ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2016-ENST-0081

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Communication et Electronique »

présentée et soutenue publiquement par

Robin Rajan THOMAS

le 09 Decembre 2016

Stratégies de détection MIMO d'ordre supérieur avec applications au relayage pour les réseaux 4G+ et 5G

Directeur de thèse : **Raymond KNOPP**
Co-encadrement de la thèse : **Sunil MAHARAJ**

Jury

M. Bruno CLERCKX, Professeur, IMPERIAL COLLEGE LONDON		Rapporteur
M. Preben MOGENSEN , Professeur, UNIVERSITÉ DE AALBORG/NOKIA-BELL LABS		Rapporteur
M. David GESBERT, Professeur, EURECOM		Examinateur
M. Luc DENEIRE , Professeur, UNIVERSITÉ DE NICE, SOPHIA ANTIPOLIS		Examinateur
M. Sheng YANG , Professeur, CENTRALESUPÉLEC		Examinateur

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

T H È S E

Higher-order MIMO Detection and Half-Duplex Relay Strategies for 4G+ and 5G Networks

Robin Rajan Thomas

A doctoral dissertation submitted to:

TELECOM Paristech

in partial fulfillment of the requirements for the degree of:

Doctorat ParisTech

Specialty : COMMUNICATIONS AND ELECTRONICS

09 December 2016

A jury committee composed of:

Reviewers:
Prof. Bruno Clerckx — Imperial College London, United Kingdom
Prof. Preben Mogensen — Univ. of Aalborg/Nokia-Bell Labs, Denmark

Examiners:
Prof. David Gesbert — Eurecom, France
Prof. Luc Deneire — Univ. of Nice, France
Prof. Sheng Yang — CentraleSupelec, France

Thesis Supervisor:
Prof. Raymond Knopp — Eurecom, France
Thesis Co-Supervisor:
Prof. Sunil Maharaj — Univ. of Pretoria, South Africa

Stratégies de détection MIMO d'ordre supérieur avec applications au relayage pour les réseaux 4G+ et 5G

Robin Rajan Thomas

Pour obtenir le grade de docteur dèlivrè par:

TELECOM Paristech

présentée pour l'obtention du grade de:

Doctorat ParisTech

Spécialité: COMMUNICATION ET ELECTRONIQUE

09 Decembre 2016

Devant le jury composé de:

Rapporteurs:
Prof. Bruno Clerckx — Imperial College London, Royaume-Uni
Prof. Preben Mogensen — Univ. de Aalborg/Nokia-Bell Labs, Danemark

Examinateurs:
Prof. David Gesbert — Eurecom, France
Prof. Luc Deneire — Univ. de Nice, France
Prof. Sheng Yang — CentraleSupelec, France

Directeurs de thèse:
Prof. Raymond Knopp — Eurecom, France
Co-encadrement de la thèse:
Prof. Sunil Maharaj — Univ. de Pretoria, Afrique du Sud

To my Family

"*Imagination will often carry us to worlds that never were. But without it we go nowhere.*"

- C. Sagan

Acknowledgments

Pursuing a PhD is a journey of self-discovery borne out of an internal passion to innovate in areas, which have been thought to be uncharted territory. Part of the journey also involves standing on the shoulders of giants whose collective knowledge and experience serve as a form of guidance along the path toward a PhD. In that very spirit, I would like to extend my sincere and deepest gratitude towards my thesis supervisor, Prof. Raymond Knopp, for initially granting me the opportunity to work with him, in the idyllic South of France. Secondly, all his thoughtful advice and guidance in enabling me to push the bounds of my wireless communication research in both a theoretic and pragmatic manner, despite all the challenges involved, are much appreciated. It has been a great privilege to have worked under his tutelage. I would also like to thank my co-supervisor, Prof. Sunil Maharaj, for his assistance and support, dating all the way back from my Master's studies. My sincerest gratitude goes out to Dr. Martina Cardone and Prof. Daniela Tuninetti for their dedicated and expert collaborative efforts, which had enabled me to make great and significant strides in my research related to half-duplex relaying. I would like to extend my appreciation to my PhD defense jury members, Prof. Bruno Clerckx, Prof. Preben Mogensen, Prof. David Gesbert, Prof. Sheng Yang and Prof. Luc Deneire for their valuable insights and feedback, which has improved the overall quality of the final manuscript.

My parents, Raju and Susan, and sister, Sabena, have always been remote passengers throughout this challenging journey. Thank you for being pillars of inspiration, motivation and support. I will always remember all my fellow friends whose shared experiences and support made pursuing the PhD, all the more worthwhile. An extra special appreciation and note of gratitude to: Giovanni Soldi, Rajeev Gangula, Anikó Mészáros, José Luis Redondo-Garcia, Samuel Kaluvuri, Tanea Zubataia, Leela Gudupudi, Panagiotis Matzakos, Kahesh Dhuness, Imran Latif, Tania Villa, Ayşe Unsal, Ankit Bhamri and Ngoc-Duy Nguyen, all of whom have been there since the beginning. A special thanks to Elena Lukashova and Kalyana Gopala for all the interesting and helpful discussions. Another acknowledgment to the 'Coffee' group whose interesting and highly entropic conversations will always be cherished, including a noteworthy thanks to Ester Gonzalez-Sosa, José Patino (Pepe), Antoine Ghorra (Tony), Pramod Bachhav, Raj Patel, Valeria Chiesa, Massimiliano Todisco (Max), Héctor Delgado, Rui Costa and Chiara Galdi. Another special thanks to my friends outside of Eurecom including Antoine Leboucher, Lorenzo Negri and Fabrizio Monticelli.

ABSTRACT

The evolution of wireless mobile communication networks has always been rapidly progressive in a manner that caters to the performance and quality-of-service requirements of today's data hungry users and content providers, while leveraging the mobile computing power available at one's fingertips. This dissertation presents two key contributions to the body of knowledge in the area of physical layer 4G+/5G wireless communication networks, particularly in the area of Heterogeneous Networks, which complements the current advancements towards next generation mobile technologies.

Interference management between uncoordinated base stations and mobile terminals is a crucial aspect of Heterogeneous Networks (HetNets) in LTE, especially if the interference can be exploited at the receiver. The initial part of this research investigates the challenge of developing higher-order MIMO detection strategies for existing and future LTE receivers while understanding the performance-complexity tradeoff, which are key performance merits. A novel pre-processing Block QR decomposition technique is proposed for a low-complexity Max-log-MAP receiver, for signal detection with four receive antennas in a HetNet interference-limited scenario with a strong dominant interferer. It is shown that for different practical SNR values, the proposed detection scheme has minimal mutual information loss for Gaussian signals in the case of the overall 4×4 and 8×8 MIMO cases in a Rayleigh channel. The Block Minimum Mean Square Error (MMSE) scheme, which assumes that the overall interference is Gaussian, is then compared with the proposed Block QR technique. The Block QR scheme is then implemented in a LTE-compliant simulation test bench with a baseband transmit and receive chain and compared to a brute-force search Max-log-MAP algorithm, which serves as an optimal benchmark in terms of performance, while being prohibitively complex for practical implementation. Furthermore, another application scenario involving a single-user MIMO point-to-point (P2P) scenario is investigated in order to perform throughput maximisation based on the limit of two codewords per transmission in LTE. A feasible performance-complexity trade-off is shown with respect to the Block QR detection scheme.

The second part of the dissertation presents a practical feasibility study of a novel two-phase three-part-message strategy for physical layer half-duplex relay network, which features superposition coding and interference-aware cancellation decoding. A key objective is to analyse the performance

of the proposed scheme in the non-asymptotic regime, and evaluate the resulting spectral efficiency with finite block-length and discrete constellation signaling and then comparing it to the theoretical performance of Gaussian codes with asymptotically large block-lengths. Similar to Part I of the dissertation, the performance evaluations are validated using the developed LTE simulation test bench with off-the-shelf blocks and practical codes. During each transmission phase, the modulation and coding scheme, as defined in the LTE standard, is adapted to the channel link qualities to enhance the overall spectral efficiency (using the developed link adaptation policy). Using a single-antenna source and relay, and a multi-antenna destination the proposed scheme is investigated using a static Gaussian and two frequency-selective channel models, viz. the EPA and ETU channel models. A spectral efficiency comparison with a baseline scheme (non-cooperative two-hop transmission, i.e., the source-destination link is absent) and with the point-to-point transmission strategy (no relay) is also presented. The results confirm that physical-layer cooperation and multi-antennas are critical for overall spectral efficiency enhancement in heterogeneous networks. Moreover, they show that the advantages of physical layer cooperation can be leveraged using existing LTE coded-modulation and interference-mitigation techniques, which are prevalent in modern user-equipments.

LIST OF ACRONYMS

16-QAM	Sixteen Quadrature Amplitude Modulation
64-QAM	Sixty-four Quadrature Amplitude Modulation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
AWGN	Additive White Gaussian Noise
AMC	Adaptive Modulation Coding
BS	Baseline
BCCH	Broadcast Control Channel
BER	Bit Error Rate
Bits/dim	Bits/dimension
BLER	Block Error Rate
CDF	Cumulative Distribution Function
CoMP	Coordinated Multipoint Transmission/Reception
CGS	Classical Gram-Schmidt
CRC	Cyclic Redundancy Check
CRS	Cell-specific Reference Signal
CSCG	Circularly Symmetric Complex Gaussian
CSI	Channel State Information
DF	Decode-and-Forward
DCI	Downlink Control Index
DL	Downlink
DL-SCH	Downlink Shared Channel
eNB	Evolved-NodeB
ETSI	European Telecommunications Standards Institute
EPA	Extended Pedestrian A
ETU	Extended Typical Urban
E-UTRAN	Evolved-UTRAN

EVD	EigenValue Decomposition
FDD	Frequency Division Multiplexing
FEC	Forward Error Correction
FLOPS	Floating-point Operations
HARQ	Hybrid Automatic Repeat Request
HD	Half-Duplex
HETNET	Heterogeneous Network
ICIC	Inter-cell Interference Coordination
IRC	Interference Rejection Combining
ITU	International Telecommunications Institute
LOS	Line-of-sight
LTE	Long Term Evolution
LLR	Log-likelihood Ratio
MAC	Medium Access Control
MAP	Maximum A Priori
MF	Matched Filter
MCS	Modulation Coding Scheme
MGS	Modified Gram-Schmidt
MIMO	Multiple-Input-Multiple-Output
ML	Maximum-Likelihood
MMSE	Minimum-Mean Square Error
MRC	Maximum Ratio Combining
NAIC	Network-Aided Interference Cancelling
NLOS	Non-Line-of-Sight
P2P	Point-to-Point
PDCCH	Physical Downlink Control Channel
PHY	Physical Layer
PRB	Physical Resource Blocks
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RV	Redundancy Version
SISO	Single-Input-Single-Output

SIMO	Single-Input-Multiple-Output
SIC	Successive Interference Cancellation
SNR	Signal-to-Noise Ratio
SNR	Signal-to-Interference-Noise Ratio
TBS	Transport Block Size
TM	Transmission Mode
SU-MIMO	Single-user MIMO
MU-MIMO	Multi-user MIMO
UE	User-Equipment
ZF	Zero Forcing

LIST OF TABLES

LIST OF FIGURES

SYMBOLS AND NOTATION

$\mathbf{A}$	Matrix $\mathbf{A}$
$\mathbf{h}$	Column vector $\mathbf{h}$
$\mathbf{H}$	Channel matrix $\mathbf{H}$
Y	Scalar Y
Y_j	A vector of length j with components $(Y_1, ..., Y_j)$
$\{\mathbf{A}\}^T$	Transpose of matrix $\mathbf{A}$
$\{\mathbf{A}\}^\dagger$	Hermitian transpose of matrix $\mathbf{A}$
$\{\mathbf{a}\}^\star$	Complex conjugate of vector $\mathbf{a}$
$[n_1 : n_2]$	Set of integers from n_1 to $n_2 \geq n_1$
$\|a\|$	Absolute value of a
$\|\|\mathbf{a}\|\|$	Norm of the vector $\mathbf{a}$
$\|\mathbf{A}\|$	Determinant of the matrix $\mathbf{A}$
$\mathbf{I}_j$	Identity matrix of dimension j
$[x]^+ := \max\{0, x\}$	$x \in \mathbb{R}$
$X \sim \mathcal{N}\left(\mu, \sigma^2\right)$	X is a proper-complex Gaussian random variable with mean μ and variance σ^2
$\mathbb{E}[\cdot]$	Expectation operator
log	Logarithms are in base 2
$\mathbf{0}$	Zero vector
$\mathscr{C}$	Gaussian Codebook

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Technology has become an integral component of modern society and the ubiquity of internet connectivity has become a key proponent of this advancement. Wireless communication systems have always undergone evolutionary advances over the years, e.g. deployment of 2G networks between the period 1990-2000 and 3G between the period 2000-2010, in order to meet the growing demands of higher throughputs and Quality of Service (QoS). Currently, fourth generation (4G) networks present a small conundrum for advocates of next generation networks, as it was designed to be adopted well beyond the traditional ten year cycle (2010-2020). Although, the standardisation of fifth generation (5G) networks and its deployment is yet to be finalised, the rate at which it may be initially adopted may be affected by the current successful worldwide adoption of 3GPP's Long-Term Evolution (LTE) networks. This will in itself brings along a host of new use cases and challenges. Whether 5G will prove to be a drastic paradigm shift or a consolidation of existing mobile technologies to improve overall network quality and performance, remains to be seen. Another challenge, is convincing stakeholders such as telecommunication operators and network vendors that the introduction of new network technologies is well worth the investment, will not negatively impact their revenues and will increase overall profit. It is well understood that current releases of the LTE standard have limitations in protocol design for machine-to-machine (M2M) communications, which has been touted as one of the key characteristic technologies of 5G. From a research point of view, it becomes an interesting challenge to investigate the performance trade-offs involved for such a redesign or whether a clean slate approach would be more suitable. This could apply not only from a M2M communication stand-

point but also in the context of other next generation technologies such as Heterogeneous Networks, Massive MIMO and mmWave to name but a few. Furthermore, it would also be of additional interest to explore how theoretical approaches are analysed with respect to practical scenarios, and then quantifying the gaps in performance, so as to satisfy the requirements of 4G+ (Release 14 and beyond) and 5G networks.

1.1 NEXT-GENERATION NETWORK ADVANCES

1.1.1 Fifth Generation (5G) Networks

Currently, key industry players, stakeholders and standardisation bodies are yet to formalise an initial standard and framework for 5G. Fragmented technological advances in next generation technologies for a hyper-connected society are already in existence. These include network function virtualisation (NFV) and software defined networks (SDNs) [5]. Anticipated advances of 5G include:

- A greater emphasis on cooperative strategies in the form of advanced co-ordinated multi-point (CoMP) techniques, advanced relay and cloud radio access network (C-RAN) architectures.
- Enhanced spectrum management through more efficient use of the sparse/disjointed radio spectrum, e.g dynamic spectrum access techniques, mmWave for proximal communications and enhanced carrier aggregation for enhanced higher data rates.
- The inclusion of protocols designed for M2M communications in relation to the Internet of Things (IoT) applications.
- Convergence for rapidly deployable emergency networks designed for public-safety, disaster-relief management and professional networks (e.g. military).
- Support for higher data rates and improved QoS for high-speed vehicular applications.
- Densification of networks for improved coverage and capacity (Heterogeneous Networks/small cells) and interoperability between different communication technologies.

- Massive MIMO systems for improved network coordination, drastic improvements in terms of capacity and throughputs related to the scalability of such systems.

Fig. 1.1 represents an overview of the enabling technologies under the 5G umbrella which has not yet been finalised as yet. A major theme of this thesis is to investigate some of the ways that Heterogeneity can bring about an increase in performance, while maintaining a level of complexity that is manageable for possible implementation in LTE (Rel. 14 and beyond) standards and future 5G systems. A driving force for Heterogeneous Networks (HetNet) deployments is the ability to improve the current performance in terms of throughput and capacity of existing mobile broadband network infrastructure [6], without necessarily adopting a 'clean-slate' approach, which would be favourable to telecom operators and network vendors whose goals may include the least disruption of the current mobile ecosystem to achieve the best performance. This is indeed a debatable and noteworthy discussion, of which certain stakeholders would rather favour a revolution over an evolution.

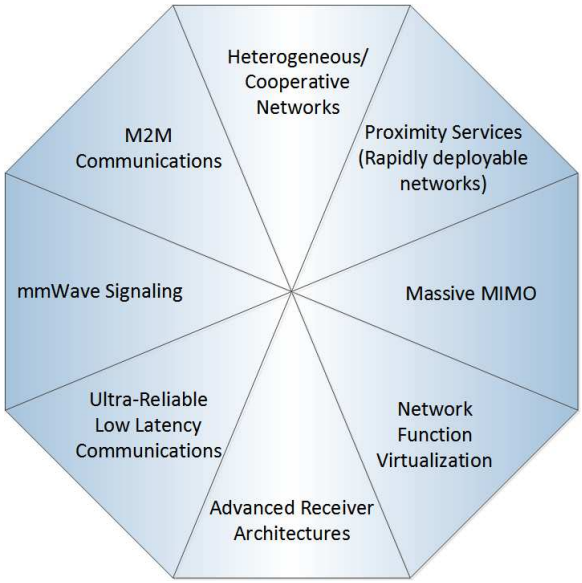


Figure 1.1: Overview of 5G conceptual technologies

1.1.2 Heterogeneous Networks (HetNets)

HetNets can be regarded as a combination of different radio technologies and coverage topologies, all working harmoniously to enhance the end user QoS. HetNets offer the capability to off-load data

traffic from a strained macrocell network on to small cells through specific inter-radio access technologies, e.g. transferring (offload) traffic from a cellular network onto a Wi-Fi access point [7]. From a coverage standpoint, these may include, transmission nodes such as femtocells, pico cells or relay nodes overlaid on to existing macrocells. Another HetNet related technology is the concept of the Remote Radio Head (RRH), which is a small scale device that houses certain functionality of a traditional base station, such as digital interfacing, processing and frequency-agile analogue functionality [6, 8]. Such devices are therefore designed to be compact and low powered, which increase the flexibility of the network through deployment in areas where a macro BS cannot be installed due to physical constraints or difficulty in site acquisition [1, 9]. RRHs can also enable C-RAN technologies by operating as distributed BSs linked with a fiber optic connection to the central BS, and is responsible for control and baseband signal processing. The aim of densifying the network is to enable a significantly higher link quality between the base station and user terminal due to the reduced distance between the two entities. The performance of the network is enhanced through the efficient re-use of the spectrum and is capable of providing higher data rates [1]. Table 1.1 outlines a list of elements that would constitute a typical HetNet with respect to coverage.

The traditional macrocellular network employed in existing networks consists of Base Stations (BSs) installed and maintained by operators, designed to serve a large number of users over a wide coverage area. A minimum data rate together with a maximum allowable delay are guaranteed, while considering outage constraints. A dedicated backhaul to the BS is provided. Picocells are low powered centralised nodes that are also installed by operators and function in a similar fashion to macrocellular networks with a dedicated backhaul. These networks are designed to supplement macrocellular networks by boosting capacity and servicing areas where cell coverage is limited. Hence only a limited number of users are served within a 300 m radius [1]. Femtocells or Home eNBs are also low powered nodes which can be deployed in an uncoordinated fashion by end users in small (homes or office) indoor environments [10]. They are meant to serve a few active users within a small radius of about 50 m. Relays are operator-deployed nodes that are strategically positioned between a macro BS and the end user and functions primarily to enhance signal reception in areas which are poorly provisioned by the macro BS such as tunnels and various environmental barriers (mountainous areas). The versatility of the relay can be enhanced through a wireless backhaul as opposed to a conventional Ethernet backhaul. The concept of HetNets has already been in existence in earlier cellular standards such as GSM, where smaller cells were separated from larger cells based on the frequency mode

Table 1.1: Overview of the different HetNet elements [1]

HetNet Element	Coverage Area (metres)	Environment	Transmit Power (dBm)
Macrocell	> 1000	Outdoor	46
Picocell	100 -300	Outdoor/Indoor	23-30
Femtocell	10-50	Indoor	< 23
Relay	300	Outdoor	30
Remote Radio Head (RRH)	> 1000	Outdoor	46

of operation. In the case of LTE, maximum utilisation of the licensed spectrum is achieved using frequency reuse [11]. Smaller-sized cells have been envisioned as far back as the early 1990's when legacy communication technologies such as CDMA [12] were in use, and analytical studies have highlighted the improvements in spectral efficiency through the reduction in cell size [13]. Femtocells (or Home eNBs) have garnered quite a bit of interest and studies have highlighted the advantages, which include gains in coverage and capacity as well as the reduced cost of deployments [14–17] for small office and home environments. The rudimentary functionalities of Home eNBs (HeNBs) or Home Node Bs (HNBs) were initially specified in Rel. 8 of the LTE standard and consist of the following assumptions [18]:

- Homes, small offices and similar sized locations are the typical environments in which HeNBs/HNBs are to be deployed as small Universal Terrestrial Radio Access Network (UTRAN) and Evolved Universal Terrestrial Radio Access Network (EUTRAN) cells.

- Resource accessibility will be managed by the operators and owners of the HNB and HeNB units.

- Interconnectivity of the HeNB/HNB with the 3G core and evolved packet core through fixed access broadband, e.g. DSL, Cable, etc.

- Support should be provided for full mobility between femtocells and neighbouring cells (macrocells, etc.) including service continuity.

Various studies are conducted before each LTE Release is frozen (or finalised) and one of these studies [19] have led to ways in which physical layer aspects for small cells could be enhanced. Some aspects include the improvement in spectral efficiency with the inclusion of the 256-QAM

(8 bits/dimension) modulation scheme on the downlink. The evaluation results show that an error vector magnitude at the transmitter of less than 4% promotes spectral efficiency gains of around 15-30% without considering receiver impairments. However, practical issues related to the femtocell core network design and certain architectural aspects [20] have impeded the full scale deployment by mobile network operators (MNOs) in current networks.

1.2 PROBLEM DOMAIN

Amongst all the advantages offered by HetNets, e.g. small cells, there are certain key challenges, which should be resolved before such networks are commercially ubiquitous and viable. The following hypothetical use case scenario is illustrated in order to understand the general research challenges associated with small cell deployments, from which two key hypotheses in relation to interference mitigation and relay networks will arise.

Consider that: “Bob lives in an apartment within a medium-sized building situated in a dense part of the city and decides to purchase a HeNB from Operator *A* after their marketing campaign about the sensationally high data rates, which can support his online gaming or virtual/augmented reality applications. After setting up the HeNB, Bob is disappointed to discover that the throughput from his femtocell is not as anticipated and in some cases worse than his previous traditional Wi-Fi router-based digital subscriber line (DSL) connection." Several issues attributed to this scenario, are highlighted below:

- The debate between open and closed access for femtocells is an ongoing topic of discussion. Proponents of security and privacy will argue that HeNBs should belong to a closed subscriber group (CSG) in order for Bob, a licensed and paying customer, to not be negatively impacted by other neighbouring public users sharing the same resource. On the other hand, CSG access will create a deadspot or coverage hole from a macrocellular point of view and degrade the performance of the network and thus an open access scheme would seem more apt. Moreover, uplink interference can be mitigated though communication of public users with the HeNB rather than the macrocell. A hybrid approach involving an adaptive resource sharing scheme could be a possible solution that extracts the best of both open and private use cases [6, 15, 21].

- Handovers between neighbouring femtocell or other macrocells/picocells will be a regular phe-

nomenon, considering the proximities of the many small cell deployments within a highly dense area, such as an apartment block. In terms of Bob's subscription agreement, it should be his prerogative on the time and location of enabling his HeNB and this poses various interference and mobility issues to Operator *A*, with regard to uncoordinated user deployment [21].

- Providing a delay resilient (low latency) backhaul, with an equivalent macrocell QoS to Bob, is a key design challenge from an operator standpoint. In the case of Bob's apartment, a dedicated (fiber) wired small cell backhaul may not be feasible due to scalability and location issues unless Operator *A* is using a shared DSL backhaul. However, this could be a possible bottleneck of the 'last mile' network causing throughput degradation. Millimeter-wave (mmWave) radio has been touted as a possible solution for 5G mobile networks, although line-of-sight (LOS) links are highly desirable to reap the full benefits and may not be available in dense urban settings [22, 23].

- Arguably, one of the most critical challenges facing HetNet deployments, and hence the subject of Part I of this thesis, is to mitigate or exploit interference caused by massive and uncoordinated (in time and space) small cell deployments using advanced higher-order MIMO receiver structures. As a result, Operator *A* has not adequately implemented the necessary interference mitigation techniques (both from a transmitter and receiver perspective) and could therefore be the root cause of Bob's performance degradation and hence poor QoS. Interference arising from neighbouring femtocells (such as Bob's surrounding interferers) and macrocells will be a key challenge to overcome and various strategies would have to be developed in order to adapt to the large amount of uncoordinated deployment scenarios [15, 21].

- If Bob's apartment happens to be in an area with poor coverage resulting in poor performance, Operator *A* may consider an enhanced solution involving the use of half-duplex relays to boost coverage as a viable solution [24, 25]. In Part II of this dissertation, a novel half-duplex physical layer relay technique will be investigated in order to provide significant spectral efficiency gains in addition to the traditional coverage benefits.

1.3 RESEARCH GOALS

The goals of the thesis can be broadly divided into two key parts, which aim to address a subset of the challenges associated with HetNets and receiver design.

- The first part of the thesis investigates the development of a scalable advanced LTE receiver for higher-order MIMO receivers with a complexity reduction approximately in the order of the previously developed low-complexity reduced dimension Max-log-MAP demodulator for a single-user and multi-user MIMO receiver. This receiver is robust in an interference-limited environment provided that knowledge of the interference channel is available. Moreover, this reduced dimension receiver considers that the dominant interferer is non-Gaussian and hence has information, which can be exploited. Currently, providing higher-order antenna configurations at the receiver is an open and challenging problem due to the constraints of the device, e.g. the complexity of decoding algorithms. This goal also falls in line with the idea of developing a higher-order interference-aware receiver capable of exploiting cross-tier/intratier interference for enhanced performance in HetNets. The proposed algorithm is applied to a point-to-point single-user MIMO scenario in order to maximise throughput in an LTE system constrained by a maximum of two codewords per transmission. The aim is to show that the design of higher-order MIMO receivers at lower complexity can promote increased data rates, while being robust in an interference-limited scenarios as exhibited in HetNet scenarios.

- The second part of the thesis is the performance evaluation and feasibility study of a Half-duplex (HD) relay strategy as part of the design of advanced relay architectures for LTE. Telecommunication operators have always been reluctant to employ relays due to the lack of any substantial gains in capacity over existing point-to-point (direct transmission) deployments, despite the state-of-the-art related to the improved network performance of full- and half-duplex relays. The relay scheme has been theoretically studied using Gaussian codes with the assumption of infinite block lengths. However, in practice it may seem that these assumptions may not hold as the block lengths are finite and signal transmissions are drawn from a discrete constellation. Bridging the gap between theory and practice presents a greater understanding into the various aspects of communication design and provides more insights into the various as-

pects of relay network design. The eventual goal is to pave the way for the deployment of high performance partial decode-and-forward (DF) relays with low implementation complexity.

1.4 CONTRIBUTIONS

The adoption of HetNets in next generation networks provides a platform to address many open areas of research, especially in the context of challenges related to interference mitigation. The theme of this dissertation revolves around the design of efficient higher-order MIMO interference-aware algorithms that can be integrated into current LTE standards while also considering LTE Rel. 14 beyond and 5G systems. Understanding the practical implications using theoretical frameworks is fundamental to good communication system design and this also forms part of investigations carried out in the dissertation. The thesis can be broadly divided into two key parts: 1) Higher-order MIMO receiver detection and 2) Advanced Half-Duplex Relay Strategies. The following research contributions resulting from the dissertation are detailed as follows:

1. Higher-order MIMO Receiver Detection

    - A novel detection technique using Block QR decomposition [26] is proposed for interference-aware receivers (with non-Gaussian interference sources) and analytically evaluated against the Block MMSE detector where is it assumed that the interference is Gaussian. These techniques are designed with the aim of implementing higher-order MIMO schemes with reduced complexity in current or near future LTE receivers. The Block QR receiver is based on the low-complexity Max-log-MAP bit metric interference awareness demodulator architecture presented in [27], which involves the exploitation of the dominant non-Gaussian interferer, primarily affecting cell-edge users.

    - The loss in mutual information of the proposed detection strategies are compared and analysed under the assumption of Gaussian signal alphabets. The mutual information of the Block QR is shown to be relatively minimal, especially in higher SNR regimes, while the Block MMSE scheme is lossless.

    - A second application scenario involving a single-user MIMO point-to-point (P2P) scen-

ario is investigated in order to perform rate maximisation across two codewords in an LTE. The feasibility of the Block QR scheme is examined in the context of providing enhanced throughputs at lower-orders of complexity.

- An evaluation of the proposed Block schemes are compared with the classic optimal brute-force Max-log-MAP algorithm using a LTE simulation test bench based on the OpenAirInterface platform [28], and serves as an optimal benchmark when comparing the performance and complexity.

The application of higher-order MIMO detection techniques at the receiver is studied in a typical HetNet scenario with a dominant interferer, where the mobile cell-edge user aims to resolve the received desired and interfering signals from two different base stations. These detection methods are also relevant in a point-to-point MIMO and can be extended to a multi-user MIMO scenario.

2. Advanced Half-duplex Relay Strategy for LTE

A novel link adaptation strategy is proposed for HD relay networks and can be implemented in existing LTE networks with off-the-shelf components and blocks. The key contributions include:

- An assessment of the theoretically optimal (up to a constant gap) two-phase three-part-message scheme by using practical channel codes as specified in the LTE standard and by considering an overall practically relevant system Block Error Rate (BLER) value of 10^{-2} is performed (LTE BLER requirement is 10^{-1}). For the static AWGN SIMO channel, i.e., when the destination is equipped with 2 antennas, it is shown that a theoretical (with equal bandwidth allocation) and practical spectral efficiency gap of 0.28 bits/dim when the source-destination and the relay-destination links are of the same strength, and of 0.67 bits/dim when the relay-destination link is 5 dB higher than the source-destination link. These values indicate that high-throughput HD relay schemes are within practical reach for de facto 4G relays and receivers of today, which already implement turbo decoders and thus incur no additional complexity. Similarly, in the SISO case it is shown

that the maximum spectral efficiency gap between the theoretical and practical implementation (with equal bandwidth allocation) is of 0.31 bits/dim when the strength of the source-destination and relay-destination links is the same and of 0.69 bits/dim when the relay-destination link is 5 dB higher than the source-destination link. The array gain is hence exploited and reveals the benefits of the proposed HD relay strategy when employing two receive antennas at the destination.

- A comparison of the rate performance of this scheme with respect to a baseline strategy is performed, where the link between the source and the destination is absent, i.e., there is no physical layer cooperation between the source and the relay to convey information to the destination. In the SIMO case, the maximum difference between the proposed cooperative strategy and the baseline scheme is 3.39 bits/dim when the channel strength of the source-destination link is the same as the relay-destination link, and 2.88 bits/dim when the relay-destination link is 5 dB higher than the source-destination link. Similarly, for the SISO scenario, spectral efficiency improvements of 3.15 bits/dim when the channel strength of the source-destination link is the same as the one of the relay-destination link, and of 2.97 bits/dim when the relay-destination link is 5 dB higher than the source-destination link, are observed. These values show that the baseline two-hop communication scheme is not beneficial in terms of rate gain, especially when the channel strength of the source-destination link is the same as the relay-destination link.

  A basic point-to-point comparison where a single-part message is transmitted from the source to the destination, is also provided. This analysis shows that enabling physical-layer cooperation among nodes is of critical importance in current and future wireless networks. The proposed scheme can, in fact, enable higher spectral efficiency performance deployment of Layer 1 relays in a manner that would be feasible from a business perspective, something that the standardized methods (what is referred to as the baseline two-hop strategy) are not currently capable of providing, mainly because operators have not found a business case due to the limited spectral efficiency benefits [29]. It has already been shown that from a coverage standpoint, the deployment of mid to high powered relay nodes can yield cost savings of at least 30% for the operators [30].

- Two different practical quasi-static fading channel models, namely the Extended Pedestrian A (EPA) and the Extended Typical Urban (ETU) models are investigated in order to assess the performance of the HD relay strategy in a realistic multipath environment. It was observed that the ETU model (root mean square delay spread of 991 ns) displayed an overall average higher spectral efficiency over the EPA model (root mean square delay spread of 43 ns) [31], thanks to the higher frequency diversity of the former model (smaller coherence bandwidth of the ETU model with respect to the EPA model); the proposed scheme is hence well suited to LTE, which benefits from the resiliency of OFDM in such multipath environments. When the source-destination and the relay-destination links are of the same strength, the rate improvement of the proposed strategy over the baseline two-hop communication scheme is of 2.20 bits/dim for the EPA model and of 1.50 bits/dim for the ETU model. When the relay-destination link is 5 dB higher than the source-destination link, the maximum improvement in rate is 1.71 bits/dim and 1.29 bits/dim for the EPA and ETU models, respectively. The rate gains are less than those in the AWGN case, but provide reasonable improvements over the existing baseline two-hop communication scheme in a more realistic channel scenario.

The outcome of the study reveals that practical implementation of HD relay techniques is possible with the modulation and coding formats already specified by the 3GPP LTE standard, and that the gap between theory and the proposed implementation is small. Optimizing the resource allocation parameters (code rates, bandwidth allocation) in relation to the proposed relay strategy and bench-marking its performance against second-order moderate block-length capacity results, in the spirit of what was initiated for the point-to-point channel in [32], may actually show the near optimality of the proposed two-phase relaying strategy.

1.5 DISSERTATION OUTLINE

The following is a broad outline of each chapter contained within this dissertation. Chapter 2 provides an overview of the existing state-of-the-art and provides the necessary conceptual background for the two key themes of the thesis. The different methods in which to handle interference from both the transmitter and receiver side are discussed, which is then followed by an overview of various popular

non-linear and linear MIMO detection techniques. Thereafter, a basic description of the different relay techniques is provided as well as the current framework in which relays are applied in LTE networks. Since both the block detection technique and HD relay strategy are evaluated in an LTE system with respect to off-the-shelf practical channel codes, a brief description of the LTE standard is given with a special emphasis on the physical layer downlink.

In Chapter 3 a Block QR pre-processing detection technique is developed and draw comparisons with the Block MMSE approach for higher-order MIMO HetNet systems from an information theoretic perspective. These scenarios are studied in an interference-limited scenario and therefore careful consideration of the interferer has to be taken into account when performing the detection of the received signal. A mutual information analysis is provided in order to understand the loss incurred between the theoretical limit and that of the Block QR approach. In addition, the developed Block QR scheme is investigated with respect to a rate optimization scheme for a SU-MIMO scenario. Part of Chapter 3 has already been published in:

- R.R. Thomas, R. Knopp , B. T. Maharaj and L. Cottatellucci, “Detection using Block QR Decomposition for MIMO HetNets", *in the 48th Asilomar Conference on Signals, Systems and Computers*, pp. 1291-1295,Pacific Grove, CA, USA, Nov. 2014.

Chapter 4 presents a complexity and performance evaluation of the aforementioned Block algorithms. Initially, a complexity analysis with respect to the number of floating point operations is analysed in order to develop some insights into the implementation constraints of such techniques. Thereafter the Block Error Rate (BLER) performance is evaluated using an LTE simulation test bench and the performance trade-offs are devised from the results. Parts of this chapter contains contributions that are yet to be submitted for publication:

- R.R. Thomas, R. Knopp , B. T. Maharaj, Novel low-complexity higher-order MIMO interference-limited detection techniques, *IEEE Transactions on Wireless Communications (To be submitted)*

Chapter 5 presents the second part of the dissertation, which entails a theoretical and practical study of a novel HD relay strategy which leverages superposition coding and Successive Interference Cancellation (SIC) decoding for enhanced spectral efficiency gains in an LTE downlink setting. The initial

aim of this study is to bridge the gap between Theory and Practice, in order to understand the limitations of practical systems. Spectral efficiency comparisons of the proposed HD relay scheme are then made with a two-hop baseline (implemented in todays LTE networks) as well as point-to-point schemes using an Additive White Gaussian Noise (AWGN) and two LTE frequency-selective channel models. The results of which have been published in:

- R.R. Thomas, M. Cardone, R. Knopp, D. Tuninetti and B. T. Maharaj, “An LTE implementation of a novel strategy for the Gaussian half-duplex relay channel" *in the 2015 IEEE International Conference on Communications (ICC)*, pp. 2209-2214, London, UK, Jun. 2015.

- R.R. Thomas, M. Cardone, R. Knopp, D. Tuninetti and B. T. Maharaj, “A Practical Feasibility Study of a Novel Strategy for the Gaussian Half-Duplex Relay Channel", *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, 101-116, Oct. 2016.

- R.R. Thomas, M. Cardone, R. Knopp, D. Tuninetti and B. T. Maharaj, “Novel Half-Duplex Relay Strategy: An LTE implementation", *Eurecom*, Tech. Rep, RR-16-325, October 2016.

Finally Chapter 6 concludes the dissertation with a summary of the key findings for the proposed higher-order MIMO detection technique as well as the half-duplex relay strategy. Prospects for future work in this line of research are also proposed.

CHAPTER 2

BACKGROUND

Chapter 1 has introduced the objectives and contributions of this dissertation with a special focus on applied interference-aware higher-order MIMO detection and advanced HD relay strategies. A key part of this research is the implementation and evaluation of the developed algorithmic techniques and strategies to existing communication standards such as LTE (including forthcoming 4G+ and 5G networks). The preliminary concepts discussed within this chapter serve to provide the contextual background and cover the current general state-of-the-art techniques in relation to interference management. Section 2.1 provides a broad overview of the different interference management approaches, interference scenarios and standardized solutions in LTE. Section 2.2 provides a description on the classical MIMO model in addition to the state-of-the-art detection techniques for MIMO receivers. The underlying concept of physical layer relays and its standardization within LTE is described in Section 2.3. An overview of the different components within a physical layer LTE system is provided, since the developed techniques are practically evaluated using an LTE link layer simulation testbench.

2.1 INTERFERENCE MANAGEMENT

The performance limits of current mobile networks in relation to spectral efficiency and reliability will be constrained by interference-limited scenarios arising from widespread HetNet deployments [33]. The received signal at the mobile terminal is interference-limited when the total neighbouring interference power is higher than that of the noise. In this case, the effect of noise-power in system

performance is considered negligible [34]. A high demand for frequency resources corresponding to the number of connected devices and UEs will intensify as the network density increases. The prevalence of many devices and cells in close proximity will inevitably lead to many interference-limited scenarios. The consequence of interference would be the restriction on spectrum re-usability leading to tight frequency re-use. There are few categories in literature, in which to manage the surrounding interference from both the transmitter and receiver side.

2.1.1 Interference Avoidance

Interference avoidance involves transmitter-based techniques to resolve the interference. This can be done by transmitting two mutually interfering signals which are orthogonal to each other in the following domains: time, frequency, location (space) and antenna spatiality. In order to avoid interference in an Orthogonal Frequency Division Multiple Access (OFDMA) scheme, the allocation of different resources in frequency/time is a possible strategy. Power control algorithms can control the amount of the interference leaked based on dynamic algorithms [35–38], although marginal benefits of such power control algorithms in LTE femtocell have been reported [39]. Another technique to avoid interference, is by utilizing a MIMO system where transmission over uncorrelated spatial paths are performed [40]. In the context of small cells, interference can be avoided by making use of two different carrier frequencies (each different for the small cell and macrocell) although this is regarded as a highly inefficient procedure resulting in bandwidth segmentation. On the other hand, the macro and small cells could share the two carrier frequencies, enabling the terminals to aggregate multiple carriers to exploit the bandwidth of the available spectrum (carrier aggregation). This could be seen as a possible solution to mitigate the coverage hole effect caused by closed HeNBs within the macrocell [21, 41]. Interference avoidance may lead to an *a priori* loss in degrees of freedom, irrespective of the strength on the interferer [42]. Another approach involves partitioning the frequency band into multiple non-overlapping segments and allocating these partitioned bands to different regions of the macrocell and femtocell [43]. In the context of cognitive radio, there is a technique known as opportunistic interference avoidance, which involves the detection of primary user (macrocell user) activities through the femto-MS/BS. The femtocell-MS/BS, can then exploit the coverage gaps (spectrum holes) for transmission in a manner that that is non-interfering with macrocell users [44]. The secondary user usually operates at a level close to the noise floor in order to avoid detrimental inter-

ference with the primary user.

2.1.2 Interference Coordination and Cooperation

Coordination and cooperation among many transmitters is another way to handle interference. This technique relies on the transmitters performing coordinated resource allocation and sharing of radio link information to reduce the overall interference provided by each base station (BS). As a result parameters such as Channel State Information (CSI) and user-specific data would have to be shared among neighbouring BSs via significant backhaul resources, e.g. a Network MIMO scenario [45], which could increase the overall cost of the network. In LTE, inter-cell interference coordination (ICIC) techniques have been introduced, with the aim of managing interference of users located at the cell-edge [46].

2.1.3 Interference Cancellation

Interference cancellation (IC) offers one of the least intrusive network approaches to deal with interference, which has been the focus of recent interference management research, especially by the 3GPP [47]. Interference cancellation techniques are based at the receiver side and involves the decoding of the desired signal (information) through active cancellation of the interference contained in the received signal. This process is void of any signaling overheads from the transmitter-side as all the cancellation strategies occur at the receiver, with the optional need of CSI. Two distinct groups of IC techniques can be distinguished, viz. Parallel Interference Cancellation (PIC) and Successive Interference Cancellation (SIC). The PIC technique operates by stripping out interfering symbols simultaneously in a parallel fashion and has also been proposed to improve the performance for massive MIMO systems [48]. PIC can be used iteratively to improve the overall system reliability and robustness with respect to the bit-error-rate (BER). SIC methods on the other hand, adopt a systematic approach towards interference cancellation in the high interference regime, whereby the first stream (with the highest interference) is decoded based on a set of ordering criteria and then subtracted from the original received signal enabling subsequent stream to be decoded in a successive manner up until the desired signal is remaining, which is free of interference [49]. In addition, SIC can be performed on a codeword or symbol basis. As the name implies codeword-SIC involves the successive detec-

tion and decoding of each codeword stream while for the symbol-SIC the detection and interference cancellation is performed independently for each subcarrier making it vulnerable to error propagation [50].

2.1.4 Application in LTE

The notion of managing interference rather than simply avoiding it, has already been tackled in [51] as early as the 1980s, in which the coverage range of BSs are computed using statistical techniques in the absence and presence of co-channel interference. Earlier analytical studies such as [52] and [53] have tackled the issue in terms of power control for spread spectrum systems and frequency reuse in small-area-coverage (small cells) radio systems dominated by co-channel interference. In LTE systems, downlink (DL) coordinated multipoint transmission (CoMP) and ICIC techniques have been standardized to mitigate interference and enhance performance in heterogeneous networks [54, 55]. CoMP essentially exploits multiple transmit and receive antennas across multiple locations (including different cells) to enhance signal quality and reduce spatial interference using dynamic coordination among BSs. CoMP has been shown to provide benefits in HetNets, particularly with mobile terminal's affected by severe interference (cell-edge user's) but at the cost of large overheads in the backhaul [56, 57]. ICIC techniques defined in LTE Releases 8 and 9, on the other hand, involve message exchanges between BSs via X2 interfaces, to mitigate dominant interference scenarios in macrocells, relays and picocells. The performance of ICIC techniques is also dependent on the type of backhaul which can influence latency and delay and is limited to controlling co-channel interference [46, 57]. As a result, different enhanced-ICIC (eICIC) solutions were proposed for LTE Rel. 10 (Long Term Evolution-Advanced) including time- and frequency-domain and power control techniques for the management of macro-femtocell interference [1]. Fig. 2.1 represents an example of a two tier network scenario that comprises of a macrocell and numerous small cells. There are generally two types of interference categories, on a coverage level:

1. Cross-tier (Inter-cell) interference refers to the interference that occurs between the terminal of two different type of coverage cells, e.g. a macro-MS may be the victim of interference from a femto-BS in a neighbouring cell. One study investigates a power allocation algorithm which has been developed, to be exploited on the DL of an OFDMA system to reduce the effect of

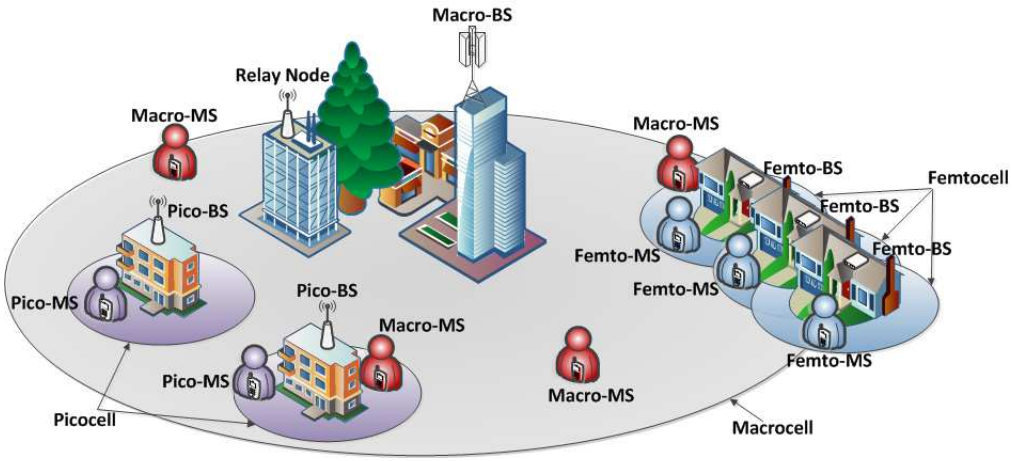


Figure 2.1: Typical Heterogeneous Network model

cross-tier interference using limited channel information of the mobile terminal in neighbouring macro cells [58, 59]. However, the system model of this particular technique only considers a sparsely populated cell where intra-tier interference is assumed to be negligible.

2. Intratier interference occurs among many co-located UEs of the same cell size (same tier), e.g. when the femto-MS of one femtocell interferes with a femto-MS of another femtocell (also depicted in Fig. 2.1) [54]. Game-theoretical approaches [60] have been also employed to achieve orthogonal resource block management among co-located femtocells. Exploiting cognitive information and spatial separation has been shown to improve the SINR performance of a typical indoor apartment femtocell mobile station (MS) scenario by 3 dB [54]. In addition, other cognitive approaches with respect to power control and spectrum access have comprehensively investigated [61].

2.2 AN OVERVIEW OF MIMO DETECTION

2.2.1 Classical Model

Point-to-point (P2P) MIMO systems have been well known to provide increased spectral efficiencies over SISO systems without corresponding increases in bandwidths at reasonable SNR values [62, 63]. Advances in MIMO wireless systems have reached a stage whereby antennas can be scaled up in

orders of magnitude resulting in what is referred in today's nomenclature as Massive MIMO systems. Such systems follow suite on the benefits of conventional MIMO systems but at a much greater scale [64]. Yet the study of practical MIMO receiver design with its already considerable state-of-the-art remains an ever present and important challenge. At the same time various open questions exist and hence there is a concerted effort to determine the most optimal tradeoff in terms of receiver algorithmic performance and complexity.

Let us consider a basic $l \times m$ P2P MIMO system model with $L = \{1,2,\ldots,l\}$ transmit antennas and $M = \{1,2,\ldots,m\}$ receive antennas. The total received signal at each M^{th} receive antenna is:

$$y_M = h_{M1}s_1 + h_{M2}s_2 + \ldots + h_{ML}s_L + n_M, \tag{2.1}$$

where h_{ML}, s_L and n_M are the complex channel coefficients and transmitted symbols from the L^{th} transmit antenna to the M^{th} receive antenna with n representing additive white Gaussian noise (AWGN). In vector form, the overall MIMO system model is:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{2.2}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1L} \\ h_{21} & h_{22} & \cdots & h_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1} & h_{M2} & \cdots & h_{ML} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_L \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_M \end{bmatrix}. \tag{2.3}$$

The system model shown in (2.2) is the basic system vector model from which all MIMO techniques both in theory and practice are developed and evaluated. The advent of increased network densification as well as higher demands for spectral resources will inevitably lead to a shift in scenarios which are *interference-limited*. Interference-limited scenarios can be further categorized according to the interference regimes [65]:

- High interference scenarios occur when the interfering signal is significantly high enough to be decoded when compared to the desired signal. In this case subtractive strategies such as successive interference cancellation (SIC) may be employed to remove the undesired effects of the interferer from the received signal.

- Weak interference scenarios result in the receiver treating the incoming interference as additive noise, although in practice this noise is considered to be non-Gaussian with a known a discrete constellation structure and thus can be exploited to enhance performance.

- The ultimate and undoubtedly most challenging interference scenario occurs in the moderate regime. The interfering user would be typically transmitting at the same rate and thus presents a challenge. This scenario would occur at the cell edge and therefore limit the overall spectral efficiency of the desired users.

As such, the interference-limited MIMO system model for decoding a packet a can be expressed as [66]:

$$\mathbf{y} = \mathbf{h}_a s_a + \overbrace{\sum_{b \neq a} \mathbf{h}_b s_b + \mathbf{n}}^{\text{interference and noise}} , \tag{2.4}$$

where $\sum_{b \neq a} \mathbf{h}_b s_b$ represents the co-layer interference (inter-symbol interference). The co-layer interference and noise component are coupled, e.g. in a moderate interference regime, modelling both the noise and interference component according to the Gaussian assumption can lead to significant performance losses at the receiver when decoding the desired signal, especially in the lower SNR regimes. In this case, various interference rejection and exploitation techniques can be considered to deal with the interference [42].

The performance and complexity of detecting the transmitted signal (**s**) based on the received signal (**y**) is an important challenge, depending on the accuracy of the available channel state information (CSI). Various estimation techniques can be used to determine the channel, **H**, in order to derive the CSI with varying degrees of accuracy and complexity. For example, the proliferation of Orthogonal Frequency Division Multiplexing (OFDM) over the past decade in various existing wireless technologies such as LTE, WiFi, DVB-T2, has led to innovative ways in which to perform channel estimation through the insertion of so-called pilot tones into the subcarriers of OFDM symbols using Least Squares (LS) or Minimum Mean Square Error (MMSE) techniques [67]. The key performance metric for detection in uncoded MIMO systems is the receive diversity gain which corresponds to the number of receive antennas, which can be asymptotically achieved through optimal detection schemes. The issue of joint detection of the transmitted symbols, in the absence or presence interfer-

ers, remains a current subject of open research. The performance benchmark of detection algorithms is Maximum-likelihood (ML) detection, which maximises the probability that the transmitted symbol has been correctly detected. For higher-order MIMO detection algorithms, the complexity of such schemes increases according to the number of transmit antennas and modulation order. Linear detection schemes such as the Zero Forcing (ZF) and Minimum Mean Square Error (MMSE) techniques, are practical for receiver implementation at the cost of lower performance especially in the higher SNR regimes [68]. A brief overview of these methods are given in the next section.

