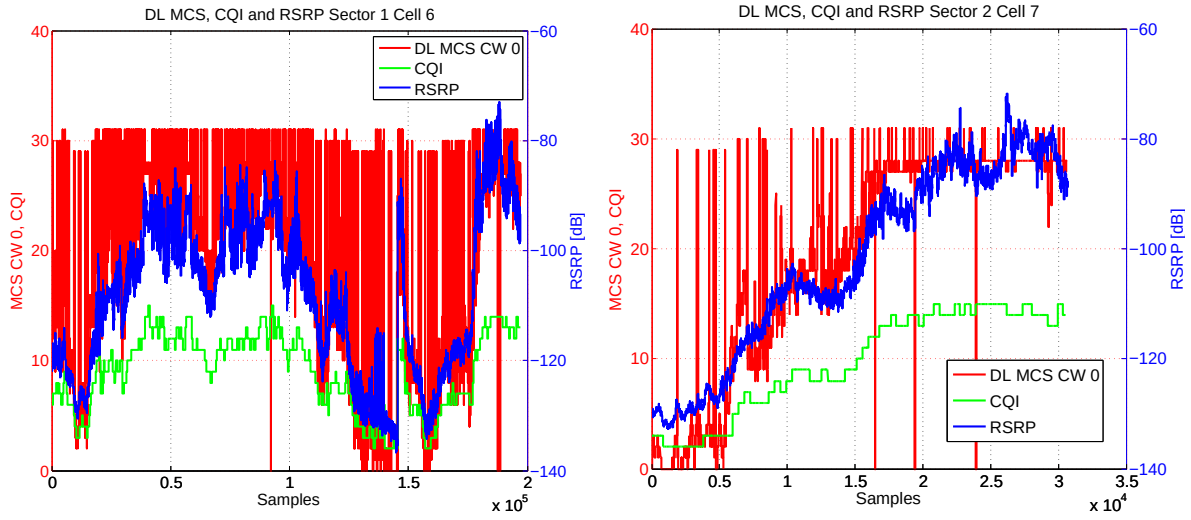
Figure A.22: MCS Statictics



(a) Sector 1 Cell 6

(b) Sector 2 Cell 7

Figure A.23: MCS vs SINR



(a) Sector 1 Cell 6

(b) Sector 2 Cell 7

Figure A.24: MCS, CQI and RSRP

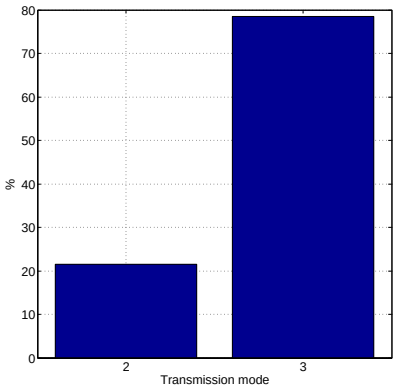


Figure A.25: Transmission Mode in Sector 1

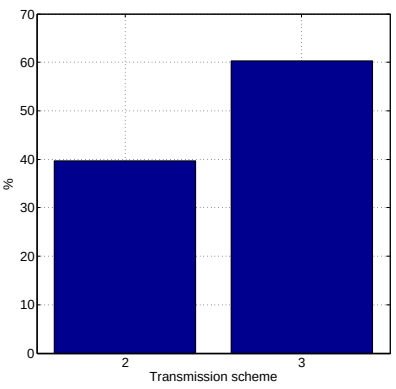


Figure A.26: Transmission Scheme in Sector 1

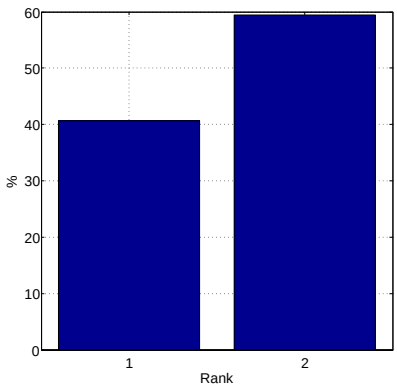
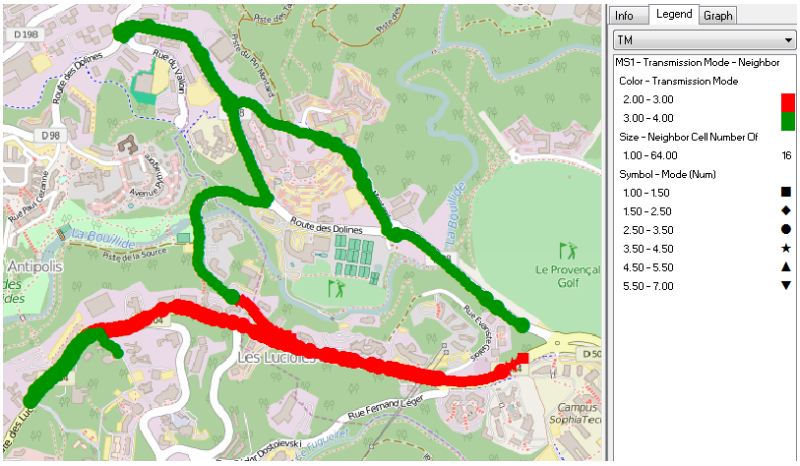
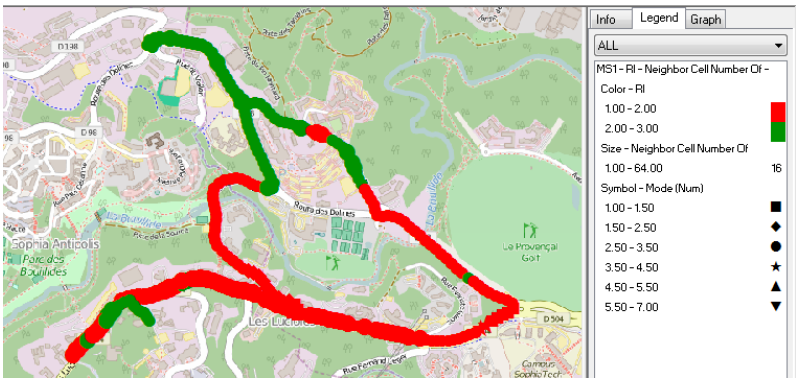


Figure A.27: Transmission Rank in Sector 1

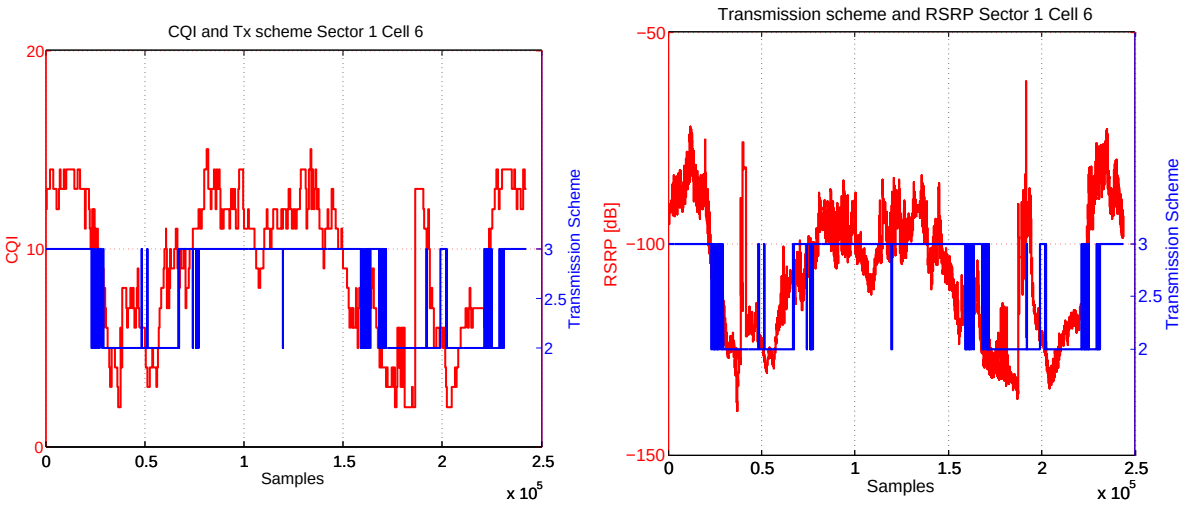


(a) Transmission mode Sector 1 Cell 6



(b) Rank Indicator Sector 1 Cell 6

Figure A.28: Transmission mode and Rank Indicator in Sector 1 Cell 6 on the track



(a) Sector 1 Cell 6

(b) Sector 2 Cell 7

Figure A.29: Transmission Scheme, CQI and RSRP in Sector 1

A.9 Conclusion

During the third campaign, the following modifications were done:

1. Areas served by Sector 1 and Sector 2 were treated individually having only one active eNodeB Sector. This solution helped to eliminate possible interference.
2. The downlink and uplink were be generated by the means of iperf with UDP protocol to benefit from fast throughput adaptation to the channel conditions.

Based on the numerical results, it is clear that those points had a valuable impact on the network performance. The UE was able to register values up to 140 Mbps on DL and up to 45 Mbps on UL, which are very close to the maximum possible throughput (150 and 50 Mbps respectively for LTE Release 8). The throughput values are in a good correlation with the assigned MCS and channel conditions, represented by CQI and RSRP.

Bibliography

(2017). Openairinterface.

3GPP (2010). Self-configuring and self-optimizing network (SON) use cases and solutions. Technical Report TR 36.902 v9.2.0 Release 9, 3GPP.