2.2.2 Non-linear Detection

2.2.2.1 Maximum-likelihood Detection (MLD)

Maximum-likelihood detection (MLD) is a optimal non-linear strategy, and aims to maximize the probability of correct detection. MLD is known to achieve optimal performance by traversing through all possible candidate set of transmitted vectors and selecting the vector, which is characterized by the smallest probability of error. The optimal detected symbol vector is given by:

$$\hat{\mathbf{s}}_{MLD} = \arg\max_{\mathbf{s}\in\chi^l} f(\mathbf{y}|\mathbf{s}), \tag{2.5}$$

$$= \arg\min_{\mathbf{s}\in\chi^l} \|\mathbf{y}-\mathbf{H}\mathbf{s}\|^2, \tag{2.6}$$

where the complete signal alphabet set is given by χ, while the candidate vector set is denoted by $|\chi|^l$. Accordingly, the complexity exponentially increases with the number of transmit antennas. The function $f(\mathbf{y}|\mathbf{s})$ to be maximized can be defined as the posterior probability and if knowledge of the transmitted signal is available, such as the equal likelihood of transmitted vectors, then the maximum a posterior (MAP) decision can be defined [68, 69]. Alternatively, in Eq. (2.6), the euclidean distance vector between the transmitted signal and the estimated transmitted signal is minimized.

2.2.2.2 Sphere Decoding

Significant efforts have been made to reduce the overall complexity of MLD techniques with well known techniques such as Sphere decoding, already reducing the overall complexity. Sphere decod-

ing was initially proposed in [70], which was inspired by the short length vector lattice computational study in [71] and computes the ML estimate by limiting the search space to lattice points that fall within a defined sphere of radius d. This inevitably leads to a reduction in the number of computations. One of the key factors that influence the performance and complexity is the initial choice of d given that a small radius results in insufficient lattice points within the sphere, while a large radius will widen the search space, and result in an increase in complexity. Sphere decoders are generally classified under a subcategory of decision feedback search tree algorithms. As such, the way in which the tree is traversed across the different dimensions (tree levels) or the number of visited nodes, directly impacts the complexity of the detector [72]. The expected complexity of sphere decoder was found to be polynomial but at the same time there exists many variations of the sphere decoder of which computing the expected complexity is a challenging endeavor [73]. In the case of block data transmission systems over dispersive channels, a reduced complexity lattice sphere decoding approach has also been proposed [74]. Few practical studies in relation to LTE [75–77] have been studied to investigate the feasibility of sphere detection in LTE receivers, although progress has been made with Fixed Complexity Sphere Decoders [77].

2.2.2.3 Lattice Reduction-Aided (LR) Detection

LR techniques [78, 79] transform a channel matrix as a basis into a near-orthogonal structure, in order to perform computationally efficient lattice decoding. LR-based MIMO detection schemes are developed with the aforementioned premise in mind. The idea involves the determination of an invertible $m \times m$ matrix $\mathbf{T}$, such that the lattice structure is preserved through integer entries as given by a new transformation matrix: $\mathbf{G} = \mathbf{HT}$. LR techniques are generally classified as non-deterministic polynomial-time NP-hard problems and heuristic methods such as the Lenstra, Lenstra and Lovász (LLL) algorithm have been employed to achieve near orthogonality of the columns of $\mathbf{H}$. In the same fashion as sphere decoding, various studies have investigated the use of LR-based detection for implementation within the 3GPP LTE framework such as the work in [80], which combines LR and list decoding to perform signal detection on the LTE downlink with near ML performance for 2×2 systems. The performance using higher-order MIMO configurations has been limited in this regard. Another study investigated the use of a hybrid LR-based detection with zero forcing (discussed in the next subsection) using SDR in the context of LTE with results provided for an uncoded BER

evaluation of a 4×4 MIMO scenario [81]. A hardware evaluation of a LR-based detection scheme was also studied to increase implementation efficiency for general MIMO-OFDM receivers [82]. LR hybrid approaches have also been considered in conjunction with SIC techniques to improve performance of MU-MIMO receivers [83].

2.2.2.4 Maximum a posterior (MAP) Detection

According to the ML criterion defined in (2.5), the vector symbol that maximizes the function $f(\mathbf{y}|\mathbf{s})$ (also referred to as the likelihood function) is to be determined. Assuming that the noise (**n**) is circularly symmetric complex Gaussian (CSCG), then the likelihood function can be represented as:

$$f(\mathbf{y}|\mathbf{s}) = \frac{\exp\left[-(\mathbf{y}-\mathbf{H}\mathbf{s})^{\dagger}\left(\sigma^2\mathbf{I}\right)^{-1}(\mathbf{y}-\mathbf{H}\mathbf{s})\right]}{\det\left[\pi\sigma^2\mathbf{I}\right]}, \tag{2.7}$$

where σ^2 is the spectral density, while **I** is the $l \times m$ identity matrix. Eq. (2.7) can be further approximated due to the independence of the noise and transmitted symbol vector:

$$f(\mathbf{y}|\mathbf{s}) \cong= \frac{1}{\pi\sigma^2\mathbf{I}} \exp\left[-\frac{1}{\pi\sigma^2\mathbf{I}}(\mathbf{y}-\mathbf{H}\mathbf{s})^{\dagger}(\mathbf{y}-\mathbf{H}\mathbf{s})\right]. \tag{2.8}$$

The MAP rule is defined as the maximum of all posteriori probabilities ($P(\mathbf{s}|\mathbf{y})$) [84]:

$$\hat{\mathbf{s}}_{MAP} = \arg\max_{\mathbf{s}\in\chi^l} P(\mathbf{s}|\mathbf{y}), \tag{2.9}$$

and after applying Bayes' theorem, it is shown that:

$$\hat{\mathbf{s}}_{MAP} = \arg\max_{\mathbf{s}\in\chi^l} \left(f(\mathbf{y}|\mathbf{s})P(\mathbf{s})\right), \tag{2.10}$$

provided that equally likely symbols are observed in the input stream, which is a realistic assumption for practical systems due to the effects of bit interleaving and scrambling, $P(s) = \frac{1}{L}\forall L, \mathbf{s} \in \chi^L$ reduces to a constant and thus the original ML criterion presented in (2.5). However, it should be noted that the MAP criterion is prohibitively complex to implement due its non-linear structure as well as large number of floating point operations, viz. additions and multiplications. The operation in the log domain significantly reduces the complexity of such MAP algorithms, but at the cost of

suboptimal performance in the low SNR regimes (due to the *max* function) [85]. This is referred to as the Max-log-MAP approximation. Applications of the Max-log-MAP approximation have included joint detection and classification of a co-scheduled user's modulation scheme for enhanced MU-MIMO detection [86] as well as joint phase noise estimation and data detection for MIMO receivers in quasi-static channels [87]. When considering practical receivers, the channel decoders require soft information measuring the likelihood that the received bit is either a 0 or 1, in order to enhance decoding performance (referred to as a bit level log likelihood ratio (LLR)) [68]. The LLR, $\Lambda(\mathbf{y}|\mathbf{s}(b_k))$, is computed by taking the logarithm of the conditional pdfs in indicating the likelihood of having transmitted a bit corresponding to a 0 or 1. The likelihood function is given as:

$$\Lambda(\mathbf{y}|\mathbf{s}(b_k)) = \ln\left(\frac{f(\mathbf{y}|b_k=0)}{f(\mathbf{y}|b_k=1)}\right), \tag{2.11}$$

$$= \max_{\mathbf{s}\in\chi^l(b_k=0)}\left\{-\frac{1}{\pi\sigma^2\mathbf{I}}(\mathbf{y}-\mathbf{Hs})\right\} - \max_{\mathbf{s}\in\chi^l(b_k=1)}\left\{-\frac{1}{\pi\sigma^2\mathbf{I}}(\mathbf{y}-\mathbf{Hs})\right\}. \tag{2.12}$$

Eq. (2.12) can be further expressed in terms of the suboptimal LLR that would minimize the euclidean distance:

$$\lambda_{k,L} = \min_{\mathbf{s}_L\in\chi_{k,L}^{(0)}}\|\mathbf{y}-\mathbf{Hs}\|^2 - \min_{\mathbf{s}_L\in\chi_{k,L}^{(1)}}\|\mathbf{y}-\mathbf{Hs}\|^2, \tag{2.13}$$

where the k^{th} bit represents the bit position of interest on the L^{th} stream, while $\chi_k^{(0)}$ and $\chi_k^{(1)}$ are sets of equal size representing all possible constellation points, where the k^{th} bit of interest is 0 and 1, respectively [88].

2.2.2.5 Low-complexity interference-aware demodulator

The detector design of receivers can be enhanced by exploiting the discrete structure of the surrounding interference sources from co-located cells or cross-tier cells, which are prevalent in HetNets or small cell networks. In practice, it becomes worthwhile to exploit the non-Gaussian structure of the interference. A low-complexity spatial interference cancellation algorithm has been earlier developed for soft decision bit metrics, which is based on the Max-log MAP detector. Traditionally, the metric consists of a summation of a large number of exponential terms and this strategy aims to simplify both the SISO and MIMO bit metrics using the Matched Filter (MF) output [89]. The basic premise entails the decoupling of the real and imaginary components of the expended MF output for all signal

constellations. Eq. (2.6) can be reformulated for a 2×2 system as:

$$\begin{aligned}\hat{\mathbf{s}}_{MLD} &= \arg\min_{\mathbf{s}\in\chi^l} \|\mathbf{y} - \mathbf{h}_0 s_0 - \mathbf{h}_1 s_1\|^2, \\ &= \arg\min_{\mathbf{s}\in\chi^l} \left[(\mathbf{y} - \mathbf{h}_0 s_0 - \mathbf{h}_1 s_1) \left(\mathbf{y}^\dagger - \mathbf{h}_0^\dagger s_0^\star - \mathbf{h}_1^\dagger s_1^\star \right) \right], \\ &= \arg\min_{\mathbf{s}\in\chi^l} \left[\|\mathbf{y}\|^2 + \left\|\mathbf{h}_0^2\right\|^2 \left\|s_0^2\right\|^2 + \left\|\mathbf{h}_1^2\right\|^2 \left\|s_1^2\right\|^2 - 2\Re\left\{\bar{y_0} s_0^\star\right\} - 2\Re\left\{\bar{y_1} s_1^\star\right\} + 2\Re\left\{\rho_{01} s_0^\star s_1\right\} \right],\end{aligned} \tag{2.14}$$

where the received matched filter vector is $\bar{\mathbf{y}} = \mathbf{H}^\dagger \mathbf{y} = [\bar{y_0}\ \bar{y_1}]^T$, $\{\ldots\}^\star$ is the conjugate, the receive antenna correlation coefficient is given as $\rho_{01} = \mathbf{h}_0^\dagger \mathbf{h}_1$. Grouping the real and imaginary components, the Max-log-MAP approximation for the LLR is given as:

$$\lambda = \max_{\substack{s_0\in\chi_{k,1} \\ s_1\in\chi_{k,2}}} \left[-\|\mathbf{h}_0\|^2 |s_0|^2 - \|\mathbf{h}_1\|^2 |s_1|^2 + 2\left[\Re\left\{\bar{y_0}\right\}\Re\left\{s_0\right\} + \Im\left\{\bar{y_0}\right\}\Im\left\{s_0\right\}\right] - 2\eta_0 \Re\left\{s_1\right\} - 2\eta_1 \Im\left\{s_1\right\} \right], \tag{2.15}$$

where the grouped terms η_0 and η_1 are given as:

$$\eta_0 = \Re(\rho_{01})\Re(s_0) + \Im(\rho_{01})\Im(s_0) - \Re(\bar{y_1}), \tag{2.16}$$

$$\eta_1 = \Re(\rho_{01})\Im(s_0) - \Im(\rho_{01})\Re(s_0) - \Im(\bar{y_1}). \tag{2.17}$$

The grouped terms η_0 and η_1 are independent of s_1 and therefore the search space over s_1 is removed. It can be observed that the if the signal alphabets belonging to s_1 (the interferer) are of equal energy such as in the case of QPSK, then the values of $\Re(s_1)$ and $\Im(s_1)$ required to maximize Eq. (2.15), imply that η_0 and η_1 are opposite in sign to that of Eq. (2.15). Hence for QPSK (equal energy alphabets) and 16-,64-QAM (unequal energy alphabets):

$$\Re(s_1) \leftarrow -\eta_0 \text{ (QPSK)}, \tag{2.18}$$

$$\Re(s_1) \leftarrow \frac{\eta_0}{\|\mathbf{h}_1\|^2} \text{ (16-,64-QAM)}, \tag{2.19}$$

$$\Im(s_1) \leftarrow -\eta_1 \text{ (QPSK)}, \tag{2.20}$$

$$\Im(s_1) \leftarrow \frac{\eta_1}{\|\mathbf{h}_1\|^2} \text{ (16-,64-QAM)}, \tag{2.21}$$

The complexity of the reduced dimension Max-log-MAP metric has shown be less than than the standard MMSE technique with respect to the number of complex multiplications and addition operations as shown in Table 2.2 [90], where ϑ is the number of comparisons within the search space of the interferer which is equivalent to a complex addition and $|\chi|$ is the discrete constellation size. Two strategies have been proposed to tackle the dominant interference scenario for cell-edge users, which correspond to the rate and magnitude of the interfering signals. The first strategy termed partial interference cancellation (PIC) is applied when the interfering signal displays a higher rate or is weaker than the desired signal hence making it near impossible to decode. On the other hand, absolute interference cancellation (AIC) is utilized when the rate of the interferer is low or the strength of the interfering stream is higher when compared to the desired stream [42, 91]. The developed bit metrics reduce the MIMO system dimensionality to that of a $2 \times m$ configuration, with hybrid extensions for higher-order MIMO systems also proposed [42]. A novel technique has been developed based on the aforementioned interference-aware detection for higher-order MIMO systems that can be applicable to MIMO HetNet and P2P link scenarios, and makes use of a Block QR decomposition to pre-process the channel for low-complexity receiver processing in an LTE setting. Block QR decomposition has already been used to reduce complexity and improve the error rate performance under certain MIMO settings [92,93], but has not yet been studied in the context of interference exploitation in higher-order MIMO detection. The theoretical and practical investigations of which are examined in Chapters 3 and 4.

2.2.3 Linear Detection

Two well known linear detection algorithms are used for practical implementation due its low-complexity of implementation: Zero-forcing and Minimum Mean Square Error (MMSE) detection (although MMSE detection offers the better performance-complexity tradeoff). Linear detectors aim to null out or minimize the effect of the interference, to facilitate the detection of the desired signal. Another technique is spatial matched filtering, which regarded as the optimal receiver for SISO and SIMO systems. In order to employ MIMO matched filtering, the channels of the receiver would ideally have to be orthogonal otherwise the receiver would be susceptible to the effects of inter-stream interference [66], which is resolved using the ZF technique.

2.2.3.1 Zero Forcing (ZF)

At the name suggests, ZF nulls out the interference through the inversion of the channel matrix:

$$\mathbf{W}_{ZF} = \left(\mathbf{H}^{\dagger}\mathbf{H}\right)^{-1}\mathbf{H}^{\dagger}, \tag{2.22}$$

and applying it to the received signal in Eq. (2.2):

$$\hat{\mathbf{s}}_{ZF} = \mathbf{W}_{ZF}\mathbf{y} = \mathbf{s} + \mathbf{W}_{ZF}\mathbf{n}, \tag{2.23}$$

It can be noted that the ZF technique is susceptible to noise enhancements. If the channel of the desired signal is colinear with respect to the interference subspace (rank deficient) [66, 69].

2.2.3.2 Minimum Mean Square Error (MMSE)

MMSE detection maximizes the post-detection SINR and thus does not suffer from the drawbacks of noise enhancement in ZF detection, although prior statistical information about the noise is required. The applied MMSE weight matrix is given as:

$$\mathbf{W}_{MMSE} = \left(\mathbf{H}^{\dagger}\mathbf{H} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{H}^{\dagger}. \tag{2.24}$$

The MMSE detection does not require the channel matrix to have full column rank but the performance degrades nonetheless [66, 69]. MMSE is advantageous in interference-limited scenarios but cannot exploit the information contained in interference to enhance performance and results in performance degradation in low SNR environments. Since the interference is assumed to be Gaussian, the spatial streams are decoupled individually. Therefore, the bit metric for the k^{th} bit of value $b \in \{0,1\}$ of symbol x_L on the L^{th} stream is given by:

$$\lambda_{k,L}(\mathbf{y}, b) \approx \max_{x_m \in \chi_{k,L}^{(b)}} \left[-\frac{\lambda_L^2}{\sigma^2} \left| \hat{x}_k - x_k \right|^2 \right], \tag{2.25}$$

where λ_L^2 is the k^{th} diagonal element of the $\Gamma = \mathbf{W}_{MMSE}\mathbf{H}$ matrix, $L = \{1, 2, ..., l\}$, $\chi_{k,L}^{(b)}$ is part of the set $x_L \in \chi_l$ in the k^{th} position with the value b. 3GPP has introduced into LTE Rel. 11, a variant of

MMSE known as the MMSE-Interference Rejection Combining (IRC) receiver in order to mitigate the effect of interference due to large scale HetNet deployments, which lead to increased levels of interference [94]. MMSE-IRC receivers exploit multi-antenna configurations to create “nulls" in the direction of interfering signals, i.e. bringing about a drop in the antenna gain and as a result improving the overall SINR. Studies [95, 96] have exploited MMSE-IRC to improve throughput and reliability in LTE systems.

2.2.4 Successive Interference Cancellation (SIC) Schemes

The main advantage of linear detectors is the low complexity of implementation at the sacrifice of performance. However, combining either of these linear receivers to form a bank of detectors can increase performance by employing SIC in conjunction with ZF or MMSE [97–99]. Each of the bank of linear of receivers detects one of the data streams, which is then successively cancelled from the remaining signal in the subsequent stages. The resulting outcome is that the final signal is interference-free. However, error propagation of the detected signal in each of the stages can negatively affect the overall performance of SIC. SIC may also be integrated with ML approaches to improve detection performance [100]. A key aspect of this technique is the selection of an optimal cancellation signal sequence based either on the SNR, SINR, column-norm or received signal ordering.

Table 2.1: MIMO Detector Complexity for WLAN 801.11n [2]

	No. of Real Additions	No. of Real Multiplications
MMSE	$\frac{10NM^2+2NM-5M^2+(4NM-2M)N_{OS}}{\sum_{i=1}^{N_s} K_i \cdot R \cdot N_{OS}}$	$\frac{11NM^2+7NM-4M^2-2M+(4NM+3M)N_{OS}}{\sum_{i=1}^{N_s} K_i \cdot R \cdot N_{OS}}$
List Sphere Decoding	$\frac{\sum_{i=1}^{M-1}[4(M-i)+5]\cdot 2^{K_i}\cdot\prod_{j=i+1}^{M} N_{ave}(j)+4\cdot 2^{K_M}}{\sum_{i=1}^{N_s} K_i \cdot R}$	$\frac{\sum_{i=1}^{M-1}[4(M-i)+4]\cdot 2^{K_i}\cdot\prod_{j=i+1}^{M} N_{ave}(j)+4\cdot 2^{K_M}}{\sum_{i=1}^{N_s} K_i \cdot R}$
List Sphere Decoding Max-log-MAP	$\frac{\prod_{i=1}^{M} N_{ave}(i)\left[\mathbf{1}_{full}+\mathbf{1}_{it}\cdot\left(\sum_{i=1}^{M} K_i-1\right)+2\sum_{i=1}^{M} K_i\right]-2\sum_{i=1}^{M} K_i}{\sum_{i=1}^{N_s} K_i \cdot R}$	$\frac{\prod_{i=1}^{M} N_{ave}(i)\left[\mathbf{1}_{full}+\mathbf{1}_{it}\cdot\left(\sum_{i=1}^{M} K_i-1\right)\right]}{\sum_{i=1}^{N_s} K_i \cdot R}$
MMSE Max-log-MAP	$\frac{4\sum_{i=1}^{M} 2^{K_i}+2\sum_{i=1}^{M} K_i\cdot 2^{K_i}-2\sum_{i=1}^{M} K_i}{\sum_{i=1}^{N_s} K_i \cdot R}$	$\frac{3\sum_{i=1}^{M} 2^{K_i}}{\sum_{i=1}^{N_s} K_i \cdot R}$

Table 2.2: MIMO Detector with complex metrics

	No. of Complex Additions	No. of Complex Multiplications
MMSE-SIC	$\mathbf{1}\left(4N^3-2N^2-N\right)+\lvert\chi\rvert$	$\mathbf{1}\left(4N^3-2N^2-N\right)+2\lvert\chi\rvert$
Reduced Dimension Max-log-MAP	$\mathbf{1}\left(5N-5\right)+\lvert\chi\rvert\left(\vartheta+4\right)$	$\mathbf{1}\left(5N\right)+4\lvert\chi\rvert$

2.2.5 Complexity Comparison

Table 2.1 is an indicative comparison of the complexities of the different MIMO detectors evaluated for a 801.11n WLAN scenario [2]. The number of receive antennas are given by N, M indicates the number of spatial transmitted streams, while for the MMSE scheme, the number of OFDM symbols is given by $N_{OS}=\left\lceil\frac{N_b}{\sum_{i=1}^{M}K_i\cdot R\cdot N_{SD}}\right\rceil$ where N_{SD} is the number of subcarriers and N_b is the total number of transmitted bits. In relation to the List Sphere Detector (LSD), the variables $N_{ave}(i)$ indicate the average number of symbol candidates that satisfy the radius criterion at level i of the tree search, K_i represents the number of bits/symbol while R is the code rate. In the case of the LSD Max-log-MAP the indicator function $\mathbf{1}_{full}$ represents the operations in the first iteration and $\mathbf{1}_{it}=1$ if priori LLRs are available from the second iteration onwards. Although the tabulated complexities are valid for the 801.11n WLAN study, a general idea can be garnered regarding the detection complexity of such schemes with the MMSE Max-log-MAP algorithm showing the least complexity in terms of number of real multiplications.

The authors in [77] investigate the use of a Fixed Complexity sphere decoder (FSD) in conjunction with the different QR decomposition Gram-Schmidt and Householder methods in an LTE setting. The performance approaches that of the MLD detector with improvements in the amount of paths explored in the FSD, while the complexity is proportional to the constellation size. Table 2.2 is an overview of the required of the complex multiplications and additions for the MMSE-SIC detector and reduced dimension Max-log-MAP discussed in Section 2.2.2.5 [90].

2.3 RELAY NETWORKS

The link budget between the base station (eNB) and the user terminal (UE) includes physical parameters such as the eNB-UE distance and affects the uplink or downlink throughput performance in a wireless network. The achievable rates exhibited in LTE systems show that the Adaptive Modulation

and Coding (AMC) schemes already achieve 75% of the throughput of the Shannon bound over the finite SNR regime with higher modulation schemes requiring a high operating SNR [101]. Further improvements in throughput can be made through network densification by reducing the aforementioned eNB-UE distance and thus enhancing the link budget to achieve the desired performance exploited using higher MCSs. The 3GPP has standardized various aspects related to HetNet deployments, one of them being the introduction of relays as deployable low powered nodes, which enables coverage extension in NLOS areas as well as increased spectral efficiency through physical layer cooperation as demonstrated in Chapter 5. Relays can also enable rapid deployment of base stations as they require no specific wired backhaul when compared to traditional eNBs [102]. The relay architecture consists of a backhaul link (source-relay link) and an access link (relay-destination link) as denoted in Fig. 2.2. There are typically two classes of relays based on the spectrum access of both the backhaul and access links. The first class referred to as outband relaying, involves the backhaul link operating in a different spectrum to that of the access link, provided adequate frequency separation can be guaranteed in order to avoid interference between these links. Inband relays on the other hand, enable operation of the backhaul and access links on the same spectrum and require additional techniques to mitigate interference between the two links, which are outlined in Section 2.3.2. There exists various types of relays investigated in literature some of which have been proven to theoretically achieve the largest known rate in a relay channel [103], while there are certain types of relays, which are deemed implementation-friendly. The different relay types are listed here below:

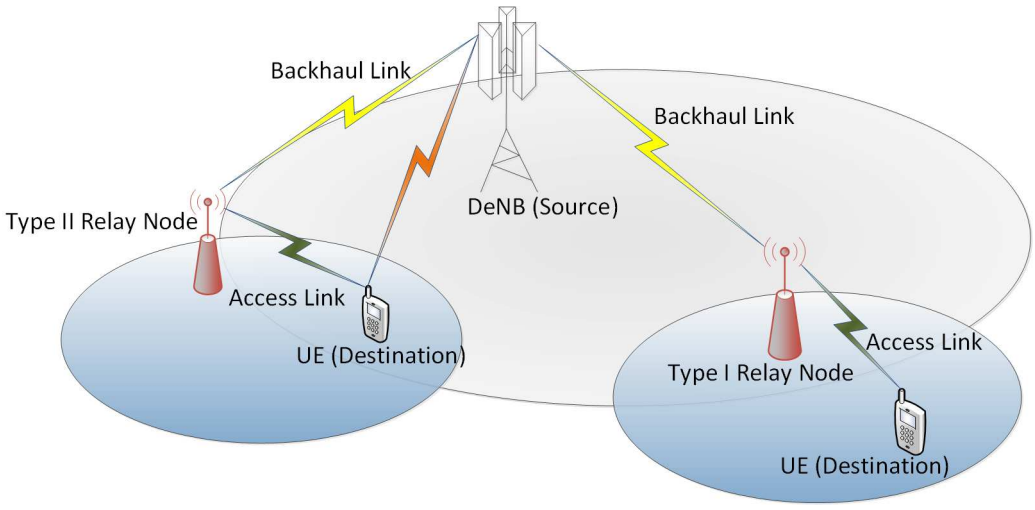


Figure 2.2: Basic relay architecture

2.3.1 Relay Types

2.3.1.1 Amplify-and-Forward (AF)

These relays operate as basic repeaters, amplifying the received signal from the base station (source) and thereafter forwarding the corresponding signal to the UE (destination). The noise and surrounding interference is also amplified in addition to the data signal, which can be detrimental in low-SNR environments. The key benefit for implementation in existing networks, is the transparency of such a relay to the UE. This ensures that the UE cannot distinguish between a connection to the eNB or relay [101].

2.3.1.2 Decoded-and-Forward (DF)

Decode-and-forward relays decode the received signal from the eNB and re-encode the signal for transmission to the UE, thus removing any effects of noise and interference, increasing performance in low-SNR environments. This enables independent rate adaptation and scheduling at the eNB and relay, which is an aspect that is actively addressed in Part II of this dissertation. There exists a processing delay due to the decoding operations, which is higher than that of the amplify-and-forward relay. This delay has to be accounted for in the lower layer LTE signalling structures. Two categories of DF relays were identified as part of an initial 3GPP study (see Fig. 2.2): Non-transparent Type 1 relays which have its own physical identity and thus transmit all the conventional physical channels and behave as a typical eNB; Transparent Type 2 relays perform initial decoding on the backhaul link of donor eNB (DeNB) transmissions to enhance overall performance [104]. However, the partial DF (PDF) strategy proposed in [105] relies on a transmitter-side, two layered superposition coding strategy (initially proposed in [106]) capable of partially decoding a transmitter's message. Due to its gains in diversity as well as robustness in low-SNR environments, the PDF has its advantages when compared to the original DF relay strategy proposed in [103]. A dynamic DF (DDF) protocol was also proposed in [107], where it was shown to achieve the optimal tradeoff for multiplexing gains. However, in practice the DDF strategy presents an interesting implementation challenge due to the non-packet based nature of the protocol [108].

2.3.1.3 Compress-and-Forward (CF)

The CF relaying strategy was also introduced in [103], where the relay compresses (and quantizes) the received message from the eNB and transmits it to the UE, without actually decoding the message. The relays make use of Wyner-Ziv source coding [109] to exploit this receiver side information at the destination. Furthermore, the work investigated in [110], quantizes the received signal using super-position coding and a corresponding practical coding strategy is developed. Another study investigated a quantize-map-and-forward (QMF) relaying technique for a diamond network scenario within a LTE framework [111].

2.3.2 Relaying in LTE Networks

This section provides an overview of the different aspects of relaying, which are considered in a practical LTE downlink scenario. As alluded to earlier on, outband relay nodes (RNs) operate using the spectrum for both the backhaul and access links and as a result no additional enhancements are required on top of the LTE Rel. 8 radio interface (physical layer) to deploy an outband RN. However inband RNs require the following considerations when planning for interference-free deployments [112]:

- Possible spatial separation between antennas on the RN, e.g. full-duplex inband RNs.
- The access and backhaul links can be time multiplexed, e.g. half-duplex inband RNs.

The key challenge is addressing the issues raised by the deployment of RNs into the existing LTE architecture in order to reap the benefits of coverage extension and improved capacity. The complexity of the algorithms required to deal with the self-interference effect of full-duplex relays, still make half-duplex relays more appealing for practical implementation in existing networks. In addition to the technical advantages, HD relay deployments are low in cost and provide an excellent business case for telecommunication operators [113]. The design of relay transmission and reception schemes should also aim for backward compatibility with earlier releases of LTE.
The following hurdles need to be overcome when considering the use of HD inband DF RNs [112]:

- Minimize the additional delay due to the decoding operation at the RN when forwarding the data.

- The RN can behave as an uplink multiplexer and downlink demultiplexer and as a result the backhaul link (a bottleneck) should operate at a high spectral efficiency.

- The radio frequency (RF) switching circuitry from reception (relay listens) to transmission phases of the RN should be as rapid as possible in order to avoid loss in physical resources, e.g. in time.

An LTE RN has functionality essentially derived from an eNB. Therefore interfaces such as S1 (base station to core network) and X2 (inter-base station) are prevalent. Attachment procedures (prach, etc.) to the network are also supported by the RN. In order to avoid any significant disruptions to the overall network structure, a proxy model was developed [112], such that the core network views the RN as a normal eNB, while the DeNB behaves like a core network from the RN's point of view.

2.3.2.1 Signalling Procedures

The procedure for a relay to startup and connect to a DeNB can be summarised as follows [112]:

1. Relay Startup: A preconfigured list of the available DeNBs to connect to is known to the RN.

2. In a similar fashion to the UE, the relay initiates an attach procedure through the management system in order to obtain the initial configuration information.

3. The RN detaches and connects to a suitable DeNB from the preconfigured list.

4. The RN transmits its unique identity (indicator) to the selected DeNB during the radio connection startup.

5. The DeNB then acts upon this information and selects a Mobility Management Entity (MME) that supports relay functionality.

6. The DeNB sets up the bearers for the S1 and X2 interfaces while the relay connects to the DeNB via these interfaces.

2.3.2.2 Radio Subframe Configuration

In Rel. 8 and 9 of the LTE standard, the proposed method for creating a transmission gap between the reception and transmission of data (half-duplex communication) at the RN involves the declaration of a multi-broadcast single frequency network (MBSFN) subframe that could be supported by legacy Rel. 8 UEs on the downlink. This ensures that the relay can stop transmissions to the UE and receive subframes from the DeNB by performing a timing advance. The UE receives these MBSFN subframes and apart from reading the first few control channel information symbols, the remaining subframe can be disregarded, implying that the RN is receiving data from the DeNB. An interesting limitation arose when the Physical Downlink Control Channel (PDCCH) carrying signaling information from the DeNB could not be processed by the relay (transmissions by DeNB) due to a timing advance of 1-2 OFDM symbols and thus a special relaying control channel in the PDSCH region was introduced in Rel. 10, known as the R-PDCCH, to account for DeNB-RN synchonization and CQI reporting [114, 115]. Fig. 2.3 represents the subframe configuration of the relaying process.

2.4 LONG TERM EVOLUTION (LTE)

2.4.1 Specification Highlights

The overall progression and successful advancement of the fourth generation of mobile communication systems has come a long way since the introduction of the Release 8 standard in 2009. This follows from the extension of earlier 3GPP specifications based on the Universal Mobile Telecommunication Systems (UMTS) standard, viz. Releases 5 and 6 termed High-Speed Packet Access (HSPA) as well as Release 7 also known as HSPA+. The UMTS group of specifications were characterized by the Code Division Multiple Access (CDMA) technology, which later evolved into Wideband CDMA (WCDMA) and then finally with the adoption of Orthogonal Frequency Division Multiplexing (OFDM) as the most suitable multiple access technology, which is adopted by most currently evolving radio standards (especially on the downlink), although it remains to be seen whether the

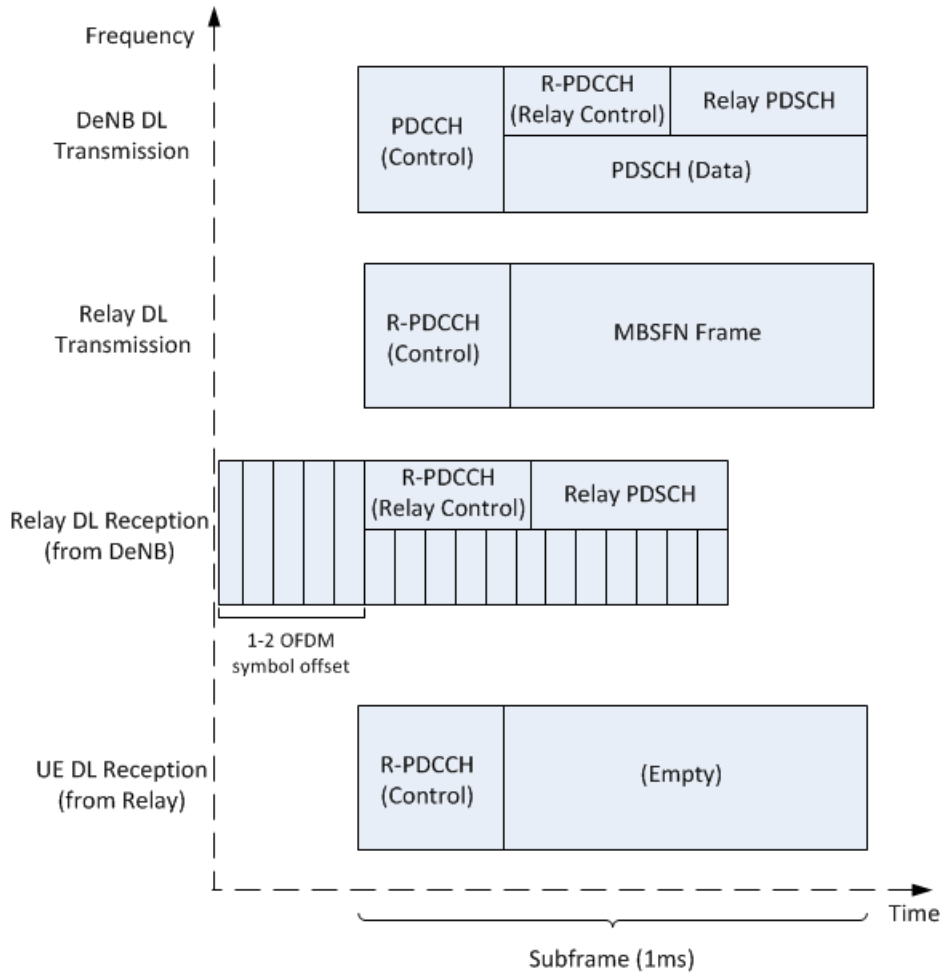


Figure 2.3: Subframe configuration for LTE relaying

forthcoming generation, i.e. 5G, will continue to adopt OFDM [31]. LTE Release 8 fulfilled the 3GPP requirements as a next generation technology due to advances in radio technology, such as multicarrier OFDM, multiple antenna systems (MIMO) and the packet-switched radio interface (a shift from the circuit-switched connection-based protocols). The next specification, i.e. Release 9, built upon the initial release and introduced enhanced several features such as femtocells (or Home eNodeBs) and LTE mobile positioning protocols in order to target a mobile device's location. The concept of self-organizing networks (SON) was also developed further from self-configuration of an eNB to self-optimization [116]. An expansion on the Broadcast Service functionality for operators was also introduced through the enhanced Multimedia Broadcast Multicast Service (eMBMS). LTE Release 10, also known as LTE-Advanced (LTE-A) provided a significant update to the standard with

advanced MIMO configurations of 8×8 on the downlink and 4×4 on the uplink. Additional improvements of Release 10 include the support for Heterogeneous Networks (HetNets), Relay Nodes, enhanced intracell interference coordination (eICIC), and carrier aggregation to exploit the fragmented spectrum to provide bandwidths of up to 100 MHz. Release 11 includes a notable cooperative strategy for interference management from the eNB side, and is referred to as co-ordinated multi point (COMP) transmission and serves to improve cell-edge network performance. The Release 12 specification was frozen in 2014 and featured improved carrier aggregation (CA) techniques between co-located TDD and FDD carriers, small cell enhancements such as dense deployments and dual connectivity, and network-aided assisted interference cancellation (NAIC) approaches for improved interference management. Currently, Release 13 has been the latest specification to be frozen with extended enhancements for Machine Type Communications (MTC)with low latency and minimal bandwidth requirements (1.4 MHz). Fig. 2.4 shows the evolutionary roadmap that LTE has undergone since Release 12.

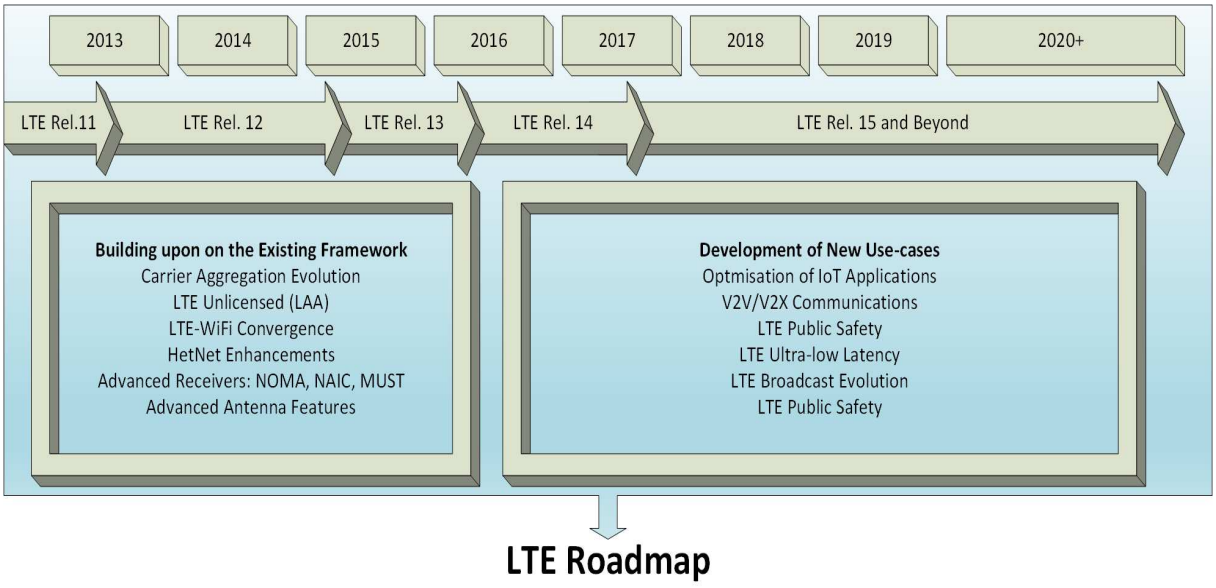


Figure 2.4: The 3GPP LTE evolutionary road map

2.4.2 State of LTE Adoption

A recent study by OpenSignal [117], which uses anonymous crowd-sourced data to build coverage maps and derive corresponding analytics has shown that countries such as South Korea are at the forefront of providing near ubiquitous connections to the LTE network, whereby customers are reported to be connected to an LTE network 97% of the time rivaling the ubiquity of 2G and 3G networks. The study considers the proportion of the time that users are connected to their respective LTE networks, i.e. network availability as opposed to signal coverage based on particular geographic areas. The crowd-sourced data also reveals that early adopters of LTE, such as the United States who have introduced LTE since its inception (2010) are falling behind in terms of average LTE throughputs, partly due to the large subscriber base heavily straining existing network resources as opposed to certain countries in South America and Europe, where LTE has recently been introduced resulting in lightly loaded networks. In addition, several countries have adopted LTE-Advanced (Release 10), in order to boost speeds through the implementation of carrier aggregation techniques [117]. Developing regions of the world, particularly the Middle East and Africa lag behind in terms of average speed but nonetheless strong growth in these markets are to be expected due to a growing subscriber base and as well as handset affordability [118]. The latest study according to [117], has shown that there exists 146 countries reported to have at least one LTE network available. The next section introduces the physical layer architecture that has made LTE, a successful fourth generation technology.

2.5 LTE PHYSICAL LAYER

The LTE physical layer (PHY) defines the underlying baseband signal processing of the protocol stack, in which the reception and transmission of the data signals are processed in order to ultimately process protocol data units (PDUs), which are forwarded to (or received from) the higher layers including the MAC and RLC layer. The uplink and downlink PHY structures are largely similar, however, both links differ in certain features based on the requirements of each link and amount of data transmitted on either links. An overview of the LTE PHY is provided in Fig. 2.5.

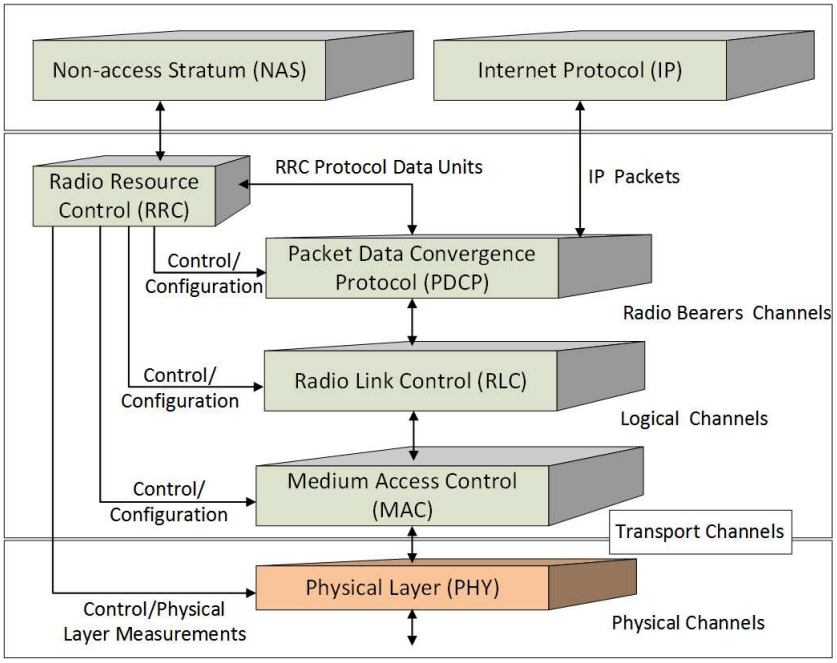


Figure 2.5: Overview of the LTE protocol stack

2.5.1 Downlink-Shared Channel (DL-SCH)

The functional processing of the downlink (from the base station to the user) will be detailed. Key design parameters such as the modulation and coding scheme and antenna transmission mode are crucial in attaining the desired reliability, quality of service and throughput for the end user (UE). The degrading nature of the mobile radio channels are characterized by the effects of multipath delays, large-scale and small-scale fading, inter-cell and intra-cell interference, Doppler spreads, signal dispersiveness and time-variability. In order to combat and mitigate the aforementioned effects, a robust design of the PHY is critical for regular operation under various scenarios, while at the same time satisfying the high throughputs as well as low latency requirements [4].

According to Fig. 2.5, it can be noted that the Transport channel provides an interface between the PHY and MAC layer in LTE. In this context a channel refers to the information flow between the different protocols in LTE and are characterized by the type of information and procedures for processing this information. There are four main types of channels in LTE, i.e. the Downlink Shared Channel (DL-SCH), Broadcast Channel (BCH), Multicast Channel (MCH) and Paging Channel (PCH). The focus of this dissertation remains on the DL-SCH transport channel where user data and dedicated

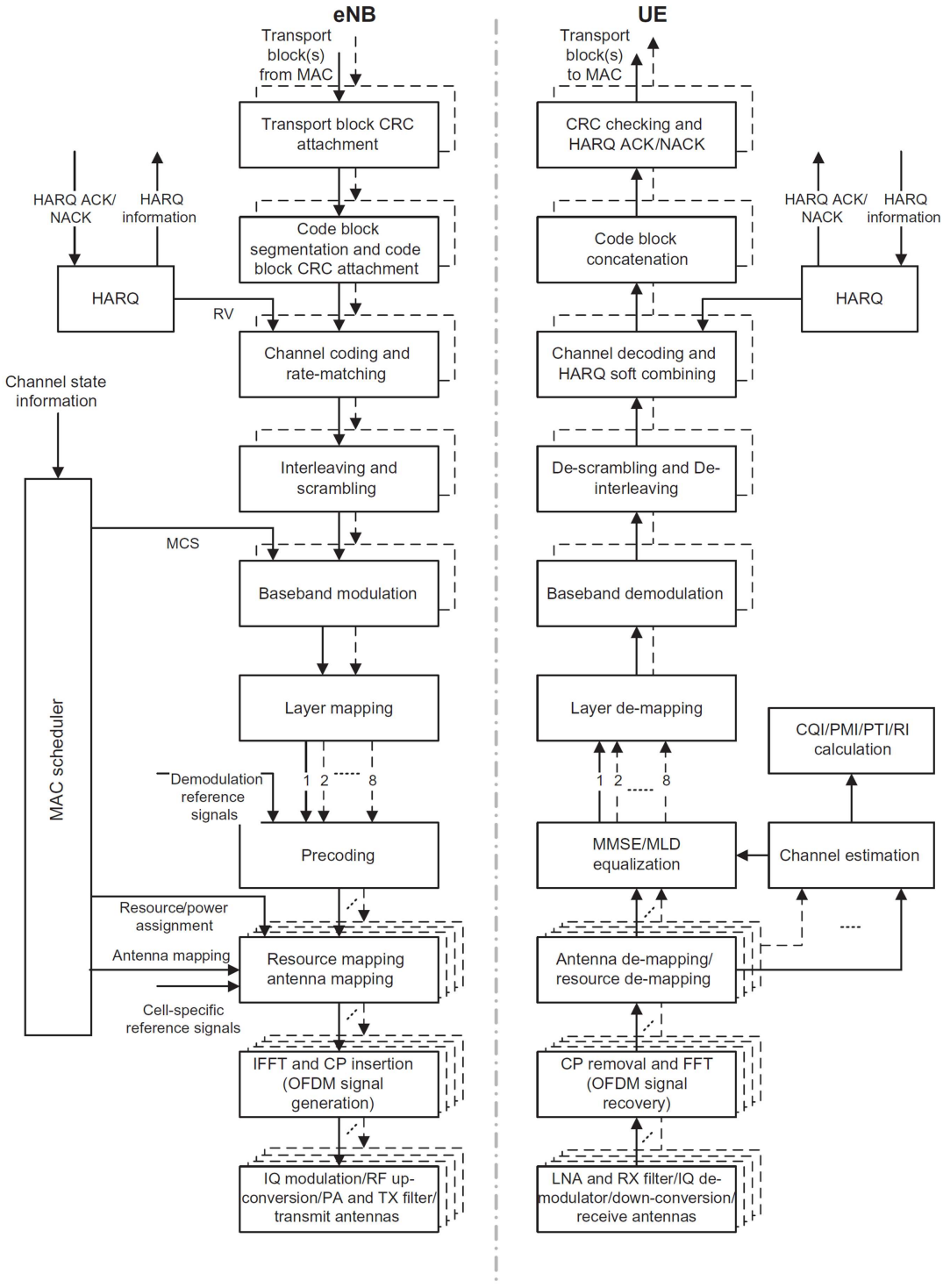


Figure 2.6: Physical layer processing component overview and flow [4]

control information as well as user-specific higher-layer information is conveyed from the eNB to UE [102].

The different PHY processing procedures of the DL-SCH are illustrated in Fig. 2.6 from the perspective of the base station (eNB) and user terminal (UE). A Transmission Time Interval (TTI) encompasses a subframe of a duration of 1 ms, with dynamically sized transport blocks which are delivered to the PHY to be transmitted over each component carrier. The multi-antenna scheme also determines the amount of transport block used within one TTI. In the case where a transmission mode enables spatial multiplexing, e.g. Transmission Mode 4 closed-loop spatial multiplexing for a single-user, the TTI consists of two transport blocks.

2.5.1.1 Transport Blocks and CRC

The transmitted data is organized into transport blocks (TBs) with an associated transport format corresponding to the transmission bandwidth and MCS. This enables the MAC layer to recognize the data rates and reliability of the data received from the PHY. The Transport Block Size (TBS) defines the message data length (in bits) and is mapped according the to the MCS and physical resource blocks (RBs). Twenty-four (24) Cyclic Redundancy Check (CRC) bits are then appended to the transport block in order to enable error detection. This mechanism is used in conjunction with the Hybrid-Automatic Repeat Request(H-ARQ) retransmission protocol.

2.5.1.2 TB Processing: Code Block Segmentation

The maximum supported code block size for the internal interleaver of the LTE turbo code is 6144 bits. If the total size of the transport blocks including CRC bits, are greater than the maximum allowable size, the code block undergoes segmentation into smaller, manageable pieces which are then ready to be fed into the Turbo encoder. The segmentation component is designed in such a manner that filler bits are added to varying TBs of different sizes, in order to match the required code block sizes of the turbo encoder. The CRC bits are appended to each code block in order to enhance advanced detection of correctly decoded code blocks, which essentially implies that the iterative turbo decoder can terminate early. As a result, there is a reduction in the terminal computational complexity as well as battery utilization. The Physical Downlink Control Channel (PDCCH) contains information

regarding the TBS and can therefore infer the code-block size and number of code-blocks. This information aides in the decoding the TBs at the receiver side.

2.5.1.3 Channel Coding

The channel coding component introduces redundancy to enable the receiver to recover from all errors resulting from the wireless channel. The parallel concatenated convolutional code (PCCC) Turbo coder [119] is employed on the key transport channels including the DL-SCH, MCH and the PCH. The channel coding block consists of a Quadrature Permutation Polynomial (QPP) interleaver and two 1/2 rate, eight state constituent convolutional encoders. The QPP interleaver distributes the data bits in manner that results in each of the two codes being affected independently in the event of an erroneous transmission. This enables easier recovery of the data bits. If the input into the interleaver is given by b_k then the output of the QPP interleaver represents a permuted version of the input denoted by $\tilde{b}_k$ and determined using:

$$\Phi(i) = \left(f_1 i + f_2 i^2\right) \bmod K, \tag{2.26}$$

where $i = [0,1,...,K-1]$ represents the bit index, f_1 and f_2 are coefficients selected based on the length of the code block b_{QPP}. The range of supported code block sizes in LTE range from 40 to 6144 bits. The QPP interleaver is referred to as maximum contention-free, implying that the decoding process can be done in parallel without the uncertainty of the different parallel process running into contention when accessing the interleaver memory. The corresponding transfer function for each constituent convolutional encoder is $[1,\ g_1(D)/g_0(D)]$, where two generating polynomials are given by:

$$g_0(D) = 1 + D^2 + D^3, \tag{2.27}$$

$$g_1(D) = 1 + D + D^3, \tag{2.28}$$

resulting in the two output sequences of parity bits [4, 102].

2.5.1.4 Rate Matching and PHY H-ARQ

The rate matching block processes the systematic, and two parity bit streams from the channel encoder by selecting the pattern of bits through puncturing and repetition, which can be transmitted in a single subframe. The number of transmitted bits are limited by the number of available physical resources and therefore support the generation of arbitrary rates using different puncturing patterns. In addition, the rate matching algorithm should also output as many unique bits as possible during each retransmission, in order to exploit the gains of the efficient Incremental Redundancy (IR) H-ARQ protocol [31]. The three bit streams consisting of the systematic, first and second parity bits are each passed through a sub-block interleaver and thereafter fed into a Circular Buffer (CB), starting with the systematic bits followed by the alternative insertion of the first and second parity bits. The CB is a key component of the rate matching algorithm responsible for matching the concatenated information bit stream output of the CB with the physical resource blocks for transmission. For a particular code rate, the coded bits for transmission are selected in a serial fashion from a specific starting point within the CB determined by the Redundancy Version (RV). As a result, different RVs correspond to different starting points, which enable different H-ARQ retransmissions of the selected data [4, 102].

2.5.1.5 Scrambling

A logical exclusive-OR (XOR) operation is performed on the output of the rate matching algorithm in conjunction with a bit-level scrambling sequence. This operation prevents the desired signal decoded at the receiver from being equally matched to the interfering signals from neighbouring cells. Characterising the neighbouring cells by a different scrambling sequence ensures that the interfering data streams are randomised after descrambling, which maximises the processing gain of the channel code. The downlink scrambling is applied to across all transport channels [102].

2.5.1.6 Modulation

The bits are mapped onto complex-valued symbols depending on the selected modulation scheme. LTE currently supports QPSK, 16-QAM and 64-QAM up till Rel. 11, with Rel. 12 onwards having

the added capability of supporting 256-QAM transmissions for spatial multiplexing (for multi-user MIMO) on the PDSCH (downlink).

2.5.1.7 Layer Mapping and Precoding

In the case of multi-antenna transmission schemes, the layer mapping component enables a few MIMO operational modes, viz. transmit diversity and spatial multiplexing for both single-user (SU) and multi-user (MU) MIMO scenarios. In order to switch between the various MIMO schemes, depending on the type of radio environment, a number of Transmission Modes (TMs) have been defined in LTE. As of Rel. 13, ten transmission modes have been identified, which are shown in Table 2.3, encompassing both single-antenna and multi-antenna transmission schemes. The PDCCH carries a Downlink Control Information (DCI) message specified by a certain format and typically includes resource block allocation, HARQ and MCS information. One codeword is required when implementing transmit diversity at the eNB and the number of layers correspond to the number of antenna ports. An antenna port can be defined as a logical entity which is distinct from a physical antenna and can be affiliated with a set of reference signals. In principle, two received signals are assumed to have experienced the same overall channel variation (inclusive of any transmit side joint processing), provided that they have been transmitted from the same antenna port. A group of supported antenna ports depends on the reference signal configuration, e.g. Antenna ports 0-3 make use of cell-specific reference signals, Antenna Port 6 is related to the positioning reference signal, Antenna ports 9-14 utilize UE-specific reference or demodulation reference signalling, etc.

In the case of single-layer transmit diversity, only one codeword is used to perform layer mapping where the number of layers (N_l) correspond to the number of antenna ports, i.e. one layer based on a single antenna port. Dual-layer transmit diversity requires that the codeword be split in an even-odd fashion. In other words, even-indexed modulation symbols are mapped to layer 0 while odd-indexed modulation symbols are mapped to layer 1. This represents a 1:2 (codeword:layers) mapping and can even be extended to a 1:4 mapping, where each of the layers are sequentially mapped from a single codeword. Spatial multiplexing offers the opportunity to perform layer mapping with each of the two codewords. Fig. 2.7 displays the mapping permutations with even-odd index grouping. The number of layers should be less than or equal to the number of antenna ports. The precoding component is essentially a pre-processing matrix (precoding matrix indicator), which aims to match the radio chan-

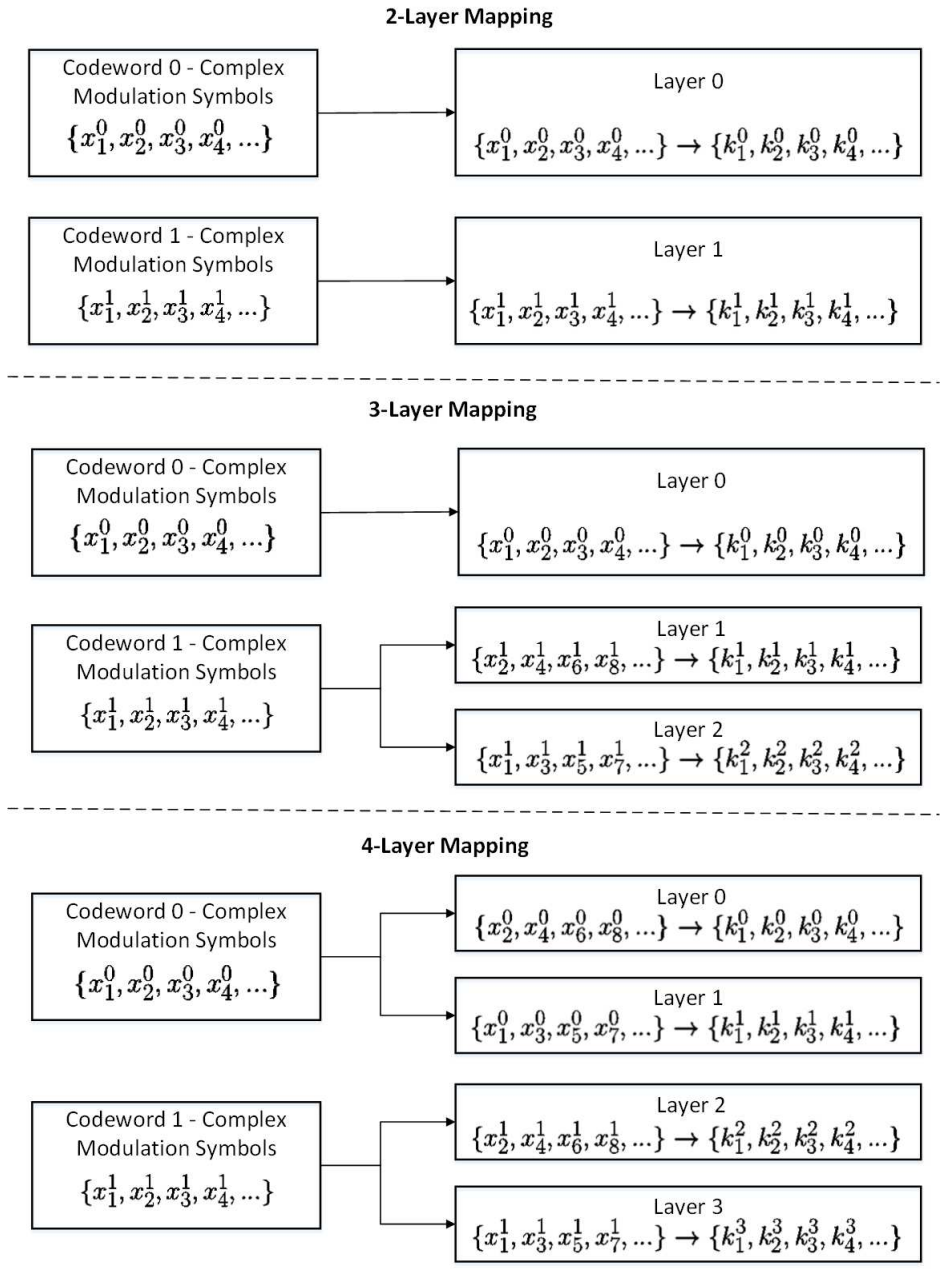


Figure 2.7: LTE codeword to layer mapping

nel's traits by maximizing the received signal power while simultaneously reducing the interference power at each layer (enhancing the receive SINR). Precoding ensures phase compensation by rotating the transmitted signal in order to account for any instantaneous channel phase shifts. In addition, various powers are allocated among the different antennas (akin to waterfilling) depending on the instantaneous channel gains. The complex modulated symbols are precoded to form a complex vector

based on whether the precoding function is associated with transmit diversity or spatial multiplexing. The eNB selects a single precoding matrix from a pre-defined codebook assuming transmission on a set of subbands. The UE can also select a limited set of precoding vectors to report back to the eNB via the uplink based on the downlink channel state information [3, 4, 102].

2.5.1.8 Resource Element and Antenna Mapping

The precoded vectors are mapped onto resource elements (REs) of a particular resource block (RB) assigned for transmission by a MAC scheduler. A resource block contains 84 resource elements, which include 12 subcarriers and 7 OFDM symbols for a normal cyclic prefix configuration (74 elements in the case of the extended cyclic prefix configuration). The resource elements comprise of transport-channel data, reference signals (e.g. CRS/UE-RS) and control signalling (e.g PDCCH). Thereafter complex-valued time domain OFDM signals are generated for each antenna port within the radio frequency (RF) component and transmitted via the physical antennas. It should be noted that if a signal (e.g. reference signal) is transmitted from multiple physical antennas, this signal originates from a single antenna port and similarly, if multiple signals are transmitted from the same set of antennas, this corresponds to multiple antenna ports.

2.5.2 OpenAirInterface Simulator

Testbeds or experimentation platforms are critical to the successful implementation and design of future generation wireless communication networks as it enables the application of novel ideas and identification of any potential problems during the deployment and testing phase on any wireless network. The various components of an experimentation platform may vary from the capability to carry out simulations, emulations or real-time hardware implementations. Ideally, in order to holistically approach the development of a novel theoretical technique/algorithm, it has to be verified and evaluated using a combination of components in order understand the performance in a practical environment under realistic conditions.

In order to obtain some insights into the practicality of the developed contributions in this dissertation, the software libraries of the OpenAirInterface simulator were exploited to carry out physical layer modem simulations for both the higher-order MIMO detection and the advanced half-duplex relay

strategies. OpenAirInterface is an open-source hardware and software experimentation platform for the development of next generation wireless networks. One of the distinguishing features of the platform is its close alignment with the 3GPP LTE specifications across three layers of the protocol stack while maintaining the dominant characteristics of a readily available radio system used in the deployment of today's commercial networks [28, 120]. Fig. 2.8 illustrates the various functionality of the OpenAirInterface, with the highlighted Baseband/PHY processing software libraries used to perform the evaluations in the simulation test bench.

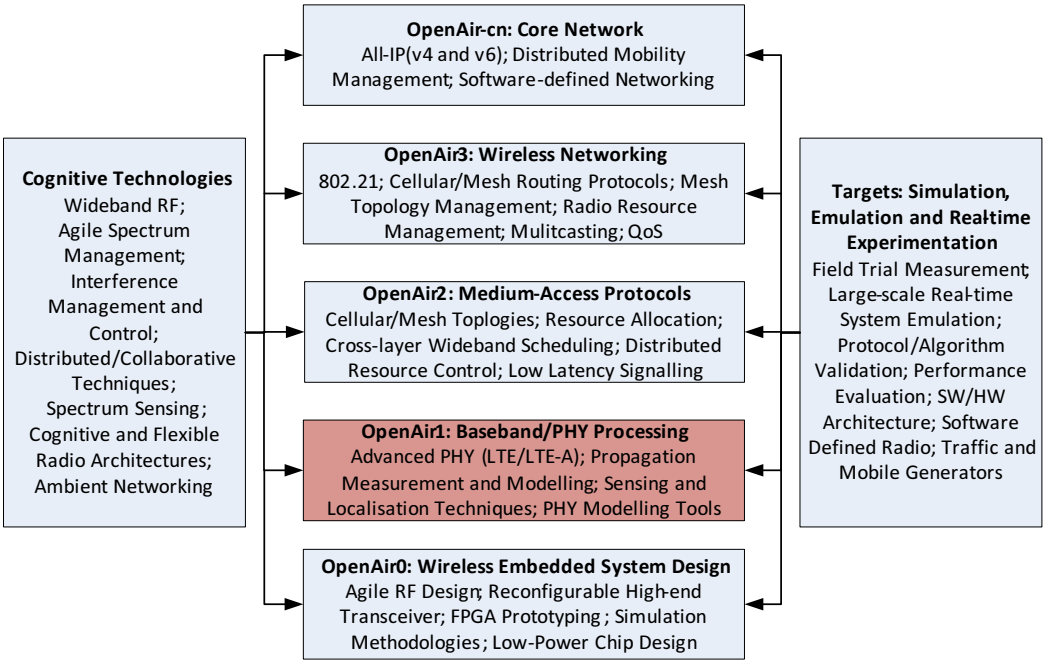


Figure 2.8: Overall OpenAirInterface framework

2.6 SUMMARY

This Chapter served to provide the broad background required to develop and understand the contributions of the dissertation, management of various interference scenarios, especially in relation to HetNets. The different detectors in literature are described in addition to the various gaps of higher-order MIMO interference-aware detection, which are also highlighted. The different relay architectures are mentioned in conjunction with the current frame structures for relays in today's LTE networks, together with a overall downlink component overview of the PHY layer in LTE.

Table 2.3: LTE Transmission Modes [3]

Transmission Mode	Description	DCI Format
1	Single-antenna, Antenna Port 0	1A
	Single-antenna, Antenna Port 0	1
2	Transmit diversity	1A
	Transmit diversity	1
3	Transmit diversity	1A
	Open-loop spatial multiplexing (large delay CDD (cyclic delay diversity))	2A
4	Transmit diversity	1A
	Closed-loop spatial multiplexing	2
5	Transmit diversity	1A
	Multi-user MIMO	1D
6	Transmit diversity	1A
	Closed-loop spatial multiplexing using a single transmission layer	1B
7	Single-antenna, Antenna Port 0 (PBCH antenna ports =1); Transmit diversity	1A
	Single-antenna, Antenna Port 5	1
8	Single-antenna, Antenna Port 0 (PBCH antenna ports =1); Transmit diversity	1A
	Dual-layer transmission, Antenna Ports 7 or 8; Single-antenna, Antenna port 7 or 8	2B

9	Single-antenna, Antenna Port 0 (non-MBSFN subframe); Single-antenna, Antenna Port 7 (MBSFN subframe)	1A
	Eight-layer transmissions, Antenna Ports 7-14; Single-antenna, Antenna Ports 7, 8, 11, 13; Single-antenna, Antenna Ports 7 or 8 (UE high-layer parameter)	2C
10	Single-antenna, Antenna Port 0 (non-MBSFN subframe); Single-antenna, Antenna Port 7 (MBSFN subframe)	1A
	Eight-layer transmissions, Antenna Ports 7-14; Single-antenna, Antenna Ports 7, 8, 11, 13; Single-antenna, Antenna Ports 7 or 8 (UE high-layer parameter)	2D

Part I

Higher-order MIMO Detection Strategies

CHAPTER 3

A THEORETIC APPROACH TO HIGHER-ORDER MIMO BLOCK DETECTION

3.1 INTRODUCTION

As described in Chapter 1, one of the goals of this research study is the development of a low-complexity detection scheme for higher-order MIMO receivers with a special focus on interference-limited scenarios. This Chapter presents a pre-processing Block QR channel decomposition and Block MMSE algorithm, which is explored analytically using information theoretic tools, in order to determine the feasibility of the proposed detection techniques in single-user interference-limited HetNet and point-to-point MIMO scenarios.

As discussed in Chapter 2, wireless networks are on the verge of shifting towards heterogeneity by leveraging various radio access technologies and multiple access node types. A basic HetNet scenario may consist of multiple uncoordinated low powered nodes (eNBs) amidst existing coordinated infrastructure. As a consequence, receivers are susceptible to increased levels of inter-cell and intra-cell interference, due to the numerous macro- or pico/femto- eNB deployments (increased densification). Advanced interference aware management techniques should therefore be implemented at the receiver side in order to manage such uncoordinated interference scenarios in an ad hoc fashion. Section 3.2 provides an overview of the state-of-the-art of Block QR and Block MMSE techniques, especially in relation to receiver design. Section 3.3 describes the investigated single-user MIMO HetNet

interference-limited scenario, while Sections 3.4 and 3.5 present the mutual information analysis and corresponding numerical results for the proposed higher-order MIMO detection techniques. Section 3.6 examines the theoretical performance of the total achievable rates for an LTE system.

3.2 RELATED WORK

Inter-cell interference coordination techniques (ICIC) for LTE Rel-8/9 and enhanced-ICIC methods for LTE Rel-10 (LTE-A) have been included in the LTE standard to resolve the interference issues experienced by HetNets using time-domain, frequency-domain and power control techniques [1]. Downlink Coordinated Multipoint (CoMP) strategies have also been defined for LTE Rel-11 and beyond, using coordinated user scheduling or beamforming techniques to improve the data rates and spectral efficiency of cell-edge users [121]. Recently, Network-Assisted Interference Cancellation (NAIC) techniques have also been examined for LTE Rel-12, with the focus on Interference Suppression (IS) and Cancellation (IC) at the receiver [47]. As the name implies, IS techniques aim to suppress the effects of interference experienced at the receiver, e.g. through the use of Linear MMSE-Interference Rejection Combining (IRC) or MMSE-Maximal Ratio Combining (MRC) techniques [122]. IC approaches correspond to Successive Interference Cancellation (SIC) detection schemes which entail the successive cancellation of the interfering streams and can be performed on a codeword or symbol level. This is performed by consecutively stripping out the interference at each stage of the decoding process, until the desired signal is obtained. Both IS and IC techniques require some degree of *a priori* knowledge about the interference channel.

3.2.1 Block QR

This chapter introduces a novel pre-processing technique using a Block QR decomposition approach for performing higher-order MIMO detection in an interference-limited and single-user (SU)-MIMO scenario. The investigated 4×4 and 8×8 scenarios (supported configurations based on the current LTE standards [11]) are analytically analysed in terms of the loss in mutual information. This receiver detection is based on pre-processing the complex channel matrix into blocks for performing an overall Block QR decomposition, which results in the isolation of the desired or interfering signal (depending on the ordering) for decoding using the interference-aware reduced dimension

demodulator elaborated upon in [89]. A key aspect of this techniqe is the exploitation of the dominant non-Gaussian interferer affecting cell-edge users in the high and moderate interference regimes. The first part of the study, considers a SU-MIMO interference-limited HetNet with two eNBs. Thereafter, a novel rate optimization scheme designed for a LTE point-to-point (P2P) SU-MIMO scenario is introduced. In [92], a blocking procedure has been developed to partition the channel matrix into 2×2 blocks to improve the packet error rate of a reduced-maximum likelihood detection (R-MLD) scheme, although in this study, the investigated scenario is not interference-limited and lacks an end-to-end standardised LTE performance evaluation. Block QR decomposition methods have also been studied in a MU-MIMO context [93], where the authors adapt the Dirty Paper Coding (DPC) scheme using Block QR decomposition to improve the performance of multiple users in the presence of inter-cell interference. The aforementioned technique is an enabler for transmitter-side interference management. The study of interference-aware detection techniques for LTE receivers in the context of higher-order MIMO systems has been limited, to the best of the authors knowledge. Initially, an information theoretic approach is undertaken to quantify the mutual information loss of the proposed Block QR technique with a 4×4 and 8×8 MIMO configuration in a Rayleigh block-fading channel. These analytical results will be used as a reference to evaluate the practical performance of the proposed Block QR scheme with the suboptimal, yet more complex (brute-force search) Max-log-MAP demodulator algorithm.

3.2.2 Block MMSE

The concept of Block Minimum Mean Square Error (MMSE) as a way to pre-process the noise and interference has been applied to equalization [123] in time varying multipath environments for OFDM as well as detection of MIMO space time coded systems [124]. A Block MMSE technique has been implemented as a computationally cost effective solution with Dirty Paper Coding (DPC) to achieve the sum-rate of the wireless MIMO Broadcast Channel in the form of a precoding strategy on the transmitter side [125]. This model adapts transmit resources to achieve capacity gains in a generalised multi-user MIMO scenario. A Block diagonalization approach has also been proposed with the aim of suppressing multi-user intertier interference (arising from users in adjacent cells), based on the MMSE criterion. Similarly, the authors distinguish an orthonormal set of vectors of each user's

precoding matrix at the transmitter and ensure that signal-to-interference-plus-noise ratio is enhanced at the receiver end of each user [126]. Ordered Block MMSE has also been shown to achieve near ML performance with low complexity for Generalised Spatial Modulation MIMO systems where selected antennas are used for transmission in each time slot [127]. It is important to highlight that the aforementioned studies do not consider the Block MMSE performance on SU-MIMO receivers with LTE practical codes that closely align with the 3GPPs specifications. The underlying premise of the Block MMSE technique assumes that the interference is Gaussian and therefore can be neglected. Furthermore, the Block MMSE is shown to be optimal in terms of the theoretical maximal achievable rate in the presence of a dual-stream interferer, under the Gaussian signal assumption. This chapter is structured as follows: Section 3.3 provides an overview of the interference-limited HetNet system model with an overview of the Block QR and Block MMSE techniques. Section 3.4 consists of an information theoretic analysis of the mutual information loss of the Block QR and Block MMSE techniques under the Gaussian signal assumption of the dominant interferer. Section 3.5 provides the numerical results of the two techniques for both 4×4 and 8×8 scenarios. Section 3.6 investigates a rate optmisation scheme designed for a LTE SU-MIMO scenario capable of transmitting a maximum of two codewords. Finally, Section 3.7 summarises the key points of the chapter.

3.3 INTERFERENCE-LIMITED HETNET SYSTEM MODEL

This section presents the system model for the Block QR scheme in a MIMO HetNet scenario. An inter-cell interference setting for a femto mobile terminal (UE) is considered. This particular model (as noted in Fig. 3.1) consists of a downlink scenario where the UE is equipped with four receive antennas and is located on the cell-edge making it highly vulnerable to interference from the neighbouring macrocell. The UE receives simultaneous transmissions from two base stations (eNBs) each equipped with two antennas, one of which is transmitting the desired signal (femto-eNB), while the neighbouring macro-cell eNB is considered as the dominant co-channel interfering source. Conversely, a proximal femto-eNB could also be an interfering source to a macro UE. Moreover, this MIMO scenario can also be extended to an 8×8 system (as supported by the LTE-A (Rel.10) standard) involving the deployment of a relay node equipped with eight antennas as opposed to a single UE (as shown later in the numerical results of Section 3.5). The received signal ($\mathbf{y}_F$) in Fig. 3.1 can be represented in vector form by grouping the desired stream $\mathbf{s}_1$ and the interfering stream $\mathbf{s}_2$ (similar

to Eq. (2.2)):

$$\mathbf{y}_F = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{3.1}$$

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix}, \tag{3.2}$$

where the received signal is given by $\mathbf{y}_F = (\mathbf{y}_1\,\mathbf{y}_2)^T$ and each corresponding received signal is represented as $\mathbf{y}_1 = (r_1\,r_2)^T$ and $\mathbf{y}_2 = (r_3\,r_4)^T$. The complex channel coefficients are given as $(\mathbf{H}_{11}\,\mathbf{H}_{12}) \in \mathbb{C}^{N_{T_1} \times N_R}$ and $(\mathbf{H}_{21}\,\mathbf{H}_{22}) \in \mathbb{C}^{N_{T_2} \times N_R}$, where N_R represents the number of receive antennas, while N_{T_1} and N_{T_2} refers to the number of transmit antennas at each eNB. The spatially multiplexed transmit data from the femto and macro eNBs are given as $\mathbf{s}_1 = (x_1\,x_2)^T$ and $\mathbf{s}_2 = (x_3\,x_4)^T$, respectively (where the matrix transpose is given by $(.)^T$). The noise vectors, $\mathbf{n}_1$ and $\mathbf{n}_2$, are zero mean circularly symmetric complex Gaussian (CSCG) variables. The aim is to diagonalise the channel using a Block QR decomposition approach, such that the interfering signal is successfully decoded and thereafter stripped to obtain the desired signal using SIC. This scheme enables the reduction of a square 4×4 channel into 2×2 blocks, thus reducing the overall complexity. Similarly, in an 8×8 system, the channel would be divided into 4×4 blocks. The interference decoding is then performed using a low-complexity Max-log-MAP demodulator [27,91], which is also applicable to a 2×2 MIMO scenario. This model assumes that perfect channel state information (CSI) is available at the receiver. Eq. (3.2) is further expanded as follows:

$$\mathbf{y}_F = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix} = \begin{bmatrix} \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} & \begin{pmatrix} h_{13} & h_{14} \\ h_{23} & h_{24} \end{pmatrix} \\ \begin{pmatrix} h_{31} & h_{32} \\ h_{41} & h_{42} \end{pmatrix} & \begin{pmatrix} h_{33} & h_{34} \\ h_{43} & h_{44} \end{pmatrix} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix} \tag{3.3}$$

3.3.1 Block QR Decomposition

Blocking algorithms generally tend to work on multiple columns and rows simultaneously and has served as an efficient way to perform matrix computations on existing hierarchical memory structures [128]. A recursive Block QR factorization has been developed, which partitions the target matrix

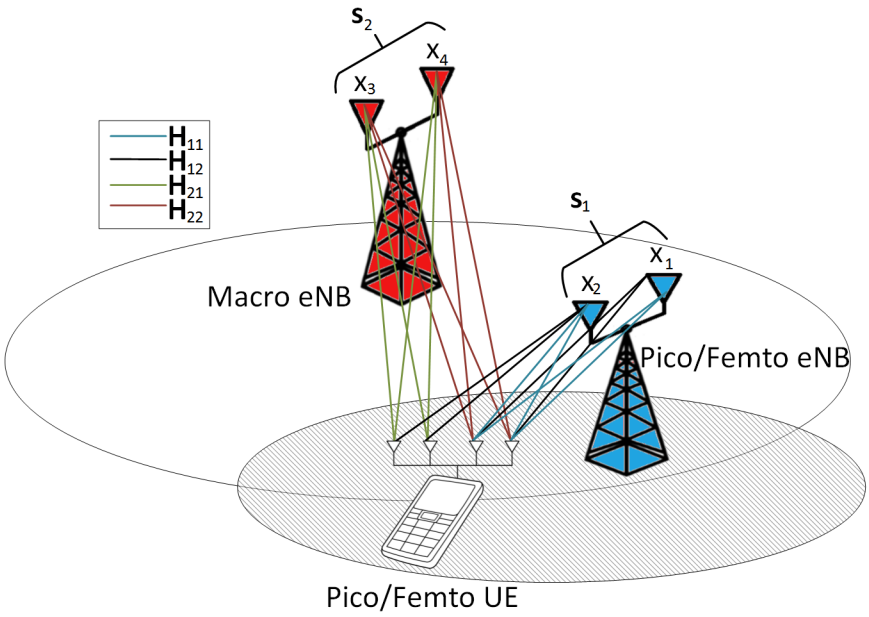


Figure 3.1: Basic HetNet model.

into a series of blocks and has been validated through mathematical induction [129, 130]. According to [129], elementary Householder transformations are used to factorize the last remaining column and in the process, inherently stops the recursion. The transformed received signal based on Eq. (3.2) is expressed as:

$$\mathbf{y}_F = \mathbf{Q}_1\mathbf{Q}_2\mathbf{R}\mathbf{s} + \mathbf{n}, \tag{3.4}$$

where $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are orthogonal matrices and are partitioned into blocks as follows:

$$\mathbf{Q}_1 = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix}; \mathbf{Q}_2 = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{Q}}_2 \end{pmatrix}, \tag{3.5}$$

where the $N_T \times N_R$ channel matrix (N_T and N_R denotes the total number of transmit and receive antennas, respectively) has been partitioned into $k \times k$ blocks, where $k = 2$ and $k = 4$ for the 4×4 and 8×8 cases, respectively. $\mathbf{I}$ represents the $(N_T - k) \times (N_T - k)$ block identity matrix and $\tilde{\mathbf{Q}}_2$ is the orthogonal transformation $(N_T - k) \times (N_T - k)$ matrix. It should be also noted that $N_T = N_R = 2^p$ for $p > 1$, where p is the blocking order for the square channel matrix. The upper triangular matrix $\mathbf{R}$ is

also block partitioned as follows:

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \mathbf{R}_{22} \end{pmatrix}. \tag{3.6}$$

The transformed received signal is:

$$\tilde{\mathbf{y}}_F = \tilde{\mathbf{H}}\mathbf{s} + \tilde{\mathbf{n}}, \tag{3.7}$$

where $\tilde{\mathbf{y}}_F = \mathbf{Q}_1^T \mathbf{y}_F$, $\tilde{\mathbf{H}} = \mathbf{Q}_2\mathbf{R}$, $\tilde{\mathbf{n}} = \mathbf{Q}_1^T\mathbf{n}$ and also $\tilde{\mathbf{Q}}_2 = \mathbf{R}_{22}\tilde{\mathbf{H}}_{22}$, given that:

$$\mathbf{Q}_1^T\mathbf{H} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \tilde{\mathbf{H}}_{22} \end{pmatrix}. \tag{3.8}$$

The Modified-Gram Schmidt (MGS), Givens rotations and Householder transformations have been investigated to determine the orthogonal $\mathbf{Q}$ and upper-triangular $\mathbf{R}$ matrix and is chosen according to the QR_{Method} as shown in Algorithm 1. The implemented Block QR algorithm is outlined in Algorithm 1, where N_{Block} represents the block size. For the analytical study, $N_{Block} = \{2,4\}$ is assumed, in line with capabilities of the LTE standard for 4×4 and 8×8 MIMO scenarios.

Algorithm 1 Block QR Detection

function $[\mathbf{Q}, \mathbf{R}, \tilde{\mathbf{H}}_{22}] = \text{RecBlockQR}(\mathbf{H}, N_R, N_{Block}, QR_{Method})$
  if $N_R \leq N_{Block}$ **then**
    Select QR_{Method}
  else
    $n_{temp} = floor(N_{Block}/2)$
    Partition: $\mathbf{H} = \left[\begin{array}{c|c} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{array}\right]$
    $[\mathbf{Q}_1,\ \mathbf{R}_{11}] = \text{RecBlockQR}(\mathbf{H}(:, 1:n_{temp}),\ n_{temp},\ N_{Block},\ QR_{Method})$
    $\begin{bmatrix} \mathbf{R}_{12} \\ \tilde{\mathbf{H}}_{22} \end{bmatrix} \leftarrow \mathbf{Q}_1^{\dagger} \begin{bmatrix} \mathbf{H}_{12} \\ \mathbf{H}_{22} \end{bmatrix}$
    $[\mathbf{Q}_2, \mathbf{R}_{22}] = \text{RecBlockQR}(\mathbf{H}(n_{temp}+1:N_R,\ :),\ N_R - n_{temp},\ N_{Block},\ QR_{Method})$
    $\mathbf{Q} = [\mathbf{Q}_1 \mid \mathbf{Q}_2],\ \mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \mathbf{R}_{22} \end{bmatrix}$

3.3.2 QR Decomposition Techniques

There are three conventional transformations in linear algebraic literature to perform a QR decomposition of a real or complex $L \times M$ matrix, depending on the characteristics of the matrix. A brief overview of these transformations is provided.

3.3.2.1 Gram-Schmidt Orthogonalisation

The Gram-Schmidt orthogonalisation process states that there exists an orthonormal basis for every finite dimensional space. This particular decomposition technique has been widely applied to pre-process the channel matrix **H**, e.g. a linear block QR decomposition method has been proposed that transforms a multi-user channel matrix into a unitary and upper triangular block matrix [131], in order to mitigate co-channel interference. This particular technique applies a generalisation of the Gram-Schmidt principle to perform Block QR decomposition. However, the classical Gram-Schmidt procedure is numerically unstable and prone to round-off errors as opposed to the modified Gram-Schmidt (MGS) procedure.

3.3.2.2 Givens QR decomposition

A matrix can also be decomposed into QR form by using plane rotations. A Givens rotation involves a 2×2 orthogonal transformation of either a two dimensional row or column vector. The plane rotations are given as:

$$G = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}, \tag{3.9}$$

where $c^2+s^2=1$, indicates that this plane rotation is orthogonal. There is a specific angle $\theta \in [0, 2\pi)$ such that $c = \cos\theta$ and $s = \sin\theta$. Consider the givens rotation of the vector $(a,b)^T$ (where $(\ldots)^T$ refers to a transpose of a matrix):

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ca+sb \\ cb-sa \end{pmatrix}, \tag{3.10}$$

Eq. (3.10) is determined by rotating the vector through the angle θ. The plane rotations can be used to introduce zeros by letting [132]:

$$c = \frac{a}{\sqrt{a^2+b^2}}, \tag{3.11}$$

and

$$s = \frac{b}{\sqrt{a^2+b^2}}, \tag{3.12}$$

such that $(a,b)^T \neq 0$. Eq. (3.10), is then transformed as follows:

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sqrt{a^2+b^2} \\ 0 \end{pmatrix}. \tag{3.13}$$

3.3.2.3 Householder QR decomposition

This particular QR decomposition technique is based on Householder transformations (elementary reflector), which also involves the orthogonal transformation of a given vector. As a result, a series of zeros are introduced in the vector through a vector product. Consider $v \in \mathbb{R}^n$ to be nonzero and an $n \times n$ matrix P can be expressed as:

$$P = I - vv^T, \tag{3.14}$$

which is also termed as a Householder reflection. The vector v is known as the Householder vector and in the case where another vector x is multiplied with P, it is then reflected in the hyperplane span $\{v\}^{\perp}$.

$$Px = \left(I - \frac{2vv^T}{v^T v}\right) x = x - \frac{2v^T x}{v^T v} v. \tag{3.15}$$

Householder transformations are symmetric and orthogonal as shown as follows [130]:

$$\begin{aligned} P^T P &= \left(I - vv^T\right)\left(I - vv^T\right), \\ &= I - 2vv^T + 2vv^T, \\ &= I. \end{aligned} \tag{3.16}$$

3.3.3 Block MMSE

The Block QR method is compared to an alternative technique known as Block MMSE, which treats the interference as noise. The received signal model (based on Eq. (3.3)) for a 4×4 model can re-written as follows:

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{H}_1 & \mathbf{H}_2 \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} + \mathbf{z}, \tag{3.17}$$

Coupling the interference and AWGN noise vectors, it is shown that the received signal model is a function of the desired transmitted signal ($\mathbf{s}_1$):

$$\mathbf{y} = \mathbf{H}_1\mathbf{s}_1 + \hat{\mathbf{z}}. \tag{3.18}$$

where the received signal is given by $\mathbf{y}$ and $(\mathbf{H}_1, \mathbf{H}_2) \in \mathbb{C}^{4\times 2}$ is the complex channel matrix, such that:

$$\mathbf{H}_1 = \begin{bmatrix} h_{11} & h_{21} & h_{31} & h_{41} \\ h_{12} & h_{22} & h_{32} & h_{42} \end{bmatrix}^T, \tag{3.19}$$

$$\mathbf{H}_2 = \begin{bmatrix} h_{13} & h_{23} & h_{33} & h_{43} \\ h_{14} & h_{24} & h_{34} & h_{44} \end{bmatrix}^T. \tag{3.20}$$

The spatially multiplexed transmit data from both eNBs is given as $\mathbf{s}_1 = (x_1, x_2)^T$ and $\mathbf{s}_2 = (x_3, x_4)^T$ where $\mathbf{s}_1$ and $\mathbf{s}_2$ are both 2 dimensional vectors, respectively (where $(.)^T$ is the matrix transpose). The noise vector $\mathbf{z}$ represents uncorrelated AWGN processes with zero mean and a covariance corresponding to $\sigma_z^2\mathbf{I}$, while $\hat{\mathbf{z}}$ represents the coloured noise component encompassing the interference and AWGN noise and therefore does not retain the statistical properties of AWGN noise. Taking the Eigenvalue Decomposition (EVD) of the coloured noise covariance dominated by the interfering signal:

$$\mathbf{R_{nn}} = \frac{\rho}{N_T}\mathbf{H}_2\mathbf{H}_2^\dagger + \sigma_z\mathbf{I}, \tag{3.21}$$

$$= \mathbf{U}\mathbf{D}\mathbf{U}^\dagger, \tag{3.22}$$

where $\mathbf{R_{nn}}$ is the noise covariance of the interference and AWGN noise, ρ is the SNR and N_T is the number of transmit antennas, $(\ldots)^\dagger$ represents the Hermitian matrix, $\mathbf{U}$ is a unitary matrix containing the eigenvectors of $\mathbf{R_{nn}}$ and $\mathbf{D}$ is a diagonal matrix containing the corresponding eigenvalues. Applying Eq. (3.22) to the received signal, it is shown that the desired signal can be detected subject to

applying the following transformation:

$$\hat{\mathbf{y}} = \sqrt{\mathbf{D}^{-1}}\mathbf{U}^{\dagger}\mathbf{y} = (\sqrt{\mathbf{D}^{-1}}\mathbf{U}^{\dagger}\mathbf{H}_1)\mathbf{s}_1 + \sqrt{\mathbf{D}^{-1}}\left((\mathbf{U}^{\dagger}\mathbf{H}_2)\mathbf{s}_2 + \mathbf{U}^{\dagger}\mathbf{z}\right) \tag{3.23}$$

$$\hat{\mathbf{y}} = (\sqrt{\mathbf{D}^{-1}}\mathbf{U}^{\dagger}\mathbf{H}_1)\mathbf{s}_1 + \hat{\mathbf{z}}, \tag{3.24}$$

It can be intuitively shown that applying the EVD to $\mathbf{R_{nn}}$, reduces it to an AWGN processes with zero mean and $\sigma_z^2\mathbf{I}$ covariance.

$$\begin{aligned}
\mathbf{R}_{\hat{\mathbf{z}}\hat{\mathbf{z}}} &= \mathbf{A}\left(\frac{\rho}{N_T}\mathbf{H}_2\mathbf{H}_2^{\dagger} + \mathbf{I}\right)\mathbf{A}^{\dagger} \\
&= \sigma_z\sqrt{\mathbf{D}^{-1}}\mathbf{U}^{\dagger}\left(\frac{\rho}{N_T}\mathbf{H}_2\mathbf{H}_2^{\dagger} + \sigma_z^2\mathbf{I}\right)\sigma_z\sqrt{\mathbf{D}^{-1}}\mathbf{U} \\
&= \sigma_z^2\sqrt{\mathbf{D}^{-1}}\mathbf{U}^{\dagger}\left(\frac{\rho}{N_T}\mathbf{H}_2\mathbf{H}_2^{\dagger} + \sigma_z^2\mathbf{I}\right)\mathbf{U}\sqrt{\mathbf{D}^{-1}} \\
&= \sigma_z^2\sqrt{\mathbf{D}^{-1}}\mathbf{D}\sqrt{\mathbf{D}^{-1}} \\
&= \sigma_z^2\mathbf{I},
\end{aligned} \tag{3.25}$$

The EVD serves as a 'noise whitening' filter on the overall coloured noise covariance. This significantly simplifies the detection procedure but does not exploit the decoding of the interference. Understanding whether the mutual information $I(\mathbf{s}_1;\hat{\mathbf{y}})$ on decoding the first stream is lossless or lossy is an important aspect of the study and is tackled in the next section.

3.4 MUTUAL INFORMATION ANALYSIS

3.4.1 Block QR

The loss in mutual information at the receiver resulting from the Block QR technique will be quantified, under the assumption that the interference is Gaussian. According to Eq. (3.2), the received signal is a function of the desired stream ($\mathbf{s}_1$) and interfering stream ($\mathbf{s}_2$) and is given as:

$$\mathbf{y}_F = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix}\begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix} \tag{3.26}$$

$$= \mathbf{H}\mathbf{s} + \mathbf{n}. \tag{3.27}$$

It should be noted that an sequential ordering scheme can be applied to the Block QR decomposed channel, in order to determine the initial stream to be decoded, e.g. the stream ordering criteria can be based on the SNR, SINR or the column-norm, however the performance becomes heavily reliant on the channel estimates in the case of imperfect receiver CSI. Applying the Recursive Block QR decomposition (originally inspired by [129]) the expression becomes:

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{Q}}_2 \end{pmatrix} \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \mathbf{R}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix} \tag{3.28}$$

$$= \mathbf{Q}_1\mathbf{Q}_2\mathbf{R}\mathbf{s} + \mathbf{n} \tag{3.29}$$

Multiplying $\mathbf{Q}_1^T$ (orthonormal) throughout Eq. (3.28):

$$\begin{pmatrix} \mathbf{Q}_{11}^T & \mathbf{Q}_{21}^T \\ \mathbf{Q}_{12}^T & \mathbf{Q}_{22}^T \end{pmatrix} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \tilde{\mathbf{Q}}_2\mathbf{R}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{Q}_{11}^T & \mathbf{Q}_{21}^T \\ \mathbf{Q}_{12}^T & \mathbf{Q}_{22}^T \end{pmatrix} \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix} \tag{3.30}$$

$$\tilde{\mathbf{y}}_\mathbf{F} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \tilde{\mathbf{Q}}_2\mathbf{R}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{Q}_{11}^T & \mathbf{Q}_{21}^T \\ \mathbf{Q}_{12}^T & \mathbf{Q}_{22}^T \end{pmatrix} \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix} \tag{3.31}$$

$$= \tilde{\mathbf{H}}\mathbf{s} + \tilde{\mathbf{n}} \tag{3.32}$$

where $\tilde{\mathbf{y}}_F = \begin{pmatrix} \tilde{\mathbf{y}}_1 \\ \tilde{\mathbf{y}}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_{11}^T\mathbf{y}_1 + \mathbf{Q}_{21}^T\mathbf{y}_2 \\ \mathbf{Q}_{12}^T\mathbf{y}_1 + \mathbf{Q}_{22}^T\mathbf{y}_2 \end{pmatrix}$, $\tilde{\mathbf{H}} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \tilde{\mathbf{Q}}_2\mathbf{R}_{22} \end{pmatrix}$ and $\tilde{\mathbf{n}} = \mathbf{Q}_1^T\mathbf{n}$.
The overall mutual information relating to the model in Eq. (3.26) according to the mutual information chain rule is [133]:

$$\begin{aligned} I(\mathbf{s}_1\mathbf{s}_2;\mathbf{y}_F|\mathbf{H}) &= I(\mathbf{s}_2;\mathbf{y}_F|\mathbf{H}) + I(\mathbf{s}_1;\mathbf{y}_F|\mathbf{H},\mathbf{s}_2) \\ &\geq I(\mathbf{s}_2;\tilde{\mathbf{y}}_2|\mathbf{H}) + I(\mathbf{s}_1;\tilde{\mathbf{y}}_F|\mathbf{H},\mathbf{s}_2). \end{aligned} \tag{3.33}$$

The aim of the initial investigation is to quantify the extent of the loss in mutual information of the first stream such that:

$$I(\mathbf{s}_2;\tilde{\mathbf{y}}_2|\mathbf{H}) \leq I(\mathbf{s}_2;\mathbf{y}_F|\mathbf{H}). \tag{3.34}$$

In the case of CSCG random vectors, the capacity of a MIMO system is the maximum of the mutual information between **s** and **y**:

$$C_{MIMO} = \max_{p(x):\mathbb{E}\{\mathbf{s}^H\mathbf{s}\}\leq P_T} I(\mathbf{s},\mathbf{y}_F|\mathbf{H}) \quad \text{bits/s/Hz}. \tag{3.35}$$

The mutual information definition that measures the dependence of two variables (**s** and **y**) by determining the entropic characteristics:

$$I(\mathbf{s},\mathbf{y}_F|\mathbf{H}) = \mathscr{H}(\mathbf{y}_F) - \mathscr{H}(\mathbf{y}_F|\mathbf{s}), \tag{3.36}$$

where $\mathscr{H}(\mathbf{y}_F)$ and $\mathscr{H}(\mathbf{y}_F|\mathbf{s})$ represent the entropy of the received signal $\mathbf{y}_F$ and conditional entropy respectively. The received signal entropy, $\mathscr{H}(\mathbf{y}_F)$, is given by:

$$\mathscr{H}(\mathbf{y}_F) = \log_2\left\{\det(\pi e \mathbf{R}_{\mathbf{y}_F\mathbf{y}_F})\right\}, \tag{3.37}$$

where $\mathbf{R}_{\mathbf{y}_F\mathbf{y}_F}$ is the covariance of the overall received signal. In this case, the received signal vector ($\mathbf{y}_F$) and noise vector (**n**) are statistically independent, resulting in:

$$\mathscr{H}(\mathbf{y}_F|\mathbf{s}) = \mathscr{H}(\mathbf{n}) = \log_2\left\{\det(\pi e \sigma_n^2 \mathbf{I}_{N_R})\right\}, \tag{3.38}$$

where $\sigma_n^2\mathbf{I}_{N_R}$ is the noise covariance. This leads to the following expanded expression of $I(\mathbf{s}_2;\mathbf{y}_F|\mathbf{H})$, using Eq. (3.34) and the aforementioned definitions:

$$\begin{aligned} I(\mathbf{s}_2;\mathbf{y}_F|\mathbf{H}) &= \mathscr{H}(\mathbf{y}_F|\mathbf{H}) - \mathscr{H}(\mathbf{y}_F|\mathbf{H},\mathbf{s}_2) && (3.39)\\ &= \log_2\left\{\det\left(\pi e \mathbf{R}_{\mathbf{y}_F|\mathbf{H}}\right)\right\} - \log_2\left\{\det\left(\pi e \mathbf{R}_{\mathbf{y}_F|\mathbf{H},\mathbf{s}_2}\right)\right\}, && (3.40)\end{aligned}$$

where $\mathbf{R}_{\mathbf{y}_F|\mathbf{H}}$ is the received signal covariance, provided that knowledge of the channel (**H**) is known. The covariance $\left(\mathbf{R}_{\mathbf{y}_F|\mathbf{H}}\right)$ is expressed as follows:

$$\begin{aligned} \mathbf{R}_{\mathbf{y}_F|\mathbf{H}} &= \mathbb{E}\left\{\mathbf{y}_F\mathbf{y}_F^\dagger\right\}, \\ &= \mathbb{E}\left\{(\mathbf{H}\mathbf{s}+\mathbf{n})\left(\mathbf{H}^\dagger\mathbf{s}^\dagger+\mathbf{n}^\dagger\right)\right\}, \\ &= \mathbf{H}\mathbf{R}_{\mathbf{ss}}\mathbf{H}^\dagger + \sigma_n^2\mathbf{I}_{N_R}, && (3.41)\end{aligned}$$

where $\mathbb{E}\{\ldots\}$ is the expectation operator and $\{\ldots\}^{\dagger}$ is the Hermitian matrix. The transmitted signal covariance is given by $\mathbf{R_{ss}} = \frac{E_s}{N_T}\mathbf{I}_{N_T}$ (where E_s is the transmit power equally allocated with the assumption of no (CSIT)). The number of the transmit and receive antennas are given by N_T and N_R. The noise spectral density is given by σ_n^2 and the differential entropy is maximised when the received signal vector $\mathbf{y}_F$ is characterised by a zero mean circular symmetric complex Gaussian (ZMCSCG) distribution. Similarly the covariance vector, $\mathbf{R}_{\mathbf{y}_F|\mathbf{H},\mathbf{s_2}}$, is given by:

$$\mathbf{R}_{\mathbf{y}_F|\mathbf{H},\mathbf{s_2}} = \mathbb{E}\left\{\mathbf{H}_1\mathbf{s}_1\mathbf{s}_1^{\dagger}\mathbf{H}_1^{\dagger} + \mathbf{n}\mathbf{n}^{\dagger}\right\} \tag{3.42}$$

$$= \mathbf{H}_1\mathbf{R}_{\mathbf{s}_1\mathbf{s}_1}\mathbf{H}_1^{\dagger} + \sigma_n^2\mathbf{I}_{N_{R_1}}, \tag{3.43}$$

where $\mathbf{H}_1 = (\mathbf{H}_{11}\,\mathbf{H}_{21})^T$, and N_{T_1} and N_{R_1} are the number of transmit antennas of the interfering BS and the number of receive antennas at the MS. Now substituting Eqs. (3.41) and (3.43) into (3.40):

$$I(\mathbf{s}_2;\mathbf{y}_F|\mathbf{H}) = \log_2\left\{\frac{\det\left(\frac{\rho}{N_T}\mathbf{H}\mathbf{H}^{\dagger} + \mathbf{I}_{N_R}\right)}{\det\left(\frac{\rho_1}{N_{T_1}}\mathbf{H}_1\mathbf{H}_1^{\dagger} + \mathbf{I}_{N_{R_1}}\right)}\right\} \tag{3.44}$$

where $\rho = \frac{E_s}{N_0}$ is the signal-to-noise ratio (SNR). Similarly, the expression for $I(\mathbf{s}_2;\tilde{\mathbf{y}}_F|\tilde{\mathbf{H}})$ can be given as:

$$I(\mathbf{s}_2;\tilde{\mathbf{y}}_F|\tilde{\mathbf{H}}) = \mathscr{H}(\tilde{\mathbf{y}}_F|\tilde{\mathbf{H}}) - \mathscr{H}(\tilde{\mathbf{y}}_F|\tilde{\mathbf{H}},\mathbf{s}_2), \tag{3.45}$$

$$= \log_2\left\{\det\left(\pi e\mathbf{R}_{\tilde{\mathbf{y}}_F|\tilde{\mathbf{H}}}\right)\right\} - \log_2\left\{\det\left(\pi e\mathbf{R}_{\tilde{\mathbf{y}}_F|\tilde{\mathbf{H}},\mathbf{s_2}}\right)\right\}, \tag{3.46}$$

where $\tilde{\mathbf{H}}$ represents the transformed Block QR channel. Similarly as in Equation (3.41) and (3.43), the transformed transmitted signal covariance is given as:

$$\mathbf{R}_{\tilde{\mathbf{y}}_F|\mathbf{H}} = \tilde{\mathbf{H}}\mathbf{R}_{\mathbf{ss}}\tilde{\mathbf{H}}^{\dagger} + \mathbb{E}\left\{\tilde{\mathbf{n}}\tilde{\mathbf{n}}^{\dagger}\right\}, \tag{3.47}$$

$$= \tilde{\mathbf{H}}\mathbf{R}_{\mathbf{ss}}\tilde{\mathbf{H}}^{\dagger} + \sigma_n^2\mathbf{I}_{N_R} \tag{3.48}$$

and similarly,

$$\mathbf{R}_{\tilde{\mathbf{y}}_F|\tilde{\mathbf{H}},\mathbf{s_2}} = \tilde{\mathbf{H}}_1\mathbf{R}_{\mathbf{s}_1\mathbf{s}_1}\tilde{\mathbf{H}}_1^{\dagger} + \sigma_n^2\mathbf{I}_{N_{R_1}}, \tag{3.49}$$

where $\tilde{\mathbf{H}}_1 = (\mathbf{R}_{11}\,\mathbf{0})^T$. Since $\mathbf{Q}_1$ is a orthonormal transformation, the resulting noise expression is a white Gaussian vector and the noise covariance remains the same as in Equation (3.41) and (3.43).