3GPP (2012). Enhanced performance requirement for lte user equipment (ue. Technical Report TR 36.829 v11.1.0 Release 11, 3GPP.

3GPP (2013). Service requirements for home node b (hnb) and home enodeb (henb). Technical Report TS 22.220 v10.3.0 Release 10, 3GPP.

3GPP (2014). Network-assisted interference cancellation and suppression for lte. Technical Report TR 36.866 v12.0.1 Release 12, 3GPP.

3GPP (2015a). Evolved Universal Terrestrial Radio Access (E-UTRA); Mmedium Access Control (MAC) protocol specification. Technical Report Technical Specification 36.321–V12.5.0 Release 12, 3GPP.

3GPP (2015b). Evolved universal terrestrial radio access (e-utra); multiplexing and channel coding. Technical Report Technical Specification 36.212–V12.7.0 Release 12, 3GPP.

3GPP (2015c). Evolved universal terrestrial radio access (e-utra); physical channels and modulation. Technical Report Technical Specification 36.211 V12.6.0 Release 12, 3GPP.

3GPP (2015d). Evolved universal terrestrial radio access (e-utra); User Equipment (UE) radio transmission and reception. Technical Report Technical Specification 36.101–V12.9.0 Release 12, 3GPP.

3GPP (2015e). Study on downlink multiuser superposition transmission for lte. Technical Report RP-150496, 3GPP.

3GPP (2016). Evolved universal terrestrial radio access (e-utra); physical layer procedures. Technical Report Technical Specification 36.213 V13.0.0 Release 13, 3GPP.

3GPP (2017). Study on new radio access technology.physical layer aspects. Technical Report TR 38.802 v14.0.0 Release 14, 3GPP.

4GAmericas (2013). MIMO and Smart Antennas for Mobile Broadband Systems.

Agilent (2009). Mimo performance and condition number in lte test. *Literature Number 5990-4759EN*.

Akay, E. and Ayanoglu, E. (2004). Low complexity decoding of bit-interleaved coded modulation for m-ary qam. In *Communications, 2004 IEEE International Conference on*, volume 2, pages 901–905 Vol.2.

Alamouti, S. M. (1998). A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 16(8):1451–1458.

Alouini, M.-S. and Goldsmith, A. J. (1999). Capacity of rayleigh fading channels under different adaptive transmission and diversity-combining techniques. *IEEE Transactions on Vehicular Technology*, 48(4):1165–1181.

Andrews, J., Choi, W., and Heath, R. (2007). Overcoming interference in spatial multiplexing MIMO cellular networks. *Wireless Communications, IEEE*, 14(6):95–104.

Aubert, S., Mohaisen, M., Nouvel, F., and Chang, K. (2010). Parallel QR decomposition in LTE-A systems. In *2010 IEEE 11th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5.

Aubert, S., Nouvel, F., and Nafkha, A. (2009). Complexity Gain of QR Decomposition Based Sphere Decoder in LTE Receivers. In *2009 IEEE 70th Vehicular Technology Conference Fall*, pages 1–5.

Azzam, L. and Ayanoglu, E. (2008). An efficient tree search for reduced complexity sphere decoding. In *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, pages 1–4.

Bai, Z., Spiegel, C., Bruck, G. H., Drewes, C., Berkmann, J., Gunzelmann, B., Scholand, T., and Jung, P. (2009). V-blast type mimo schemes for utra lte. *Procedia Earth and Planetary Science*, 1(1):1489 – 1497.

Bai, Z., Spiegel, C., Bruck, G. H., Jung, P., Horvat, M., Berkmann, J., Drewes, C., and Gunzelmann, B. (2010a). Closed loop transmission with precoding selection in lte/lte-advanced system. In *2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010)*, pages 1–5.

Bai, Z., Spiegel, C., Bruck, G. H., Jung, P., Horvat, M., Berkmann, J., Drewes, C., and Gunzelmann, B. (2010b). Dynamic transmission mode selection in lte/lte-advanced system. In *2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010)*, pages 1–5.

Berrou, C. and Glavieux, A. (1996). Near optimum error correcting coding and decoding: turbo-codes. *IEEE Transactions on Communications*, 44(10):1261–1271.

Berrou, C., Glavieux, A., and Thitimajshima, P. (1993). Near shannon limit error-correcting coding and decoding: Turbo-codes. 1. In *Communications, 1993. ICC '93 Geneva. Technical Program, Conference Record, IEEE International Conference on*, volume 2, pages 1064–1070 vol.2.

Bilel, B. R., Navid, N., Raymond, K., and Christian, B. (2011). Openairinterface large-scale wireless emulation platform and methodology. In *In PM2HW2N 6th ACM International Workshop on Performance Monitoring, Measurement and Evaluation of Heterogeneous Wireless and Wired Networks*, pages 109–112.

Bjerke, B. A., Ketchum, J., Walton, R., Nanda, S., Medvedev, I., Wallace, M., and Howard, S. (2005). Packet error probability prediction for system level simulations mimo-ofdm based 802.11n wlans. In *IEEE International Conference on Communications, 2005. ICC 2005. 2005*, volume 4, pages 2538–2542 Vol. 4.

Bjorck, A. (1994). Numerics of gram-schmidt orthogonalization. *Linear Algebra and its Applications*, 197:297 – 316.

Brueninghaus, K., Astély, D., Salzer, T., Visuri, S., Alexiou, A., Karger, S., and Seraji, G.-A. (2005). Link performance models for system level simulations of broadband radio access systems. In *in Proceedings IEEE International Symposium PIMRC*, volume 4, pages 2306–2311.

Butler, M. R. G. and Collings, I. B. (2004). A zero-forcing approximate log-likelihood receiver for mimo bit-interleaved coded modulation. *IEEE Communications Letters*, 8(2):105–107.

Bölcskei, H., Borgmann, M., and Paulraj, A. J. (2002). Performance of space-frequency coded broadband ofdm under real-world propagation conditions. In *2002 11th European Signal Processing Conference*, pages 1–4.

Caire, G., Taricco, G., and Biglieri, E. (1998). Bit-interleaved coded modulation. *IEEE Transactions on Information Theory*, 44(3):927–946.

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

Catreux, S., Erceg, V., Gesbert, D., and Heath, R. W. (2002). Adaptive modulation and mimo coding for broadband wireless data networks. *IEEE Communications Magazine*, 40(6):108–115.

Chen, X., Yi, H., Luo, H., Yu, H., and Wang, H. (2011a). A novel cqi calculation scheme in ltelte-a systems. In *2011 International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–5.

Chen, X., Yi, H., Luo, H., Yu, H., and Wang, H. (2011b). A novel cqi calculation scheme in ltelte-a systems. In *2011 International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–5.

Chen, Z., Li, H., Cui, G., and Rangaswamy, M. (2014). Adaptive transmit and receive beamforming for interference mitigation. *Signal Processing Letters, IEEE*, 21(2):235–239.

Cheng, J. F., Nimbalker, A., Blankenship, Y., Classon, B., and Blankenship, T. K. (2008). Analysis of circular buffer rate matching for lte turbo code. In *2008 IEEE 68th Vehicular Technology Conference*, pages 1–5.

Cho, Y., Kim, J., Yang, W., and Kang, C. (2010). *MIMO-OFDM Wireless Communications with MATLAB*. Wiley - IEEE. Wiley.

Chun, Y. J. and Kim, S. W. (2008). Log-likelihood-ratio ordered successive interference cancellation in multi-user, multi-mode MIMO systems. *Communications Letters, IEEE*, 12(11):837–839.

Conway, J. H., Sloane, N. J. A., and Bannai, E. (1987). *Sphere-packings, Lattices, and Groups*. Springer-Verlag New York, Inc., New York, NY, USA.

Damen, M. O., Gamal, H. E., and Caire, G. (2003). On maximum-likelihood detection and the search for the closest lattice point. *IEEE Transactions on Information Theory*, 49(10):2389–2402.

de Jong, Y. L. C. and Willink, T. J. (2005). Iterative tree search detection for mimo wireless systems. *IEEE Transactions on Communications*, 53(6):930–935.

Dekorsy, A. (2005). A cutoff rate based cross-layer metric for MIMO-HARQ transmission. In *2005 IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 4, pages 2166–2170 Vol. 4.

Dhagle, S. (2009). Lte resource grid.

Donthi, S. and Mehta, N. (2011). An accurate model for eesm and its application to analysis of cqi feedback schemes and scheduling in lte. *Wireless Communications, IEEE Transactions on*, 10(10):3436–3448.

ElBamby, M. S., Bennis, M., Saad, W., and Latva-aho, M. (2014). Dynamic uplink-downlink optimization in tdd-based small cell networks. *CoRR*, abs/1402.7292.

Ericsson (2016). Ericsson Mobility Report. Technical Report On the pulse of the networked society, Ericsson.

Fincke, U. and Pohst, M. (1985). Improved methods for calculating vectors of short length in a lattice, including a complexity analysis. 44(170):463–471.

Foschini, G. J. and Gans, M. J. (1998). On limits of wireless communications in a fading environment when using multiple antennas. *Wireless personal communications*, 6(3):311–335.

Fujino, T., Wakazono, S., and Sasaki, Y. (2009). A gram-schmidt based lattice-reduction aided mmse detection in mimo systems. In *GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference*, pages 1–8.

Galiotto, C., Crowley, H., Marchetti, N., and Doyle, L. (2015). Accuracy vs. complexity trade-off in simulations of future wireless networks. In *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1687–1691.