Therefore:

$$I(\mathbf{s}_2;\tilde{\mathbf{y}}_F|\tilde{\mathbf{H}}) = \log_2\left\{\frac{\det\left(\frac{\rho}{N_T}\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\dagger + \sigma_n^2\mathbf{I}_{N_R}\right)}{\det\left(\frac{\rho_1}{N_{T_1}}\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^\dagger + \sigma_n^2\mathbf{I}_{N_{R_1}}\right)}\right\}. \tag{3.50}$$

3.4.2 Block MMSE

It is has been well-established that MMSE detection techniques are mutual information lossless due to the sufficiently available statistics to detect **s**, in addition to maximising the post-detection SNR [133]. In this section, for the sake of completeness, it is also shown that by using first principles, the mutual information between the transmitted and received signal for the Block MMSE detection technique is lossless, provided that the transmitted signal vector and noise are assumed to be Gaussian in nature. Following from Eq. (3.17), the generalized received signal model for the Block MMSE case is expressed as:

$$\mathbf{y} = \mathbf{H}_1\mathbf{s}_1 + \mathbf{H}_2\mathbf{s}_2 + \mathbf{z}, \tag{3.51}$$

where the post-detection noise includes the interference (coloured noise process), $\hat{\mathbf{z}} = \mathbf{H}_2\mathbf{s}_2 + \mathbf{z}$, which is indicated by Eq. (3.18). The chain rule of mutual information based on Eq. (3.51) is shown to be:

$$I(\mathbf{s}_1,\mathbf{s}_2;\mathbf{y}|\mathbf{H}) = I(\mathbf{s}_1;\mathbf{y}|\mathbf{H}_1\mathbf{H}_2) + I(\mathbf{s}_2;\mathbf{y}|\mathbf{H}_1\mathbf{H}_2,\mathbf{s}_1). \tag{3.52}$$

The mutual information $I(\mathbf{s}_1;\mathbf{y}|\mathbf{H}_1\mathbf{H}_2)$ is classically defined in terms of the signal entropy:

$$I(\mathbf{s}_1;\mathbf{y}|\mathbf{H}_1\mathbf{H}_2) = \mathscr{H}(\mathbf{y}) - \mathscr{H}(\mathbf{y}|\mathbf{s}_1). \tag{3.53}$$

Furthermore,

$$\mathscr{H}(\mathbf{y}) = \log_2\left\{\det\left(\pi e\mathbf{R}_{\mathbf{yy}}\right)\right\}, \tag{3.54}$$

$$\begin{aligned}\mathscr{H}(\mathbf{y}|\mathbf{s}_1) &= \mathscr{H}(\hat{\mathbf{z}}) \\ &= \log_2\left\{\det\left(\pi e\mathbf{R}_{\mathbf{zz}}\right)\right\},\end{aligned} \tag{3.55}$$

The covariance of the received signal (assuming Gaussian signals) can be expanded as follows:

$$\begin{aligned}\mathbf{R}_{\mathbf{yy}} &= \mathbb{E}\left\{\mathbf{y}\mathbf{y}^{\dagger}\right\} \\ &= \mathbb{E}\left\{\left(\mathbf{H}_1\mathbf{s}_1+\hat{\mathbf{z}}\right)\left(\mathbf{H}_1^{\dagger}\mathbf{s}_1^{\dagger}+\hat{\mathbf{z}}^{\dagger}\right)\right\}, \\ &= \mathbf{H}_1\mathbf{R}_{\mathbf{s}_1\mathbf{s}_1}\mathbf{H}_1^{\dagger}+\mathbf{R}_{\hat{\mathbf{z}}\hat{\mathbf{z}}}, \end{aligned} \tag{3.56}$$

while the coloured noise covariance can be represented as:

$$\begin{aligned}\mathbf{R}_{\hat{\mathbf{z}}\hat{\mathbf{z}}} &= \mathbb{E}\left\{\hat{\mathbf{z}}\hat{\mathbf{z}}^{\dagger}\right\} \\ &= \mathbb{E}\left\{\left(\mathbf{H}_2\mathbf{s}_2+\mathbf{z}\right)\left(\mathbf{H}_2\mathbf{s}_2+\mathbf{z}\right)^{\dagger}\right\} \\ &= \mathbf{H}_2\mathbf{R}_{\mathbf{s}_2\mathbf{s}_2}\mathbf{H}_2^{\dagger}+\sigma_z^2\mathbf{I}. \end{aligned} \tag{3.57}$$

As a result the mutual information is shown to be:

$$\begin{aligned} I(\mathbf{s}_1;\mathbf{y}|\mathbf{H}) &= \log_2\left\{\det\left(\pi e\left(\frac{\rho_1}{N_{T_1}}\mathbf{H}_1\mathbf{H}_1^{\dagger}+\mathbf{R}_{\hat{\mathbf{z}}\hat{\mathbf{z}}}\right)\right)\right\}-\log_2\left\{\det\left(\pi e\mathbf{R}_{\hat{\mathbf{z}}\hat{\mathbf{z}}}\right)\right\}, \\ &= \log_2\left\{\frac{\det\left(\frac{\rho_1}{N_{T_1}}\mathbf{H}_1\mathbf{H}_1^{\dagger}+\left(\frac{\rho_2}{N_{T_2}}\mathbf{H}_2\mathbf{H}_2^{\dagger}+\sigma_z^2\mathbf{I}\right)\right)}{\det\left(\frac{\rho_2}{N_{T_2}}\mathbf{H}_2\mathbf{H}_2^{\dagger}+\sigma_z^2\mathbf{I}\right)}\right\}, \\ &= \mathbb{E}_{\mathbf{H}}\left[\log_2\left\{\det\left(\frac{\rho_1}{N_{T_1}}\mathbf{H}_1\mathbf{H}_1^{\dagger}\left(\frac{\rho_2}{N_{T_2}}\mathbf{H}_2\mathbf{H}_2^{\dagger}+\sigma_z^2\mathbf{I}\right)^{-1}+\mathbf{I}\right)\right\}\right], \end{aligned} \tag{3.58}$$

which is effectively information lossless.

3.5 NUMERICAL RESULTS

The maximal achievable rate for the different SNR values are examined (in terms of the cumulative distribution functions (CDFs)) for the theoretical, Block QR, Block MMSE and traditional MMSE schemes for interference-limited higher-order MIMO systems. Figs. 3.2 and 3.3 represent the theoretical CDF curves for the 4×4 (the first detected two-dimensional stream) and 8×8 (the first detected four-dimensional stream) channels, respectively as described by Eq. (3.34). For an outage probability of 10^{-1} (target BLER for the LTE standard [4, 102]) and at an SNR of 15 dB the achievable rate difference for the 4×4 model between the Block QR and theoretical case is 0.55 bits/s/Hz, while at a higher SNR of 25 dB, the difference reduces to 0.3 bits/s/Hz. As a result, the Block QR scheme

approaches the theoretical limit in the higher SNR regime for the first stream, while the MMSE rate in the presence of interference degrades significantly, which is especially prominent in the 8×8 case as shown in Fig. 3.3. It must be noted that statistics of the channel are Rayleigh (block fading) with

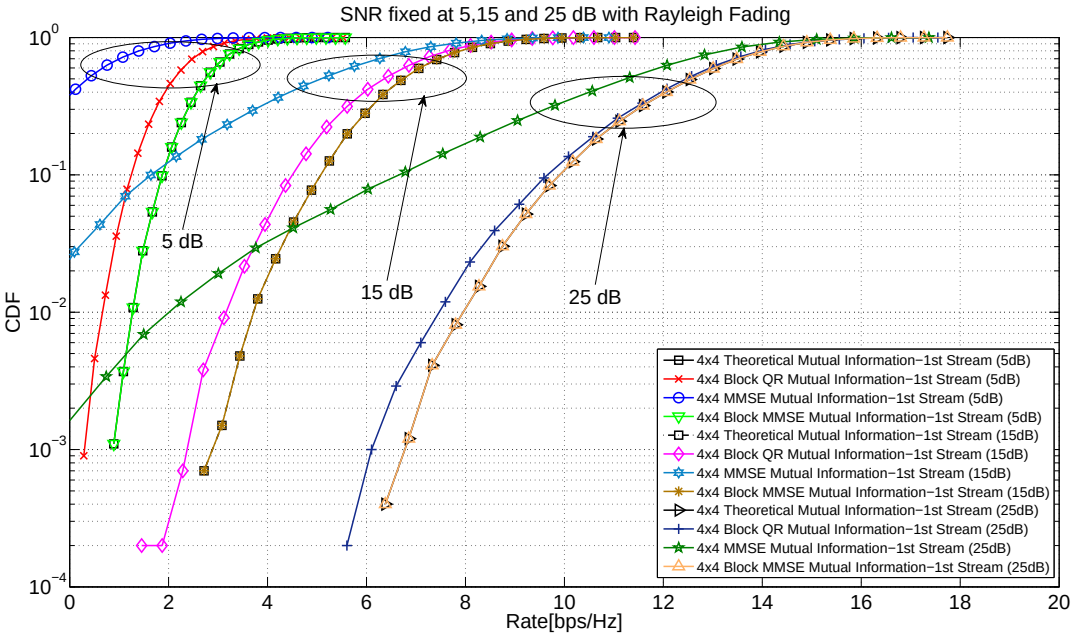


Figure 3.2: Mutual information loss at various SNR for the 4×4 MIMO model.

equal power allocation at the transmitter. Moreover, the Gram-Schmidt QR decomposition method was selected for the implementation of the Block algorithm, of which the maximal achieved rate can be noted in Figs. 3.2 and 3.3, although the Householder and Givens QR techniques display the same performance only differing in complexity. These numerical results are then studied with respect to the implementation in a LTE simulation test bench, consisting of a full baseband transmit and receive chain, which is based on the software libraries in the OpenAirInterface simulator [28]. This would involve a comparative performance evaluation of the LTE physical layer downlink simulation of the low-complex Max-log-MAP detector [27, 91] and suboptimal brute-force search Max-log-MAP detector [85] (see Section 2.2.2.4). The practical evaluation of an 8×8 system is out of the scope of this study since a novel low-complexity soft-decision demodulating metric has to be developed for the detection of the four-dimensional stream.

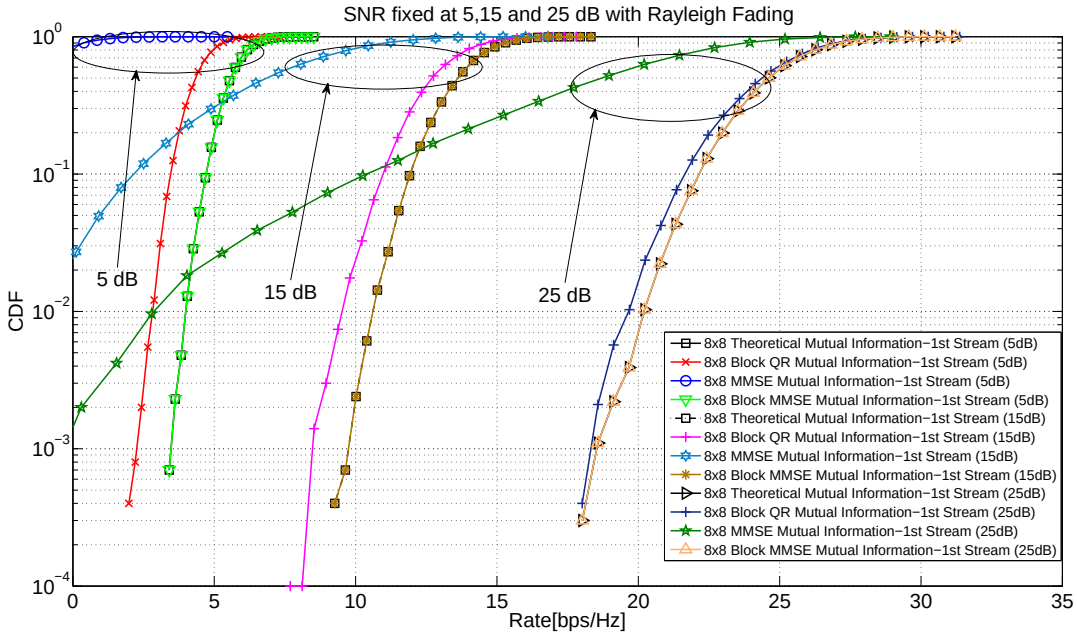


Figure 3.3: Mutual information loss at various SNR for the 8×8 MIMO model.

3.6 LTE SINGLE-USER MIMO (SU-MIMO) RATE OPTIMIZATION

In this section, the focus is shifted towards the maximisation of the overall rate in a 4×4 and 8×8 point-to-point SU-MIMO system by considering the detection of the two transmitted streams ($\mathbf{s}_1$ and $\mathbf{s}_2$). In LTE, a UE can only be configured to receive a maximum of only two codewords within a single transmission interval. Each of the two codewords, R_1 and R_2 can be maximised to yield a maximum total achievable rate for a single-user. Earlier work in [134] investigated a per-antenna-rate-control (PARC) scheme to maximise the throughput of a HSDPA WCDMA MIMO system, provided that the number of transmitted data substreams are less than the number of transmit antennas. The authors show that using a SIC-GRAKE receiver, the throughput can be maximised by adaptively selecting the number of transmitted data substreams according to the radio operating conditions. Unlike the aforementioned study, this particular study focuses on the rate optmisation using higher-order MIMO detection schemes in line with the current LTE standard. Referring to the earlier expression regarding the chain rule of mutual information for Gaussian signal alphabets (Eq. (3.33)) for the 4×4 and 8×8 MIMO schemes:

$$I(\mathbf{s}_1\mathbf{s}_2;\mathbf{y}|\mathbf{H}) = I(\mathbf{s}_1;\mathbf{y}|\mathbf{H}) + I(\mathbf{s}_2;\mathbf{y}|\mathbf{H},\mathbf{s}_1), \tag{3.59}$$

where $\mathbf{s}_1$ is the first detected stream, $\mathbf{s}_2$ is the interference-free stream (after application of SIC) and $\mathbf{H} = [\mathbf{h}_1\ \mathbf{h}_2]$, where $\{\mathbf{h}_1, \mathbf{h}_2\} \in \mathbb{C}^{4\times 2}$, $\{\mathbf{s}_1, \mathbf{s}_2\} \in \mathbb{C}^{2\times 1}$ and $\mathbf{n} \in \mathbb{C}^{4\times 1}$ represents the CSCG noise vector. The channel has been block divided in blocks of two for the 4×4 case and blocks of four for the 8×8. The total rate is defined by the sum of the first detected stream ($I(\mathbf{s}_1;\mathbf{y}|\mathbf{H})$) in the presence of co-layer interference and the resulting interference-free stream ($I(\mathbf{s}_1;\mathbf{y}|\mathbf{H},\mathbf{s}_2)$). The mutual information of the first stream is given by [42]:

$$\begin{aligned} I(\mathbf{s}_1;\mathbf{y}|\mathbf{H}) &= \log_2\left[\frac{\det\left\{\mathbf{I}+\rho_1\mathbf{h}_1\mathbf{h}_1^\dagger+\rho_2\mathbf{h}_2\mathbf{h}_2^\dagger\right\}}{\mathbf{I}+\rho_2\mathbf{h}_2\mathbf{h}_2^\dagger}\right], \\ &= \mathbb{E}_{\mathbf{H}}\left(\log_2\left[\det\left\{\mathbf{I}+\rho_1\mathbf{h}_1\mathbf{h}_1^\dagger\left(\sigma_n^2\mathbf{I}+\rho_2\mathbf{h}_2\mathbf{h}_2^\dagger\right)^{-1}\right\}\right]\right), \end{aligned} \tag{3.60}$$

where ρ represents the received SNR of each stream. The mutual information of the second stream is given by:

$$\begin{aligned} I(\mathbf{s}_1;\mathbf{y}|\mathbf{H},\mathbf{s}_2) &= \log_2\left[\frac{\det\left\{\mathbf{I}+\rho_2\mathbf{h}_2\mathbf{h}_2^\dagger\right\}}{\mathbf{I}}\right], \\ &= \mathbb{E}_{\mathbf{h_2}}\left(\log_2\left[\mathbf{I}+\rho_2\left\|\mathbf{h}_2\right\|^2\right]\right). \end{aligned} \tag{3.61}$$

The eNB can adaptively transmit the two codewords at different MCS values governed by the following achievable rates:

$$R_1 \leq I(\mathbf{s}_1;\mathbf{y}|\mathbf{H}), \tag{3.62}$$

$$R_2 \leq I(\mathbf{s}_1;\mathbf{y}|\mathbf{H},\mathbf{s}_2). \tag{3.63}$$

The overall maximised rate of an LTE single-user can be obtained via maximising the total achievable rate over R_1 and R_2, respectively.

3.6.1 Block QR SU-MIMO Rate Maximisation

The conventional total achievable rate (R_T) for a MIMO receiver where the two codewords have to be jointly decoded is given as:

$$T_{Joint}(R_T) = \max_{R_T} \left[R_T (1 - P_{out}(R_T)) \right], \tag{3.64}$$

$$= \max_{R_1, R_2} \left[(R_1 + R_2)(1 - P_{out}(R_T)) \right], \tag{3.65}$$

where P_{out} represents the outage probability at a target Rate R_T. The next strategy is the joint SIC decoding scheme, where in order to achieve the desired maximum total rate, both codewords would have to be decoded successfully and therefore a decoding failure (outage), i.e. $P_{out}(R_1) = 1$ of either codewords will cause the overall system to be in outage.

$$T_{SIC} = \max_{R_T} \left[(R_T)(1 - P_{out}(R_1))(1 - P_{out}(R_2)) \right] \tag{3.66}$$

However, the total achievable rate ($T_{LTE}(R_T)$) in LTE is given as:

$$T_{LTE}(R_T) = \max_{R_1, R_2} \left[\overbrace{R_1(1 - P_{out}(R_1))}^{1st\ Codeword} + \overbrace{R_2(1 - P_{out}(R_2))(1 - P_{out}(R_1))}^{2nd\ Codeword} \right] \tag{3.67}$$

It can be noted that if the decoding of the first codeword is in outage then the total rate is in outage, while if the second codeword is in outage then the total rate is equivalent to the rate of the first codeword. Figs. 3.4 and 3.5 show the total achievable rate as well as the rate of the first (R_1) and second (R_2) streams based on Eq. (3.67) for a 4×4 MIMO system over the practical SNR regime in a Rayleigh (non-line-of-sight (NLOS)) and Rician Fading (strong line-of-sight (LOS)) environment for Gaussian signal alphabets, respectively. The channel is therefore modelled as a function of both LOS and NLOS components, which is based on the following model:

$$\mathbf{H} = \sqrt{\frac{K}{K+1}} \bar{\mathbf{H}}_{LOS} + \sqrt{\frac{1}{K+1}} \hat{\mathbf{H}}_{NLOS}, \tag{3.68}$$

where $\bar{\mathbf{H}}_{LOS}$ is a deterministic matrix, corresponding to the specular path of the LOS component, while $\hat{\mathbf{H}}_{NLOS}$ consists of the NLOS component with independent and identically distributed (i.i.d)

CSCG random variables, unit variance and zero mean [133]. Similarly, numerical results were developed for the 8×8 MIMO system, as observed in the Figs. 3.6 and 3.7. The Joint SIC performance with respect to the maximised achievable rates are marginally less than the LTE achievable throughput.

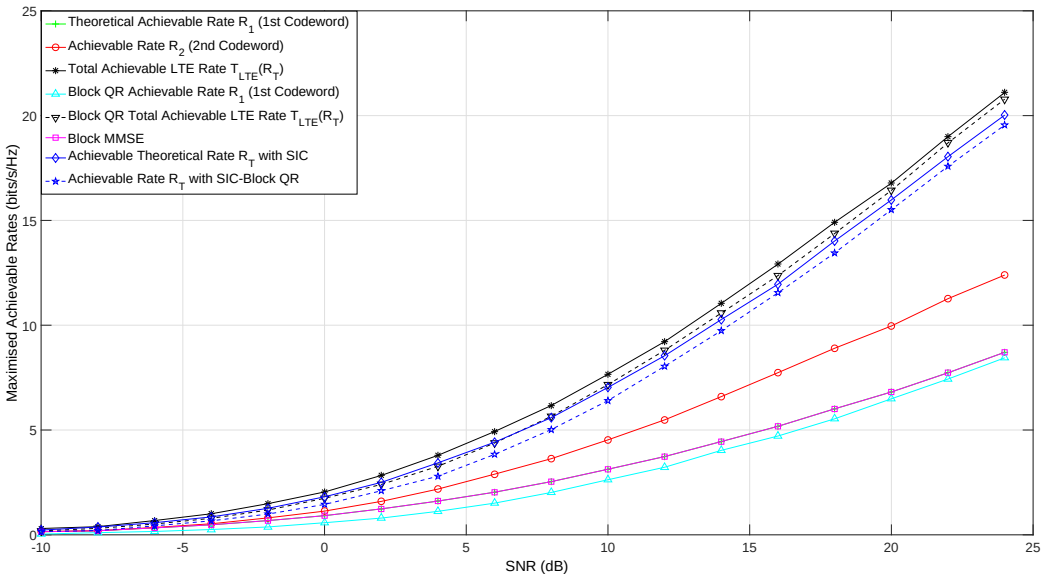


Figure 3.4: Achievable rate as a function of SNR in a 4×4 Rayleigh Channel

The rate loss of the Block QR technique occurs only on the first stream and therefore affects the total rate. This shows the importance of first stream detection with respect to the overall single-user LTE rate. A known improvement in the total LTE rate for the 8×8 scenario can be observed when compared to the 4×4 scenario. The rates evaluated in the Rayleigh channel also outperform the strong LOS scenario (Rician channel) due to the rank deficiency of the MIMO channel due to the strong LOS component in the latter case. Table 3.1 shows a comparison between the total LTE theoretical and Block QR system rate for the three different SNR values in the practical regime.

It is shown that the 8×8 configuration incurs a greater difference in spectral efficiency between the Rayleigh and Ricean channel in the higher operating SNR point, when compared to the 4×4 case.

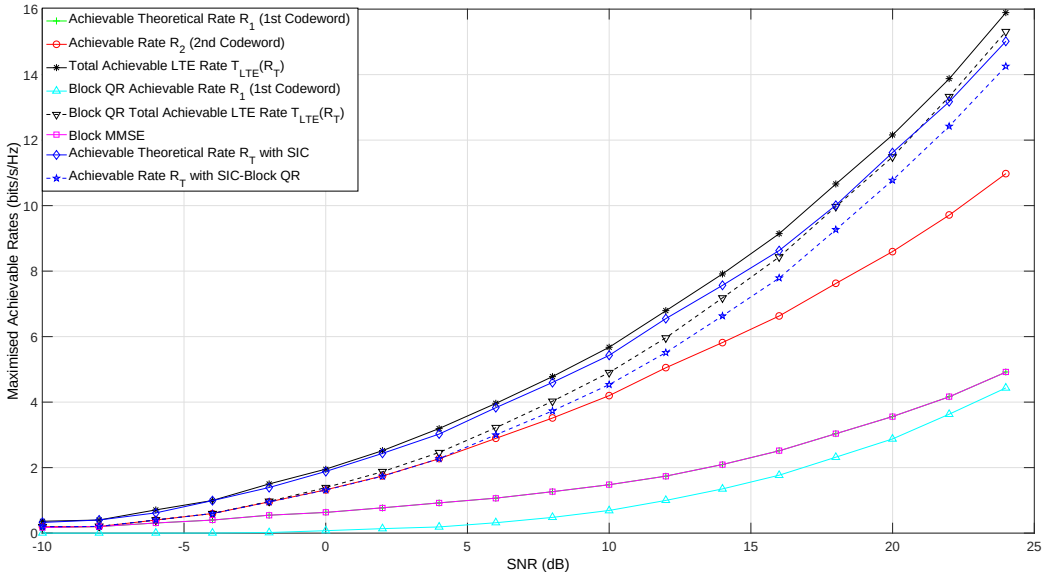


Figure 3.5: Achievable rate as a function of SNR in a 4×4 Rician Channel with K=10

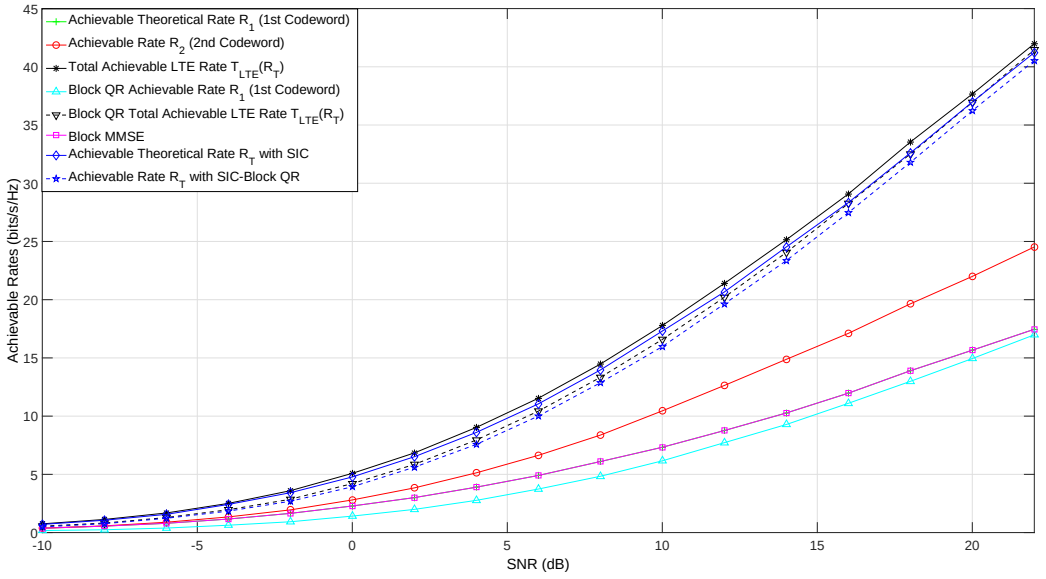


Figure 3.6: Achievable rate as a function of SNR in a 8×8 Rayleigh Channel

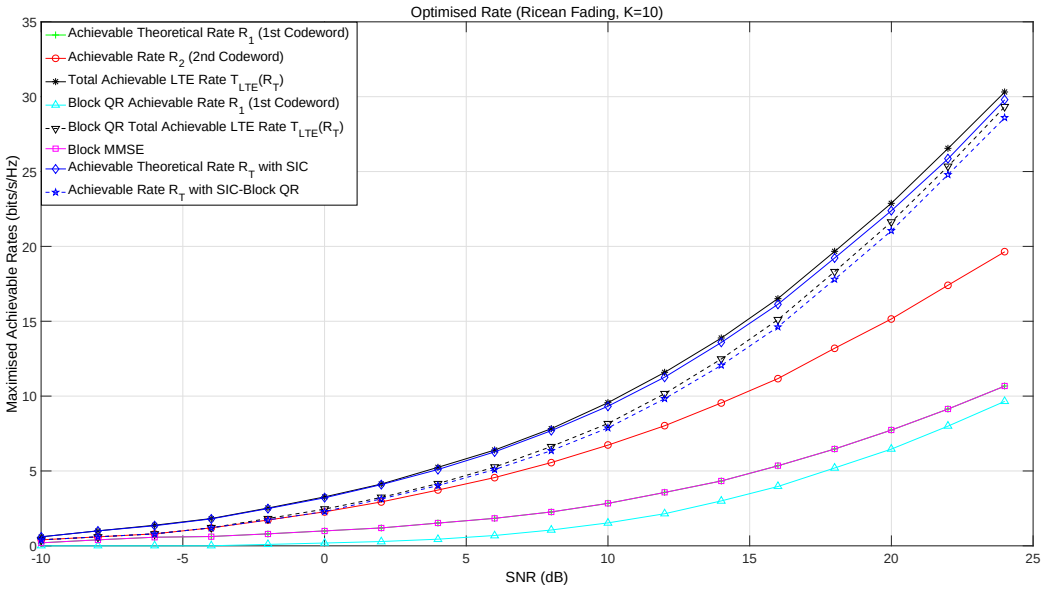


Figure 3.7: Achievable rate as a function of SNR in a 8×8 Rician Channel with K=10

Table 3.1: Difference in Total LTE Theoretical and LTE Block QR rate

SNR (dB)	Channel Type	4×4 (bits\s\Hz)	8×8 (bits\s\Hz)
0	Rayleigh	0.38	0.90
	Rician ($K = 10$)	0.56	0.90
10	Rayleigh	0.65	1.06
	Rician ($K = 10$)	0.84	1.41
20	Rayleigh	0.38	0.50
	Rician ($K = 10$)	0.62	1.30

3.7 CONCLUSIONS

A novel channel preprocessing MIMO detection scheme based on a recursive Block QR decomposition approach for a single-user HetNet scenario has been proposed and theoretically evaluated in relation to the Block MMSE technique, which assumes that the interference is Gaussian. The corresponding information theoretic results show that in the case of the 4×4 and 8×8 MIMO configurations, the first stream detection mutual information loss between the theoretical and Block QR outage probability is minimal, for signals characterised by Gaussian alphabets, especially in higher SNR regimes. A scheme has also been introduced whereby the LTE system rate has been optimised

for a SU-MIMO point-to-point link. This Chapter has provided analytical insights of the proposed Block QR scheme for higher-order MIMO systems in both interference-limited and point-to-point SU-MIMO scenarios.

CHAPTER 4

A PRACTICAL EVALUATION OF HIGHER-ORDER MIMO BLOCK DETECTION

4.1 INTRODUCTION

In this chapter, the proposed Block detection schemes are evaluated in a practical context, having laid the foundation of the algorithmic performance on an analytical level in Chapter 3. The complexity is analysed in relation to the number of floating-point operations (flops), which serves an important merit from a hardware implementation point of view. The reliability in terms of the Block Error Rate (BLER) is also evaluated with the aid of the developed LTE simulation test bench consisting of functional blocks for the baseband transmit and receive chain, which are based on the OpenAirInterface software libraries.

4.1.1 From Algorithm to Bit-level Design

Modern receiver design hinges on the use of fixed-point arithmetic in order to compromise on computational complexity and energy efficiency. The developed receiver algorithms can be evaluated using a floating-point functional model, from an initial design perspective. In this case, the simulation model has to conform as much as possible with existing standard specifications (such as the LTE standard considered in this dissertation) and performance in order to reap the benefits as well as understand the

limitations in a practical setting. However, the initial functional simulation model is not without its drawbacks as hardware-related behavioral aspects such as precision, hardware complexity, throughput and computational latency are not yet fully understood [135]. In this work, the functional components of the simulation test bench are leveraged using the OpenAirInterface physical layer/baseband processing software libraries (as shown in the highlighted portion of Fig. 2.8) and are programmed using C and evaluated in MATLAB using the MEX interface. Since the base link level simulator has been developed in C, the signal processing algorithms are designed using fixed-point arithmetic and therefore constraints are placed on certain aspects such as dynamic range and precision. This is in contrast to the development using floating-point arithmetic, which increases precision at the cost of hardware complexity.

4.2 ON THE COMPLEXITY OF THE BLOCK QR AND BLOCK MMSE SCHEMES

The trade-off between performance and complexity is an important figure of merit for receiver design. The search for the careful balance between the two aspects presents an ongoing challenge for designers. The complexity of a given algorithm can be broadly described in terms of:

- *Computational performance*: Quantifies the computational resources required to run a particular algorithm/program. These may include parameters such as (but not limited to) time, memory and CPU usage. It also depends on various other factors such as the hardware specifications, type of compiler and structure of the code.

- *Scalability*: This aspect sheds insights into the scalability of the algorithm, e.g. if the algorithm expands, does the algorithm scale well for the larger problem? In relation to MIMO receiver design, this metric would also quantify the scalability of the algorithm if the number of transmit or receive antennas are increased.

4.2.1 Block QR

In this analysis, the number of floating operations (e.g. complex addition and multiplication, division, etc.) required to transform a square channel matrix, $\mathbf{H} \in \mathbb{C}^{L \times M}$ ($L = M$), are determined. The generalised implementation complexity $\mathcal{O}(n)$ of the QR decomposition algorithms are also discussed. Block

QR provides a more flexible approach to blocking, which uses a divide-and-conquer approach to perform the factorisation. This algorithm is suitable for parallel computations since it is largely based on matrix multiplications that can be executed concurrently, and is hence adaptable for multi-thread programming of matrix decompositions. The following analysis outlines the additional approximate complexity introduced by the Block QR front-end transformation:

$$\begin{aligned} T(M) &= 2\log_{\frac{M}{k}} M + 2M^3 + 19, \\ &= 2M^3 + 2\log_2 M + 19, \\ &\cong \mathscr{O}\left(n^3\right), \end{aligned} \tag{4.1}$$

where M and k represent the number of receive antennas and block size, respectively. The recursive nature of the function call in the algorithm introduces a $\log_2 M$ factor in the number of operations. As a result, an overall additional complexity of $2M^3 + 2\log_2 M + 19$ operations is introduced, in order to perform the linear transformation of the channel. It should be however, emphasized that the complexity of interest is the flops required to decode the interference-free signal:

$$\tilde{\mathbf{y}}_2 = \left(\mathbf{Q}_{21}^{\dagger} + \mathbf{Q}_{22}^{\dagger}\right)\mathbf{y}_2 = \tilde{\mathbf{Q}}_2^{\dagger}\mathbf{R}_{22}\mathbf{s}_2 + \tilde{\mathbf{n}}_2, \tag{4.2}$$

where the received transformed signal is $\{\tilde{\mathbf{y}}_2, \mathbf{y}_2\} \in \mathbb{C}^{2\times 1}$, the unitary matrices represented by$\{\mathbf{Q}_{21}, \mathbf{Q}_{22}\} \in \mathbb{C}^{2\times 2}$, $\left\{\tilde{\mathbf{Q}}_2, \mathbf{R}_{22}\right\} \in \mathbb{C}^{2\times 2}$, the transmitted signal is given by $\mathbf{s}_2 \in \mathbb{C}^{2\times 1}$ and the transformed CSCG noise vector is $\tilde{\mathbf{n}}_2 \in \mathbb{C}^{2\times 1}$. As a result, Eq. (4.2) requires $2M^3 + 3M^2$ flops. For completeness, the generalised overall complexity for the different QR techniques are also presented in Table 4.1 [130] for a $\mathbf{A} \in \mathbb{R}^{L\times M}$ matrix, provided $L \geq M$. Since the performance is identical, the least complex Modified Gram-Schmidt technique would be sufficient for implementation and offers good numerical stability but not greater than that of the Householder or Givens techniques [136].

4.2.2 Block MMSE

The number of flops required to perform the Block MMSE on the received signal will be analysed, in a similar fashion to Section 4.2.1. The main goal behind the Block MMSE technique is the computation of the eigenvalue decomposition (EVD) of the noise covariance matrix (coloured noise). According to

Table 4.1: Complexity comparison of different QR techniques

QR Method	Generalised Flop Count
Householder	If **Q** is not required: $2M^2\left(L-\frac{M}{3}\right)$
	If **Q** is required: $4\left(L^2M-LM^2+\frac{M^3}{3}\right)$
Givens	If **Q** is not required: $3M^2\left(L-\frac{M}{3}\right)$
Classical Gram-Schmidt (CGS)	$2LM^2$
Modified Gram-Schmidt (MGS)	$2LM^2$

[137], an EVD requires M^3 flops for a $M \times M$ matrix. The different computational complexities across the different stages of the procedure is determined in order to ascertain the required number of flops for this particular scenario. The procedure is divided into three stages, which involve: 1) Computation of the noise covariance matrix, 2) Computation of corresponding eigenvalue decomposition and 3) Transformation of the received signal.

4.2.2.1 Computing the Coloured Noise Matrix

The symmetric noise covariance matrix is shown in Eq. (3.57) and is given as:

$$\mathbf{A} = \frac{\rho}{N_T}\mathbf{H}_2\mathbf{H}_2^{\dagger} + \sigma_z\mathbf{I}, \quad \mathbf{A} \in \mathbb{C}^{M\times N}, \mathbf{H}_2 \in \mathbb{C}^{M\times P}, \mathbf{I} \in \mathbb{C}^{M\times N} \tag{4.3}$$

As a result, the total number of flops for the given matrix, **A**, is :

$$\begin{aligned} T_1 &= 2M^2L - M^2 + M^2 + M^2, \\ &= 2M^2L + M^2 \text{ flops.} \end{aligned} \tag{4.4}$$

4.2.2.2 EVD Evaluation

The computation of the EVD represents a significant portion of the complexity of the Block MMSE scheme and consists of three stages [138]:

1. Reduction of matrix A to tridiagonal form [Sec 8.3 [130]] in the form of: $A = Q_A T Q_A^T$

2. Compute the EVD of T using a set of techniques to obtain: $T = Q_T D Q_T^T$

3. Back-transform the eigenvectors to obtain: $A = Q_A Q_T D Q_T^T Q_A^T$, where $Q = Q_A Q_T$

The first step involves the matrix reduction into tridiagonal form and makes use of Householder transformations to achieve that goal. If the matrix is real, the complexity cost of reduction into tridiagonal form is $\frac{4}{3}M^3$ flops, while for complex matrices the complexity cost is $4 \times \frac{4}{3}M^3$ flops. Half of these operations consist of matrix-vector multiplications Level-2 Basic Linear Algebra Subprograms (L2-BLAS) and the other half of the other operations consist of k Hermitian rank updates (which are L3-BLAS operations). Therefore, the number of flops required to reduce the matrix into tridiagonal form is:

$$T_{EVD1} = \frac{16}{3}M^3 \text{ flops.} \tag{4.5}$$

There are four techniques to compute the spectral decomposition of T, and include the following [138]: QR Algorithm, Cuppen's Divide and Conquer Approach, Bisectional Inverse Iteration (BII) and Multiply Relatively Robust Representations (MRRR). It has been hinted that MATLAB implements a QR Algorithmic operation that has been proven to be more accurate [139]. The cost estimate using the QR approach is given as:

$$T_{EVD2} = 2 \times 3fM^3, \tag{4.6}$$

where f refers to the number of Francis steps required for the eigenvalue to converge. If the QR algorithm is employed to perform the back transformations then:

$$\begin{aligned} T_{EVD3} &= 4 \times \frac{4}{3}M^3 \\ &= \frac{16}{3}M^3 \end{aligned} \tag{4.7}$$

Therefore, the overall cost of computing the EVD of the noise covariance matrix is:

$$\begin{aligned} T_2 &= 6fM^3 + \frac{32}{3}M^3 \quad \text{flops (complex)} \\ &= 3fM^3 + \frac{16}{3}M^3 \quad \text{flops (real).} \end{aligned} \tag{4.8}$$

4.2.2.3 Application of the EVD on received signal

The cost of the following operation has to be determined:

$$\sqrt{\mathbf{D}^{-1}}\mathbf{U}^{\dagger}\mathbf{y} = (\sqrt{\mathbf{D}^{-1}}\mathbf{U}^{\dagger}\mathbf{H}_1)\mathbf{s}_1 + \sqrt{\mathbf{D}^{-1}}(\mathbf{U}^{\dagger}\mathbf{H}_2)\mathbf{s}_2 + \sqrt{\mathbf{D}^{-1}}\mathbf{U}^{\dagger}\mathbf{z}, \tag{4.9}$$

where $\mathbf{y} \in \mathbb{C}^{M\times 1}, \mathbf{D} \in \mathbb{R}^{M\times M}, \mathbf{U} \in \mathbb{C}^{M\times M}, \mathbf{H}_1, \mathbf{H}_2 \in \mathbb{C}^{M\times L}, \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{C}^{(M/2)\times 1}, \mathbf{z} \in \mathbb{C}^{M\times 1}$. Computing $\sqrt{\mathbf{D}^{-1}}$ incurs $4M$ flops for division (inverse) and $6M$ flops for square-root, which yields a total of $24M^2$ flops. As a result, the number of flops required to transform the received signal is:

$$\begin{aligned} T_3 &= 2(2M^3 - M^2 + 3M^2L - ML - M) + 4M^3 - 2M^2 + 2ML, \\ &= 8M^3 - 4M^2 + 6M^2L - 4ML - 2M. \end{aligned} \tag{4.10}$$

The total number of flops required to apply the overall Block MMSE technique is given as:

$$\begin{aligned} \mathscr{O}_{MMSE}(n) &= T_1 + T_2 + T_3, \\ &= (2M^2L + M^2) + (6fM^3 + \frac{32}{3}M^3) + (8M^3 - 4M^2 + 6M^2L - 4ML - 2M), \\ &= 8M^3 - 3M^2 + 8M^2L - 4ML - 2M + 6fM^3 + \frac{32}{3}M^3 \quad \text{flops.} \\ &\simeq \mathscr{O}(n^3) \end{aligned} \tag{4.11}$$

Fig. 4.1 is a comparison of the complexity between the Block QR technique and Block MMSE EVD computation, where $N_T = L$ and $N_R = M$ represent the number of transmit and receive antennas, respectively. The MGS Block QR technique requires approximately 64% less flops than the Block MMSE technique in the 4×4 case, which represents a significant decrease in complexity, partly due to the expensive inversion operation required in the Block MMSE scheme.

4.3 LTE INTERFERENCE-LIMITED PERFORMANCE

The performance evaluations of the proposed Block QR decomposition technique are then validated using the OpenAirInterface software platform [120]. This platform serves as an open source hardware and software development framework, which among many features, includes the end-to-end

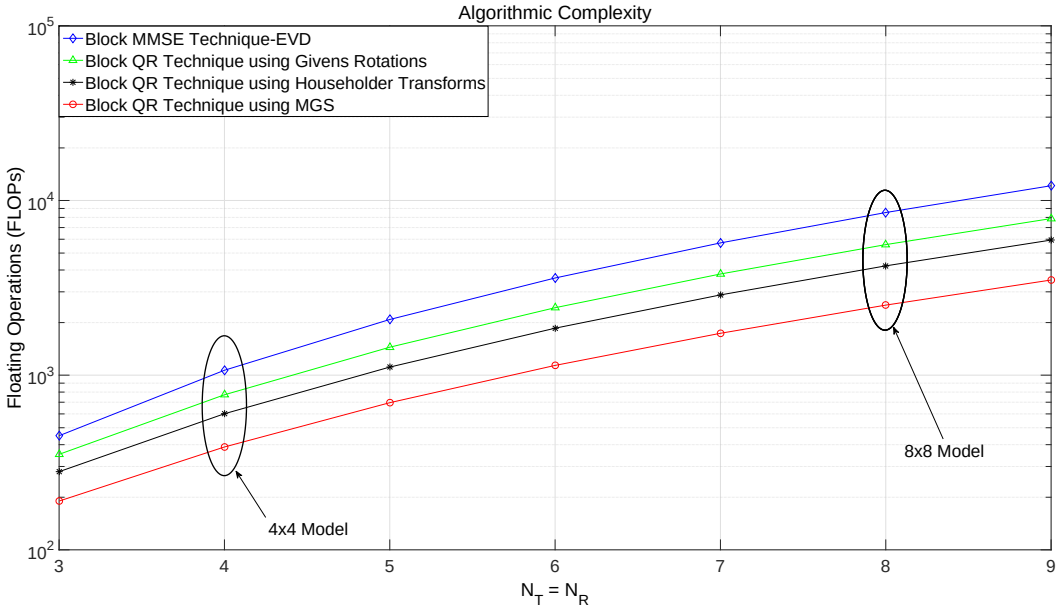


Figure 4.1: Complexity comparison among the different Block QR and Block MMSE EVD schemes

experimentation of the physical layer uplink and downlink channels in accordance with the 3GPP LTE specifications. In this particular scenario, the baseband simulations are carried out on the LTE downlink-shared channel (DL-SCH), which is essentially responsible for the transmission of user pertinent (user-plane) data (including control information) to the MAC layer. Fig. 4.2 shows the physical layer processing blocks implemented in the LTE simulation test bench. Messages in LTE are transmitted in the form of Transport Blocks, where error detection is enabled by appending 24 Cyclic Redundancy Check (CRC) bits at the end of each transport block. The Transport Block Size (TBS) depends on the Modulation and Coding Scheme (MCS), which characterises the modulation order and coding rate of a codeword and can be referenced from a lookup table depending on the number of transmitted layers [140]. In an LTE transmission time interval, a maximum of 2 codewords can only be transmitted in order to achieve the balance between receiver processing (including signaling requirements of CQI and H-ARQ) and performance. Therefore, in the case of a 4×4 MIMO system, the modulated symbols of the two codewords are mapped according to an even and odd permutation (See Fig. 2.7) to create a total of four layers, i.e. each codeword is mapped onto two layers. The implemented LTE channel codes comprise of a 1/3 rate Turbo encoder, which follows the structure of a parallel concatenated convolutional code with two 8-state constituent encoders and one turbo code

internal interleaver [141]. During channel compensation, the matched filter channel outputs are computed and channel coefficients are appropriately quantized to fixed-point values. The compensated signals are combined using maximal ratio combining (MRC) [142] and thereafter, the soft-decision outputs of the first detected stream are obtained. The optimal brute-force Max-log-MAP demodulator for 4 streams is then compared with the suboptimal reduced metric Max-log-MAP algorithms with Block QR processing. The LLRs are then demapped for decoding by the iterative turbo decoder. Table 4.2 outlines the key simulation parameters.

Table 4.2: Simulation specifications for the Block QR scheme

Parameter	Value
Transmission Mode	4
Channel Bandwidth	5 MHz
Number of Resource Blocks	50
Modulation and Coding Scheme	9 (QPSK)
Channel Coding	1/3 rate Turbo encoder
Channel Type	Rayleigh
Transport Block Size (4-layer mapping)	7992

4.3.0.4 Discussion

The Block Error Rate (BLER) performances have been evaluated for the moderate to higher SNR regimes for both the brute-force Max-log-MAP and Block QR receiver algorithms in an LTE scenario. Fig. 4.3 shows the corresponding LTE performance for QPSK (MCS=9) signals for both practical demodulators. The difference between the two practical demodulators (brute-force (red) and Block

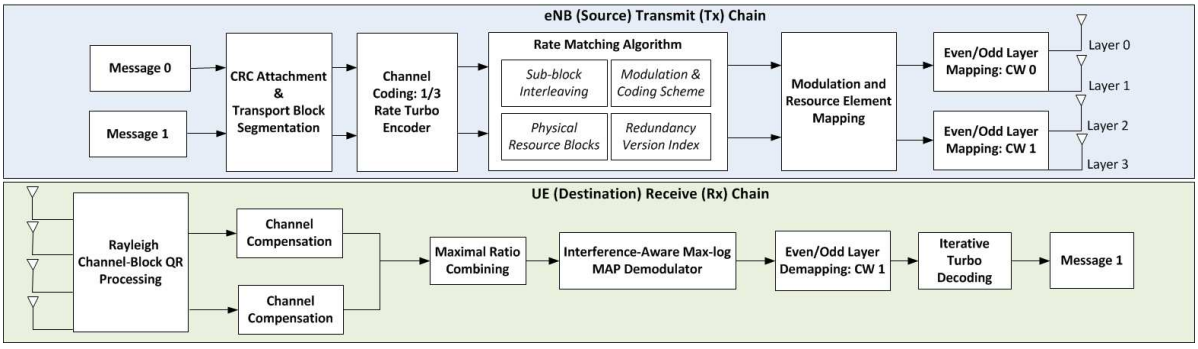


Figure 4.2: Block diagram of the simulation setup

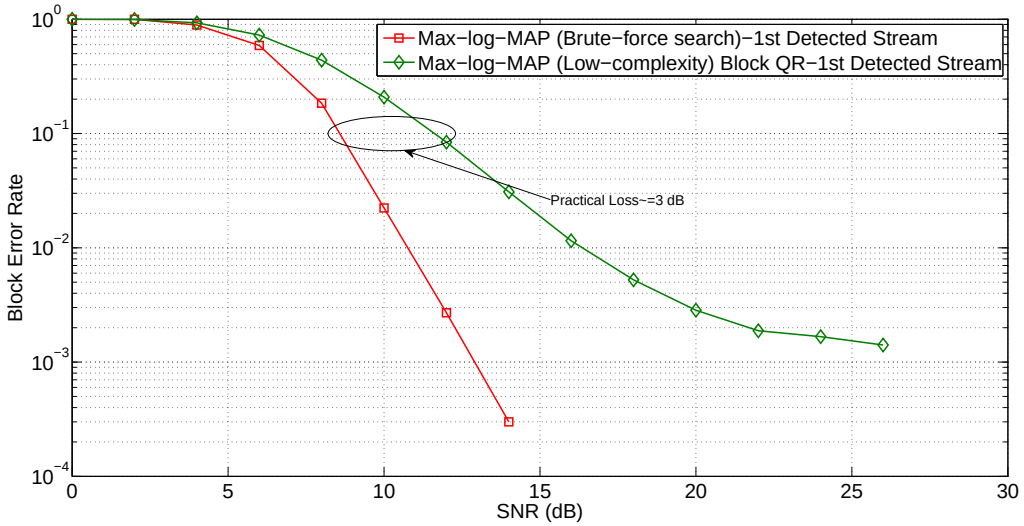


Figure 4.3: BLER performance of the low-complex Block QR Max-log-MAP and Brute-force Max-log-MAP search algorithm using LTE practical codes.

QR Max-log-MAP (green) curves) at 10^{-1} is 3 dB, respectively. It was observed that the Block QR Max-log-MAP demodulator was significantly less complex in terms of practical simulation time (measured in hours) when compared to the brute-force search Max-log-MAP algorithm by a factor of approximately 12. This supports the feasibility of such a scheme in future advanced receivers in increasingly interference-limited scenarios. The inclusion of the Block MMSE LTE performance was not yet supported by the simulation test bench and is a subject of ongoing work.

Fig. 4.4 shows the LTE performance of the Block QR detection technique across a range of different MCS values ranging from 0 to 20 at a BLER of 10^{-1} for a flat-fading channel. The required practical SNR for each of the modulation schemes, i.e. QPSK $(\text{MCS} = \{0,\ldots,9\})$, 16-QAM $(\text{MCS} = \{10,\ldots,16\})$ and 64-QAM $(\text{MCS} = \{17,\ldots,20\})$ can be distinguished. The increase in error floor as the constellation size increases can be attributed to the saturation of the LLRs and thus not falling within the required input dynamic range into the iterative turbo decoder. In order to avoid saturation, the 'output shift' parameter, which is responsible of ensuring that the LLR is scaled appropriately, should be individually calibrated.

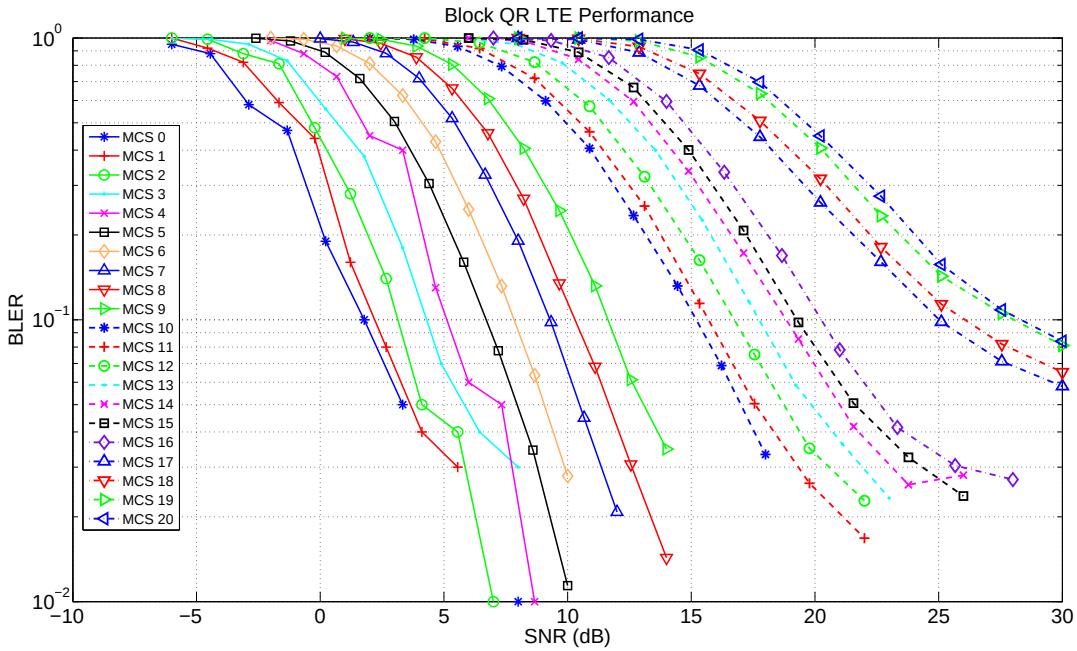


Figure 4.4: Block QR Performance with different MCS at 10% BLER

4.4 CONCLUSIONS

This technique exploits the low-complexity Max-log-MAP algorithm, to perform interference detection for higher order MIMO receivers. Furthermore, the proposed Block QR scheme is validated in a practical scenario, using a simulation test bench based on a full LTE downlink physical layer scenario with discrete constellations and finite codeword lengths. The practical results reveal the feasibility of the Block QR detection for enhancing network aided interference cancelling (NAIC) receivers.