Gesbert, D., Kountouris, M., Jr., R. W. H., b. Chae, C., and Salzer, T. (2007). Shifting the mimo paradigm. *IEEE Signal Processing Magazine*, 24(5):36–46.

Ghaffar, R. (2010). *Interference mitigation in multi-antenna systems*. PhD thesis, Thesis.

Ghaffar, R. (2012). Lte-advanced multi-user MIMO: Improved feedback and precoding design. In *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, pages 1–5.

Ghaffar, R. and Knopp, R. (2009). Linear precoders for multiuser mimo for finite constellations and a simplified receiver structure under controlled interference. In *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, pages 1431–1435.

Ghaffar, R. and Knopp, R. (2010a). Low complexity metrics for bicm siso and MIMO systems. In *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*, pages 1–6.

Ghaffar, R. and Knopp, R. (2010b). Making multiuser mimo work for lte. In *21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 625–628.

Ghaffar, R., Knopp, R., and Ho, P. H. (2015). Low Complexity BICM MIMO OFDM Demodulator. *IEEE Transactions on Wireless Communications*, 14(1):558–569.

Gilbert, S., Needham, M., and Crisler, K. (1996). Communication of data reception history information. US Patent 5,559,810.

Godara, L. C. (1997). Application of antenna arrays to mobile communications. ii. beamforming and direction-of-arrival considerations. *Proceedings of the IEEE*, 85(8):1195–1245.

Golden, G. D., Foschini, C. J., Valenzuela, R. A., and Wolniansky, P. W. (1999). Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture. *Electronics Letters*, 35(1):14–16.

Goldenbaum, M., Akl, R. A., Valentin, S., and Stańczak, S. (2011). On the effect of feedback delay in the downlink of multiuser OFDM systems. In *2011 45th Annual Conference on Information Sciences and Systems*, pages 1–6.

Goldsmith, A. J. and Chua, S. G. (1998). Adaptive coded modulation for fading channels. *IEEE Transactions on Communications*, 46(5):595–602.

Golub, G. H. and Van Loan, C. F. (2012). Matrix computations.

Guey, J.-C., Fitz, M. P., Bell, M. R., and Kuo, W.-Y. (1999). Signal design for transmitter diversity wireless communication systems over rayleigh fading channels. *IEEE Transactions on Communications*, 47(4):527–537.

Hanzaz, Z. and Schotten, H. (2013). Analysis of effective sinr mapping models for MIMO OFDM in lte system. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2013 9th International*, pages 1509–1515.

Hassibi, B. and Hochwald, B. M. (2002). High-rate codes that are linear in space and time. *IEEE Transactions on Information Theory*, 48(7):1804–1824.

Hayes, J. (1968). Adaptive feedback communications. *IEEE Transactions on Communication Technology*, 16(1):29–34.

Hedayat, A. and Nosratinia, A. (2007). Outage and diversity of linear receivers in flat-fading mimo channels. *IEEE Transactions on Signal Processing*, 55(12):5868–5873.

IEEE (2009). 802.16m Evaluation Methodology Document (EMD).

Jalden, J. and Ottersten, B. (2005). On the complexity of sphere decoding in digital communications. *IEEE Transactions on Signal Processing*, 53(4):1474–1484.

Jang, E. W., Lee, J., Lou, H. L., and Cioffi, J. M. (2007). Optimal combining schemes for MIMO systems with hybrid arq. In *2007 IEEE International Symposium on Information Theory*, pages 2286–2290.

Kaltenberger, F., Knopp, R., Nikaein, N., Nussbaum, D., Gauthier, L., and Bonnet, C. (2015). OpenAirInterface: Open-source software radio solutions for 5G. In *EUCNC 2015, European Conference on Networks and Communications, 29 June-02 July 2015, Paris, France*, Paris, FRANCE.

Kim, J., Moon, S. H., Sung, C. K., and Lee, I. (2011). A new snr prediction method for MIMO-ofdm systems with maximum likelihood detector. In *Communications (ICC), 2011 IEEE International Conference on*, pages 1–5.

Kim, J. S., Moon, S. H., and Lee, I. (2010). A new reduced complexity ml detection scheme for MIMO systems. *IEEE Transactions on Communications*, 58(4):1302–1310.

Kim, S. W. and Kim, K. P. (2006). Log-likelihood-ratio-based detection ordering in v-blast. *IEEE Transactions on Communications*, 54(2):302–307.

Kwon, H., Lee, J., and Kang, I. (2012). Interference mitigation via interference-aware successive decoding. *CoRR*, abs/1209.3824.

Kwon, H., Lee, J., and Kang, I. (2013). Symbol-level combining for hybrid arq on interference-aware successive decoding. In *2013 IEEE Global Communications Conference (GLOBECOM)*, pages 3650–3654.