Part II

Advanced Half-Duplex Relay Networks

CHAPTER 5

NOVEL HALF-DUPLEX RELAY STRATEGY FOR LTE NETWORKS

5.1 INTRODUCTION

The benefits of cooperative communications, as a means to enable single-antenna terminals to cooperatively operate with efficiency and diversity gains usually reserved to multi-antenna systems, have been extensively studied [143]. The different cooperative communication techniques and relay strategies available in the literature are largely based on the seminal information theoretic work by Cover and El Gamal [103]. These advances have led to technical studies of practical relay architectures by 3GPP for inclusion in the LTE Release 9 standard [104, 144], which then led to the incorporation of physical layer relaying node operations in the LTE Release 10 (LTE-Advanced) specifications [145]. In this chapter a practical LTE-based implementation [146] of the novel two-phase, three-part-message relay strategy proposed in [147] for the Gaussian Half-Duplex Relay Channel (HD-RC) is provided, which is known to be to within a constant gap of the cut-set upper bound on the capacity of the network. The two-phase, three-part-message scheme proposed in [147] employs superposition of Gaussian codebooks at the source, Successive Interference Cancellation (SIC) both at the relay and at the destination, and Decode-and-Forward (DF) at the relay. The two-phase, three-part-message scheme proposed in [147] employs superposition of Gaussian codebooks at the source, Successive Interference Cancellation (SIC) both at the relay and at the destination, and Decode-and-

Forward (DF) at the relay. The channel model has a direct link between the source and the destination, through which the source continuously sends information to the destination at a rate close to the capacity of that link. At the same time, the source leverages the relay to convey extra information to the destination at a rate that, in the high-SNR regime, can be interpreted as the minimum capacity of the source-relay (for the relay-listen phase) and relay-destination (for the relay-send phase) links minus the capacity of the source-destination link. The relative duration of the relay-listen and relay-send phases is determined so that the amount of information decoded in the former can be reliably conveyed in the latter. The channel model has a direct link between the source and the destination, through which the source continuously sends information to the destination at a rate close to the capacity of that link. At the same time the source leverages the relay to convey extra information to the destination at a rate that is, roughly speaking, the minimum capacity of the source-relay (for the relay-listen phase) and relay-destination (for the relay-send phase) links minus the capacity of the source-destination link. The relative duration of the relay-listen and relay-send phases is determined so that the amount of information decoded in the former can be reliably conveyed in the latter. In this investigation, the model of [147] is initially extended to the case of a destination equipped with multi-antennas and then a practical implementation of the scheme compliant with the LTE standard is proposed. A single- and a two-antenna destination, as well as, the Additive White Gaussian Noise (AWGN) and two LTE frequency selective channel models are considered. Half-Duplex (HD) relay schemes, which exploit physical layer cooperation, are within practical reach for near-future high-spectral efficiency Heterogeneous Network (HetNet) deployments. Furthermore, it is shown that enabling physical-layer cooperation among nodes and using SIMO technology is of critical importance in existing and future wireless HetNets.

Section 5.2 presents the current state-of-the-art in practical HD relay strategies. Section 5.3 introduces the system model and overviews the scheme proposed in [147], by adapting it to the case when the destination is equipped with multiple antennas. Section 5.4 presents the simulation test bench and discusses some additional design considerations that are crucial for a more realistic analysis. Section 5.5 and Section 5.6 evaluates the BLER performance of the proposed scheme for the static AWGN and the frequency-selective fading channels, respectively. In particular, Section 5.5 and Section 5.6 consider the single- and two-antenna destination cases and compare the achieved spectral efficiency with a baseline two-hop communication scheme that does not allow for direct link source-destination transmission. Section 5.5 also compares the achieved spectral efficiency with the theoretical case.

Finally, Section 5.7 concludes the Chapter. Proofs may be found in Appendix B.

5.2 PRACTICAL HALF-DUPLEX RELAYS: A STATE-OF-THE-ART

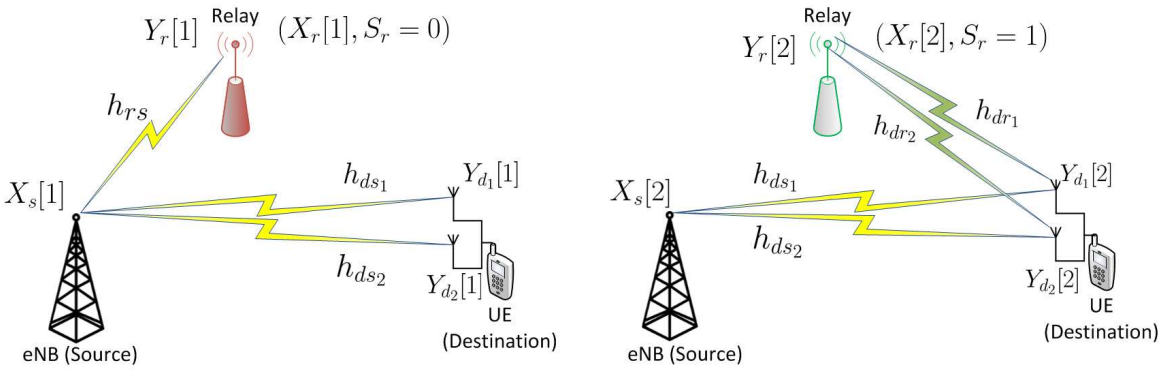
Wireless relay networks have long been conceptualized to cooperatively form virtual antenna arrays to improve spatial diversity and serve as a medium for multi-hop communications. The implementation of the types of relay strategies in 3G and 4G cellular standards have been classified according to the Layers in the stack at which user-plane data is forwarded at the relay, i.e., either Layer 1, 2 or 3 relaying. For example, Layer 1 (physical layer) relaying is limited to the amplification of the desired signal, which is forwarded to the mobile terminal with power control supported by control signaling features; Layer 2 consists of a DF relay, which decodes and re-encodes the user data to be forwarded on Layer 2, in the absence of direct-link transmission; Layer 3 relays function in a similar fashion to Layer 2 relays, but can operate as a self-backhauling base station, which forwards the IP data packets on Layer 3 [148–150].

However, the implementation of relay strategies in 3G and 4G cellular standards have been limited to layers 2 and 3 with only rudimentary relay functions in the physical layer, e.g., as a wireless repeater with Amplify-and-Forward (AF) capabilities [148, 149]. It is well known that AF relay networks are prone to noise enhancements at the relay side. Low-powered relay nodes can aid densification of existing networks in high traffic areas and are foreseen to improve the overall network performance in many aspects, such as throughput and coverage. Their employment in HetNets is crucial and this aspect has already been standardised in LTE-Advanced [145]. In the context of the proposed HD relay strategy, various practical studies have examined the implementation and performance of relay networks with respect to LTE frameworks as well as particular channel codes. In [151], work was carried out to analyze the coverage performance of a DF relay algorithm in a typical LTE-Advanced system, while in [152], a spectral efficiency analytical comparison was made between AF and DF modes of operation at the relay station. These studies were limited to an analytical study aimed to prove the benefits in terms of coverage and throughput of existing relay types in an LTE framework, but lacked the end-to-end performance analysis of a full physical layer LTE scenario, which can be used as a reference for future practical deployments of relay networks. In contrast, this study seeks to evaluate the spectral efficiency performance of a two-phase, three-part-message relay strategy on a physical layer simulation test bench by using practical channel codes as specified in the LTE standard.

Various relay schemes using channel codes have been extensively studied under a generalized setting. Recently, in [153], the authors showed that the capacity of the line network with one HD relay (i.e., without a direct link between the source and the destination) is achieved by a discrete input, which stresses the importance of random switching (between transmission and reception phases) at the relay. The proposed HD strategy described in this Chapter uses a transmission scheme that, although it does not achieve the capacity, can be implemented within the LTE standard and is is proven to be optimal to within a constant gap, in addition to having an easy closed-form expression. In [154], a distributed turbo coding scheme was designed for a DF relay channel; this study showed diversity and coding gains, especially in low-SNR regimes. Similarly, the work proposed in [155], used Dynamic DF (DDF) and turbo codes to increase the coding gain at the relay as well as achieve full diversity at the destination in a HD relay channel. In [156], the authors investigated the performance of fountain codes in a basic cooperative relay scheme over a block fading channel. Comparisons were made between two protocols, i.e., distributed space-time coded and time-division, revealing a performance gain of the latter in terms of outage. In [157], a Reed-Solomon convolutional coded system was designed to validate a coded opportunistic AF and DF relay scheme for a quasi-static channel. The results showed that the coded opportunistic DF scheme outperforms the AF scheme due to the capability of decoding/re-encoding at the relay, hence eliminating any noise enhancement to which the AF scheme is susceptible. In [158], a physical layer network coding scheme for a two-message HD relay scenario was presented, which was limited to an uncoded BPSK scenario with hard-decision decoding. Contrary to the works in [154–157] where no rate splitting is performed over the two time slots (phases), the proposed relay strategy splits the message into three parts, superposes them for transmission and performs SIC with perfect channel estimation. Moreover, although these studies considered the use of practical codes to show coding and diversity gains of relay communications, the use of multiple receive antennas at the destination (which is supported by LTE), adaptive link transmission techniques (the work in [155]) examines a similar adaption in terms of channel quality at the relay in a different context, and the study over frequency selective channels, are not taken into account. In contrast, the investigated study aims to take a practical step toward implementing relays with an overall improvement in capacity. Few studies have tackled the performance of relay strategies with respect to current LTE standards. In an LTE-Advanced HD relay study [111], the authors proposed the integration of a Quantize-Map-Forward (QMF) - a network generalization of the

classical Compress-and-Forward (CF) relay strategy-for a two-relay diamond network[1] These results paved the way for the implementation of diamond relay networks in the LTE standard. However, this study did not perform rate splitting and was limited to an AWGN channel with a single-antenna destination. Differently, in this work a theoretically optimal (up to a constant gap) scheme which splits the message into three parts and also exploits direct transmission (from the source to the destination) is considered and its performance on an AWGN channel as well as on two LTE frequency channel models with a 2-antenna destination assessed. Another example is the work performed in [159], where the Bit Error Rate (BER) performances of various memoryless and full-memory relay schemes were compared in a convolutional and LTE Turbo coded system. Contrary to [159], where the direct source-destination link is absent (i.e., no physical cooperation between the source and the relay), the proposed HD relay scheme exploits physical cooperation between the source and the relay as a means to enhance the system performance. A two-layered demodulation strategy was proposed in [160] to improve the DF BLock Error Rate (BLER) performance of an asymmetric relay channel link using a soft-combining of the Log-Likelihood Ratios (LLRs) during the broadcast and relaying phases. The proposed two-phase, three-part-message scheme considers the source transmitting in both phases, which differs from [160], where the source stays silent in the second phase. The authors in [161] used a software-defined radio framework, based on the open source GNU radio platform, to implement a scalar two-user Gaussian broadcast channel; their main merit was to show rate improvements from the use of SIC with respect to Time Division Multiplexing (TDM). The aforementioned scheme artificially introduces interference among user signals on the transmitter side, which are then decoded using SIC; the corresponding performance shows the benefits of superposition coding in terms of spectral efficiency. Similarly, the benefits of single- and multi-relay physical layer cooperation have been demonstrated with the GNU radio testbed [162]. The Wireless Open-Access Radio (WARP) testbed has also been used to validate the benefits of physical layer cooperation [163]. In [163], the authors compared the performance of the AF, DF and QMF relay schemes in an indoor environment using the IEEE 802.11 (WiFi) protocol. This particular empirical study exploits the single message HD relay network as a distributed 2×1 Multiple-Input-Single-Output (MISO) system to enhance the throughput. The work in [164] proposes a hybrid decoding approach, which represents a generalization of [163], for cooperative physical layer relaying involving the opportunistic use of QMF and DF

[1] An N-relay diamond network is a relay network topology where the source can communicate with the destination only through N non-interfering relays.



(a) Phase I- $S_r = 0$ (b) Phase II- $S_r = 1$

Figure 5.1: Two-phase relay system model.

relaying, which adapts according to the underlying network configuration. In the proposed HD relay strategy, the message is divided into three parts and exploit the benefits of superposition encoding and SIC decoding to enhance spectral efficiency in a single user HD LTE relay scenario, which was not considered in [161–164].

5.3 HD RELAY SYSTEM MODEL

A Half-Duplex Relay Channel (HD-RC) consists of three nodes: the source, the relay, and the destination. The source has a uniformly distributed message $w \in \left[1 : 2^{NR}\right]$ for the destination where N denotes the codeword length and R the transmission rate. At time $i, i \in [1 : N]$, the source maps its message w into a channel input symbol $X_{s,i}(w)$ and the relay, if in transmission mode of operation, maps its past channel observations into a channel input symbol $X_{r,i}(Y_r^{i-1})$. At time N, the destination outputs an estimate of the message w based on all its channel observations Y_d^N as $\hat{w}(Y_d^N)$. A rate R is said to be ε-achievable if, for some block length N, there exists a code such that $\mathbb{P}[\hat{w} \neq w] \leq \varepsilon$ for any $\varepsilon > 0$. The capacity C is the largest nonnegative rate that is ε-achievable for $\varepsilon \in (0,1)$.

The SIMO Gaussian HD-RC is shown in Fig. 5.1, where the three nodes are the eNodeB (source), the relay, and the UE (destination), which is equipped with $n_d = 2$ antennas. For the static case, the input/output relationship is

$$Y_r = h_{rs} X_s \, (1 - S_r) + Z_r \in \mathbb{C}, \tag{5.1a}$$

$$\mathbf{Y}_d = \mathbf{h}_r X_r S_r + \mathbf{h}_s X_s + \mathbf{z}_d \in \mathbb{C}^2, \tag{5.1b}$$

and let $\mathbf{Y}_d = \begin{bmatrix} Y_{d_1} & Y_{d_2} \end{bmatrix}^T$, $\mathbf{h}_r = \begin{bmatrix} h_{dr1} & h_{dr2} \end{bmatrix}^T$, $\mathbf{h}_s = \begin{bmatrix} h_{ds1} & h_{ds2} \end{bmatrix}^T$ and $\mathbf{z}_d = \begin{bmatrix} Z_{d_1} & Z_{d_2} \end{bmatrix}^T$. The channel parameters $(h_{ds_i}, h_{dr_i}, h_{rs}), i \in [1:2]$ are fixed for the whole transmission duration (i.e., time needed to transmit one codeword) and assumed known to all nodes (i.e., full Channel State Information (CSI) - for the fading case, the CSI assumptions will be discussed in Section 5.6), the inputs are subject to unitary power constraints, S_r is the switch random binary variable which indicates the state of the relay, i.e., when $S_r = 0$ the relay is receiving while when $S_r = 1$ the relay is transmitting, and the noises form independent proper-complex white Gaussian noise processes with zero-mean and unit-variance. The model is without loss of generality because non-unitary power constraints or noise variances can be incorporated into the channel gains. It is also worth noting that the scheme designed in this section as well as its derived performance guarantee hold for any value of $n_d \geq 1$. However, the focus will be on $n_d = 2$ since this is the case considered in the practical implementation. For the SIMO Gaussian HD-RC, the following is a generalization of [147, Proposition 5].

Proposition 1. *For the static/non-fading SIMO Gaussian HD-RC the following rate is achievable*

$$R = \log\left(1 + \|\mathbf{h}_s\|^2\right) + \frac{\underline{a}\,[\underline{b}]^+}{\underline{a} + [\underline{b}]^+}, \tag{5.2a}$$

$$\underline{a} := \log\left(1 + \frac{\|\mathbf{h}_r\|^2 + \|\mathbf{h}_r\|^2\|\mathbf{h}_s\|^2\left(1 - |v|^2\right)}{1 + \|\mathbf{h}_s\|^2}\right),\ v := \frac{\mathbf{h}_s^H \mathbf{h}_r}{\|\mathbf{h}_s\|\|\mathbf{h}_r\|}, \tag{5.2b}$$

$$\underline{b} := \log\left(1 + \frac{|h_{rs}|^2}{1 + \|\mathbf{h}_s\|^2}\right) - \log\left(1 + \frac{\|\mathbf{h}_s\|^2}{1 + \|\mathbf{h}_s\|^2}\right). \tag{5.2c}$$

Moreover, R in Eq. (5.2a) *is to within* 3.51 *bits/dim from the cut-set upper bound to the capacity, irrespectively of the number of antennas at the destination and of the channel gains.*

Proof. Subsequently, a sketch of the proof of Proposition 1 is given. The complete proof can be derived by obvious modifications from the proof of [147, Proposition 5].

Codebooks A scheme with the following four unit average power Gaussian codebooks to transmit a message that is rate-split into three parts is studied:

$$\mathscr{C}_{a1} = \{X_{a1}^{N_1}(w_0) : w_0 \in [1 : M_0]\}, \quad \mathscr{C}_{a2} = \{X_{a2}^{N_2}(w_0) : w_0 \in [1 : M_0]\},$$
$$\mathscr{C}_b = \{X_b^{N_1}(w_1) : w_1 \in [1 : M_1]\}, \quad \mathscr{C}_c = \{X_c^{N_2}(w_2) : w_2 \in [1 : M_2]\},$$

by which the following rate is achieved:

$$R = R_a + R_b + R_c = \frac{\log(M_0)}{N_1+N_2} + \frac{\log(M_1)}{N_1+N_2} + \frac{\log(M_2)}{N_1+N_2} = \frac{\log(M_0 M_1 M_2)}{N_1+N_2} \text{ [bits/dim]}.$$

The transmission is divided into two phases: the first phase (i.e., relay receiving) lasts N_1 channel uses, and the second phase (i.e., relay transmitting) lasts N_2 channel uses. The overall transmission duration for a codeword is $N = N_1 + N_2$ channel uses. In the following, in order to simplify the notation, the length N_1 and N_2 in the superscript of the codewords are omitted.

Phase I During this phase the relay is listening, i.e., $S_r = 0$ (see Fig. 5.1(a)). The source selects uniformly at random $w_0 \in [1 : M_0]$ (sent cooperatively with the relay to the destination) and $w_1 \in [1 : M_1]$ (sent directly to the destination). The transmitted signals are

$$X_s[1] = \sqrt{1-\delta} X_b(w_1) + \sqrt{\delta} X_{a1}(w_0), \tag{5.3a}$$
$$X_r[1] = 0, \tag{5.3b}$$
$$\delta = \frac{1}{1+\|\mathbf{h}_s\|^2}, \tag{5.3c}$$

where $\delta \in [0,1]$ is the scaling parameter that allows for superposition coding. It is chosen in such a way that the message that is treated as noise at the destination is received at most at the power of the noise as in [147].

The relay applies successive decoding of $X_b(w_1)$ followed by $X_{a1}(w_0)$ from

$$Y_r[1] = h_{rs}\sqrt{1-\delta} X_b(w_1) + h_{rs}\sqrt{\delta} X_{a1}(w_0) + Z_r[1],$$

which is possible with arbitrary high reliability for sufficiently large N if

$$R_b \leq \gamma \log\left(1+|h_{rs}|^2\right) - \gamma \log\left(1+\frac{|h_{rs}|^2}{1+\|\mathbf{h}_s\|^2}\right),$$
$$R_a \leq \gamma \log\left(1+\frac{|h_{rs}|^2}{1+\|\mathbf{h}_s\|^2}\right), \tag{5.4}$$

where $\gamma = \frac{N_1}{N_1+N_2}$. The destination, by using Maximal-Ratio Combining (MRC), decodes $X_b(w_1)$ by treating $X_{a1}(w_0)$ as noise from

$$\mathbf{Y}_d[1] = \mathbf{h}_s\sqrt{1-\delta}X_b(w_1) + \mathbf{h}_s\sqrt{\delta}X_{a1}(w_0) + \mathbf{Z}_d[1],$$

which is possible with arbitrary high reliability for sufficiently large N if

$$R_b \leq \gamma \log\left(1+\|\mathbf{h}_s\|^2\right) - \gamma \log\left(1+\frac{\|\mathbf{h}_s\|^2}{1+\|\mathbf{h}_s\|^2}\right). \tag{5.5}$$

The expression, $\log\left(1+(1-\delta)\mathbf{h}_s^\dagger \Sigma_1^{-1}\mathbf{h}_s\right)$, was computed, in order to obtain R_b in Eq. (5.5) where $\Sigma_1 \in \mathbb{C}^{2\times 2}$ is the covariance matrix of the equivalent noise

$$\tilde{\mathbf{Z}}_1 = \mathbf{h}_s\sqrt{\delta}X_{a1}(w_0) + \mathbf{Z}_d[1],$$

and where the final expression follows as an application of the matrix inversion lemma[2]. Notice that the same rate is obtained if the receiver computes the scalar $\frac{\mathbf{h}_s^\dagger}{\|\mathbf{h}_s\|}\mathbf{Y}_d[1] = \|\mathbf{h}_s\|\sqrt{1-\delta}X_b(w_1) + \|\mathbf{h}_s\|\sqrt{\delta}X_{a1}(w_0) + Z$, $Z \sim \mathcal{N}(0,1)$ and decodes X_b. Finally, by assuming $\|\mathbf{h}_s\|^2 < |h_{rs}|^2$ (it is shown later that this assumption is without loss of generality), Phase I is successful if Eq. (5.4) and Eq. (5.5) are satisfied.

Phase II During this phase the relay is transmitting, i.e., $S_r = 1$ (see Fig. 5.1(b)). The source selects uniformly at random $w_2 \in [1 : M_2]$ (sent directly to the destination) and the relay forwards its estimation of w_0 from Phase I, indicated as $\widehat{w}_0$. The transmitted signals are

$$X_s[2] = X_c(w_2),$$
$$X_r[2] = X_{a2}(\widehat{w}_0).$$

[2]Matrix inversion lemma $(A+XBX^T)^{-1} = A^{-1} - A^{-1}X(B^{-1}+X^TA^{-1}X)^{-1}X^TA^{-1}$.

The destination uses MRC and applies successive decoding based on

$$\mathbf{Y}_d[2] = \mathbf{h}_s X_c(w_2) + \mathbf{h}_r X_{a2}(\widehat{w}_0) + \mathbf{Z}_d[2].$$

In particular, it first decodes w_0 by using the received signal from both phases and by assuming that $\widehat{w}_0 = w_0$. This is true given the rate constraints found for Phase I. It then decodes $X_c(w_2)$, after having subtracted the contribution of its estimated w_0. Successful decoding is possible with arbitrary high reliability for sufficiently large N if

$$R_a \leq (1-\gamma)\log\left(1+\frac{\|\mathbf{h}_r\|^2+\|\mathbf{h}_r\|^2\|\mathbf{h}_s\|^2\left(1-|v|^2\right)}{1+\|\mathbf{h}_s\|^2}\right)+\gamma\log\left(1+\frac{\|\mathbf{h}_s\|^2}{1+\|\mathbf{h}_s\|^2}\right), \tag{5.6}$$

$$R_c \leq (1-\gamma)\log\left(1+\|\mathbf{h}_s\|^2\right), \tag{5.7}$$

where v is defined in Eq. (5.2b). The expression, $\log\left(1+\mathbf{h}_r^{\dagger}\Sigma_2^{-1}\mathbf{h}_r\right)$, was also computed in order to obtain R_a in Eq. (5.6), where $\Sigma_2 \in \mathbb{C}^{2\times 2}$ is the covariance matrix of

$$\tilde{\mathbf{Z}}_2 = \mathbf{h}_s X_c(w_2) + \mathbf{Z}_d[2],$$

and where the final expression follows as an application of the matrix inversion lemma. By imposing that the rate R_a is the same in both phases, i.e., that Eq. (5.4) and Eq. (5.6) are equal, it has been observed that γ should be chosen equal to γ^*, which is defined as follows:

$$\gamma^* = \frac{\underline{a}}{\underline{a}+\underline{b}}, \tag{5.8}$$

where $\underline{a}$ and $\underline{b}$ are defined in Eq. (5.2b) and Eq. (5.2c), respectively. Note that because of the assumption $\|\mathbf{h}_s\|^2 < |h_{rs}|^2$, it has been shown that $\underline{b} > 0$, i.e., $\underline{b} = [\underline{b}]^+$.

The rate sent directly from the source to the destination, that is, the sum of Eq. (5.5) and Eq. (5.7), is

$$R_b + R_c = \log\left(1+\|\mathbf{h}_s\|^2\right) - \underbrace{\gamma^*\log\left(1+\frac{\|\mathbf{h}_s\|^2}{1+\|\mathbf{h}_s\|^2}\right)}_{\in[0,1]}. \tag{5.9}$$

Therefore the total rate after the two phases of transmission is $R = R_b + R_c + R_a$ as in Eq. (5.2a),

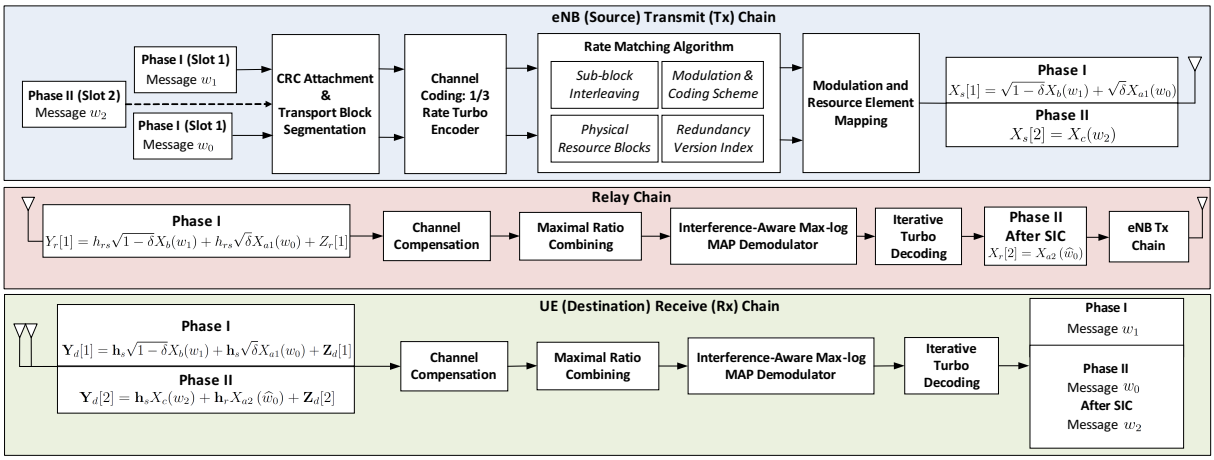


Figure 5.2: Overall simulation block diagram for the transmit and receive chains.

which implies that

$$\mathsf{C} \geq R_b + R_c + R_a \geq \log(1 + \|\mathbf{h}_s\|^2) + \frac{\underline{a}\,\underline{b}}{\underline{a}+\underline{b}}. \tag{5.10}$$

The rate expression for R in Eq. (5.2a), with $[b]^+$ rather than b, holds since for $\|\mathbf{h}_s\|^2 \geq |h_{rs}|^2$ it reduces to a direct transmission from the source to the destination. The proposed scheme is optimal to within 3.51 bits/dim as shown in Appendix B, irrespectively of the number of antennas n_d at the destination and of the channel gains. Note that the single-antenna result in [147, Proposition 5] is obtained as a special case of Proposition 1 by setting $v = 1$. □

Next a practical LTE-based implementation to achieve the rate in Proposition 1 is proposed.

5.4 SIMULATION TEST BENCH

The scheme described in Section 5.3 uses four Gaussian codebooks $\mathscr{C}_{a1}$, $\mathscr{C}_{a2}$, $\mathscr{C}_b$ and $\mathscr{C}_c$ to transmit the three sub-messages w_0, w_1 and w_2. In a practical implementation, the codes would not be Gaussian, but will contain symbols from a finite constellation. The performance of these practical codes both in Gaussian and non-Gaussian noise has to be understood, since SIC is employed in the decoding operations both at the relay and at the destination, In particular, the performance of $\mathscr{C}_b$ in non-Gaussian noise (first decoding step of Phase I) and of $\mathscr{C}_{a1}$ in Gaussian noise (second decoding

step of Phase I) at the relay has to be quantified; in addition to the performance of $\mathscr{C}_{a1}$, $\mathscr{C}_{a2}$ and $\mathscr{C}_b$ in non-Gaussian noise ($\mathscr{C}_b$ in the first decoding step in Phase I and $\mathscr{C}_{a1}$, $\mathscr{C}_{a2}$ in the first decoding step in Phase II) and the performance of code $\mathscr{C}_c$ in Gaussian noise (second decoding operation in Phase II with no error propagation) at the destination. In the decoding stages where a message is treated as noise, a demodulator has been implemented, which specially accounts for the fact that the overall noise is non-Gaussian. Different choices for the codebooks $(\mathscr{C}_{a1}, \mathscr{C}_{a2}, \mathscr{C}_b, \mathscr{C}_c)$ are considered. As with each choice, it has been ensured that a BLER at the destination is below a given threshold for all the decoding stages (in order to maintain the overall system BLER at the destination below 10^{-2} and hence the BLER corresponding to each of the three decoding operations is set to be at most $1/3 \cdot 10^{-2}$). Although the analysis is focused on the case when the destination is equipped with two antennas, the single-antenna case is also presented for completeness. The key question at hand, is the following: how does the spectral efficiency of practical codes compare to the theoretical performance in Proposition 1? A simulation testbed using the OAI (a platform for wireless communication experimentation) software libraries in order to evaluate the performance of the aforementioned scheme with practical codes (see Fig. 5.2) is developed. The software platform is based on 3GPP's evolving standard of LTE which consists of the essential features of a practical radio communication system, which closely align with the standards in commercially deployed networks. Fig. 5.2 shows the key functional units of the simulation design. The simulations were carried out on the Downlink Shared Channel (DL-SCH), which is the primary channel for transmitting user-data (or control information) from the eNodeB (eNB) to the UE [31]. The data messages are transported in units known as Transport Blocks (TBs) to convey the sub-messages w_0, w_1 and w_2. The TB Size (TBS) depends on the choice of the Modulation and Coding Scheme (MCS), which describes the modulation order and the coding rate of a particular transmission.

Processing: The TBs undergo a series of processing stages prior to modulation before the codeword can be mapped into the Resource Elements (REs) in the Physical DL-SCH (PDSCH). Error detection at the receiver is enabled by appending 24 Cyclic Redundancy Check (CRC) bits to the TB. The code block (comprising the TB and the CRC bits) has a minimum and a maximum size of 40 and 6144 bits, respectively, as required by the Turbo encoder. Filler bits are added if the code block is too small in size and the code block is segmented into blocks of smaller size if the maximum size is exceeded. The subsequent bit sequence is then fed into the $1/3$ rate Turbo encoder.

Channel Coding: The channel coding scheme comprises of a 1/3 rate Turbo encoder, which follows the structure of a parallel concatenated convolutional code with two 8-state constituent encoders, and one Turbo code internal interleaver [141]. A single set of systematic bits and two sets of parity bits are produced at the output of the encoder as detailed in [141], which follows the structure of a parallel concatenated convolutional code with two 8-state constituent encoders and one turbo code internal interleaver:

$$g_0(D) = 1 + D^2 + D^3$$
$$g_1(D) = 1 + D + D^3,$$

where D represents the number of encoded bits per stream as defined in [141].

Rate Matching The rate matching component ensures, through puncturing or repetition of the bits, that the output bits of the Turbo encoder match the available physical resources using the MCS, the Redundancy Version (RV) index and the Physical Resource Blocks (PRBs). For the numerical evaluations, an equal bandwidth allocation is chosen between the two phases of the relay strategy, i.e., the number of PRBs allocated in the first (relay listening) and second (relay transmitting) phases is the same. In other words, with reference to Eq. (5.8), $\gamma = 0.5$ has been set, which may not be the optimal choice. The message w_0, transmitted by the source and the relay over the two phases, corresponds to two different RV indices with equal resource allocations. The selection is made possible through puncturing or repetition of the bits at the output of the encoder. The Circular Buffer (CB) generates puncturing patterns depending on the allocated resources, and the sub-block interleaver (which forms part of the CB) facilitates the puncturing of the three outputs of the encoder [31]. Furthermore, the code block is concatenated if segmentation was required prior to channel coding.

Modulation and REs Mapping: During this stage, complex-valued symbols are generated according to the chosen modulation scheme, i.e., QPSK, 16-QAM or 64-QAM, which are supported in LTE. In this study, the performance of QPSK as well as higher-order modulations such as 16-QAM and 64-QAM is investigated. In particular, the same modulation order at the source and at the relay is employed. Such a scheme performs well when the channel quality between the relay and the destination is not much better than that of the source-destination link. However, when the relay-destination link

is significantly stronger than the source-destination link, a better performance or higher spectral efficiency could be attained with higher modulation schemes at the relay. The complex-valued symbols are then mapped to the antenna ports for transmission on the PDSCH.

Channel Compensation and MRC: The channel compensation block is responsible for computing the Matched Filtered (MF) outputs and the effective channel magnitudes of the received signal. These parameters are required for the soft-decoding of the desired message using the interference-aware demodulator. The MRC block utilizes the MF outputs to constructively add the two received signals to maximize the post-processing SNR (notice that the MRC block is not needed in the case when the destination is equipped with a single-antenna).

Interference-Aware Demodulator: The demodulator comprises of a discrete constellation interference-aware receiver designed to be a low-complexity version of the max-log MAP detector. The main idea is to decouple the real and imaginary components through a simplified bit-metric using the MF output and thus reduce the search space by one complex dimension [91]. As a result, it is possible to decode the required codeword in the presence of an interfering codeword of the same (or different) modulation scheme. Thereafter, it is possible to remove the decoded signal from the received signal and then decode the remaining signal in an interference-free channel in the case of no error propagation. The sub-message w_0 is decoded at the destination at the end of Phase II, i.e., X_{a1} (received in Phase I) and X_{a2} (received in Phase II) are combined to obtain the sub-message w_0.

Further Design Considerations for towards a Realistic Analysis: From a purely practical standpoint, two main issues related to low-layer signaling need to be addressed, in order for the proposed HD relay scheme to be viable for future (e.g., 5G) standardization within the LTE framework. The first pertains to timing advance of the relay transmission. In a conventional OFDM system, the receiver synchronizes its reception to the start of OFDM symbols and this synchronization must be accurate enough so that the entire channel response falls within the duration of the cyclic-prefix. Otherwise, inter-symbol interference will significantly degrade the performance of the receiver. In the context of the physical-layer relaying scheme, the received signal in the second phase of the protocol is jointly transmitted from the eNB and the relay station. These two signal components must arrive more or

Table 5.1: MCS mapping for each decoding operation with $I/S = 0$ dB for a SIMO scheme.

Phase I - X_b					**Phase II** - (X_{a1}, X_{a2})		**Phase II** - X_c		**Theoretical**	**Practical**	**Theoretical BS**	**Practical BS**
MCS	S [dB]	C_{X_b} [dB]	$C_{X_{a1}}$ [dB]	TBS [bits]	MCS	TBS [bits]	MCS	TBS [bits]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]
9	1.89	5.48	9.54	4008	11	4968	13	5736	2.44	2.26	1.01	1.12
14	5.84	9.07	12.10	6456	14	6456	17	7736	3.41	3.36	1.56	1.46
20	11.02	14.60	18.2	9912	20	9912	22	11448	5.10	5.09	2.36	1.93

less synchronized in order for the receiver to be able to receive the OFDM waveform without much difficulty. This clearly requires some form of timing adjustment since the path from the source to the destination going through the relay is longer than the direct path. The relay will adjust its timing to the received signal from the eNB (Phase I) and must advance the signal for each UE that it is serving according to the difference in path delay between the direct path and the path going through it to the destination. This will clearly require additional signaling. The second main issue is related to signaling in support of H-ARQ. In the first phase of the protocol, separate ACK/NAK signals need to be conveyed to the source from the relay (for two messages) and from the destination (for one message). Similarly, in the second phase, the destination must send two ACK/NAK signals to the source and one to the relay. This will clearly require significant modifications in low-layer signaling, but a feasible solution is well within reach. These design considerations represent an interesting research direction, which is currently under investigation.

5.5 PERFORMANCE EVALUATION: STATIC AWGN CHANNEL

With reference to the system model in Eq. (5.1), let $|h_{rs}|^2 = C$, $|h_{ds_1}|^2 = |h_{ds_2}|^2 = S$ (i.e., the two source-destination links are of the same strength), $|h_{dr_1}|^2 = |h_{dr_2}|^2 = I$ (i.e., the two relay-destination links are of the same strength), and the phase of the channel gains is set to some random value that is kept constant during the whole transmission duration (i.e., $N_1 + N_2$ channel uses). Perfect receive CSI is assumed at all nodes. For each of the decoding operations during Phases I and II, the BLER performances at the relay and destination are validated for different values of C (at the relay) and evaluated with respect to S (channel quality for the source-destination link). Furthermore, $\delta = \frac{1}{1+2S}$ has been set, as the superposition parameter in the SIMO scenario. An equal bandwidth allocation between the two phases of the strategy is also assumed, i.e., $\gamma = 0.5$.

Decoding at the relay At the end of Phase I, the relay first decodes w_1, then it strips it out from its received signal and finally decodes w_0, in an interference-free link. Error propagation, which results from feeding back incorrectly decoded symbols, is not considered here. The reason is that a coded system with CRC bits (that is implemented in the 4G receivers of today) for which the residual error (misdetection) at the output occurs with a very low probability (e.g., 10^{-9}) is considered. The source-relay channel is assumed to be strong enough to guarantee a system BLER that is less than 10^{-3} at the relay (e.g., see the third and fourth columns of Table A.3) (A complete list of all the Tables for the different I/S ratios and antenna configurations can be found in Appendix A). In other words, in order to have successful decoding operations (of both w_0 and w_1) at the relay at the end of Phase I, a value of C such that $\text{BLER} \leq \text{BLER}[X_b] + \text{BLER}[X_{a1}] \leq 5 \cdot 10^{-4} + 5 \cdot 10^{-4} = 10^{-3}$ is assumed. It is shown (e.g., see the third and fourth columns of Table A.3) that the source-relay link is stronger than the source-destination link because of the LOS proximity between the source and relay. Under this assumption, it is worth noting that the relay successfully decodes the message w_0 with $\text{BLER}[X_{a1}] \leq 5 \cdot 10^{-4}$, which is a more stringent constraint than the one imposed at the destination of $\text{BLER}[X_{a1}] \leq 1/3 \cdot 10^{-2}$. Finally, at the beginning of Phase II, the sub-message w_0 is re-encoded and forwarded by the relay to the destination.

Decoding at the destination The SISO and SIMO BLER performance of w_1 for Phase I, are shown in Fig. 5.3(a) and Fig. 5.6(a), respectively, for MCS values ranging in the interval $[0:20]$ to encompass QPSK, 16-QAM and 64-QAM transmissions with different code rates. Similarly for Phase II, the BLER performances of w_0 are shown in Fig. 5.3(b) (SIMO) and Fig. 5.6(b) (SISO) for $I/S = 0$ dB and in Fig. 5.3(c) (SIMO) and Fig. 5.6(c) (SISO), for $I/S = 5$ dB. Finally, Fig. 5.3(d) (SIMO) and Fig. 5.6(d) (SISO) show the BLER performances of the third transmitted message w_2, at the end of Phase II, which is decoded after the perfect stripping of w_0 once it has been successfully decoded. It is worth noting that in the proposed model, the relay-destination link is assumed to be stronger than the source-destination link so that using the relay indeed boosts the rate performance with respect to direct transmission.

The overall SIMO (two antennas at the destination) spectral efficiency of the HD relay strategy at the end of Phase II is shown in Fig. 5.4 and Fig. 5.5 (see the blue curves). Each of the markers on the blue curves - practical rates achieved by the scheme described in Section 5.3 - represents an MCS, which in total ranges from 0 to 20. Each MCS characterizes the type of signal transmission (i.e., QPSK,

16-QAM or 64-QAM) with a certain code rate. The rates are plotted for two different ratios of the relay-destination link and source-destination link, namely $I/S = 0$ dB (see Fig. 5.4) and $I/S = 5$ dB (see Fig. 5.5). This is shown to demonstrate that utilising a relay-destination link which is stronger than the source-destination link indeed boosts the rate performance with respect to direct transmission (see green curves) or conventional relay schemes (see black curves) - the baseline strategy, which will be later elaborated upon on this Section. Similarly, Fig. 5.7 and Fig. 5.8 shows the practical rate performance of the SISO case, i.e., when the destination is equipped with a single antenna.

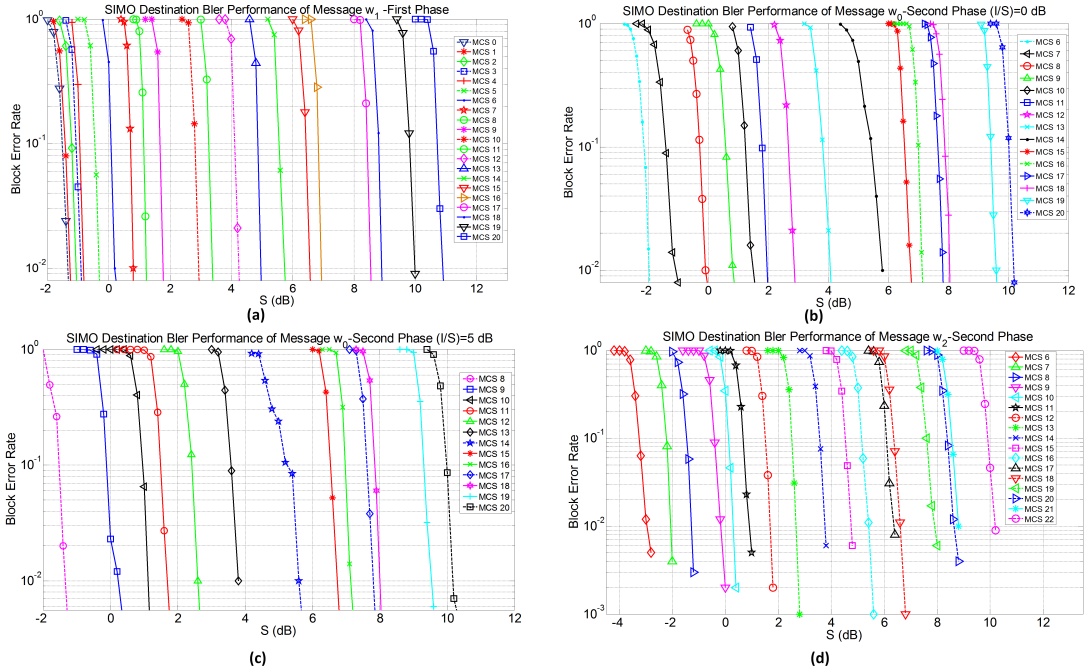


Figure 5.3: AWGN SIMO BLER performances of w_0, w_1 and w_2 at the destination versus different strengths of the direct source-destination link

Table A.3 displays the link adaptation of the proposed strategy for the SIMO case when $I/S = 0$ dB with three different MCS values at the end of each phase; this illustrates the overall selection strategy in order to compute the practical spectral efficiencies[3]. An overall system decoding probability of 10^{-2} at the destination, i.e., BLER $\leq$ BLER$[X_b]$ + BLER$[X_{a1}, X_{a2}]$ + BLER$[X_c] \leq 1/3 \cdot 10^{-2} + 1/3 \cdot 10^{-2} + 1/3 \cdot 10^{-2} = 10^{-2}$ is assumed. For each value of the MCS of X_b (first column of Table A.3), a value of S (second column of Table A.3), which ensures that BLER$[X_b] \leq 1/3 \cdot 10^{-2}$ was assumed. Thereafter, for each value of the ratio I/S, the MCS of (X_{a1}, X_{a2}) is selected for which each value of S

[3]Although Table A.3 has only been reported, which considers the SIMO case and $I/S = 0$ dB, a similar table was made for $I/S = 5$ dB and for the SISO case.

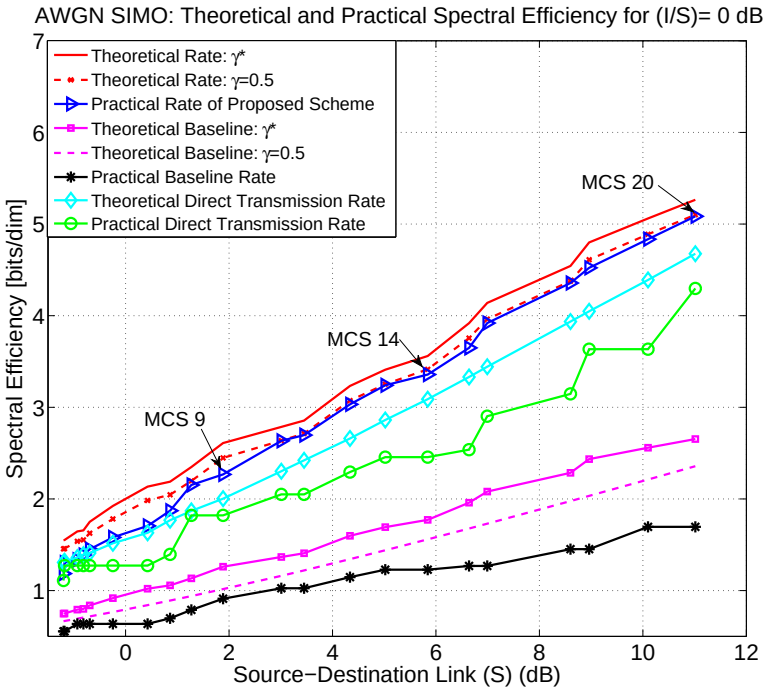


Figure 5.4: AWGN SIMO comparison between theoretical and practical spectral efficiencies for $(\mathrm{I}/\mathrm{S}) = 0$ dB

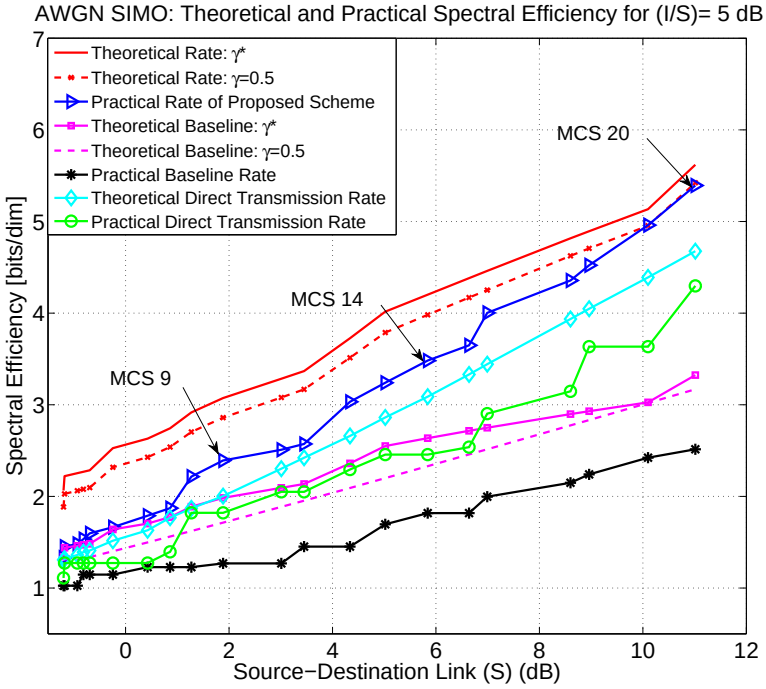


Figure 5.5: AWGN SIMO comparison between theoretical and practical spectral efficiencies for $(\mathrm{I}/\mathrm{S}) = 5$ dB

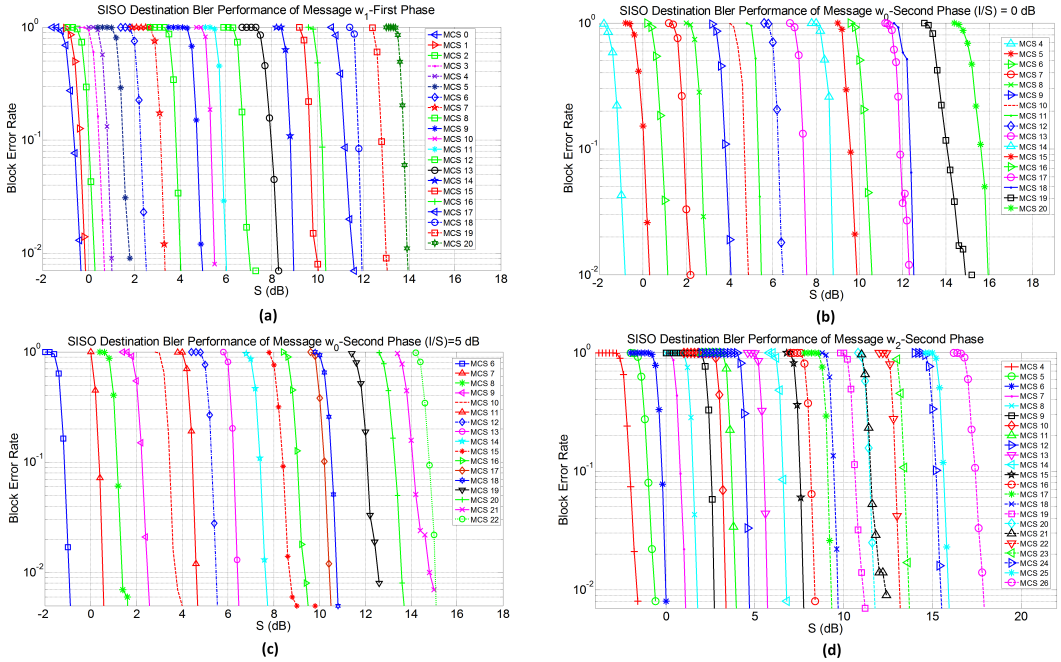


Figure 5.6: AWGN SISO BLER performances of w_0, w_1 and w_2 at the destination versus different strengths of the direct source-destination link

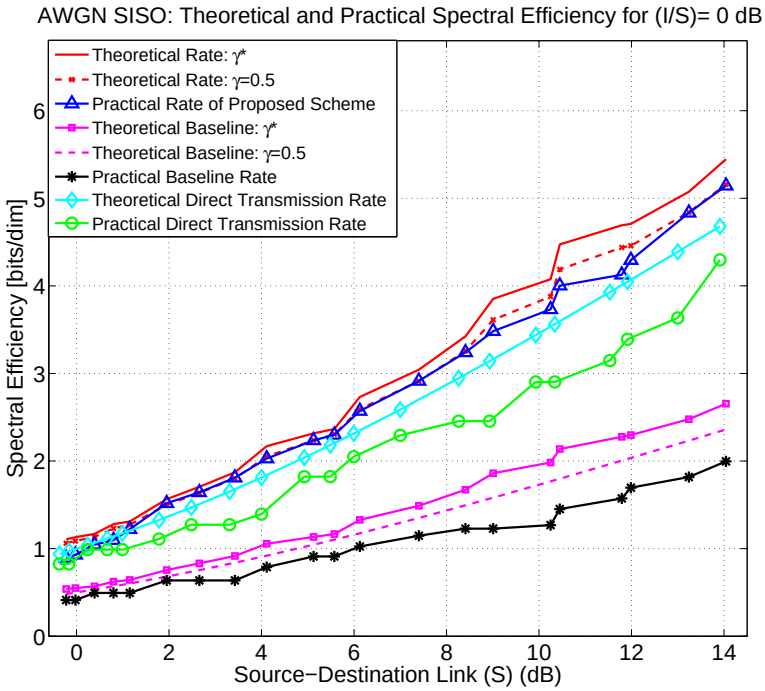


Figure 5.7: AWGN SISO comparison between theoretical and practical spectral efficiencies for $(\mathrm{I}/\mathrm{S}) = 0$ dB

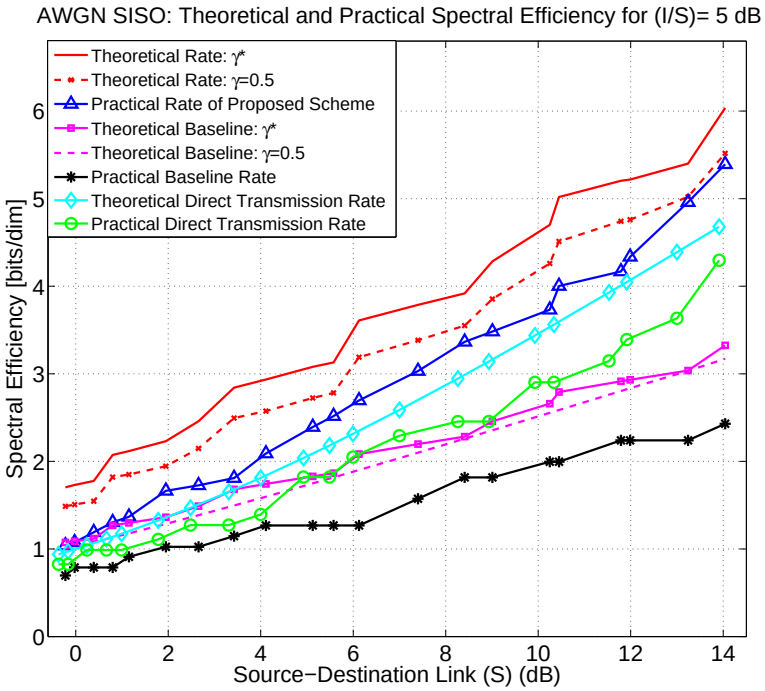


Figure 5.8: AWGN SISO comparison between theoretical and practical spectral efficiencies for $(\mathrm{I}/\mathrm{S}) = 5$ dB

(second column of Table A.3), ensures BLER$[X_{a1}, X_{a2}] \leq 1/3 \cdot 10^{-2}$. These MCS values are reported in the sixth column of Table A.3. Similarly, the MCS of X_c (eighth column of Table A.3) is selected, such that BLER$[X_c] \leq 1/3 \cdot 10^{-2}$. The TBSs of each MCS (at the source and at the relay) are defined according to the 3GPP LTE standard [140] and are also reported in Table A.3.

Comparison with theoretical performance One of the major goals of this work is to compare the theoretical and practical spectral efficiency performance of the proposed strategy. To this end, the following procedure was followed: For both the SISO and SIMO cases, two curves are drawn, corresponding to the theoretical achieved average (averaged over 10^4 different realizations of the phases of the channel parameters) spectral efficiency. In particular,

- Theoretical rate with γ^* (solid red curve): this curve is drawn by applying Eq. (5.2a) which includes the values of (S, I, C) used to obtain the practical spectral efficiency. In other words, this rate is computed by using the optimal γ^* defined in Eq. (5.8).

- Theoretical rate with $\gamma = 0.5$ (dashed red line): this curve is drawn by considering the values of (S, I, C) used to obtain the practical spectral efficiency and by computing

$$R = \mathbb{E}\left[\min\left\{\text{Eq.}(5.4), \text{Eq.}(5.6)\right\}\right] + \text{Eq.}(5.7) + \text{Eq.}(5.5),$$

  with $\gamma = 0.5$. Hence, this rate is computed with the value of γ used in the simulations.

Clearly, it is also observed from Fig. 5.4, Fig. 5.5, Fig. 5.7 and Fig. 5.8, that the theoretical rate with the optimal γ is no less than the theoretical rate with $\gamma = 0.5$. This is because in the latter case, the value of $\gamma = 0.5$ is fixed, which might not be the optimal value. The spectral efficiency of the proposed practical scheme - blue curve in Fig. 5.4, Fig. 5.5, Fig. 5.7 and Fig. 5.8 - (see for example the eleventh column of Table A.3 for the SIMO case with $I/S = 0$ dB) was determined by using the ratio of the TBS (useful message length) with respect to the number of soft-bits (G-codeword size) together with the modulation order, which does not include the overhead bits such as the cyclic prefix, pilots and control channel information (PDCCH symbols). In particular,

$$R = \frac{\text{TBS}(X_b) + \text{TBS}\left(X_{a1}, X_{a2}\right) + \text{TBS}(X_c)}{\left(\frac{\text{G}_1}{\text{Q}_{\text{mod1}}} + \frac{\text{G}_2}{\text{Q}_{\text{mod2}}}\right)} \text{ bits/dim}, \tag{5.11}$$

where G_1 is the number of soft-bits used to decode (X_b, X_{a1}) and G_2 to decode (X_{a2}, X_c), and Q_{mod1} and Q_{mod2} are the corresponding modulation orders.[4]

Simulation analysis For the SIMO case, from Fig. 5.4 and Fig. 5.5 it is noticed that the maximum difference between the theoretical rate (with $\gamma = 0.5$) and the achieved rate by the proposed scheme is 0.28 bits/dim at MCS $= 6$ when $I/S = 0$ dB and 0.67 bits/dim at MCS $= 7$ when $I/S = 5$ dB. Similarly, for the SISO case, the maximum difference between the theoretical rate (with $\gamma = 0.5$) and the achieved rate by the proposed scheme is 0.31 bits/dim at MCS $= 17$ when $I/S = 0$ dB (see Fig. 5.7) and 0.69 bits/dim at MCS $= 7$ when $I/S = 5$ dB (see Fig. 5.8). These values indicate that high-throughput HD relay schemes are within practical reach for de facto 4G relays and receivers of today. These rate gaps between theory and practice can be mostly attributed to two key factors: (i) the TBs used are of finite length, differently from the theoretical assumption of an infinite block length and (ii) the coding and modulation schemes were chosen to match existing standards rather than being

[4] TBS is the number of information bits, G is the number of coded bits and $\text{Q} \equiv \log M$ where M is the modulation order.

optimized for this setup. Furthermore, the fact that the difference when $I/S = 5$ dB is larger than when $I/S = 0$ dB is due to the fact that when the ratio I/S increases it becomes more critical to choose higher MCS values for the relay (with respect to the source) in order to fully exploit the strength of the relay-destination link. Adapting the modulation order and the number of PRBs across different rounds may reduce the gap between the theoretical and practical performances. Furthermore, it should be noted that the difference between theoretical and practical rates might be decreased by performing an optimization of the parameters δ (superposition factor) and γ (fraction of time the relay listens to the channel) in the interval $[0,1]$, instead of considering them as fixed values (which has been deemed out of the scope of this study).

5.5.1 Baseline Relay Scheme

For comparisons with existing relay structures, a baseline two-hop communication scheme (BS) is considered, which mimics the relay structure of today's LTE networks, where the UE does not have a direct connection with the eNB, i.e., the source-destination link is absent and the eNB can communicate with the UE only through the relay (no physical layer cooperation). For both the SIMO and the SISO scenarios, the BS practical rate (black curve) is compared to the theoretical one (magenta dashed curve) in Fig. 5.4, Fig. 5.5, Fig. 5.7 and Fig. 5.8. The theoretical BS achievable rate with $\gamma = 0.5$ (twelfth column of Table A.3) is given by:

$$R_{\text{Theor}-\text{BS}} = \min\{\gamma \log(1+C), (1-\gamma)\log(1+\alpha I)\}, \tag{5.12}$$

with $\alpha = 2$ in the SIMO case and $\alpha = 1$ in the SISO case. The optimal theoretical γ^* is obtained by equating the two terms within the min in Eq. (5.12), i.e.,

$$\gamma^* = \frac{\log(1+\alpha I)}{\log(1+\alpha I) + \log(1+C)}. \tag{5.13}$$

Both the curves when $R_{\text{Theor}-\text{BS}}$ in Eq. (5.12) is evaluated in γ^* in Eq. (5.13) (solid magenta curve) and the curve when $R_{\text{Theor}-\text{BS}}$ in Eq. (5.12) is evaluated in $\gamma = 0.5$ (dashed magenta curve) are plotted for the baseline two-hop communication strategy in a similar fashion to the proposed three-part-message scheme. Similar comments as those drawn for the proposed scheme also hold in this case. In the SIMO case, it is observed that from Fig. 5.4 and Fig. 5.5, the maximum difference in spectral

efficiency between the practical strategy and the practical BS rates (thirteenth column of Table A.3) is of 3.39 bits/dim for $I/S = 0$ dB, and of 2.88 bits/dim for $I/S = 5$ dB, which is a significant improvement in spectral efficiency of the cooperative relay strategy over the currently employed DF basic relay scheme. In the SISO case, it was also noticed that the maximum difference between the practical rate and the practical BS rate is of 3.15 bits/dim and of 2.97 bits/dim, respectively. The fact that the difference is larger when $I/S = 0$ dB than when $I/S = 5$ dB is due to the fact that when the ratio I/S is small, the presence of the source-destination link plays a significant role in enhancing the rate performance. Moreover, it was noticed that the curve representing the spectral efficiency of the scheme described in Section 5.3 is almost twice the one of the baseline strategy. This means that the scheme here proposed, not only provides an array gain with respect to the baseline strategy, but also a generalised Degrees-of-Freedom (gDoF) gain (pre-log factor). This can be explained as follows. The values of (S,I,C) here considered are more or less comparable; this implies, that in the high-SNR regime $R_{\text{Theor}-\text{BS}}$ in Eq. (5.12) becomes $R_{\text{Theor}-\text{BS}} \approx \frac{1}{2}\log(1+\text{SNR})$, while R in Eq. (5.2) becomes $R \approx \log(1+\text{SNR})$, i.e., $R \approx 2R_{\text{Theor}-\text{BS}}$.

5.5.2 Direct transmission scheme (No Relay)

The theoretical (cyan curve) and practical (green curve) rates for direct communication between the source and the destination transmitting a single-part-message (basic point-to-point scenario), are also provided in Figs. 5.4, 5.5, 5.7 and 5.8. According to Figs. 5.4 and 5.5, the maximum difference in spectral efficiency between the three-part-message practical scheme and the single-part-message practical point-to-point scheme is of 1.21 bits/dim for $I/S = 0$ dB, and of 1.33 bits/dim for $I/S = 5$ dB. In the SISO case, a maximum difference of 1.20 bits/dim ($I/S = 0$ dB) and 1.33 bits/dim ($I/S = 5$ dB) is observed. These values show that the proposed three-part-message scheme has a higher spectral efficiency when compared to the direct transmission scheme, while maintaining the same slope. In particular, in order to observe a gDoF gain (in addition to the increased power efficiency) of the proposed three-part-message scheme with respect to the point-to-point transmission, it would be possible to consider $I = S^{\alpha}$ and $C = S^{\beta}$ with $\min\{\alpha,\beta\} > 1$, which would be interesting from an information theoretic perspective (similar to [147]), in order to understand the asymptotics of the proposed scheme. However, when considering practical codes, the operating practical SNR regime of the LTE codebooks would be eventually be exceeded. For example, if $\alpha = 2$, S would be limited to around

10 dB at the point where the links to and from the relay would be at the highest spectral efficiency. In this case, MCS formats beyond 6 bits/dim (64-QAM) would be required on the source-relay and relay-destination links, which are not part of the performance evaluation tools which are currently available in the test bench. The resulting numerical evaluations show that the three-part-message scheme brings rate advantages (with respect to the direct transmission scheme) in the finite-SNR regime, thus not requiring the use of asymptotic performance indicators. The theoretical (cyan curve) and practical (green curve) point-to-point rates also further strengthen the notion that the conventional relay scheme (as indicated by the baseline scenario) does not provide any quantitative benefits, in terms of spectral efficiency, over the proposed scheme.

5.6 PERFORMANCE EVALUATION: FREQUENCY-SELECTIVE FADING CHANNEL

The AWGN channel modeling of the proposed relay strategy, analyzed in Section 5.5, represents an idealistic scenario. In an effort to model a practical scenario, the the spectral efficiency of the strategy using two well-known low mobility frequency-selective channel models defined by 3GPP, i.e., the EPA and ETU models [144] is evaluated. In particular, the focus is on the scenario where the destination is equipped with two antennas (SIMO). Channel state information at all nodes is assumed. For a more practically relevant scenario, the channel state information should be estimated or learned and this is currently part of ongoing work. The power delay profiles (amplitude and path delay) of the EPA and ETU channel models are derived from [31, 144]. The EPA and ETU models consist of seven and nine discrete multipath components, each with a coherence bandwidth of 2.43 MHz and 0.2 MHz, respectively. The amplitude distribution for each tap in the EPA and ETU models is described by a Rayleigh fading process. The complex channel coefficients for both the source-destination (h_{ds_1} and h_{ds_2}) and relay-destination (h_{dr_1} and h_{dr_2}) links are generated according to the generalized channel transfer function (in the frequency domain) given by $h_{i,n} = \sum_{l=1}^{L} \alpha_l \exp(-j2\pi\tau_l n f_{sub})$, where i represents the symbol index, $n \in [1:300]$ represents the subcarrier index for a bandwidth of 5 MHz, α_l represents the complex path amplitude, l is the path index, L represents the total number of paths, τ_l is the path delay and f_{sub} represents the periodic subcarrier spacing of 15 kHz (as specified in LTE) [31]. Furthermore, a zero Doppler shift is assumed for both channel models in line with the low-mobility assumption of the destination (UE). The link between the source and the relay is modeled

to be a static AWGN channel with the aim of deploying a relay in a location with a high link quality from the source. An analysis of the proposed scheme over a fading channel would involve an evaluation of the achievable rate under a given outage probability, which would hold for the infinite block length and block-fading channel case. This would be performed by extending Eq. (5.2) to a vector channel with channel coefficients governed by the statistics of the EPA and ETU models and by evaluating the achievable rates at which the the desired outage probability is obtained. The aforementioned analysis presents an interesting challenge for future research directions. The major goal of the comparison carried out in the rest of this section, was to show the rate advantage of the proposed scheme in realistic channel models with respect to the existing baseline and point-to-point transmission schemes. Due to the fact that the system was not designed and optim-

Table 5.2: LTE delay spread profile.

Extended Pedestrian A			Extended Typical Urban		
	Excess Tap Delay [ns]	Relative Power [dB]		Excess Tap Delay [ns]	Relative Power [dB]
1	0	0.0	1	0	-1.0
2	30	-1.0	2	50	-1.0
3	70	-2.0	3	120	-1.0
4	90	-3.0	4	200	0.0
5	110	-8.0	5	230	0.0
6	190	-17.2	6	500	0.0
7	410	-20.8	7	1600	-3.0
-	-	-	8	2300	-5.0
-	-	-	9	5000	-7.0

ized for these frequency selective channels, it was a challenge to guarantee an overall system BLER of 10^{-2} at the destination at reasonable SNR values for higher MCSs, with an interfering codeword from the same discrete constellation. Hence, it was decided to relax the BLER constraint such that $\text{BLER} \leq \text{BLER}[X_b] + \text{BLER}[X_{a1}, X_{a2}] + \text{BLER}[X_c] \leq 10^{-1}$, only considering MCS values corresponding to QPSK transmissions.

Fig. 5.9 displays the complete BLER results for each of the decoding operations at the destination for the EPA channel. The results in Fig. 5.10 and Fig. 5.11 show the performance of the three-part-message relay strategy (see blue curve) using the EPA channel model for $I/S = 0$ dB and $I/S = 5$ dB, respectively. Fig. 5.12 displays the complete BLER results for each of the decoding operations at the destination for the ETU channel. Similarly, the performance of the ETU channel model, presented in Fig. 5.13 and Fig. 5.14, has also been investigated.

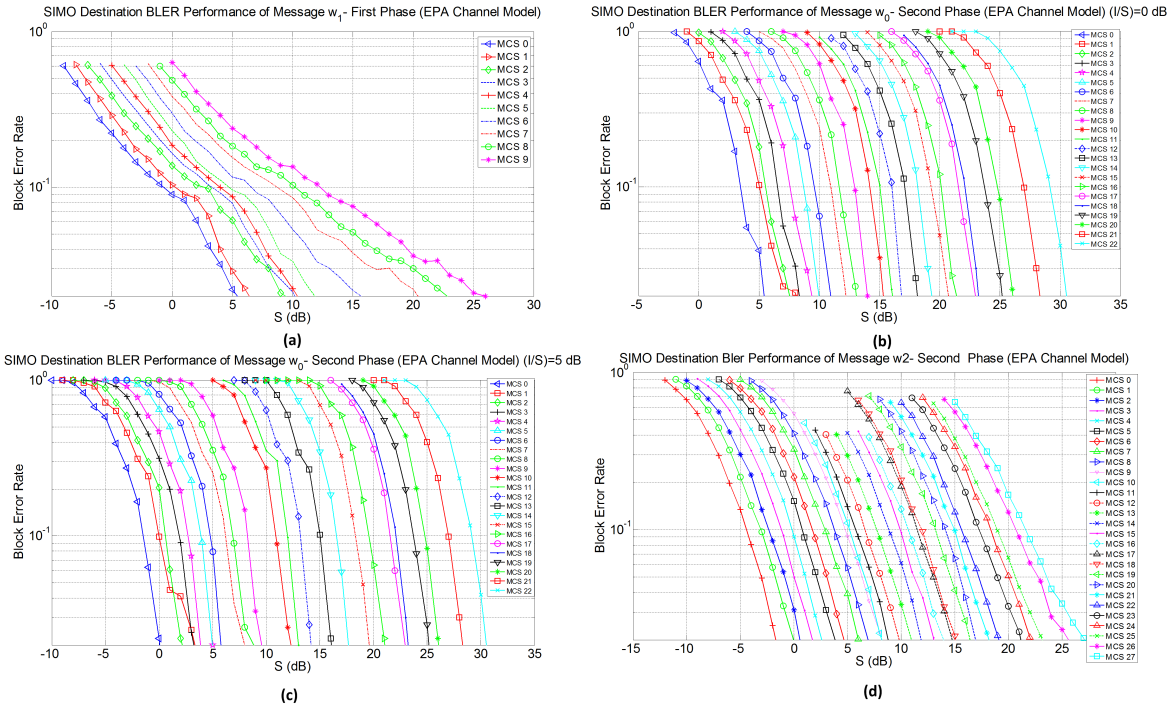


Figure 5.9: SIMO BLER performances of w_0, w_1 and w_2 at the destination versus different strengths of the direct source-destination link for the EPA channel model

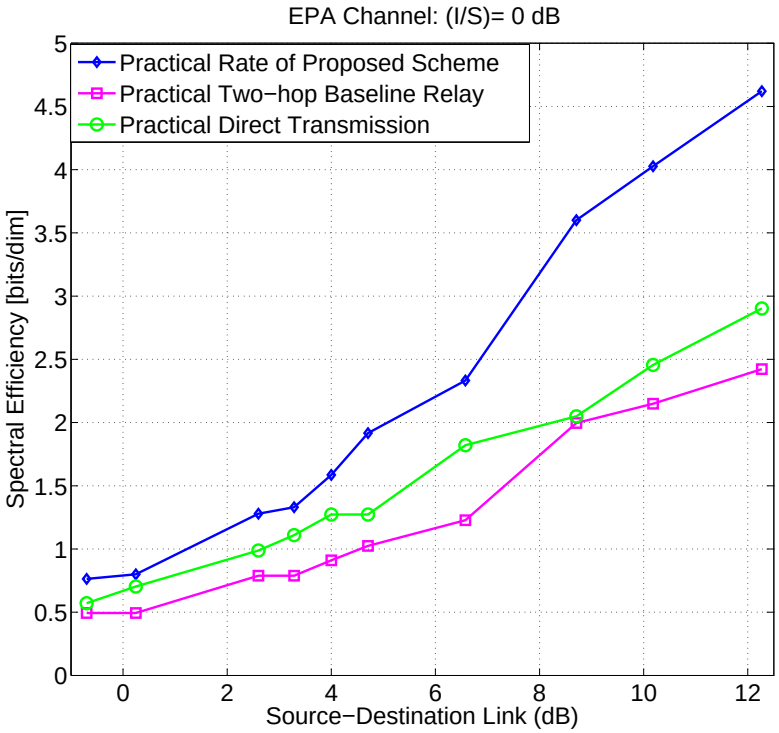


Figure 5.10: EPA practical rate comparison for $(\mathrm{I/S}) = 0$ dB

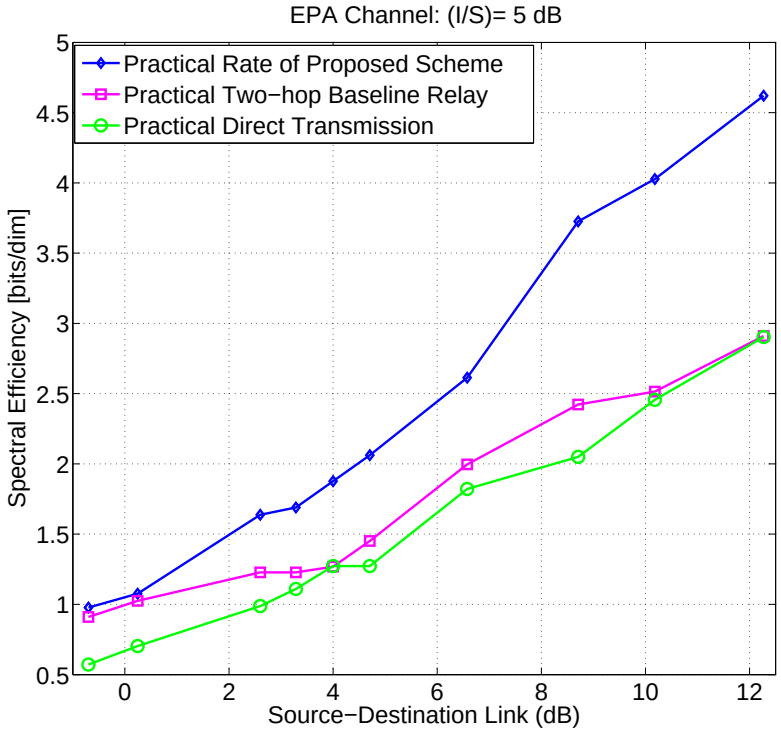


Figure 5.11: EPA practical rate comparison for $(\mathrm{I}/\mathrm{S}) = 5$ dB

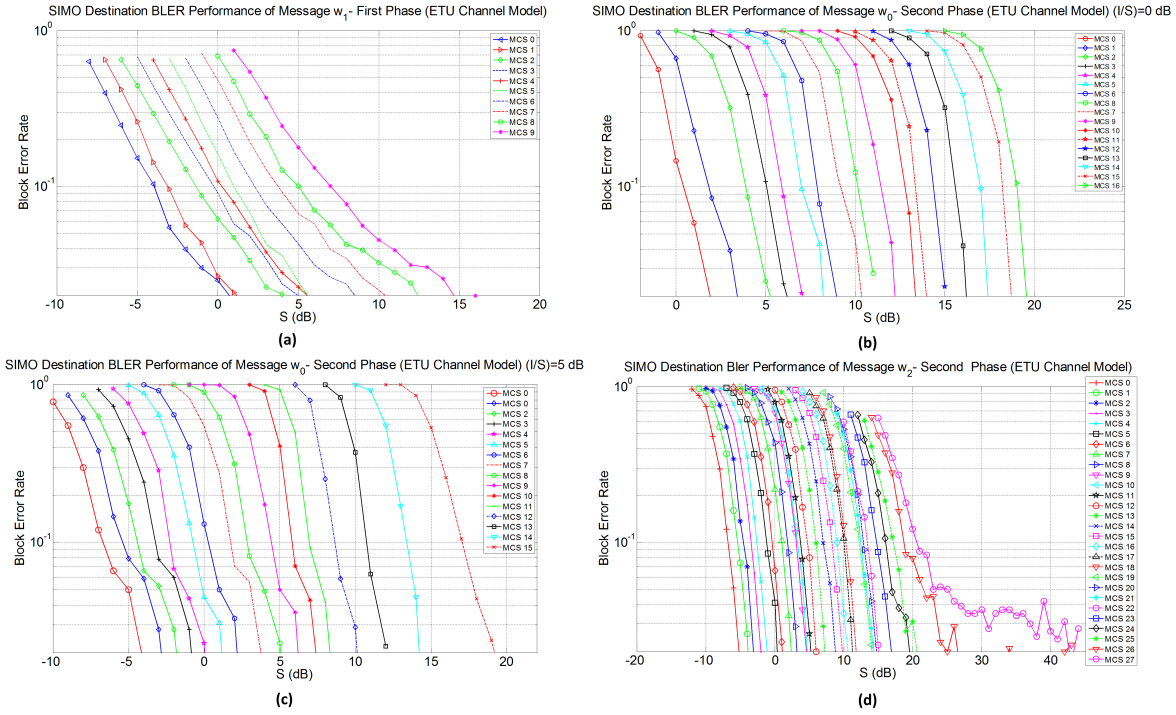


Figure 5.12: SIMO BLER performances of w_0, w_1 and w_2 at the destination versus different strengths of the direct source-destination link for the ETU channel model

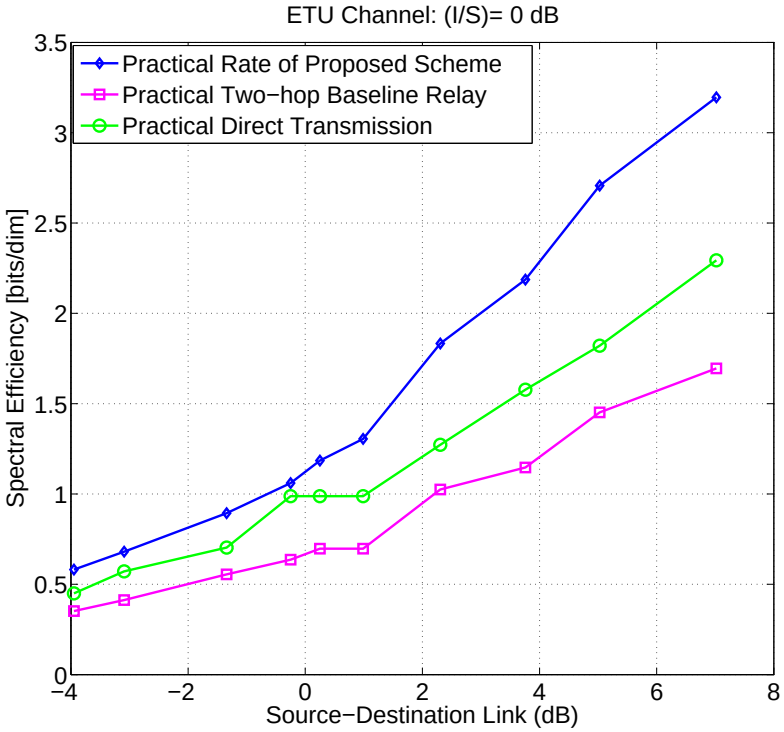


Figure 5.13: ETU practical rate comparison for $(\mathrm{I}/\mathrm{S}) = 0$ dB

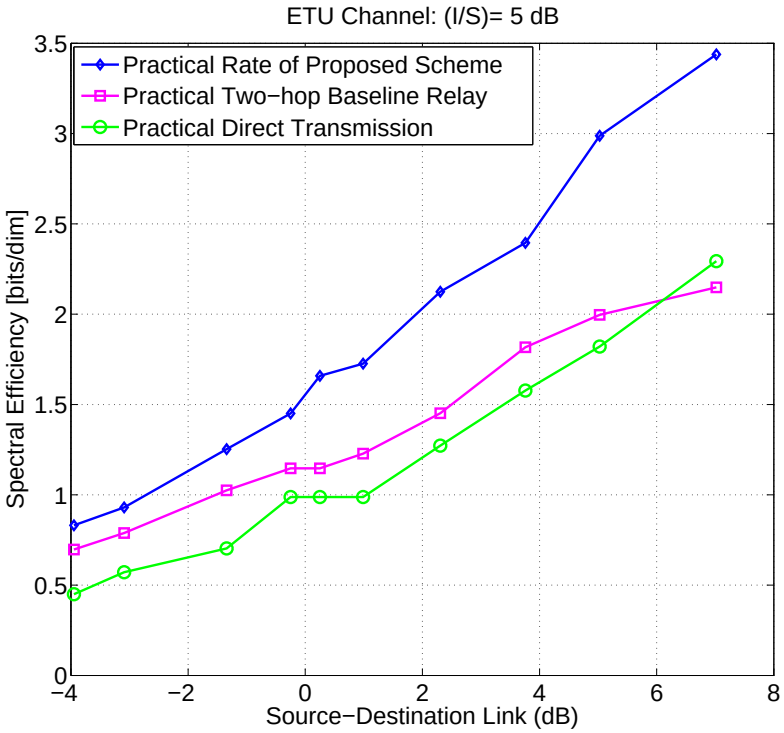


Figure 5.14: ETU practical rate comparison for $(\mathrm{I}/\mathrm{S}) = 5$ dB

In the case of the EPA channel, the maximum difference between the practical strategy (blue curve) and the BS rates (magenta line) is of 2.20 bits/dim for $I/S = 0$ dB and of 1.71 bits/dim for $I/S = 5$ dB. In the case of the ETU channel model, the maximum difference in spectral efficiency between the practical strategy and the BS rates is 1.50 bits/dim for $I/S = 0$ dB and 1.29 bits/dim for $I/S = 5$ dB. The proposed relay scheme also performs well in comparison to the point-to-point transmission strategy, justifying the feasibility of the proposed relay scheme. A maximum difference of 1.71 bits/dim can be observed between the practical and the point-to-point scheme (green curve) for the EPA case, when $I/S = 0$ dB and $I/S = 5$ dB. Similarly, for the ETU case, a maximum difference of 0.90 bits/dim when $I/S = 0$ dB and 1.16 bits/dim when $I/S = 5$ dB can be noted. However, it is interesting to observe that the benefit - in terms of spectral efficiency - of utilizing a relay (baseline two-hop communication scheme) over the classical point-to-point communication scheme can be realized when considering realistic fading channel models in which the strength of the relay-destination link is 5 dB stronger than the source-destination link as shown in Fig. 5.11 and Fig. 5.14. The difference in spectral efficiency for the practical LTE channel model is noticeably less than the one in the AWGN case, highlighting the degrading effects of multipath on the proposed relay strategy. Nonetheless, even for these two practically relevant LTE channel models, the proposed strategy still provides remarkable improvements in spectral efficiency over the basic BS and point-to-point transmission schemes.

5.7 CONCLUSIONS

This chapter has presented a practical transmission strategy for the Gaussian half-duplex relay channel by using codes as specified in the LTE standard and by running simulations on an LTE test bench. The scheme uses superposition encoding, decode-and-forward relaying and successive interference cancellation in order to rate split the message into three parts and send them in two time slots from a source to a destination with the help of a relay (which forwards one of the three sub-messages). Comparisons between the theoretical achievable rate with (point-to-point capacity achieving) Gaussian codes and the rate achieved in a practical scenario were provided for an overall system BLER of 10^{-2} for the two-antenna case (and single-antenna case) at the destination. The simulation results shed light on the benefits of physical layer cooperation with respect to direct-link transmission and to the case of a baseline two-hop communication scheme. The half-duplex relay strategy was also

shown to provide robust spectral efficiency gains over the proposed baseline two-hop communication scheme using the EPA and ETU channel models. The outcomes presented in this Chapter should be used to reconsider the case of relay stations for future 5G deployments where new layers 1 and 2 signaling can be envisaged to reap the benefits from physical-layer relaying.

CHAPTER 6

CONCLUSIONS AND FUTURE OUTLOOK

This dissertation has investigated techniques related to the area of higher-order MIMO Block receiver detection and HD relay networks, with a focus towards implementation in future 4G+ and 5G networks. One of the motivations of this research was the viability of such techniques in today's LTE networks and possible application in the forthcoming generation of mobile networks. Chapter 1 provides an initial overview of the fifth generation of mobile networks as well as the role of HetNets in network densification. The problem domain is defined by examining the two key research challenges, which form the focal points of this study. The first challenge is the design of higher-order MIMO detectors in interference-limited scenarios, of which high levels of interference will be an inevitable consequence of dense networks. Therefore, the trade-offs between complexity and performance are key merits to consider during the design process. The second research challenge involved providing a viable use-case for relay deployment in LTE networks. Telecommunication operators have shown a lack of interest in such networks due to the limited spectral efficiency characteristics of traditional HD relays and Part II of this research has shown otherwise, by demonstrating the feasibility of an advanced HD relay strategy in an LTE deployment scenario.

Chapter 2 presented a background overview of the various aspects related to MIMO detection, interference management, relay networks as well as the LTE communication standard. These themes represent the collective state-of-the-art and serve as a precursor to the contributions of this research study.

6.1 SUMMARY OF FINDINGS

6.1.1 Part I:Higher-order MIMO Detection Strategies

Part I of the thesis investigated the use of a Blocking QR technique to perform low-complexity higher-order MIMO detection in an interference-limited scenario and rate maximisation in an LTE SU-MIMO point-to-point setting. An information theoretic analysis and practical evaluation of the Block QR detection technique is presented in order to gain insights into the algorithmic performance and complexity.

The maximal achievable rates for the first stream detection using a novel Block QR scheme for the 4×4 and 8×8 interference-limited scenarios revealed a loss in mutual information of 0.55 bits/s/Hz and 0.98 bits/s/Hz respectively, at an SNR of 15 dB using a flat-fading (Rayleigh) channel model with an outage of 10%. According to these observations, this loss seems to decrease accordingly when shifting to a higher SNR regime while at the same time increase with the number of antennas. The Block MMSE detection of the first stream was validated to be information lossless and is therefore considered to be the optimal detection when performing first stream detection, with a higher computational cost.

A rate optimisation scheme for a SU-MIMO point-to-point link is also investigated using the Block QR technique, in order to maximise the overall rate. The system rate is maximised over the transmission of the two codewords as specified in LTE, using the rates, R_1 and R_2 for both the 4×4 and 8×8 MIMO configurations. Therefore, intuitively the total rate loss of the Block QR scheme has been shown to arise only from the detection loss occurring in the first stream. The minimal rate loss in the practical SNR regime demonstrates the theoretical feasibility of maximising the overall rate using the Block QR scheme in the context of a SU-MIMO scenario, especially in a Rayleigh Channel.

The next phase of the research study was to evaluate the complexity and performance of the Block QR scheme in a LTE simulation test bench in order to gain meaningful insights into the BLER performance in a Rayleigh channel. For the 4×4 case, it was observed that the Block MMSE scheme was approximately twice as complex as the MGS-Block QR algorithm and for the 8×8 case, the complexity quadrupled with the number of required flops. An optimal but highly complex brute-force Max-log-MAP detector is also implemented as a performance benchmark against the proposed Block QR algorithm. The results show that a performance loss of 3 dB is obtained between the brute-force

Max-log-MAP detector and Block QR Max-log-MAP detector, which approximately corresponds to the same loss in the case of the theoretical outage probability (2.5 dB). This validates the idea that the theoretical and practical performance are comparable.

6.1.2 Part II: Advanced HD Relay Strategy

Part II of the thesis provides insights into the feasibility of a physical-layer co-operative link adaptation strategy for HD relays, which leverages superposition coding and SIC decoding to transmit a 3-part-message over a single LTE subframe. One of the goals of this part of the research, was to bridge the gap between theory and practice in the context of codeword length and signal assumptions. As such, a comparison is drawn between the the theoretical achievable rate with Gaussian codes and the rate achieved in a practical scenario for an overall system BLER of 10^{-2} for a destination (UE) equipped with a single and dual antenna configuration. These results provided an understanding into the benefits of physical-layer cooperation with respect to existing two-hop relay schemes and point-to-point links. The relay strategy was evaluated to be robust and provided improved spectral efficiency gains, using two practical LTE channel models.

The findings presented in this feasibility study strengthen the case for future 5G HD relay deployments where the benefits of physical-layer relaying can be realised to bring about spectral efficiency improvements. This may prove to be even more effective in millimeter-wave deployments where predominantly line-of-sight links are required and wired access-points will be extremely costly for large scale deployment. In this case, a few donor eNBs with an optical backbone to the core network and many wireless relay stations using physical layer relaying to their proximity can be considered. For UEs in line-of-sight with both the destination and relay, there will be substantial benefits as described by these results. The results from this study represent a strong enabler for HD relays to be deployed in the upcoming releases of the LTE standard or alternatively, 5G systems which could exploit such advanced relay architectures.

6.2 RECOMMENDATIONS FOR FUTURE WORK

An information theoretic analysis that shifts away from the assumption of Gaussian interferers to discrete alphabets would be considered as a next step in determining the analytical performance

of the Block QR scheme. This is especially applicable to the rate maximisation scheme, where it could help understand the rates (MCS) at which each of the codewords can be transmitted to achieve maximum throughput at a particular SNR. The analysis could also be extended to include the performance in fading channels, with rich multipath environments. The inclusion of the Block MMSE BLER practical performance is a subject of ongoing work and will enable a complete understanding regarding the performance-complexity trade-off of the Block MMSE detection scheme. Another future research direction with regard to the practical evaluation of the Block detection schemes, would include a fixed-point complexity analysis in order to understand the hardware implications such as any losses in precision that may affect the decoding performance. A limitation with the BLER comparison was that only QPSK signals were considered, i.e. MCS=9, and an extension to 16-QAM and 64-QAM signaling could be performed, although the computational time for the brute-force Max-log-MAP algorithm would be significant. An efficient way to realise the comparisons of higher-order modulation schemes would also be recommended as future work. The impact of imperfect CSI on the rate performance is also a suggested future direction with regard to the practical performance.

In relation to the advanced HD Relay strategy, a subject of future investigations would be the analysis of the proposed three-part-message when channel state information is not available and hence has to be estimated or learned as well an investigation of resource allocation strategies for dynamic bandwidth assignment. In Part II of this study, an equal duration of the two phases and a fixed value for the superposition factor was considered; for inclusion in real-time systems, these parameters have to be adaptive and adjusted accordingly. More general dimensioning of resources can be made over a number of H-ARQ rounds. In addition, higher-order MIMO configurations can be considered at the relay and source, which are supported in LTE. An information theoretic analysis of the achievable rate, given an outage constraint in a fading scenario would also present an interesting avenue for future research.

CHAPTER 7

RÉSUMÉ [FRANÇAIS]

L'évolution des réseaux de communication mobile sans fil a toujours été progressive notamment à cause des demandes croissantes de la part des utilisateurs. Ce travail présente deux contributions clés dans le domaine des réseaux de communication sans fil 4G+/5G, en particulier dans le domaine de la détection MIMO d'ordre supérieur et le design en réseau relai semi-duplex.

La première partie de ce travail de recherche s'intéresse au développement d'une stratégie de détection MIMO d'ordre supérieur pour les terminaux 4G+/5G existants et futurs, d'un point de vue à la fois théorique et pratique. Une nouvelle technique de décomposition QR Bloc de prétraitement est proposée pour un récepteur LTE dans un scénario limité à un seul utilisateur, interférence limitée, ainsi qu'un scénario point par point dont les résultats surlignent les avantages et inconvénients en performance et complexité.

La deuxième partie de cette étude comprend une étude de faisabilité d'une stratégie de message novatrice à deux phases et trois parties pour un réseau relai semi-duplex à couche physique, qui comporte un codage de superposition et un décodage d'annulation d'interférence successif, conscient des interférences. Un point clé de cette étude était d'analyser la performance du schéma d'adaptation des liaisons proposées dans le régime non asymptotique (longueur de bloc finie et de signalisation de constellation discrète) et d'évaluer l'efficacité spectrale (ES) par rapport aux hypothèses théoriques des blocs d'asymptotiquement grandes longueurs. Une comparaison ES supplémentaire est également présentée, avec une transmission de relais à deux étapes non coopérative et une stratégie de

transmission point à point (sans relais). Les résultats obtenus révèlent les gains d'ES qui peuvent être obtenus en exploitant la coopération de couche physique entre le relais et la station de base.

7.1 INTRODUCTION

La technologie est devenue une partie intégrante de la société moderne et l'omniprésence de la connectivité Internet est devenu un promoteur clé de ce progrès. Les systèmes de communication sans fil ont toujours fait l'objet d'avancées évolutives au fil des ans, par exemple déploiement de réseaux 2G entre la période 1990-2000 et 3G entre la période 2000-2010, afin de répondre à la demande croissante des débits plus élevés et la qualité de service (QoS). Actuellement, la quatrième génération de réseaux (4G) présente un petit casse-tête pour les défenseurs des réseaux de prochaine génération, comme il a été conçu pour être bien adoptée au-delà du cycle de dix ans traditionnelle (2010-2020). Bien que, la normalisation de la cinquième génération (5G) réseaux et son déploiement est encore à finaliser, qui sera en elle-même apporte le long d'une foule de nouveaux cas et les défis d'utilisation, la vitesse à laquelle il peut être initialement adopté peut être affectée par la courant adoption mondiale réussie de Long Term Evolution (LTE) des réseaux de 3GPP. Si 5G se révélera être un changement de paradigme drastique ou une consolidation des technologies mobiles existantes pour améliorer la qualité du réseau et les performances globales, reste à voir. Un autre défi, est de convaincre les parties prenantes telles que les opérateurs de télécommunications et les fournisseurs de réseau que l'introduction de nouvelles technologies de réseau vaut bien l'investissement, ne sera pas un impact négatif sur leurs revenus et augmentera bénéfice global. Il est bien entendu que les versions actuelles de la norme LTE ont des limitations dans la conception de protocole pour (M2M) les communications de machine à machine, qui a été présentée comme l'une des technologies caractéristiques clés de 5G. D'un point de vue de la recherche, il devient un défi intéressant d'étudier les compromis de performances impliquées pour une telle restructuration ou si une approche de la table rase serait plus approprié. Cela pourrait concerner non seulement du point de vue de la communication M2M, mais aussi dans le contexte d'autres technologies de la prochaine génération, tels que les réseaux hétérogènes, MIMO Massive et ondes millimétriques pour ne nommer que quelques-uns. En outre, il serait également un intérêt supplémentaire pour explorer la façon dont les approches théoriques sont analysées par rapport à des scénarios pratiques, puis quantifier les écarts de performance, de manière à satisfaire aux exigences de la 4G+ (Release 14 et au-delà) et les réseaux 5G.

7.1.1 Objectifs de recherche

L'adoption de HetNets dans les réseaux de la prochaine génération offre une plate-forme pour traiter de nombreuses zones ouvertes de la recherche, en particulier dans le contexte des défis liés à l'atténuation des interférences. Le thème de cette thèse tourne autour de la conception d'algorithmes de brouillage-courant d'ordre supérieur MIMO efficace qui peut être intégré dans les normes LTE actuels tout en tenant compte LTE Rel. 14 au-delà et les systèmes 5G. Comprendre les implications pratiques en utilisant des cadres théoriques est fondamental à une bonne conception du système de communication et de ce fait également partie des enquêtes menées dans le mémoire. La thèse peut être divisée en deux parties principales: 1) D'ordre supérieur MIMO détection du récepteur; 2) Coopérative Half-Duplex Stratégies relais. Les contributions de recherche suivantes sont détaillées comme suit:

1. D'ordre supérieur MIMO Récepteur de détection

   - Une technique de détection de roman en utilisant décomposition QR Bloc [26] est proposée pour les récepteurs d'interférence-courant (avec des sources d'interférence non-Gaussien) et évaluée par rapport au détecteur MMSE Bloc où est-il supposé que l'interférence est Gaussien. Ces techniques sont conçues dans le but de mettre en œuvre des systèmes MIMO d'ordre supérieur avec une complexité réduite dans les futurs récepteurs LTE existants / près. Ce récepteur est basé sur le bit interférence métrique sensibilisation l'architecture à faible complexité présentée dans [27], ce qui implique l'exploitation du brouilleur dominant avec une structure non-Gaussien, affectant principalement les utilisateurs cellulaire de pointe.

   - La perte de l'information mutuelle des stratégies de détection proposées sont comparés et analysées dans un scénario caractéristique de Hetnet d'interférence limitée. L'information mutuelle du QR Bloc se révèle être relativement minime, en particulier dans les régimes de SNR plus élevés, alors que le schéma MMSE Bloc est sans perte.

   - Un autre scénario d'application impliquant un MIMO mono-utilisateur point à point (P2P) scénario est étudié afin d'effectuer la maximisation débit en utilisant la limite de deux

mots de code par transmission en LTE. La faisabilité des schémas QR Bloc et MMSE Bloc sont examinés dans le contexte de fournir des débits améliorés au plus bas-ordres de complexité.

- Une évaluation des systèmes de blocs proposés sont comparées à la force brute optimale algorithme Max-log-MAP classique en utilisant un banc d'essai de simulation LTE basé sur la plate-forme OpenAirInterface [28], et sert de référence optimal lorsque l'on compare la performance et la complexité.

L'application des techniques de détection MIMO d'ordre supérieur est en relation avec un scénario typique d'une Hetnet brouilleuse dominante, où le terminal mobile a pour but de résoudre les signaux interférents souhaités, et à partir de deux stations de base différentes. Ces méthodes de détection sont également pertinents dans un MIMO point à point et multi-utilisateurs scénario MIMO.

2. Stratégie de relais coopératif élément semi-duplex pour le LTE
   Une stratégie d'adaptation de lien roman est proposée pour les réseaux HD relais et peut être mise en œuvre dans les réseaux LTE existants avec des composants et des blocs off-the-shelf. Les contributions clés comprennent:

   - Une évaluation des deux phases théoriquement optimales (jusqu'à un écart constant) trois partiel message système en utilisant des codes de canal pratiques tel que spécifié dans la norme LTE et en considérant une pratique Bloquer système concerné Error Rate global (BLER) valeur de 10^{-2} est effectuée (exigence LTE BLER est de 10^{-1}). Pour le canal AWGN SIMO statique, ie, lorsque la destination est équipée de 2 antennes, il est montré qu'une théorique (avec allocation de bande passante égale) et écart d'efficacité spectrale pratique de 0.28 bits/dim lorsque la source-destination et les liens relais-destination sont de la même force et de 0.67 bits/dim lorsque le lien relais-destination est de 5 dB plus élevé que le lien source-destination. Ces valeurs indiquent que les systèmes à haut débit HD relais sont à la portée pratique pour de facto les relais et les récepteurs d'aujourd'hui 4G, qui mettent en œuvre déjà des décodeurs turbo et donc encourent aucune complexité supplémentaire. De même, dans le cas de SISO il est démontré que l'écart d'efficacité

spectrale maximale entre la mise en œuvre théorique et pratique (avec allocation de bande passante égale) est de 0.31 bits/dim lorsque l'intensité de la source-destination et relais-destination des liens est la même et de 0.69 bits/dim lorsque le lien relais-destination est de 5 dB plus élevé que le lien source-destination. Le gain du réseau est donc exploité et révèle les avantages de la stratégie de relais HD proposé lors de l'emploi deux antennes de réception à la destination.

- Une comparaison de la performance du taux de ce régime par rapport à une stratégie de base est effectuée, où le lien entre la source et la destination est absent, soit il n'y a pas couche coopération physique entre la source et le relais pour transmettre des informations à la destination. Dans le cas SIMO, la différence maximale entre la stratégie de coopération proposée et le régime de base est de 3.39 bits / dim lorsque la force du canal de la liaison source-destination est la même que le lien relais-destination, et 2.88 bits/dim lorsque la liaison relais-destination est 5 dB plus élevée que le lien source-destination. De même, pour le scénario SISO, amélioration de l'efficacité spectrale de 3.15 bits/dim lorsque la force du canal de la liaison source destination est la même que de celui de la liaison relais-destination, et de 2.97 bits/dim lorsque le lien relais-destination est 5 dB plus élevé que le lien source de destination, sont observés. Ces valeurs montrent que le système de communication à deux sauts de base est pas bénéfique en termes de gain de taux, en particulier lorsque la force du canal de la liaison source-destination est la même que le lien relais-destination.

  Une comparaison de base point à point où une seule pièce de message est transmise de la source à la destination, est également fournie. Cette analyse montre que l'activation de la coopération de la couche physique entre les nœuds est d'une importance critique dans et l'avenir d'aujourd'hui des réseaux sans fil. Le schéma proposé peut, en effet, permettre un déploiement plus élevé de performance de l'efficacité spectrale de la première couche relais d'une manière qui serait réalisable d'un point de vue commercial, quelque chose que les méthodes normalisées (ce qui est appelé la stratégie à deux sauts de base) ne sont actuellement pas en mesure de fournir, principalement parce que les opérateurs n'ont pas trouvé un cas d'affaires en raison des avantages de l'efficacité spectrale limitée [29]. Il a déjà été montré que du point de vue de la couverture, le déploiement de mi à des noeuds

de relais haute puissance peut des économies de coûts de rendement d'au moins 30% pour les opérateurs [30].

- Deux pratiques différentes évanouissements quasi-statique modèles de canaux, à savoir le Piéton Extended A (EPA) et les urbains (ETU) modèles typiques étendus sont étudiés afin d'évaluer la performance de la stratégie HD relais dans un environnement multitrajet réaliste. Il a été observé que le modèle de ETU (racine signifie retard carré propagation de 991 ns) affiché en moyenne une plus grande efficacité spectrale globale sur le modèle EPA (racine délai moyen carré propagation de 43 ns) [31], grâce à la diversité de fréquence plus élevée de l'ancien modèle (largeur de bande de cohérence inférieure du modèle de l'ETU par rapport au modèle de l'EPA); le schéma proposé est donc bien adapté à la technologie LTE, qui bénéficie de la résilience de l'OFDM dans de tels environnements multivoies. Lorsque la source-destination et les liens relais-destination sont de la même force, l'amélioration du taux de la stratégie proposée sur la ligne de base schéma de communication de deux-hop est de 2.20 bits/dim pour le modèle EPA et de 1.50 bits/dim pour le modèle de ETU. Lorsque le lien relais-destination est de 5 dB plus élevé que le lien source de destination, l'amélioration maximale du taux est de 1.71 bits/dim et 1.29 bits/dim pour l'EPA et ETU modèles, respectivement. Les gains de taux sont inférieurs à ceux dans le cas de AWGN, mais apportent des améliorations raisonnables sur le régime existant de communication à deux sauts de base dans un scénario de canal plus réaliste.

Le résultat de l'étude révèle que la mise en œuvre pratique des techniques de relais HD est possible avec la modulation et formats de codage déjà spécifié par la norme 3GPP LTE, et que l'écart entre la théorie et la mise en œuvre de notre projet est faible. Optimisation des paramètres d'allocation des ressources (taux de code, l'allocation de bande passante) par rapport à la stratégie et le banc de marquage relais proposé son rendement par rapport second ordre modéré résultats de capacité bloc de longueur, dans l'esprit de ce qui a été lancé pour le point à point de canal dans [32], peut effectivement montrer proche optimalité de la stratégie relayant deux phases proposée.

7.2 CHAPITRE 2

Ce chapitre présente l'état de l'art des différents thémes de la thése. Les techniques de gestion des interférences, les stratégies de détection du récepteur, les techniques de relais semi-duplex et une vue d'ensemble de LTE sont présentés.