Lan, Y., Benjebboiu, A., Chen, X., Li, A., and Jiang, H. (2014). Considerations on downlink non-orthogonal multiple access (noma) combined with closed-loop su-MIMO. In *2014 8th International Conference on Signal Processing and Communication Systems (ICSPCS)*, pages 1–5.

Latif, I., Kaltenberger, F., and Knopp, R. (2012). Link abstraction for multi-user MIMO in lte using interference-aware receiver. In *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, pages 842–846.

Latif, I., Kaltenberger, F., Nikaein, N., and Knopp, R. (2013). Large scale system evaluations using PHY abstraction for lte with openairinterface. In *Proceedings of the 6th International ICST Conference on Simulation Tools and Techniques*, pages 24–30. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Lee, H., Lim, J., Park, W., and Kim, T. (2014). Link performance abstraction for interference-aware communications (iac). In *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*, pages 1–5. IEEE.

Lee, H., Lim, J. H., Cho, S., and Kim, S. (2015). Interference cancellation based on blindly-detected interference parameters for lte-advanced ue. In *2015 IEEE International Conference on Communications (ICC)*, pages 3143–3148.

Lee, M. W., Yoon, J. H., and Park, J. (2012). High-speed tournament givens rotation-based qr decomposition architecture for mimo receiver. In *2012 IEEE International Symposium on Circuits and Systems*, pages 21–24.

Lei, X. and Jin-bao, Z. (2010). An accurate mean mutual information computational approach of link performance abstraction. In *Signal Processing (ICSP), 2010 IEEE 10th International Conference on*, pages 1552–1556.

Li, J., Letaief, K. B., and Cao, Z. (2004). A reduced-complexity maximum-likelihood method for multiuser detection. *IEEE Transactions on Communications*, 52(2):289–295.

Lin, S., Costello, D. J., and Miller, M. J. (1984). Automatic-repeat-request error-control schemes. *IEEE Communications Magazine*, 22(12):5–17.

Lindström, M. (2009). LTE-Advanced Radio Layer 2 and RRC Aspects. Technical report, Ericsson 3GPP TSG-RAN WG2.

Love, D. J. and Heath, R. W. (2003). Limited feedback precoding for spatial multiplexing systems. In *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, volume 4, pages 1857–1861 vol.4.

Maltsev, A., Pestretsov, V., Maslennikov, R., and Khoryaev, A. (2006). Triangular systolic array with reduced latency for QR-decomposition of complex matrices. In *2006 IEEE International Symposium on Circuits and Systems*, page 4 pp.

Manchon, C. N., Deneire, L., Mogensen, P., and Sorensen, T. B. (2008). On the design of a MIMO-SIC receiver for LTE downlink. In *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*, pages 1–5.

McKay, M. R., Smith, P. J., Suraweera, H. A., and Collings, I. B. (2008). On the mutual information distribution of OFDM-based spatial multiplexing: Exact variance and outage approximation. *IEEE Transactions on Information Theory*, 54(7):3260–3278.

Miridakis, N. I. and Vergados, D. D. (2013). A survey on the successive interference cancellation performance for single-antenna and multiple-antenna OFDM systems. *IEEE Communications Surveys Tutorials*, 15(1):312–335.

Moon, S.-H., Lee, K.-J., Kim, J., and Lee, I. (2012). Link performance estimation techniques for MIMO-OFDM systems with maximum likelihood receiver. *Wireless Communications, IEEE Transactions on*, 11(5):1808–1816.

Nanda, S., Balachandran, K., and Kumar, S. (2000). Adaptation techniques in wireless packet data services. *IEEE Communications Magazine*, 38(1):54–64.

Nanda, S. and Rege, K. M. (1998). Frame error rates for convolutional codes on fading channels and the concept of effective eb/n0. *IEEE Transactions on Vehicular Technology*, 47(4):1245–1250.

Olmos, J., Ruiz, S., García-Lozano, M., and Martín-Sacristán, D. (2010). Link abstraction models based on mutual information for lte downlink. In *COST*, volume 2100, pages 1–18.

Onggosanusi, E. N., Dabak, A. G., Hui, Y., and Jeong, G. (2003). Hybrid ARQ transmission and combining for MIMO systems. In *Communications, 2003. ICC '03. IEEE International Conference on*, volume 5, pages 3205–3209 vol.5.

Onggosanusi, E. N., Dabak, A. G., Schmidl, T., and Muharemovic, T. (2002). Capacity analysis of frequency-selective mimo channels with sub-optimal detectors. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages III–2369–III–2372.

Paulraj, A. and Kailath, T. (1994). Increasing capacity in wireless broadcast systems using distributed transmission/directional reception (dtdr). US Patent 5,345,599.

Pohst, M. (1981). On the computation of lattice vectors of minimal length, successive minima and reduced bases with applications. *SIGSAM Bull.*, 15(1):37–44.