7.3 CHAPITRE 3

Ce chapitre introduit une nouvelle technique de prétraitement utilisant une approche de décomposition QR Bloc pour effectuer une détection MIMO d'ordre supérieur dans un scénario MIMO à limitation d'interférences et mono-utilisateur (SU). Les scénarios étudiés 4×4 et 8×8 (configurations supportées basées sur les standards LTE actuels [11]) sont analysés analytiquement en termes de perte d'information mutuelle. Cette détection de récepteur est basée sur le prétraitement de la matrice de canaux complexes en blocs pour effectuer une décomposition globale de QR Bloc, ce qui conduit à l'isolement du signal souhaité ou interférant (en fonction de la commande) pour le décodage en utilisant le démodulateur à dimension réduite dans [89]. Un aspect clé de cette technique est l'exploitation du brouilleur non-Gaussien dominant affectant les utilisateurs de bordures de cellules dans les régimes de brouillage élevé et modéré. La première partie de l'étude, considère un HetNet limité à une perturbation SU-MIMO avec deux eNB. Ensuite, un nouveau modèle d'optimisation de taux conçu pour un scénario LTE point à point (P2P) SU-MIMO est introduit. Dans [92], une procédure de blocage a été développée pour partitionner la matrice de canaux en 2×2 blocs pour améliorer le taux d'erreur de paquets d'une détection de maximum de vraisemblance réduite. Bien que dans cette étude, le scénario étudié ne soit pas limité par l'interférence et ne de bout en bout LTE évaluation de la performance. Des méthodes de décomposition QR séquentielles ont également été étudiées dans un contexte MU-MIMO [93], où les auteurs adaptent le système Dirty Paper Coding (DPC) Utilisant la décomposition QR Bloc pour améliorer les performances de plusieurs utilisateurs en présence de l'interférence entre cellules. La technique mentionnée ci-dessus est un activateur de l'interférence côté émetteur la gestion. L'étude des techniques de détection de perturbations pour les récepteurs LTE dans le contexte des systèmes MIMO d'ordre supérieur a été limitée, au meilleur de la connaissance des auteurs. Initialement, une approche de la théorie de l'information est entreprise pour quantifier la perte d'information mutuelle de la technique QR Bloc proposée avec une configuration 4×4 et 8×8

MIMO dans un Rayleigh bloc-décoloration canal. Ces résultats analytiques serviront de référence pour évaluer la performance pratique du QR Bloc proposé avec la sous-optimale, mais plus complexe (recherche de force brute) Max-log-MAP démodulateur algorithme.

7.3.1 MMSE Bloc

Le concept d'erreur quadratique minimale de bloc (MMSE) comme moyen de prétraiter le bruit et l'interférence a été appliqué à l'égalisation [123] dans des environnements multitrajets variant dans le temps pour OFDM ainsi que la détection des systèmes MIMO spatio-temporels [124]. Une technique MMSE en bloc a été mise en œuvre en tant que solution rentable du point de vue du calcul avec DPC (Dirty Paper Coding) pour obtenir la somme du canal de diffusion sans fil MIMO sous la forme d'une stratégie de précodage côté transmetteur [125]. Ce modèle adapte les ressources d'émission pour réaliser des gains de capacité dans un scénario MIMO multi-utilisateurs généralisé. Une approche de diagonisation en blocs a également été proposée dans le but de supprimer les interférences inter-utilisateurs multiples (provenant des utilisateurs des cellules adjacentes), sur la base du critère MMSE. De même, les auteurs distinguent un ensemble orthonormal de vecteurs de la matrice de précodage de chaque utilisateur au niveau de l'émetteur et s'assurent que le rapport signal sur interférence plus bruit est amélioré à l'extrémité réceptrice de chaque utilisateur [126]. Le module commandé MMSE a également été montré pour obtenir des performances proches du ML avec une complexité réduite pour les systèmes MIMO de modulation spatiale généralisée où des antennes sélectionnées sont utilisées pour la transmission dans chaque intervalle de temps [127]. Il est important de souligner que les études précitées ne tiennent pas compte de la performance MMSE Block sur les récepteurs SU-MIMO avec des codes pratiques LTE qui s'alignent étroitement avec les spécifications 3GPPs. Le principe sous-jacent de la technique MMSE Block suppose que l'interférence est Gaussien et peut donc être négligée. De plus, le MMSE de bloc est représenté comme étant optimal en termes de taux maximum théorique possible en présence d'un interférent à double flux, sous l'hypothèse du signal Gaussian.

7.3.2 Modèle de système HetNet limité aux interférences

Cette section présente le modèle de système pour le schéma QR Bloc dans un scénario HetNet MIMO. Un scénario interférences entre les cellules pour un terminal femto mobile (UE) est considéré. Ce modèle particulier (comme indiqué sur la Figure 7.1) consiste en un scénario de liaison descendante dans lequel l'UE est équipé de quatre antennes de réception et est situé sur le limite de cellule, ce qui le rend très vulnérable aux interférences provenant de la macrocellule voisine. L'UE reçoit des transmissions simultanées à partir de deux stations de base (eNB) équipées chacune de deux antennes, dont l'une transmet le signal souhaité (femto-eNB), tandis que la macrocellule voisine eNB est considérée comme la source interférant co-canal dominante. Inversement, une femto-eNB proximale pourrait également être une source d'interférence pour une macro UE. Ce scénario MIMO peut également être étendu à un système 8×8 (tel que supporté par la norme LTE-A (Rel.10)) impliquant le déploiement d'un noeud relais équipé de huit antennes par opposition à un seul UE. Le modèle de signal reçu est la suivante:

$$\mathbf{y}_F = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{7.1}$$

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix}, \tag{7.2}$$

Le signal reçu est donné par $\mathbf{y}_F = (\mathbf{y}_1\, \mathbf{y}_2)^T$, coefficients complexes pour canal être représentée par $(\mathbf{H}_{11}\, \mathbf{H}_{12}) \in \mathbb{C}^{N_{T_1} \times N_R}$ and $(\mathbf{H}_{21}\, \mathbf{H}_{22}) \in \mathbb{C}^{N_{T_2} \times N_R}$, N_R représente le nombre d'antennes de réception, N_{T1} et N_{T2} désignent le nombre d'antennes d'émission à chaque eNB. Les vecteurs de bruit sont des variables Gaussien complexes de symétrie circulaire avec une moyenne nulle ($\mathbf{n}_1$ and $\mathbf{n}_2$).

7.3.3 Block QR Decomposition

Les algorithmes de blocage tendent généralement à travailler sur plusieurs colonnes et lignes simultanément et ont servi de moyen efficace pour effectuer des calculs matriciels sur des structures de mémoire hiérarchiques existantes [128]. Une factorisation QR séquentielle récursive a été développée, qui divise la matrice cible en une série de blocs et a été validée par l'induction mathématique [129, 130]. Selon À [129], les transformations élémentaires des ménages sont utilisées pour factoriser

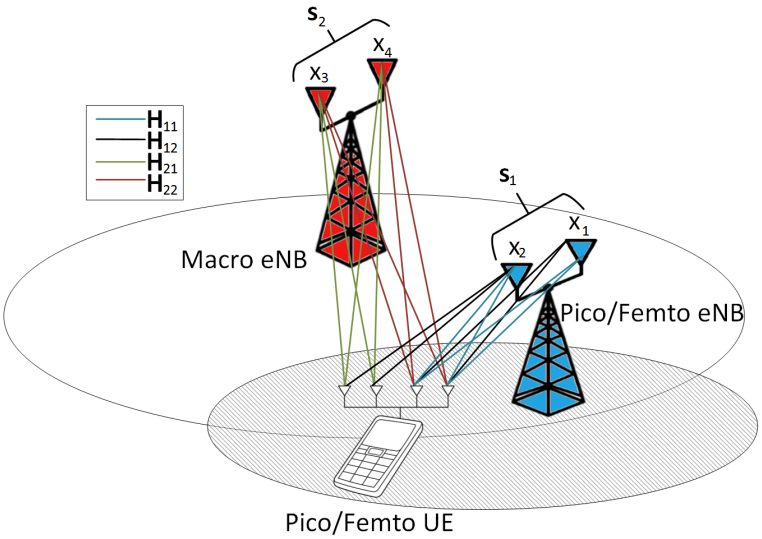


Figure 7.1: Modèle HetNet de système.

la dernière colonne restante et, dans le processus, arrête automatiquement la récursivité. Le signal reçu transformé basé sur l'équation (7.2) est exprimé comme suit:

$$\mathbf{y}_F = \mathbf{Q}_1\mathbf{Q}_2\mathbf{R}\mathbf{s} + \mathbf{n}, \tag{7.3}$$

$\mathbf{Q}_1$ et $\mathbf{Q}_2$ sont des matrices orthogonales et sont divisées en blocs comme suit:

$$\mathbf{Q}_1 = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix}; \mathbf{Q}_2 = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{Q}}_2 \end{pmatrix}, \tag{7.4}$$

La matrice triangulaire supérieure $\mathbf{R}$ est également partitionnée en bloc comme suit:

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \mathbf{R}_{22} \end{pmatrix}. \tag{7.5}$$

Le signal reçu transformé est:

$$\tilde{\mathbf{y}}_F = \tilde{\mathbf{H}}\mathbf{s} + \tilde{\mathbf{n}}, \tag{7.6}$$

et $\tilde{\mathbf{y}}_F = \mathbf{Q}_1^T \mathbf{y}_F$, $\tilde{\mathbf{H}} = \mathbf{Q}_2 \mathbf{R}$, $\tilde{\mathbf{n}} = \mathbf{Q}_1^T \mathbf{n}$ et aussi $\tilde{\mathbf{Q}}_2 = \mathbf{R}_{22} \tilde{\mathbf{H}}_{22}$, étant donné que:

$$\mathbf{Q}_1^T \mathbf{H} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \tilde{\mathbf{H}}_{22} \end{pmatrix}. \tag{7.7}$$

7.3.4 Résultats Numériques

L'efficacité spectrale maximale réalisable pour les différentes valeurs de SNR (fonctions de distribution cumulative (CDF)) est discutée pour les techniques théoriques, QR Bloc, MMSE Bloc et MMSE conventionnelles pour les systèmes MIMO d'ordre supérieur limitées aux perturbations. Figure 7.2 et Figure 7.3 représentent les courbes CDF théoriques pour la première détection de flux du modèle 4×4 et 8×8 MIMO. Ceci est représenté par l'équation suivante:

$$I(\mathbf{s}_2; \tilde{\mathbf{y}}_2 | \mathbf{H}) \leq I(\mathbf{s}_2; \mathbf{y}_F | \mathbf{H}). \tag{7.8}$$

Pour une probabilité d'interruption de 10^{-1} (ciblé taux d'erreur de bloc pour la norme LTE [4,102]) et à un SNR de 15 dB la différence de taux réalisable pour les 4×4 entre le QR Bloc et le cas théorique est 0.55 bits/s/Hz, alors qu'à un SNR supérieur de 25 dB, la différence se réduit à 0.3 bits/s/Hz. En conséquence, le schéma QR Bloc approche la limite théorique dans le régime SNR plus élevé pour le premier flux, tandis que le taux MMSE en présence d'interférence se dégrade de façon significative, ce qui est particulièrement important dans le cas 8×8 comme le montre la Figure 7.3. Il faut noter que les statistiques du canal sont Rayleigh (décoloration de bloc) avec une allocation de puissance égale à l'émetteur. De plus, la méthode de décomposition QR de Gram-Schmidt a été sélectionnée pour la mise en œuvre de l'algorithme Bloc, de sorte que le taux maximal atteint peut être noté sur la Figure 7.2 et 7.3, bien que les techniques Householder et Givens QR affichent la même performance et ne diffèrent que par leur complexité. Ces résultats numériques sont ensuite étudiés par rapport à la mise en œuvre dans un banc d'essai de simulation LTE, constitué d'une chaîne d'émission et de réception pleine bande de base, basée sur les bibliothèques logicielles du simulateur OpenAirInterface [28]. Cela impliquerait une évaluation comparative de la performance de la simulation de liaison descendante de la couche physique LTE du détecteur de Max-log-MAP à faible complexe [27, 91] et de la détection de la force brute de la sous-optimale Max-log-MAP detector [85]. L'évaluation pratique

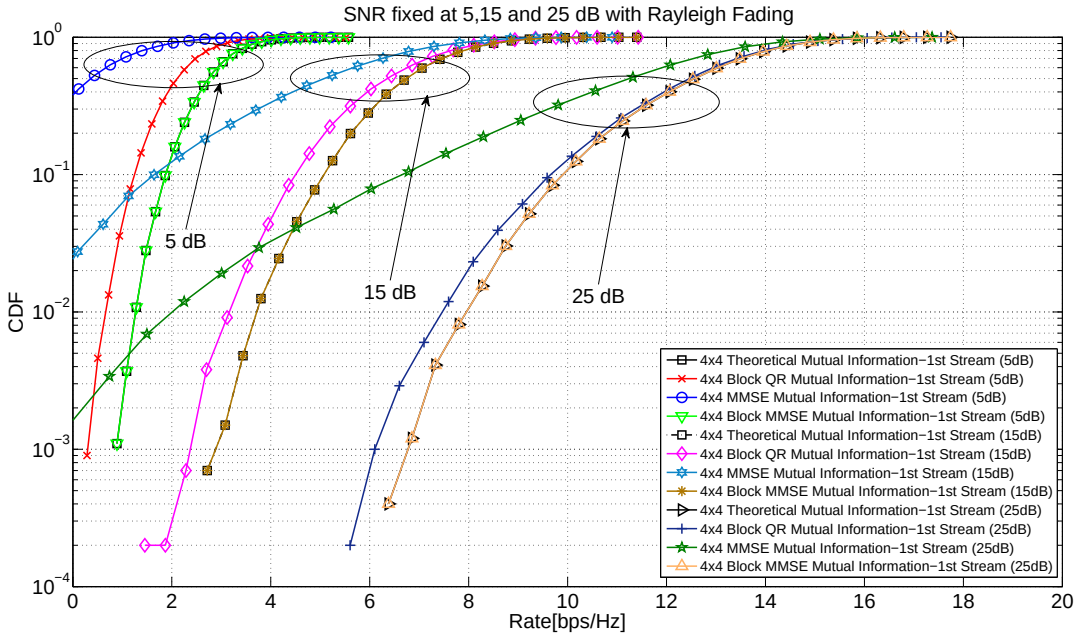


Figure 7.2: Perte d'information mutuelle à différents SNR pour le modèle MIMO 4×4.

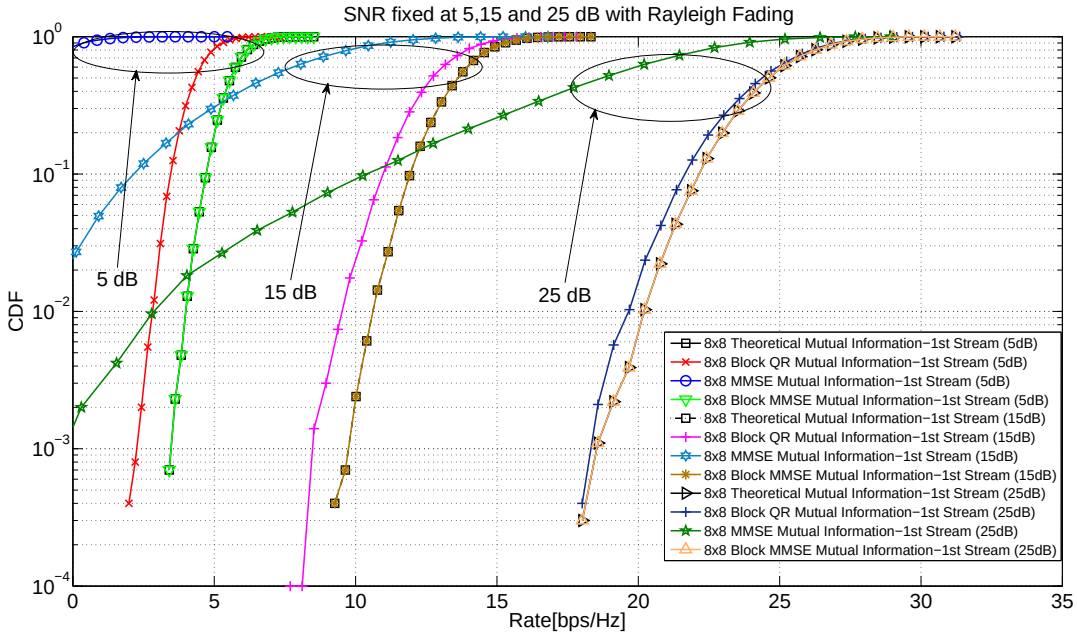


Figure 7.3: Perte d'information mutuelle à différents SNR pour le modèle MIMO 8×8.

d'un modèle MIMO 8×8 sort de la portée de cette étude parce qu'une nouvelle métrique Max-log-MAP décision souple doit être développée pour la détection d'un flux à quatre dimensions.

7.3.5 Total Taux atteignable pour un seul utilisateur MIMO

Dans cette section, l'objectif est de maximiser le taux global dans un système MIMO seul utilisateur point à point 4×4 et 8×8 en considérant la détection des deux flux transmis ($\mathbf{s}_1$ et $\mathbf{s}_2$). Dans LTE, un UE ne peut être configuré que pour recevoir un maximum de seulement deux mots de code dans un intervalle de transmission unique. Chacun des deux mots de code, R_1 et R_2 peut être maximisé pour donner un taux maximum total réalisable pour un utilisateur unique. La règle de la chaîne d'information mutuelle pour les alphabets de signaux Gaussien pour les systèmes MIMO 4×4 et 8×8 est donnée comme suit:

$$I(\mathbf{s}_1\mathbf{s}_2;\mathbf{y}|\mathbf{H}) = I(\mathbf{s}_1;\mathbf{y}|\mathbf{H}) + I(\mathbf{s}_2;\mathbf{y}|\mathbf{H},\mathbf{s}_1), \tag{7.9}$$

Le premier flux détecté est $\mathbf{s}_1$, $\mathbf{s}_2$ est le flux sans interférence (après application de successif et d'annulation d'interférence) et $\mathbf{H} = [\mathbf{h}_1\ \mathbf{h}_2]$, $\{\mathbf{h}_1, \mathbf{h}_2\} \in \mathbb{C}^{4\times 2}$, $\{\mathbf{s}_1, \mathbf{s}_2\} \in \mathbb{C}^{2\times 1}$ et $\mathbf{n} \in \mathbb{C}^{4\times 1}$ représente le vecteur AWGN. Le taux total convenable (R_T) pour un récepteur MIMO où les deux mots de code doivent être décodés conjointement est donné comme suit:

$$T_{Joint}(R_T) = \max_{R_T} \left[R_T\left(1 - P_{out}(R_T)\right)\right], \tag{7.10}$$

$$= \max_{R_1,R_2} \left[\left(R_1 + R_2\right)\left(1 - P_{out}(R_T)\right)\right], \tag{7.11}$$

P_{out} représente la probabilité de panne à un taux cible R_T. La stratégie suivante est le schéma de décodage SIC commun, dans lequel, pour atteindre le taux total maximal souhaité, les deux mots de code devraient être décodés avec succès et donc une défaillance de décodage (panne). Par conséquent, $P_{out}(R_1) = 1$ de l'un ou l'autre des mots de code provoquera l'interruption du système global.

$$T_{SIC} = \max_{R_T} \left[\left(R_T\right)\left(1 - P_{out}(R_1)\right)\left(1 - P_{out}(R_2)\right)\right]. \tag{7.12}$$

Le taux total réalisable ($T_{LTE}(R_T)$) dans LTE est donné comme suit:

$$T_{LTE}(R_T) = \max_{R_1,R_2} \left[\overbrace{R_1(1-P_{out}(R_1))}^{1st\ Codeword} + \overbrace{R_2(1-P_{out}(R_2))(1-P_{out}(R_1))}^{2nd\ Codeword} \right] \tag{7.13}$$

Le débit réalisable dans les équations (7.11), (7.12) et (7.13) est représenté sur les Figures 7.4 et 7.5 pour un scénario 4×4 et 8×8 pour des canaux sujets aux évanouissements de Rayleigh.

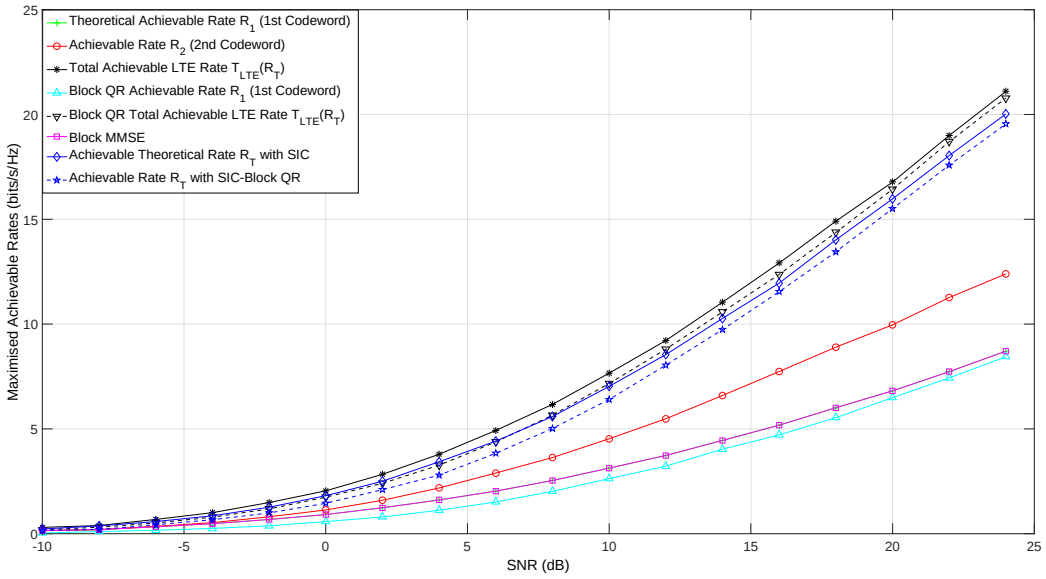


Figure 7.4: Taux réalisable en fonction de SNR dans un canal 4×4 Rayleigh

7.4 CHAPITRE 4

Dans ce chapitre, les schémas de détection de bloc proposés sont évalués dans un contexte pratique, ayant posé les bases de la performance algorithmique au niveau analytique dans le Chapitre 3. La complexité est étudiée en termes de nombre d'opérations à virgule flottante, ce qui est un atout important du point de vue matériel. La performance en termes de taux d'erreur de bloc (BLER) est également évaluée à l'aide du banc d'essai de simulation LTE développé constitué de blocs fonctionnels pour la chaîne d'émission et de réception en bande de base, basés sur les bibliothèques de logiciels OpenAirInterface. La Figure 7.6 est une comparaison de la complexité entre la technique QR

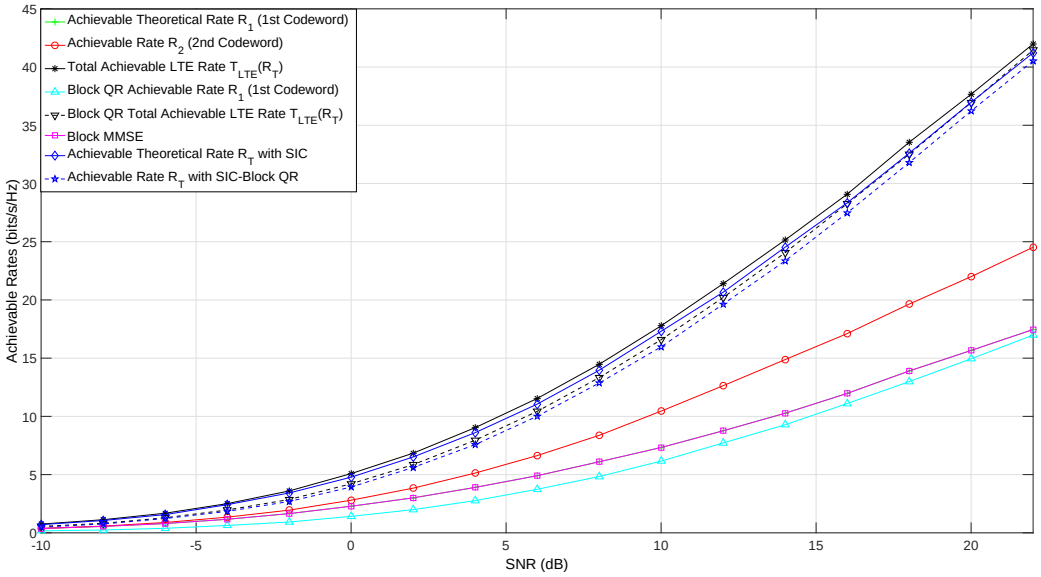


Figure 7.5: Taux réalisable en fonction de SNR dans un canal 8×8 Rayleigh

de bloc et le calcul EVD MMSE Bloc, où $N_T = L$ et $N_R = M$ représentent respectivement le nombre d'antennes d'émission et de réception. La technique QR de séquence MGS nécessite environ 64% moins de flops que le calcul EVD de la technique MMSE Bloc dans le cas 4×4, ce qui représente une diminution significative de la complexité, en partie due à l'opération d'inversion coûteuse requise par la technique MMSE Bloc. Les performances du taux d'erreur de bloc (BLER) ont été évaluées pour les régimes de SNR modérés à supérieurs pour les algorithmes de la force brute Max-log-MAP et du récepteur de QR Bloc dans un scénario LTE. La Figure 7.7 montre la performance LTE correspondante pour les signaux QPSK (MCS=9) pour les deux démodulateurs pratiques.

7.5 CHAPITRE 5

Les avantages des communications coopératives, qui permettent aux terminaux à antenne unique de fonctionner en coopération avec les gains d'efficacité et de diversité habituellement réservés aux systèmes à antennes multiples, ont été largement étudiés [143]. Les différentes techniques de communication coopérative et les stratégies de relais disponibles dans la littérature sont largement basées sur le travail de théorie de l'information séminale de Cover et El Gamal [103]. Ces avancées ont conduit à des études techniques d'architectures de relais pratiques par 3GPP pour inclusion dans la

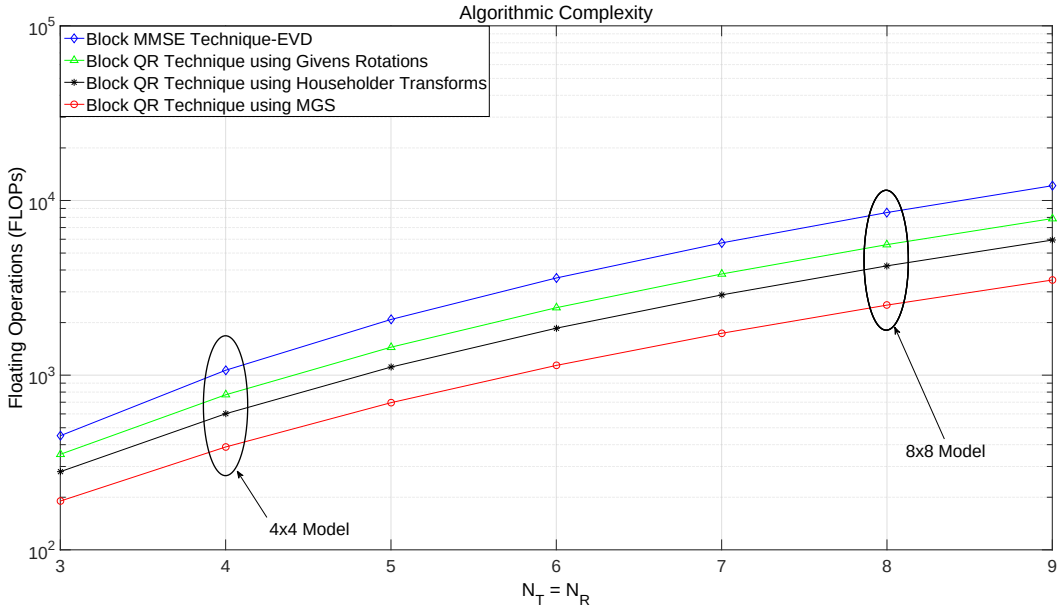


Figure 7.6: Comparaison de la complexité entre les différents algorithmes QR Bloc et MMSE EVD bloc

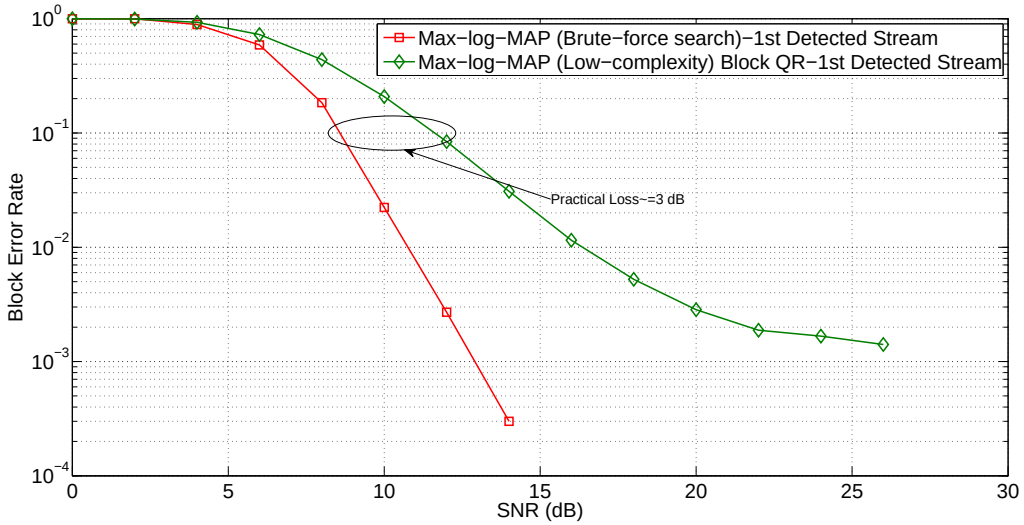


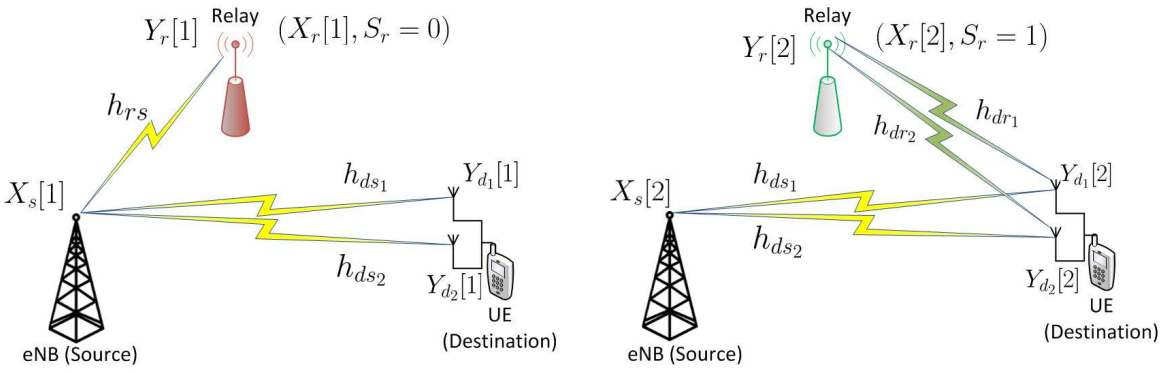
Figure 7.7: BLER de l'algorithme de recherche Max-log-MAP à faible complexe QR Bloc et à force brute Max-log-MAP utilisant des codes pratiques LTE.

norme LTE Release 9 [104, 144], qui a ensuite conduit à l'incorporation des opérations de noeud de relais de couche physique dans les spécifications LTE Release 10 (LTE-Advanced) [145]. Dans ce

chapitre, une mise en œuvre pratique de la nouvelle stratégie relais pour le canal de relais semi-duplex Gaussien (HD-RCC) est fourni, ce qui est connu pour se trouver à l'intérieur d'un intervalle constant de la limite supérieure de découpage sur la capacité du réseau [147]. Le schéma à deux phases trois-partie-message proposé dans [147] emploie la superposition des codes de Gaussien à la source, Sucessive Interference Cancellation (SIC) à la fois au relais et à la destination et Decode -et-Forward (DF) au relais. Le modèle de canal a un lien direct entre la source et la destination, par lequel la source envoie continuellement des informations vers la destination à un taux proche de la capacité de cette liaison. Dans le même temps, la source exploite le relais pour transmettre des informations supplémentaires à la destination à une vitesse qui, dans le régime SNR élevé, peut être interprétée comme la capacité minimale du relais source (pour la phase relais-écoute) et relais-destination (pour la phase relais-envoi) moins la capacité de la liaison source-destination. La durée relative des phases de relais-écoute et d'émission de relais est déterminée de sorte que la quantité d'informations décodées dans la première peut être transmise de manière fiable dans celle-ci. Le modèle de canal a un lien direct entre la source et la destination, par lequel la source envoie continuellement des informations vers la destination à un taux proche de la capacité de cette liaison. Dans le même temps, la source exploite le relais pour transmettre des informations supplémentaires à la destination à une vitesse qui est à peu près la capacité minimale du relais de source (pour la phase d'écoute de relais) et de relais de destination (pour le relais- Envoyer phase) moins la capacité de la liaison source-destination. La durée relative des phases de relais-écoute et d'émission de relais est déterminée de sorte que la quantité d'informations décodée dans la première peut être transmise de manière fiable dans celle-ci. Dans ce travail, nous étendons d'abord le modèle de [147] au cas de la destination multi-antenne et proposons ensuite une implémentation pratique conforme à la norme LTE.

7.5.1 Modèle de système de relais

Un canal de relais semi-duplex (HD-RC) se compose de trois nœuds: la source, le relais et la destination. Le message transmis source uniformément distribué pour la destination, $w \in \left[1 : 2^{NR}\right]$, la longueur du mot de code est notée N et le taux de transmission est donné par R. Au moment $i, i \in [1 : N]$, la source mappe son message w dans un symbole d'entrée de canal $X_{s,i}(w)$ et le relais lors de la transmission, ses observations de canal passées en un symbole d'entrée de canal $X_{r,i}(Y_r^{i-1})$. Le SIMO Gaussien HD-RC est représenté sur la Figure 5.1, où les trois nœuds sont l'eNodeB (source), le relais



(a) Phase I- $S_r = 0$ (b) Phase II- $S_r = 1$

Figure 7.8: Modèle de système de relais.

et l'UE (destination), équipé de $n_d = 2$ antennes. La relation entrée/sortie est:

$$Y_r = h_{rs} X_s \, (1 - S_r) + Z_r \in \mathbb{C}, \tag{7.14a}$$

$$\mathbf{Y}_d = \mathbf{h}_r X_r S_r + \mathbf{h}_s X_s + \mathbf{z}_d \in \mathbb{C}^2, \tag{7.14b}$$

Permettre $\mathbf{Y}_d = \begin{bmatrix} Y_{d_1} & Y_{d_2} \end{bmatrix}^T$, $\mathbf{h}_r = \begin{bmatrix} h_{dr1} & h_{dr2} \end{bmatrix}^T$, $\mathbf{h}_s = \begin{bmatrix} h_{ds1} & h_{ds2} \end{bmatrix}^T$ et $\mathbf{z}_d = \begin{bmatrix} Z_{d_1} & Z_{d_2} \end{bmatrix}^T$. Les paramètres de canal sont fixés pour toute la durée de transmission et supposés connus de tous les noeuds, $(h_{ds_i}, h_{dr_i}, h_{rs}), i \in [1:2]$. Le modèle est sans perte de généralité parce que des contraintes de puissance non unitaires ou des variances de bruit peuvent être incorporées dans les gains de canal. Il convient également de noter que le système conçu dans cette section ainsi que sa garantie de performance dérivée sont valables pour toute valeur de $n_d \geq 1$. Nous nous concentrerons sur $n_d = 2$ car c'est le cas considéré dans la mise en œuvre pratique. Ensuite, nous proposons une implémentation LTE pratique de la stratégie de relais HD proposée.

7.5.2 Modèle de simulation

Le schéma décrit dans Section 7.5.1 utilise quatre dictionnaires Gaussien $\mathscr{C}_{a1}$, $\mathscr{C}_{a2}$, $\mathscr{C}_b$ et $\mathscr{C}_c$ pour transmettre les trois sous-messages w_0, w_1 et w_2. Dans une mise en œuvre pratique, les codes ne seraient pas Gaussien, mais contiendront des symboles d'une constellation finie. Nous avons besoin de comprendre la performance de ces codes pratiques à la fois dans le bruit Gaussien et non Gaussien parce que SIC est employé dans les opérations de décodage du relais et de la destination. En particulier, au relais, nous devons comprendre la performance de $\mathscr{C}_b$ en bruit non-Gaussien (première

phase de décodage de Phase I) et de $\mathscr{C}_{a1}$ en bruit Gaussien; à la destination nous devons comprendre la performance de $\mathscr{C}_{a1}$, $\mathscr{C}_{a2}$ et $\mathscr{C}_b$ dans le bruit non-Gaussien et la performance du code $\mathscr{C}_c$ dans le bruit Gaussien. Dans les étages de décodage où un message est traité comme un bruit, on met en œuvre un démodulateur qui tient compte spécialement du fait que le bruit global est non-Gaussien. Nous consideronsdes choix différents pour les livres de code $(\mathscr{C}_{a1}, \mathscr{C}_{a2}, \mathscr{C}_b, \mathscr{C}_c)$. Comme pour chaque choix, on s'assure qu'un BLER à la destination est inférieur à un seuil donné pour tous les étages de décodage (afin de maintenir le système global BLER à la destination au-dessous de 10^{-2} et donc le BLER correspondant à chacune des trois opérations de décodage est fixée à au plus $1/3 \cdot 10^{-2}$). Bien que l'analyse se concentre sur le cas où la destination est équipée de deux antennes, le boîtier à antenne unique est également présenté pour l'exhaustivité. La question clé que nous cherchons à répondre est la suivante: comment l'efficacité spectrale des codes pratiques se compare-t-elle à la performance théorique? Un banc d'essai de simulation utilisant les bibliothèques logicielles OAI (plate-forme pour l'expérimentation de communication sans fil) afin d'évaluer les performances du schéma mentionné ci-dessus avec des codes pratiques est développé. La plate-forme logicielle est basée sur la norme évolutive de LTE de 3GPP qui se compose des caractéristiques essentielles d'un système pratique de communication par radio, qui s'alignent étroitement avec les normes des réseaux commercialement déployés. Les simulations ont été effectuées sur le canal partagé de liaison descendante (DL-SCH), qui est le canal principal pour transmettre des données d'utilisateur (ou informations de contrôle) de l'eNodeB (eNB) à l'UE [31]. Les messages de données sont transportés dans des unités connues sous le nom de Transport Blocks (TBs) pour transmettre les sous-messages w_0, w_1 et w_2. La taille du bloc de transport (TBS) dépend du choix du schéma de modulation et de codage (MCS), qui décrit l'ordre de modulation et le taux de codage d'une transmission particulière.

7.5.3 Évaluation de la performance: Statique canal AWGN

Selon le modèle de système dans l'équation (7.14): $|h_{rs}|^2 = C$, $|h_{ds_1}|^2 = |h_{ds_2}|^2 = S$ (les deux liens source-destination sont de la même force), $|h_{dr_1}|^2 = |h_{dr_2}|^2 = I$ (les deux liens relay-destination sont de la même force), et la phase des gains de canal est réglée sur une valeur aléatoire qui est maintenue constante pendant toute la durée de transmission ($N_1 + N_2$ utilisations des canaux). La réception parfaite CSI est assumée à tous les noeuds. Pour chacune des opérations de décodage pendant les phases I et II, les performances BLER au relais et à la destination sont validées pour différentes

Table 7.1: MCS pour chaque opération de décodage avec $I/S = 0$ dB pour un schéma SIMO.

Phase I - X_b					**Phase II** - (X_{a1}, X_{a2})		**Phase II** - X_c		**Theoretical**	**Practical**	**Theoretical BS**	**Practical BS**
MCS	S [dB]	C_{X_b} [dB]	$C_{X_{a1}}$ [dB]	TBS [bits]	MCS	TBS [bits]	MCS	TBS [bits]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]
9	1.89	5.48	9.54	4008	11	4968	13	5736	2.44	2.26	1.01	1.12
14	5.84	9.07	12.10	6456	14	6456	17	7736	3.41	3.36	1.56	1.46
20	11.02	14.60	18.2	9912	20	9912	22	11448	5.10	5.09	2.36	1.93

valeurs de C (au relais) et évaluées par rapport à S (qualité de canal pour la liens source-destination). Le $\delta = \frac{1}{1+2S}$ varaible est le paramètre de superposition dans le scénario SIMO. Une répartition égale de la largeur de bande entre les deux phases de la stratégie est également supposée, c'est-à-dire $\gamma = 0.5$.

Décodage au relais A la fin de la Phase I, le relais décode d'abord w_1, puis il le retire de son signal reçu et finalement décode w_0, dans un lien sans interférences. La propagation d'erreur, qui résulte de l'alimentation de symboles incorrectement décodés, n'est pas considérée ici. La raison est que nous considérons un système codé avec des bits CRC (qui est implémenté dans les récepteurs 4G d'aujourd'hui) pour lequel l'erreur résiduel (erreur de détection) à la sortie se produit avec une très faible probabilité (par exemple 10^{-9}). Le canal source-relais est supposé être assez fort pour garantir un système BLER inférieur à 10^{-3} au relais (par exemple, voir les troisième et quatrième colonnes du Tableau 7.1). En d'autres termes, pour avoir des opérations de décodage réussies (à la fois w_0 et w_1) au relais à la fin de Phase I, on suppose une C telle que $\text{BLER} \leq \text{BLER}[X_b] + \text{BLER}[X_{a1}] \leq 5 \cdot 10^{-4} + 5 \cdot 10^{-4} = 10^{-3}$. Il est montré (voir les troisième et quatrième colonnes de Tableau 7.1) que la liaison source-relais est plus forte que la liaison source-destination en raison de la proximité LOS entre la source et le relais. Dans cette hypothèse, il convient de noter que le relais décode avec succès le message w_0 avec $\text{BLER}[X_{a1}] \leq 5 \cdot 10^{-4}$, qui est une contrainte plus stricte que celle imposée à la destination. Enfin, au début de la Phase II, le message w_0 est recodé et transmis par le relais à la destination.

Décodage à la destination Il est intéressant de noter que dans notre modèle, la liaison relais-destination est supposée être plus forte que la liaison source-destination de sorte que l'utilisation du relais augmente effectivement la performance de la vitesse par rapport à la transmission directe. L'efficacité spectrale globale de la stratégie de relais HD à la fin de la Phase II est montrée dans la Figure 7.9 et la Figure 7.10 (voir les courbes bleues). Chacun des marqueurs sur les courbes bleues-

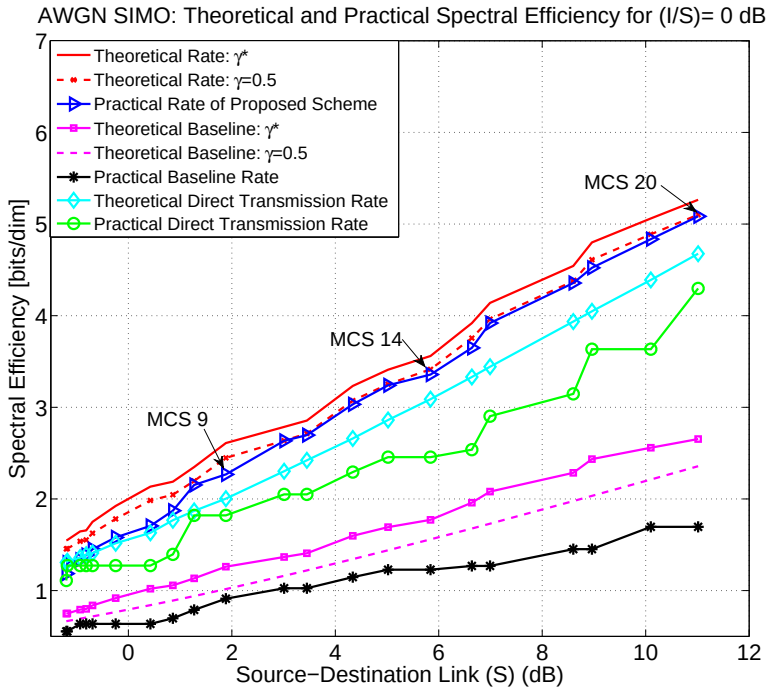


Figure 7.9: AWGN SIMO comparaison entre théorique et pratique efficacités spectrales pour $(\mathrm{I}/\mathrm{S}) = 0$ dB

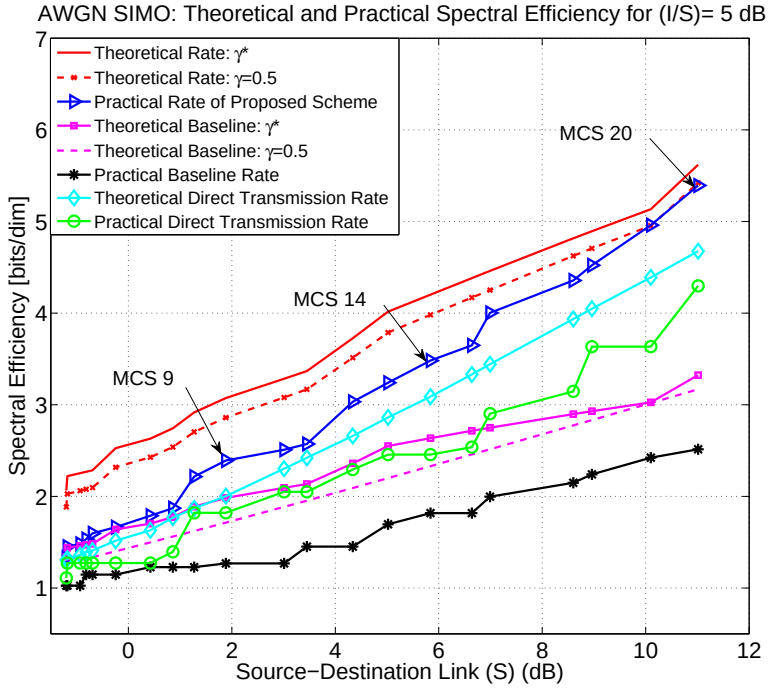


Figure 7.10: AWGN SIMO comparaison entre théorique et pratique efficacités spectrales pour $(\mathrm{I}/\mathrm{S}) = 5$ dB

taux pratiques atteints par le schéma représente un MCS qui varie en total de 0 à 20. Chaque MCS caractérise le type de transmission de signal À-dire QPSK, 16-QAM ou 64-QAM) avec un certain taux de code. Les taux sont représentés graphiquement pour deux rapports différents entre le lien relais-destination et le lien source-destination pour $I/S = 0$ dB (voir Figure 7.9) et $I/S = 5$ dB (voir Figure 7.10). Cela montre que l'utilisation d'une liaison relais-destination qui est plus forte que la liaison source-destination augmente la performance de la vitesse par rapport à la transmission directe (voir les courbes vertes) ou aux systèmes de relais conventionnels (voir les courbes noires) stratégie.

7.6 CONCLUSION

Cette thèse a étudié les techniques liées à la zone d'ordre supérieur MIMO bloc de détection du récepteur et des réseaux de relais HD, avec une attention vers la mise en œuvre dans les futurs réseaux 4G + et 5G. L'une des motivations de cette recherche était la viabilité de ces techniques dans les réseaux LTE d'aujourd'hui ou les réseaux de prochaine génération. Chapitre 1 donne un aperçu initial de la cinquième génération de réseaux mobiles, ainsi que le rôle de HetNets dans la densification du réseau. Le domaine de problème est défini en examinant les deux défis clés de recherche étudiés dans cette étude. Le premier défi est la conception d'un ordre supérieur récepteurs MIMO dans des scénarios d'interférence limitée, qui sera une conséquence inévitable des réseaux denses. Par conséquent, les compromis entre la complexité et la performance des mérites essentiels à prendre en considération. Le deuxième défi de recherche fournit un cas d'utilisation pour le déploiement de relais dans les réseaux LTE. Les opérateurs de télécommunications ont montré un manque d'intérêt en raison des caractéristiques d'efficacité spectrale limitées de relais HD traditionnels et de la partie II de cette recherche a démontré le contraire, en démontrant la faisabilité d'une stratégie HD relais avancée dans un scénario de déploiement LTE.

Chapitre 2 a présenté un aperçu des différents aspects liés à la détection MIMO, de gestion des interférences, des réseaux de relais, ainsi que la norme de communication LTE fond. Ces thèmes représentent l'état de la art collectif et servent comme un précurseur de la contribution de cette étude.

7.6.1 Résumé des résultats

Partie I de la thèse a étudié l'utilisation d'une technique QR Bloc pour effectuer à faible complexité d'ordre supérieur détection MIMO dans un scénario et un taux de maximisation d'interférence limitée dans un cadre LTE SU-MIMO de point à point. Une analyse théorique de l'information et de l'évaluation pratique de la technique de détection QR Bloc est présenté dans un aperçu de gain de l'ordre dans la performance algorithmique et de la complexité.

Les taux réalisables maximaux pour la première détection de flux en utilisant un schéma QR Bloc roman pour 4×4 et 8×8 scénarios d'interférence limitée a révélé une perte de l'information mutuelle de 0.55 bits/s/Hz et 0.98 bits/s/Hz, respectivement, à un SNR de 15 dB en utilisant un (Rayleigh) modèle de canal plat décoloration avec une panne de 10%. Selon ces observations, cette perte semble augmenter avec un réglage de MIMO d'ordre supérieur, et diminuer en conséquence lors du passage à un régime de SNR supérieur. La détection MMSE Bloc du premier flux a été validé pour être sans perte d'informations et est donc considérée comme la détection optimale lorsque vous effectuez la première détection de flux.

Un schéma d'optimisation de tarifs pour un lien SU-MIMO point-à-point est également étudié en utilisant la technique QR Bloc, afin de maximiser le taux global. Le taux de système est maximisé sur la transmission des deux mots de code comme dans LTE, R_1 et R_2 pour le montant de 4×4 et 8×8 cas. Par conséquent, intuitivement la perte de vitesse totale du régime QR Bloc a été montré pour résulter que de la perte de détection se produisant dans le premier courant. La perte de vitesse minimale dans le régime de SNR pratique démontre la faisabilité théorique de maximiser le taux global en utilisant le schéma QR Bloc dans le contexte d'un scénario SU-MIMO.

La prochaine phase de l'étude était d'évaluer la complexité et la performance du QR Bloc dans un banc d'essai de simulation LTE afin d'obtenir un aperçu significatif de la performance de BLER dans un canal de Rayleigh. Pour le cas 4×4 , il a été observé que le régime MMSE Bloc était environ deux fois plus complexe que l'algorithme MGS QR Bloc et pour le cas 8×8 la complexité a quadruplé le nombre de flops requis. Une force brute optimale, mais très complexe Max-log-MAP détecteur est également mis en œuvre comme une référence de performance contre l'algorithm susmentionnés. Les résultats montrent que la perte de 3 dB performance est obtenue entre la force brute de détecteur Max-log-MAP et QR Bloc détecteur Max-log-MAP, ce qui correspond à peu près à la même perte dans le cas de la probabilité de coupure théorique (2.5 dB). Ceci valide l'idée que la performance

théorique et pratique sont comparables.

Partie II de la thèse donne un aperçu de la faisabilité d'un lien stratégie d'adaptation de la couche physique coopérative pour les relais HD, qui tire parti de codage de superposition et de SIC décodage pour transmettre un trois-part-message sur un seul berceau LTE. L'un des objectifs de cette partie de la recherche, est de combler le fossé entre la théorie et la pratique dans le contexte de la longueur et de signal hypothèses codeword. En tant que tel, une comparaison est faite entre le taux réalisable théorique avec les codes de Gaussien et le taux obtenu dans un scénario pratique pour un BLER de système global de 10^{-2} pour une destination (UE) équipée d'une antenne simple et double configuration. Ces résultats ont fourni une compréhension sur les avantages de la coopération de la couche physique par rapport aux systèmes de relais à deux sauts existants et des liaisons point-à-point. La stratégie de relais a été évaluée pour être robuste et à condition d'améliorer les gains d'efficacité spectrale, en utilisant deux modèles de canaux LTE pratiques.
Les résultats présentés dans cette étude de faisabilité de renforcer le cas pour les futurs déploiements de relais 5G HD où les avantages de la couche physique relais peuvent être réalisés pour apporter des améliorations d'efficacité spectrale. Cela peut se révéler être encore plus efficace dans les déploiements à ondes millimétriques où principalement la ligne de mire des liens sont nécessaires et points d'accès filaires seront extrêmement coûteux pour le déploiement à grande échelle. Dans ce cas, quelques donateurs eNBs avec une dorsale optique au réseau de base et de nombreuses stations de relais sans fil en utilisant la couche physique relayer leur proximité peuvent être envisagées. Pour UEs dans la ligne de mire à la fois avec la destination et le relais, il y aura des avantages substantiels tels que décrits par ces résultats. Les résultats de cette étude représentent un catalyseur fort pour les relais HD à déployer dans les prochaines versions de la norme LTE ou encore, des systèmes 5G qui pourraient exploiter ces architectures de relais avancées.