Polyanskiy, Y., Poor, H. V., and Verdu, S. (2010). Channel coding rate in the finite blocklength regime. *IEEE Transactions on Information Theory*, 56(5):2307–2359.

Polyanskiy, Y., Poor, H. V., and Verdu, S. (2011). Feedback in the non-asymptotic regime. *IEEE Transactions on Information Theory*, 57(8):4903–4925.

Ramesh, R., Koorapaty, H., Cheng, J.-F., and Balachandran, K. (2009). A physical layer abstraction for maximum likelihood demodulation of MIMO signals. In *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, pages 1–5.

Rogalin, R., Bursalioglu, O. Y., Papadopoulos, H., Caire, G., Molisch, A. F., Michaloliakos, A., Balan, V., and Psounis, K. (2014). Scalable synchronization and reciprocity calibration for distributed multiuser mimo. *IEEE Transactions on Wireless Communications*, 13(4):1815–1831.

Roy, R. H. (1997). Spatial division multiple access technology and its application to wireless communication systems. In *1997 IEEE 47th Vehicular Technology Conference. Technology in Motion*, volume 2, pages 730–734 vol.2.

Sandanalakshmi, R., Palanivelu, T., and Manivannan, K. (2007). Effective snr mapping for link error prediction in OFDM based systems. In *Information and Communication Technology in Electrical Sciences (ICTES 2007), 2007. ICTES. IET-UK International Conference on*, pages 684–687. IET.

Sandhu, S. and Paulraj, A. (2000). Space-time block codes: a capacity perspective. *IEEE Communications Letters*, 4(12):384–386.

Sayana, K., Zhuang, J., and Stewart, K. (2007). Link performance abstraction based on mean mutual information per bit (mmib) of the llr channel. *IEEE 802.16 Broadband Wireless Access Working Group*.

Schmidt, E. (1908). *Über die Auflösung Linearer Gleichungen mit Unendlich Vielen Unbekannten*. Rend. Circ. Mat. Palermo (1) 25:53-77.

Schnorr, C. P. and Euchner, M. (1993). Lattice basis reduction: Improved practical algorithms and solving subset sum problems. In *Math. Programming*, pages 181–191.

Schwarz, S., Mehlfuhrer, C., and Rupp, M. (2010). Calculation of the spatial preprocessing and link adaption feedback for 3gpp umts/lte. In *2010 Wireless Advanced 2010*, pages 1–6.

Schwarz, S. and Rupp, M. (2011). Throughput maximizing feedback for MIMO OFDM based wireless communication systems. In *2011 IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications*, pages 316–320.

Seethaler, D., Matz, G., and Hlawatsch, F. (2004). An efficient mmse-based demodulator for mimo bit-interleaved coded modulation. In *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, volume 4, pages 2455–2459 Vol.4.

Sesia, Stefania; Toufik, I. B. M. (2009). *"LTE, The UMTS Long Term Evolution : From theory to practice"*. Wiley, February 2009.

Shabany, M. and Gulak, P. G. (2008). The application of lattice-reduction to the k-best algorithm for near-optimal mimo detection. In *2008 IEEE International Symposium on Circuits and Systems*, pages 316–319.

Shah, A. and Haimovich, A. M. (2000). Performance analysis of maximal ratio combining and comparison with optimum combining for mobile radio communications with cochannel interference. *IEEE Transactions on Vehicular Technology*, 49(4):1454–1463.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423.

Shepard, C., Yu, H., Anand, N., Li, E., Marzetta, T., Yang, R., and Zhong, L. (2012). Argos: Practical many-antenna base stations. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, Mobicom '12, pages 53–64, New York, NY, USA. ACM.

Skillicorn, D. B. (1990). Architecture-independent parallel computation. *Computer*, 23(12):38–50.

Srinivasan, R., Zhuang, J., Jalloul, L., Novak, R., and Park, J. (2008). Ieee 802.16 m evaluation methodology document (emd). *IEEE 802.16 Broadband Wireless Access Working Group*.

Stancanelli, E., Cavalcanti, F., and Silva, Y. (2011). Revisiting the effective sinr mapping interface for link and system level simulations of wireless communication systems. In *Communications (LATINCOM), 2011 IEEE Latin-American Conference on*, pages 1–6.

Steinhardt, A. O. (1988). Householder transforms in signal processing. *IEEE ASSP Magazine*, 5(3):4–12.

Stuber, G. L., Barry, J. R., McLaughlin, S. W., Li, Y., Ingram, M. A., and Pratt, T. G. (2004). Broadband mimo-ofdm wireless communications. *Proceedings of the IEEE*, 92(2):271–294.

Sung, H., Lee, K. B. E., and Kang, J. W. (2003). A simplified maximum likelihood detection scheme for mimo systems. In *2003 IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No.03CH37484)*, volume 1, pages 419–423 Vol.1.

Szczecinski, L., Correa, C., and Ahumada, L. (2010). Variable-Rate Transmission for Incremental Redundancy Hybrid ARQ. In *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, pages 1–5.

Tao, T. and Czylwik, A. (2011). Performance analysis of link adaptation in lte systems. In *2011 International ITG Workshop on Smart Antennas*, pages 1–5.

Tarokh, V., Seshadri, N., and Calderbank, A. R. (1998). Space-time codes for high data rate wireless communication: performance criterion and code construction. *IEEE Transactions on Information Theory*, 44(2):744–765.

Telatar, E. (1999). Capacity of multi-antenna gaussian channels. *European Transactions on Telecommunications*, 10(6):585–595.

Thomas, R., Knopp, R., Maharaj, B., and Cottatellucci, L. (2014). Detection using block QR decomposition for MIMO hetnets. In *Signals, Systems and Computers, 2014 48th Asilomar Conference on*, pages 1291–1295.

Toumpakaris, D., Lee, J., Matache, A., and Lou, H. L. (2008). Performance of MIMO HARQ under receiver complexity constraints. In *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, pages 1–5.

Tuomaala, E. and Wang, H. (2005). Effective sinr approach of link to system mapping in OFDM/multi-carrier mobile network. In *Mobile Technology, Applications and Systems, 2005 2nd International Conference on*.

Vanka, S., Srinivasa, S., Gong, Z., Vizi, P., Stamatiou, K., and Haenggi, M. (2012). Superposition coding strategies: Design and experimental evaluation. *IEEE Transactions on Wireless Communications*, 11(7):2628–2639.

Verdu, S. (2012). Non-asymptotic achievability bounds in multiuser information theory. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1–8.

Villa, T., Merz, R., Knopp, R., and Takyar, U. (2012). Adaptive modulation and coding with hybrid-ARQ for latency-constrained networks. In *European Wireless, 2012. EW. 18th European Wireless Conference*, pages 1–8.

Viterbo, E. and Boutros, J. (1999). A universal lattice code decoder for fading channels. *IEEE Transactions on Information Theory*, 45(5):1639–1642.

Wagdy, A., Khattab, T., and Sourour, E. A. (2013). Modified QR-D and MMSE PMI selection technique for MIMO closed loop spatial multiplexing in LTE/LTE-advanced. In *2013 7th IEEE GCC Conference and Exhibition (GCC)*, pages 93–97.

Wagner, S. and Kaltenberger, F. (2014). Interference-aware receiver design for MU-MIMO in LTE: Real-time performance measurements. *Intel Technology Journal, Volume 18, N°1, 2014*.

Wang, F., Wang, X., and Zhu, Y. (2014). Transmit beamforming for multiuser downlink with per-antenna power constraints. In *Communications (ICC), 2014 IEEE International Conference on*, pages 4692–4697.

Wang, Z. and Giannakis, G. B. (2004a). Outage mutual information of space-time MIMO channels. *IEEE Transactions on Information Theory*, 50(4):657–662.

Wang, Z. and Giannakis, G. B. (2004b). Outage mutual information of space-time mimo channels. *IEEE Transactions on Information Theory*, 50(4):657–662.

Winters, J. H. (1984). Optimum combining in digital mobile radio with cochannel interference. *IEEE Transactions on Vehicular Technology*, 33(3):144–155.

Wolniansky, P., Foschini, G., Golden, G., and Valenzuela, R. (1998). V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel. In *Signals, Systems, and Electronics, 1998. ISSSE 98. 1998 URSI International Symposium on*, pages 295–300.

Wubben, D., Bohnke, R., Kuhn, V., and Kammeyer, K. D. (2003). MMSE Extension of V-BLAST Based on Sorted QR Decomposition. In *2003 IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No.03CH37484)*, volume 1, pages 508–512 Vol.1.

Yan, C., Harada, A., Benjebbour, A., Lan, Y., Li, A., and Jiang, H. (2015). Receiver design for downlink non-orthogonal multiple access (noma). In *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pages 1–6.

Yu, C. H., Hellsten, A., and Tirkkonen, O. (2010). Rate adaptation of amc/harq systems with cqi errors. In *2010 IEEE 71st Vehicular Technology Conference*, pages 1–5.

Zehavi, E. (1992). 8-psk trellis codes for a rayleigh channel. *IEEE Transactions on Communications*, 40(5):873–884.

Zhang, W. and Fang, Y. (2015). Dynamic adjustment of downlink/uplink allocation ratio in tdd wireless systems. US Patent 8,937,891.

Zheng, L. and Tse, D. N. C. (2003). Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *IEEE Transactions on Information Theory*, 49(5):1073–1096.

Zhong, T. M., Madhukumar, A. S., and Chin, F. (2003). QRD-RLS Adaptive Equalizer and its CORDIC-Based Implementation for CDMA Systems. *International Journal on Wireless Optical Communications*, 01(01):25–39.