7.7 RECOMMANDATIONS POUR LES TRAVAUX FUTURS

Une analyse théorique de l'information qui se déplace loin de l'hypothèse de brouilleurs Gaussien à discrets alphabets serait considérée comme une prochaine étape dans la détermination de la performance analytique du système QR Bloc. Cela est particulièrement applicable au régime de

maximisation du taux, où il pourrait aider à comprendre les taux (MCS) au cours de laquelle chacun des mots de code peut être transmis pour atteindre un débit maximal à un SNR particulier. Une future orientation de la recherche à l'égard de l'évaluation pratique des schémas de détection de bloc, comprendrait une analyse de la complexité de point fixe afin de comprendre les implications matérielles tels que des pertes de précision. Une limitation de la comparaison de BLER pour inclure uniquement les signaux QPSK, à savoir MCS=9, était le temps de calcul nécessaire pour évaluer la force brute du détecteur Max-log-MAP pour 16-QAM et la signalisation 64-QAM. Un moyen efficace de réaliser les comparaisons d'ordre supérieur constellations serait également recommandé que les travaux futurs.

En ce qui concerne la stratégie de relais HD avancé, un objet d'enquêtes futures serait l'analyse du projet en trois parties message lorsque les informations d'état de canal ne sont pas disponibles et donc doit être estimée / apprises ainsi que l'enquête de la ressource stratégies d'allocation pour l'affectation dynamique de bande passante. Dans la Partie II de cette étude, une durée égale des deux phases et une valeur fixe pour le facteur de superposition a été considéré; pour l'inclusion dans les systèmes en temps réel, ces paramètres doivent être adaptatifs et ajusté en conséquence. Plus dimensionnement général des ressources peut être faite sur tours H-ARQ. configurations MIMO d'ordre supérieur peuvent être considérés au relais et la source ainsi. Une analyse du taux réalisable, étant donné une contrainte d'arrêt dans un scénario fading serait également présenter une avenue intéressante pour la recherche future.

REFERENCES

[1] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 22–30, Jun. 2011.

[2] I. Medvedev, B. A. Bjerke, R. Walton, J. Ketchum, M. Wallace, and S. Howard, "A Comparison of MIMO Receiver Structures for 802.11N WLAN - Performance and Complexity," in *IEEE Intl. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2006, pp. 1–5.

[3] European Telecommunication Standards Institute, "Physical Layer for relaying operation (Rel. 13)," ETSI, Tech. Rep. ETSI TS 136 213 V13.0.0, May 2016.

[4] S. Ahmadi, *LTE-Advanced: A practical systems approach to understanding LTE releases 10 and 11 radio access technologies*, 1st ed. ELSEVIER, 2014.

[5] GSMA, "The Mobile Economy 2016," Jan. 2016, White Paper, pp. 1–56. [Online]. Available: http://www.gsmamobileeconomy.com

[6] H. Holma, A. Toskala, and J. Reunanen, *LTE Small Cell Optimization: 3GPP Evolution to Release 13*, 1st ed. Wiley, 2016.

[7] S. Singh, H. S. Dhillon, and J. G. Andrews, "Offloading in Heterogeneous Networks: Modeling, Analysis, and Design Insights," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2484–2497, May 2013.

[8] C. F. Lanzani, G. Kardaras, and D. Boppana, “Remote radio-heads and the evolution towards 4G networks,” Feb. 2009, Altera Corporation-White Paper, pp. 1–5.

[9] Fujitsu, “High Capacity Indoor Wireless Solutions: Picocell or Femtocell?” Last Accessed Sept. 2016, White Paper, pp. 1–9. [Online]. Available: https://www.fujitsu.com/us/Images/High-Capacity-Indoor-Wireless.pdf

[10] M. Condoluci, M. Dohler, G. Araniti, A. Molinaro, and K. Zheng, “Toward 5G densenets: architectural advances for effective machine-type communications over femtocells,” *IEEE Commun. Mag.*, vol. 53, no. 7, pp. 134–141, Jan. 2015.

[11] J. Wannstrom and K. Mallinson, “Heterogeneous Networks in LTE,” Last Accessed Sept. 2016 2013, Article. [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/1576-hetnet

[12] W. C. Y. Lee, “Overview of cellular CDMA,” *IEEE Trans. Veh. Technol.*, vol. 40, no. 2, pp. 291–302, May 1991.

[13] M. S. Alouini and A. J. Goldsmith, “Area spectral efficiency of cellular mobile radio systems,” *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1047–1066, Jul. 1999.

[14] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, “Femtocells: Past, Present, and Future,” *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.

[15] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, “Femtocell networks: A Survey,” *IEEE Trans. Wireless Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sept. 2008.

[16] M. Yavuz, F. Meshkati, S. Nanda, A. Pokhariyal, N. Johnson, B. Raghothaman, and A. Richardson, “Interference Management and Performance Analysis of UMTS/HSPA+ Femtocells,” *IEEE Commun. Mag.*, vol. 47, no. 9, pp. 102–109, Sept. 2009.

[17] H. Claussen, L. T. W. Ho, and L. G. Samuel, “Financial Analysis of a Pico-Cellular Home Network Deployment,” in *IEEE Intl. Conf. on Commun. (ICC)*, Jun. 2007, pp. 5604–5609.

[18] Third Generation Partnership Project, “Overview of 3GPP Release 9 ,” 3GPP, Tech. Rep. Ver 0.3.4, Sept. 2014.

[19] ——, “Small cell enhancements for E-UTRA and E-UTRAN-Physical Layer aspects (Release 12),” 3GPP, Tech. Rep. Ver 12.1.0, Dec. 2013.

[20] C. Peng, Y. Li, Z. Li, J. Zhao, and J. Xu, “Understanding and diagnosing real-world Femtocell performance problems,” in *IEEE Infocomm*, Apr. 2016, pp. 1–9.

[21] A. Damnjanovic *et al.*, “A survey on 3GPP heterogeneous networks,” *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 10–21, Jun. 2011.

[22] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, “Millimeter Wave Beamforming for Wireless Backhaul and Access in Small Cell Networks,” *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[23] T. S. Rappaport *et al.*, “Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!” *IEEE Access*, vol. 1, pp. 335–349, Aug. 2013.

[24] T. Nakamura, S. Nagata, A. Benjebbour, Y. Kishiyama, T. Hai, S. Xiaodong, Y. Ning, and L. Nan, “Trends in small cell enhancements in LTE advanced,” *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 98–105, Feb. 2013.

[25] S. Chia, M. Gasparroni, and P. Brick, “The next challenge for cellular networks: backhaul,” *IEEE Microw. Mag.*, vol. 10, no. 5, pp. 54–66, Aug. 2009.

[26] R. Thomas., R. Knopp, B. Maharaj, and L. Cottatellucci, “Detection using block QR decomposition for MIMO HetNets,” in *48th Asilomar Conference on Signal, Systems and Computers*, Nov. 2014, pp. 1291–1295.

[27] R. Ghaffar and R. Knopp, “Interference Suppression for Next Generation Wireless Systems,” in *IEEE VTC Spring 2009.*, Apr. 2009, pp. 1–5.

[28] F. Kaltenberger, “OpenAirInterface Project,” Eurecom, Last accessed Sept. 2016, ht-

tps://gitlab.eurecom.fr/oai/openairinterface5g/wikis/home.

[29] N. Networks, “LTE-Advanced The advanced LTE toolbox for more efficient delivery of better user experience,” Last Accessed Sept. 2016, White Paper, pp. 1–19. [Online]. Available: http://resources.alcatel-lucent.com/asset/200177

[30] E. Lang, S. Redana, and B. Raaf, “Business Impact of Relay Deployment for Coverage Extension in 3GPP LTE-Advanced,” in *IEEE ICC Workshops 2009*, Jun. 2009, pp. 1–5.

[31] S. Sesia, I. Toufik, and M. Baker, *LTE The UMTS Long Term Evolution: From Theory to Practice*, 2nd ed. Wiley, 2011.

[32] Y. Polyanskiy, “Channel coding: non-asymptotic fundamental limits,” Ph.D. dissertation, Princeton University, Nov. 2010.

[33] H. Zhang, S. Chen, X. Li, H. Ji, and X. Du, “Interference management for heterogeneous networks with spectral efficiency improvement,” *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 101–107, Apr. 2015.

[34] A. Goldsmith, *Wireless Communications*, 1st ed. Cambridge University Press, 2004.

[35] S. V. Hanly, “An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity,” *IEEE J. Sel. Areas in Commun.*, vol. 13, no. 7, pp. 1332–1340, Sept. 1995.

[36] D. Lopez-Perez, X. Chu, A. V. Vasilakos, and H. Claussen, “Power Minimization Based Resource Allocation for Interference Mitigation in OFDMA Femtocell Networks,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 333–344, Apr. 2014.

[37] F. J. Martin-Vega, G. Gomez, M. C. Aguayo-Torres, and M. D. Renzo, “Analytical Modeling of Interference Aware Power Control for the Uplink of Heterogeneous Cellular Networks,” *IEEE Trans. on Wireless Commun.*, vol. 15, no. 10, pp. 6742–6757, Apr. 2016.

[38] Y.-W. Kuo, L.-D. Chou, and Y.-M. Chen, “Adaptive Smart Power Control Algorithm for LTE

downlink cross-tier interference avoidance," in *IEEE VTC-Fall*, Aug. 2015, pp. 24–30.

[39] A. M. Voicu, L. Simic, and M. Petrova, "Interference mitigation in two-tier LTE networks: Does power control pay off for femtocells?" in *IEEE Intl. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2014, pp. 1332–1337.

[40] H. Venkataraman and G.-M. Muntean, *Cognitive Radio and its Application for Next Generation Cellular and Wireless Networks*, 1st ed. Springer, 2012.

[41] Third Generation Partnership Project, "Carrier aggregation enhancements: User Equipment (UE) and Base Station (BS) radio transmission and reception (Release 11)," 3GPP, Tech. Rep. 3GPP TR 36.823 v11.0.1, Sept. 2013.

[42] R. Ghaffar, "Interference Mitigation in Multi-Antenna Systems," Doctor of Philosophy: Communications and Electronics, Telecom Paristech, Biot, France, Nov. 2010.

[43] D. Chen, T. Jiang, and Z. Zhang, "Frequency Partitioning Methods to Mitigate Cross-Tier Interference in Two-Tier Femtocell Networks," *IEEE Trans. on Veh. Technol.*, vol. 64, no. 5, pp. 1793–1805, May 2015.

[44] L. Huang, G. Zhu, and X. Du, "Cognitive femtocell networks: an opportunistic spectrum access for future indoor wireless coverage," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 44–51, Apr. 2013.

[45] K. Hosseini, W. Yu, and R. S. Adve, "Large-Scale MIMO Versus Network MIMO for Multicell Interference Mitigation," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 930–941, Oct. 2014.

[46] Third Generation Partnership Project, "Summary of the description of candidate eICIC solutions," 3GPP, Tech. Rep. 3GPP TSG-WG1 R1-104968, Aug. 2010.

[47] ——, "Study on Network-Assisted Interference Cancellation and Suppression (NAIC) for LTE (Release 12)," 3GPP, Tech. Rep. 3GPP TR 36.866 v12.0.0, Mar. 2014.

[48] Y. Hama and H. Ochiai, “A low-complexity matched filter detector with parallel interference cancellation for massive MIMO systems,” in *IEEE WiMob*, Oct. 2016, pp. 1–6.

[49] N. I. Miridakis and D. D. Vergados, “A Survey on the Successive Interference Cancellation Performance for Single-Antenna and Multiple-Antenna OFDM Systems,” *IEEE Commun. Surveys Tutorials*, vol. 15, no. 1, pp. 312–335, Apr. 2013.

[50] C. N. i Manchon, L. Deneire, P. Mogensen, and T. B. Sorensen, “On the Design of a MIMO-SIC Receiver for LTE Downlink,” in *IEEE VTC-Fall*, Sept. 2008, pp. 1–5.

[51] G. L. Schrenk, “Interference management in cellular system design,” in *Vehicular Technology Conference, 1984. 34th IEEE*, vol. 34, May 1984, pp. 148–155.

[52] R. W. Nettleton and H. Alavi, “Power control for a spread spectrum cellular mobile radio system,” in *IEEE Intl. Conf. on Commun.*, vol. 33, May 1983, pp. 242–246.

[53] D.Cox, “Cochannel Interference Considerations in Frequency Reuse Small-Coverage-Area Radio Systems,” *IEEE Trans. Commun.*, vol. 30, no. 1, pp. 135–142, Jan. 1982.

[54] S.-M. Cheng, S.-Y. Lien, F.-S. Chu, and K.-C. Chen, “On exploiting cognitive radio to mitigate interference in macro/femto heterogeneous networks,” *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 40–47, Jun. 2011.

[55] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, “LTE-advanced: next-generation wireless broadband technology,” *IEEE Wireless Commun.*, vol. 17, no. 3, pp. 10–22, Jun. 2010.

[56] Third Generation Partnership Project, “Coordinated multi-point operation for LTE physical layer aspects (Release 11),” 3GPP, Tech. Rep. 3GPP TR 36.819 v11.1.0, Dec. 2011.

[57] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, “Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges,” *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 148–155, Feb. 2012.

[58] M.-S. Kim, H. Je, and F. A. Tobagi, “Cross-Tier Interference Mitigation for Two-Tier OFDMA Femtocell Networks with Limited Macrocell Information,” in *IEEE GLOBECOM*, Dec. 2010, pp. 1–5.

[59] L. Huang and G. Zhu, “Cognitive Femtocell Networks: An opportunistic spectrum access for future indoor wireless coverage,” *IEEE Wireless Communications Magazine*, vol. 20, no. 2, pp. 44–51, Apr. 2013.

[60] C. Xu, M. Sheng, X. Wang, C. X. Wang, and J. Li, “Distributed Subchannel Allocation for Interference Mitigation in OFDMA Femtocells: A Utility-Based Learning Approach,” *IEEE Trans. on Veh. Technol.*, vol. 64, no. 6, pp. 2463–2475, Jun. 2015.

[61] H. O. Kpojime and G. A. Safdar, “Interference Mitigation in Cognitive-Radio-Based Femtocells,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1511–1534, Oct. 2015.

[62] E. Telatar, “Capacity of Multi-antenna Gaussian Channels,” *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Apr. 1999.

[63] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, “V-blast: an architecture for realizing very high data rates over the rich-scattering wireless channel,” in *Proc. of URSI Intl. Symp. on Signals, Systems and Electronics (ISSSE)*, Sep. 1998, pp. 295–300.

[64] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[65] Eurecom and Orange and TCS, “Device-to-device communications: innovative concepts and performance evaluation,” D5.2 Project Sharing, Tech. Rep., Dec. 2014.

[66] T. Brown, E. D. Carvalho, and P. Kyritsi, *Practical Guide to the MIMO Radio Channel with MATLAB Examples*, 1st ed. Wiley, 2012.

[67] S. Coleri, M. Ergen, A. Puri, and A. Bahai, “Channel estimation techniques based on pilot arrangement in OFDM systems,” *IEEE Transactions on Broadcast.*, vol. 48, no. 3, pp. 223–229, Sept. 2002.

[68] L. Bai, J. Choi, and Q. Yu, *Low Complexity MIMO Receivers*. Springer, 2014.

[69] Y. S. Cho, J. Kim, W. Y. Yang, and C.-G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*, 1st ed. WILEY, 2010.

[70] H. Vikalo, B. Hassibi, and T. Kailath, "A universal lattice code decoder for fading channels," *IEEE Trans. Inform. Theory.*, vol. 45, no. 5, pp. 1639–1642, Nov. 1997.

[71] U.Fincke and M.Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Mathematics of Computation*, vol. 44, no. 4, pp. 463–471, Apr. 1985.

[72] J. Fink, S. Roger, A. Gonzalez, V. Almenar, and V. Garciay, "Complexity assessment of sphere decoding methods for MIMO detection," in *IEEE Intl. Symp. on Signal Processing and Information Technology (ISSPIT)*, Dec 2009, pp. 9–14.

[73] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm i. expected complexity," *IEEE Trans. Sig. Proc*, vol. 53, no. 8, pp. 2806–2818, Aug. 2005.

[74] M. A. M. Albreem, "An Efficient Lattice Sphere Decoding Technique for Multi-Carrier Systems," *Wireless Personal Communications*, vol. 82, no. 3, pp. 1825–1831, Jan. 2015.

[75] G. Zhou, W. Xu, and G. Bauch, "A Low-Complexity Interference-Aware Sphere Detector for Imperfect Channel Estimates," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5337–5348, Jul. 2016.

[76] M. A. Shah, B. Mennenga, and G. Fettweis, "Iterative Soft-In Soft-Out Sphere Detection for 3GPP LTE," in *IEEE VTC-Spring 2010*, May 2010, pp. 1–5.

[77] S. Aubert, F. Nouvel, and A. Nafkha, "Complexity gain of QR Decomposition based Sphere Decoder in LTE Receivers," in *IEEE VTC-Fall*, Sept. 2009, pp. 1–5.

[78] A. Lenstra, H. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Math. Ann.*, vol. 261, pp. 515–534, 1982.

[79] Y. H. Gan and W. H. Mow, "Complex lattice reduction algorithms for low-complexity MIMO detection," in *IEEE GLOBECOM*, Dec. 2005, pp. 2954–2957.

[80] H. X. Nguyen, "An Efficient Signal Detection Algorithm for 3GPP LTE Downlink," in *Fifth Intl. Conf Digital Telecommunications (ICDT)*, Jun. 2010, pp. 77–81.

[81] U. Ahmad, M. Li, S. Pollin, A. Amin, L. V. der Perre, and R. Lauwereins, "Hybrid lattice reduction algorithm and its implementation on an SDR baseband processor for LTE," in *19th European Signal Processing Conference (EUSIPCO)*, Aug. 2011, pp. 91–95.

[82] W.-Y. Chen, C.-F. Liao, and Y.-H. Huang, "Hardware-Efficient Interpolation-Based QR Decomposition and Lattice Reduction Processor for MIMO-OFDM Receivers," *Journal of Signal Processing Systems*, vol. 82, no. 3, pp. 1–13, Sept. 2016.

[83] L. Arevalo, R. C. d. Lamare, and R. Sampaio-Neto, "Iterative Multi-Branch Lattice Reduction-Aided Successive Interference Cancellation for Multiuser MIMO Systems," in *Intl. ITG Workshop on Smart Antennas (WSA)*, Mar. 2015, pp. 1–5.

[84] S. Yang and L. Hanzo, "Fifty Years of MIMO Detection: The Road to Large-Scale MIMOs," *IEEE Trans. Signal Process*, vol. 17, no. 4, pp. 1941–1988, Sept. 2015.

[85] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *IEEE Intl. Conf. on Commun. (ICC)*, vol. 2, Jun. 1995, pp. 1009–1013.

[86] A. Gomaa, L. M. A. Jalloul, M. M. Mansour, K. Gomadam, and D. Tujkovic, "Max-Log-MAP Optimal MU-MIMO Receiver for Joint Data Detection and Interferer Modulation Classification," *IEEE Commun. Lett*, vol. 20, no. 7, pp. 1389–1392, Jul. 2016.

[87] R. Krishnan, G. Colavolpe, A. G. i Amat, and T. Eriksson, "Algorithms for Joint Phase Estimation and Decoding for MIMO Systems in the Presence of Phase Noise and Quasi-Static Fading Channels," *IEEE Trans. Signal Process*, vol. 63, no. 13, pp. 3360–3375, Jul. 2015.

[88] Y. Lomnitz and D. Andelman, "Efficient maximum likelihood detector for MIMO systems with

small number of streams," *Electronics Letters*, vol. 43, no. 22, Oct. 2007.

[89] R. Ghaffar and R. Knopp, "Low Complexity Metrics for BICM SISO and MIMO Systems," in *IEEE VTC Spring 2010*, May 2010, pp. 1–6.

[90] ——, "Low complexity BICM demodulation for MIMO transmission," in *IEEE Workshop on Signal Process. Adv. in Wireless Commun. (SPAWC)*, Jul. 2008, pp. 396–400.

[91] R. Ghaffar and R. Knopp, "Spatial Interference Cancellation Algorithm," in *IEEE WCNC*, Apr. 2009, pp. 1–5.

[92] S. Kinjo, A. Liu, and S. Ohno, "A study on reduced MLD utilizing the QR decomposition for MIMO communication systems," in *Intl. Conf. on Green Circuits and Systems (ICGCS)*, Jun. 2010, pp. 259–262.

[93] A. Yousafzai and M. Nakhai, "Block QR decomposition and near-optimal ordering in intercell cooperative MIMO-OFDM," *IET Commun.*, vol. 4, no. 12, pp. 1452–1462, Aug. 2010.

[94] Y. Sagae and Y. Ohwatari and Y. Sano, "Improved Interference Rejection and Suppression Technology in LTE Release 11 Specifications," NTT Docomo, Inc., Tech. Rep. 2, Oct. 2013.

[95] E. Majeed, S. Iwelski, Z. Bai, G. H. Bruck, and P. Jung, "Advanced receiver design for interfering small cell deployments in LTE networks," in *IEEE Conf. on Standards for Commun. and Netw. (CSCN)*, Oct. 2015, pp. 294–299.

[96] A. Davydov and G. Morozov, "Enhanced Interference Cancellation of Cell-Specific Reference Signals for LTE-A," in *IEEE VTC-Fall*, Sept. 2015, pp. 1–5.

[97] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *URSI Intl. Symp. Signals, Systems, and Electronics, 1998*, Sep. 1998, pp. 295–300.

[98] F. S. Tseng and C. Y. Hsu, "A Transceiver Design for Limited Feedback AF MIMO Relay System with MMSE-SIC Receiver," in *Intl. Conf. on Networking and Network Applications*

(NaNA), Jul. 2016, pp. 126–130.

[99] D. Kong, J. Zeng, X. Su, L. Rong, and X. Xu, "Multiuser detection algorithm for PDMA uplink system based on SIC and MMSE," in *IEEE/CIC Intl. Conf. on Commun. in China (ICCC)*, Jul. 2016, pp. 1–5.

[100] J. W. Choi, J. Lee, H. L. Lou, and J. Park, "Improved MIMO SIC Detection Exploiting ML Criterion," in *IEEE VTC-Fall*, Sept. 2011, pp. 1–5.

[101] Third Generation Partnership Project, "Radio Frequency (RF) system scenarios (Release 13)," 3GPP, Tech. Rep. 3GPP TR 36.942 V13.0.0, Jan. 2016.

[102] S. P. E. Dahlman and J. Skold, *4G LTE/LTE-A for mobile broadband*, 1st ed. ELSEVIER, 2011.

[103] T. M. Cover and A. A. E. Gamal, "Capacity Theorems for the Relay Channel," *IEEE Trans. Info. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.

[104] Third Generation Partnership Project, "Further advancements for E-UTRA physical layer aspects(Release 9)," 3GPP, Tech. Rep. 3GPP TR 36.814 v9.0.0, Mar. 2010.

[105] M. Yuksel and E. Erkip, "Broadcast strategies for the fading relay channel," in *Proc. IEEE Mil. Commun. Conf*, vol. 2, Oct. 2004, pp. 1060–1065.

[106] T. Cover, "Broadcast Channels," *IEEE Trans. Inform. Theory.*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.

[107] K. Azarian, H. E. Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[108] M. N. Khormuji and E. G. Larsson, "Cooperative transmission based on decode-and-forward relaying with partial repetition coding," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1716–1725, Apr. 2009.

[109] A. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Trans. Inform. Theory.*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[110] Y. Liu, W. Xu, K. Niu, Z. He, and B. Tian, “A practical compress-and-forward relay scheme based on superposition coding,” in *IEEE 14th Intl. Conf. Commun. Techn. (ICCT)*, Nov. 2012, pp. 1286–1290.

[111] E. Atsan, R. Knopp, S. Diggavi, and C. Fragouli, “Towards integrating Quantize-Map-Forward relaying into LTE,” in *IEEE Information Theory Workshop (ITW), 2012*, Sept. 2012, pp. 212–216.

[112] C. Hoymann, W. Chen, J. Montojo, A. Golitschek, C. Koutsimanis, and X. Shen, “Relaying operation in 3GPP LTE: challenges and solutions,” *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 156–162, Feb. 2012.

[113] M. Kottkamp and A. Roessler and J. Schlienz, “LTE-Advanced Technology Introduction,” Rohde & Schwarz, Tech. Rep. 08.2012-1MA1693E, Aug. 2012.

[114] ETSI and Third Generation Partnership Project, “Physical layer for relaying operation (Rel. 12),” ETSI, Tech. Rep. ETSI 136216 v12.0.0, Oct. 2014.

[115] Third Generation Partnership Project, “Further advancements for E-UTRA physical layer aspects (Release 9),” 3GPP, Tech. Rep. 3GPP TR 36.814 V9.0.0, Mar. 2010.

[116] ——, “Self-configuring and self-optimizing network (SON) use cases and solutions (Release 9),” 3GPP, Tech. Rep. 3GPP TR 136 902 v9.3.1, May. 2011.

[117] S. Westwood, “The State of LTE (January 2016),” OpenSignal, Tech. Rep., Jan. 2016. [Online]. Available: https://opensignal.com/reports/2016/02/state-of-lte-q4-2015/

[118] P. Cerwall, “Ericsson Mobility Report,” Ericsson, Tech. Rep. EAB-16:006659, Jun. 2016.

[119] S. Benedetto and G. Montorsi, “Design of Parallel Concatenated Convolutional Codes,” *IEEE Trans. on Commun.*, vol. 44, no. 5, pp. 591–600, May 1996.

[120] F. Kaltenberger, R. Ghaffar, R. Knopp, H. Anouar, and C. Bonnet, “Design and Implementation of a Single-frequency Mesh Network Using OpenAirInterface,” *EURASIP J. Wirel. Commun. Netw.*, vol. 2010, no. 19, pp. 1–12, Apr. 2010.

[121] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H. P. Mayer, L. Thiele, and V. Jungnickel, “Coordinated multipoint: Concepts, performance, and field trial results,” *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 102–111, Feb. 2011.

[122] F. M. L. Tavares, G. Berardinelli, N. H. Mahmood, T. B. Sorensen, and P. Mogensen, “On the Potential of Interference Rejection Combining in B4G Networks,” in *IEEE VTC-Fall*, Sept. 2013, pp. 1–5.

[123] L. Rugini, P. Banelli, and G. Leus, “Simple Equalization of Time-Varying Channels for OFDM,” *IEEE Commun. Letters*, vol. 9, no. 7, pp. 619–621, Jul. 2005.

[124] C.-L. Ho, J.-Y. Wu, and T.-S. Lee, “Block based symbol detection for high rate space-time coded systems,” in *IEEE VTC Spring*, May 2004, pp. 375–379.

[125] J. Kim, H. Nam, and H. Kim, “Performance of adaptive precoding for wireless MIMO broadcast channels with limited feedback,” in *IEEE VTC-Fall*, Sept. 2005, pp. 487–491.

[126] H. Sung, K.-J. Lee, and I. Lee, “An MMSE Based Block Diagonalization for Multiuser MIMO Downlink Channels with Other Cell Interference,” in *IEEE VTC-Fall*, Sept. 2009, pp. 1–5.

[127] Y. Xiao, Z. Yang, L. D., P. Yang, L. Yin, and W. Xiang, “Low-Complexity Signal Detection for Generalized Spatial Modulation,” *IEEE Commun. Letters*, vol. 18, no. 3, pp. 406–406, Mar. 2014.

[128] E. Elmroth, F. Gustavson, I. Jonsson, and B. Kagstrom, “Recursive Blocked Algorithms and Hybrid Data Structures for Dense Matrix Library Software ,” *SIAM Review*, vol. 46, no. 1, pp. 3–45, Mar. 2004.

[129] E. Elmroth and F. Gustavson, “Applying recursion to serial and parallel QR factorization leads to better performance,” *IBM J. RES. DEVELOP.*, vol. 44, no. 4, pp. 605–624, Jul. 2000.

[130] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 4th ed. John Hopkins University Press, 2013.

[131] M. A. Jun and D. Wang, "A low complexity uplink multiuser MIMO detecting algorithm based on BQRD," in *Proc. 2nd Intl. CSSE Conf.*, Paris, Jan. 2013, pp. 1–4.

[132] G. W. Stewart, *Matrix Algorithms Volume I: Basic Decompositions*, 1st ed. Society for Industrial and Applied Mathematics, 1998.

[133] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, 1st ed. Cambridge University Press, 2005.

[134] A. M. Voicu, L. Simic, and M. Petrova, "Per-Antenna-Rate-Control (PARC) in Frequency Selective Fading with SIC-GRAKE Receiver," in *IEEE VTC-Fall*, vol. 2, Sept. 2004, pp. 1458–1462.

[135] T. D. Chiueh and P. Y. Tsai, *OFDM Baseband Receiver Design for Wireless Communications*. John Wiley and Sons, 2007.

[136] B. N. Datta, *Numerical Linear Algebra and Applications*, 2nd ed. SIAM, 2010.

[137] H. Jamali-Rad, T. van Waterschoot, and G. Leus, "Anchorless cooperative localization for mobile wireless sensor networks," in *Proc. 32nd WIC/IEEE SP Symp. on Information Theory and Signal Processing in the Benelux (WICSP)*, Jun. 2011, pp. 1–8.

[138] F. G. V. Zee, R. A. van de Geijn, and G. Quintana-Ortí, "Restructuring the Tridiagonal and Bidiagonal QR Algorithms for Performance," *ACM Trans. Math. Softw.*, vol. 40, no. 3, pp. 18:1–18:34, Apr. 2014.

[139] J. W. Demmel, O. A. Marques, B. N. Parlett, and C. Vömel, "Performance and Accuracy of LAPACK's Symmetric Tridiagonal Eigensolvers," *SIAM Journal on Scientific Computing*, vol. 30, no. 3, pp. 1508–1526, Apr. 2008.

[140] Third Generation Partnership Project, "Phyisical Layer Procedures (Release 11)," 3GPP, Tech.

Rep. 3GPP TS 36.213 v11.3.0, Jun. 2013.

[141] ——, “Multiplexing and channel coding (Release 11),” 3GPP, Tech. Rep. 3GPP TS 36.212 v11.3.0, Jun. 2013.

[142] C. Johnson, *Long Term Evolution in Bullets*, 1st ed. CreateSpace Independent, 2010.

[143] A. Nosratinia, T. Hunter, and A. Hedayat, “Cooperative communication in wireless networks,” *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 74–80, Oct. 2004.

[144] Third Generation Partnership Project, “Relay architectures for E-UTRA (LTE-Advanced)(Release 9),” 3GPP, Tech. Rep. 3GPP TR 36.806 v2.0.0, Feb. 2010.

[145] ——, “Base Station (BS) radio transmission and reception,” 3GPP, Tech. Rep. 3GPP TR 36.104 v8.2.0, May 2008.

[146] R. Thomas, M. Cardone, R. Knopp, D. Tuninetti, and B. Maharaj, “An LTE implementation of a Novel Strategy for the Gaussian Half-Duplex Relay Channel,” in *IEEE Intl. Conf. on Commun. (ICC)*, Jun. 2015, pp. 1–5.

[147] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, “On the Gaussian Half-Duplex Relay Channel,” *IEEE Trans. Info. Theor.*, vol. 60, no. 5, pp. 2542–2562, May 2014.

[148] M. Iwamura, H. Takashi, and S. Nagata, “Relay technology in LTE-Advanced,” *NTT Docomo Technical Journal*, vol. 12, no. 2, pp. 29–36, Sept. 2010.

[149] Z. Juan, P. Sartori, and B. Wei, “Performance Analysis of Layer 1 Relays,” in *IEEE Intl. Conf. on Commun. Communication Workshop (ICC)*, Jun. 2009, pp. 1–6.

[150] Ericsson, “A discussion on some technology components for LTE-Advanced,” 3GPP/Ericsson, Tech. Rep. 3GPP R1-082084, May. 2008.

[151] T. Beniero, S. Redana, J. Hamalainen, and B. Raaf, “Effect of Relaying on Coverage in 3GPP LTE-Advanced,” in *IEEE VTC-Spring 2009*, April. 2009, pp. 1–5.

[152] A. B. Saleh and S. Redana and B. Raaf and J. Hamalainen, “Comparison of Relay and Pico eNB Deployments in LTE-Advanced,” in *IEEE VTC-Fall*, Sept. 2009, pp. 1–5.

[153] N. Zlatanov, V. Jamali, and R. Schober, “On the Capacity of the Two-Hop Half-Duplex Relay Channel,” in *IEEE GLOBECOM*, Dec. 2015, pp. 1–7.

[154] B. Zhao and M. Valenti, “Distributed turbo coded diversity for relay channel,” *Electronic Letters*, vol. 39, no. 10, pp. 786–787, May 2003.

[155] K. Ishibashi, K. Ishii, and H. Ochiai, “Dynamic Coded Cooperation Using Multiple Turbo Codes in Wireless Relay Networks,” *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 197–207, Feb. 2011.

[156] X. Liu and T. Lim, “Fountain codes over fading relay channels,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3278–3287, Jun. 2009.

[157] L. Chen, “Opportunistic cooperative communications with reed-solomon convolutional concatenated codes,” in *IEEE Intl. Conf. on Commun. Tech. (ICCT)*, Sept. 2011, pp. 22–26.

[158] H. Yang and W. Meng and B. Li and G. Wang, “Physical layer implementation of network coding in two-way relay networks,” in *IEEE Intl. Conf. on Commun. (ICC)*, Jun. 2012, pp. 671–675.

[159] G.J. Bradford and J.N. Laneman, “Low latency relaying schemes for next-generation cellular networks,” in *IEEE Intl. Conf. on Commun. (ICC)*, Jun. 2012, pp. 4294–4299.

[160] Y. Meng, M. You, S. Zhao, and L. Zhang, “A Distributed Layered Modulation for Relay Downlink Cooperative Transmission,” in *IEEE Intl. Symp. on Personal, indoor and mobile radio communications (PIMRC)*, Sept. 2010, pp. 1–5.

[161] S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou, and M. Haenggi, “Superposition Coding Strategies: Design and Experimental Evaluation,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2628–2639, July 2012.

[162] J. Zhang, J. Jia, Q. Zhang, and E. M. K. Lo, “Implementation and Evaluation of Cooperative Communication Schemes in Software-Defined Radio Testbed,” in *IEEE Proc. INFOCOM, 2010*, Mar. 2010, pp. 1–9.

[163] M. Duarte, A. Sengupta, S. Brahma, C. Fragouli, and S. Diggavi, “Quantize-map-forward (QMF) Relaying: An Experimental Study,” in *Proc. of the Fourteenth ACM Intl. Symp. on Mobile Ad Hoc Networking and Computing*, Jul. 2013, pp. 227–236.

[164] S. Brahma, M. Duarte, A. Sengupta, I. H. Wang, C. Fragouli, and S. Diggavi, “QUILT: A Decode/Quantize-Interleave-Transmit approach to cooperative relaying,” in *IEEE Proc. INFOCOM, 2014*, Apr. 2014, pp. 2508–2516.

APPENDIX A

HD RELAY STRATEGY: MCS MAPPING

The Tables A.1, A.2, A.3, A.4, A.5, A.6, A.7 and A.8 display Phase I and Phase II of the link adaptation procedure at the destination for each of the investigated scenarios in terms of link strength and channel models. In Tables A.1-A.4, a ‡ indicates the MCS at which the maximum difference between the theoretical and the practical rates occurs while a $\star$ indicates the MCS at which the maximum difference between the rates of the practical and baseline schemes occurs. The link strength at the relay is given as the $C = \max\{C_{X_b}, C_{X_{a1}}\}$, which can satisfy an overall system BLER of 10^{-3} at the relay. The Theoretical rate and Theoretical Baseline rate are also computed for a $\gamma = 0.5$ based on the triplet parameters of S, I, C. In Tables A.5-A.8, a $\star$ indicates the MCS at which the maximum difference between the rates of the practical and baseline schemes occurs.

Table A.1: MCS mapping for each decoding operation with $I/S = 0$ dB for a SISO scheme.

Phase I - X_b				**Phase II** - (X_{a1}, X_{a2})		**Phase II** - X_c		**Theory**	**Practical**	**Theory BS**	**Practical BS**
MCS	S_{X_b} [dB]	C [dB]	TBS [bits]	MCS	TBS [bits]	MCS	TBS [bits]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]
0	-0.22	1.25	680	5	2216	6	2600	1.06	0.89	0.60	0.53
1	-0.02	1.25	904	5	2216	6	2600	1.09	0.93	0.63	0.53
2	0.39	1.25	1096	5	2216	7	2600	1.13	1.04	0.67	0.63
3	0.80	2.02	1416	5	2216	7	3112	1.23	1.10	0.71	0.63
4	1.15	2.02	1800	6	2600	7	3112	1.27	1.22	0.75	0.63
5	1.95	3.51	2216	7	3112	9	2216	1.50	1.52	0.84	0.80
6	2.66	4.17	2600	8	3496	9	4008	1.63	1.64	0.92	0.80
7	3.43	4.83	3112	9	4008	10	4008	1.80	1.81	1.03	0.79
8	4.11	6.59	3496	9	4008	12	4968	2.05	2.02	1.12	0.98
9	5.13	6.59	4008	10	4008	13	5736	2.24	2.24	1.25	1.11
10	5.58	6.59	4008	11	4392	13	5736	2.30	2.30	1.33	1.11
11	6.13	8.64	4392	12	4968	14	6456	2.60	2.57	1.41	1.25
12	7.40	9.54	4968	13	5736	15	7224	2.92	2.92	1.56	1.37
13	8.41	11.11	5736	14	6456	16	7736	3.26	3.24	1.75	1.46
14	9.01	13.40	6456	15	7224	17	7736	3.61	3.48	1.86	1.45
15	10.25	13.40	7224	16	7736	18	7992	3.88	3.73	2.02	1.49
16	10.45	15.75	7736	16	7736	19	9144	4.12	4.00	2.08	1.69
17‡	11.79	15.75	7736	17	7736	20	9912	**4.44**‡	**4.13**‡	2.28	1.82
18	11.99	15.75	7992	17	7736	21	10680	4.46	4.29	2.34	1.96
19	13.24	16.67	9144	19	9144	22	11448	4.84	4.84	2.52	2.08
20*	14.04	18.2	9912	19	9144	23	12576	5.16	**5.14***	2.67	**2.28***

Table A.2: MCS mapping for each decoding operation with $I/S = 5$ dB for a SISO scheme.

Phase I - X_b				**Phase II** - (X_{a1}, X_{a2})		**Phase II** - X_c		**Theory**	**Practical**	**Theory BS**	**Practical BS**
MCS	S_{X_b} [dB]	C [dB]	TBS [bits]	MCS	TBS [bits]	MCS	TBS [bits]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]
0	-0.22	6.01	680	5	2600	6	2600	1.49	1.03	0.79	0.90
1	-0.02	6.01	904	5	2600	6	2600	1.51	1.07	0.91	1.04
2	0.39	6.01	1096	5	2600	7	3112	1.55	1.19	0.91	1.04
3	0.80	8.08	1416	6	3112	7	3112	1.82	1.30	0.91	1.17
4	1.15	8.08	1800	6	3112	7	3112	1.85	1.37	1.02	1.17
5	1.95	8.08	2216	7	3496	9	4008	1.96	1.66	1.02	1.30
6	2.66	9.20	2600	8	3496	9	4008	2.15	1.73	1.15	1.39
7‡	3.43	11.22	3112	9	4008	10	4008	**2.49**‡	**1.81**‡	1.23	1.39
8	4.11	11.22	3496	10	4008	12	4968	2.58	2.09	1.23	1.63
9	5.13	11.22	4008	11	4392	13	5736	2.72	2.39	1.45	1.63
10	5.58	11.22	4008	12	5736	13	5736	2.78	2.51	1.45	1.76
11	6.13	13.70	4392	12	4968	14	6456	3.12	2.69	1.57	1.88
12	7.40	13.70	4968	13	5736	15	7224	3.38	3.03	1.70	2.21
13	8.41	13.70	5736	14	6456	16	7736	3.55	3.36	2.00	2.21
14	9.01	15.34	6456	15	7224	17	7736	3.85	3.48	2.00	2.19
15	10.25	16.60	7224	16	7736	18	7992	4.26	3.73	2.00	2.19
16	10.34	18.20	7736	16	7736	19	9144	4.51	4.00	2.00	2.34
17	11.79	18.20	7736	17	7992	20	9912	4.74	4.13	2.15	2.34
18	11.99	18.20	7992	17	7992	21	10680	4.76	4.17	2.15	2.34
19	13.24	18.20	9144	19	9144	22	11448	5.02	4.96	2.15	2.34
20*	14.04	21.00	9912	19	9912	23	12576	5.52	**5.39***	2.42	**2.63***

Table A.3: MCS mapping for each decoding operation with $I/S = 0$ dB for a SIMO scheme.

Phase I - X_b				**Phase II -** (X_{a1}, X_{a2})		**Phase II -** X_c		**Theory**	**Practical**	**Theory BS**	**Practical BS**
MCS	S [dB]	C [dB]	TBS [bits]	MCS	TBS [bits]	MCS	TBS [bits]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]
0	-1.26	3.56	680	8	3112	8	3496	1.46	1.19	0.75	0.70
1	-1.18	3.56	904	7	3112	9	4008	1.46	1.30	0.75	0.70
2	-0.93	4.15	1096	7	3112	9	4008	1.53	1.34	0.79	0.80
3	-0.81	4.15	1416	7	3112	9	4008	1.55	1.39	0.80	0.80
4	-0.69	4.82	1800	7	3112	9	4008	1.63	1.45	0.83	0.80
5	-0.25	5.88	2216	8	3496	9	4008	1.78	1.58	0.92	0.80
6‡	0.43	7.02	2600	8	3496	11	4392	**1.98**‡	**1.71**‡	1.02	0.80
7	0.86	7.02	3112	9	4008	11	4392	2.05	1.87	1.06	0.86
8	1.27	7.93	3496	10	4008	12	4968	2.20	2.15	1.13	0.97
9	1.89	9.54	4008	11	4968	13	5736	2.44	2.26	1.01	1.12
10	3.01	9.54	4008	13	5736	14	6456	2.64	2.63	1.37	1.24
11	3.45	9.54	4392	13	5736	14	6456	2.71	2.70	1.41	1.24
12	4.33	11.59	4968	14	6456	15	7224	3.07	3.03	1.60	1.38
13	5.02	12.10	5736	14	6456	16	7736	3.25	3.24	1.69	1.45
14	5.84	12.10	6456	14	6456	17	7736	3.41	3.36	1.56	1.46
15	6.64	14.00	7224	15	7224	18	7992	3.76	3.65	1.96	1.49
16	7.00	15.58	7736	15	7224	19	9144	3.96	3.92	2.08	1.49
17	8.60	16.20	7736	19	9144	20	9912	4.38	4.36	2.29	1.68
18	8.96	18.20	7992	19	9144	21	10680	4.61	4.52	2.43	1.68
19	10.10	18.20	9144	20	9912	21	10680	4.89	4.84	2.56	1.94
20*	11.02	18.20	9912	20	9912	22	11448	5.10	**5.09***	2.36	**1.93***

Table A.4: MCS mapping for each decoding operation with $I/S = 5$ dB for a SIMO scheme.

Phase I - X_b				**Phase II** - (X_{a1}, X_{a2})		**Phase II** - X_c		**Theory**	**Practical**	**Theory BS**	**Practical BS**
MCS	S [dB]	C [dB]	TBS [bits]	MCS	TBS [bits]	MCS	TBS [bits]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]	Rate [bits/dim]
0	-1.26	7.93	680	9	4008	8	3496	1.88	1.33	1.34	1.17
1	-1.18	9.54	904	9	4008	9	4008	2.03	1.45	1.44	1.17
2	-1.06	9.54	1096	9	4008	9	4008	2.06	1.48	1.46	1.17
3	-0.81	9.54	1416	9	4008	9	4008	2.08	1.53	1.47	1.30
4	-0.69	9.54	1800	9	4008	9	4008	2.10	1.60	1.48	1.30
5	-0.25	11.59	2216	9	4008	9	4008	2.32	1.66	1.64	1.30
6	0.43	11.59	2600	10	4008	11	4392	2.43	1.79	1.70	1.39
7‡	0.86	12.10	3112	10	4008	11	4392	**2.54**‡	**1.87**‡	1.77	1.39
8	1.27	13.40	3496	11	4392	12	4968	2.70	2.21	1.89	1.39
9	1.89	14.00	4008	12	4968	13	5736	2.86	2.40	1.98	1.42
10	3.01	14.00	4008	13	5736	14	6456	3.08	2.51	2.09	1.42
11	3.45	14.00	4392	13	5736	14	6456	3.17	2.57	2.13	1.62
12	4.33	16.20	4968	13	5736	15	7224	3.51	3.03	2.36	1.62
13	5.02	18.20	5736	14	6456	16	7736	3.79	3.24	2.55	1.88
14	5.84	18.20	6456	15	7224	17	7736	3.98	3.48	2.64	2.00
15	6.64	18.20	7224	15	7224	18	7992	4.17	3.65	2.72	2.00
16	7.00	18.20	7736	16	7736	19	9144	4.25	4.00	2.75	2.20
17	8.60	18.20	7736	19	9144	20	9912	4.63	4.36	2.90	2.35
18	8.96	18.20	7992	19	9144	21	10680	4.70	4.53	2.93	2.45
19	10.10	18.20	9144	20	9912	22	11448	4.95	4.96	3.03	2.64
20*	11.01	21.00	9912	21	10680	23	12576	5.42	**5.39***	3.32	**2.73***

Table A.5: MCS mapping for each decoding operation with $I/S = 0$ dB for the EPA model.

Phase I - X_b			**Phase II** - (X_{a1}, X_{a2})		**Phase II** - X_c		**Practical**	**Practical BS**
MCS	S [dB]	TBS [bits]	MCS	TBS [bits]	MCS	TBS [bits]	Rate [bits/dim]	Rate [bits/dim]
0	4.60	680	1	904	7	3112	0.76	0.49
1	5.56	904	1	904	7	3112	0.80	0.49
2	8.50	1096	4	1800	12	4968	1.29	0.79
3	9.12	1416	4	1800	12	4968	1.33	0.79
4	9.55	1800	5	2216	13	5736	1.59	0.91
5	10.93	2216	7	3112	14	6456	1.92	1.02
6	14.20	2600	10	4008	17	7736	2.33	1.23
7	18.93	3112	14	6456	23	12576	3.60	2.00
8	21.22	3496	16	7736	24	13536	4.03	2.15
9*	24.02	4008	19	9144	26	15264	**4.62***	**2.42***

Table A.6: MCS mapping for each decoding operation with $I/S = 5$ dB for the EPA model.

Phase I - X_b			**Phase II** - (X_{a1}, X_{a2})		**Phase II** - X_c		**Practical**	**Practical BS**
MCS	S [dB]	TBS [bits]	MCS	TBS [bits]	MCS	TBS [bits]	Rate [bits/dim]	Rate [bits/dim]
0	4.60	680	5	2216	7	3112	0.97	0.91
1	5.56	904	6	2600	7	3112	1.08	1.02
2	8.50	1096	9	4008	12	4968	1.64	1.23
3	9.12	1416	9	4008	12	4968	1.69	1.23
4	9.55	1800	10	4008	13	5736	1.88	1.27
5	10.93	2216	10	4008	14	6456	2.06	1.45
6	14.20	2600	13	5736	17	7736	2.61	2.00
7	18.93	3112	15	7224	23	12576	3.73	2.42
8	21.22	3496	17	7736	24	13536	4.03	2.51
9*	24.02	4008	19	9144	26	15264	**4.62***	**2.91***

Table A.7: MCS mapping for each decoding operation with $I/S = 0$ dB for the ETU model.

Phase I - X_b			**Phase II** - (X_{a1}, X_{a2})		**Phase II** - X_c		**Practical**	**Practical BS**
MCS	S [dB]	TBS [bits]	MCS	TBS [bits]	MCS	TBS [bits]	Rate [bits/dim]	Rate [bits/dim]
0	0	680	0	680	5	2216	0.58	0.35
1	0.25	904	0	680	6	2600	0.68	0.41
2	2.73	1096	1	904	8	3496	0.89	0.55
3	3.85	1416	2	1096	9	4008	1.06	0.64
4	4.51	1800	2	1096	9	4008	1.19	0.70
5	4.94	2216	3	1416	11	4392	1.30	0.70
6	7.40	2600	5	2216	14	6456	1.83	1.02
7	9.19	3112	7	3112	15	7224	2.19	1.15
8	11.72	3496	9	4008	19	9144	2.71	1.45
9*	14.05	4008	12	4968	21	10680	**3.20***	**1.67***

Table A.8: MCS mapping for each decoding operation with $I/S = 5$ dB for the ETU model.

Phase I - X_b			**Phase II** - (X_{a1}, X_{a2})		**Phase II** - X_c		**Practical**	**Practical BS**
MCS	S [dB]	TBS [bits]	MCS	TBS [bits]	MCS	TBS [bits]	Rate [bits/dim]	Rate [bits/dim]
0	0	680	5	2216	5	2216	0.83	0.70
1	0.25	904	5	2216	6	2600	0.93	0.79
2	2.73	1096	7	3112	8	3496	1.25	1.02
3	3.85	1416	8	3496	9	4008	1.45	1.15
4	4.51	1800	9	4008	11	4392	1.66	1.15
5	4.94	2216	9	4008	11	4392	1.73	1.23
6	7.40	2600	10	4008	14	6456	2.12	1.45
7	9.19	3112	11	4392	15	7224	2.40	1.82
8	11.72	3496	13	5736	19	9144	2.99	2.00
9*	14.05	4008	14	6456	21	10680	**3.43***	**2.15***

APPENDIX B

PERFORMANCE GUARANTEE OF THE PROPOSED SCHEME

A well-known upper bound on the capacity C of the relay channel is given by the cut-set bound [103] for which the channel model in (5.1), reads as

$$
\begin{aligned}
\mathsf{C} &\leq \max_{P_{X_s,X_r,S_r}} \min\left\{I\left(X_s,X_r,S_r;\mathbf{Y}_d\right), I\left(X_s;Y_r,\mathbf{Y}_d|X_r,S_r\right)\right\} \\
&\overset{(a)}{\leq} \max_{P_{X_s,X_r,S_r}} \min\left\{I\left(X_s,X_r;\mathbf{Y}_d|S_r\right)+H\left(S_r\right), I\left(X_s;Y_r,\mathbf{Y}_d|X_r,S_r\right)\right\} \\
&\overset{(b)}{=} \max_{\gamma\in[0,1],\ P_{X_s,X_r|S_r=0},\ P_{X_s,X_r|S_r=1}} \min\Big\{ \\
&\quad \gamma I\left(X_s,X_r;\mathbf{h}_s X_s+Z_d|S_r=0\right)+(1-\gamma) I\left(X_s,X_r;\mathbf{h}_r X_r+\mathbf{h}_s X_s+Z_d|S_r=1\right)+\mathscr{H}\left(\gamma\right), \\
&\quad \gamma I\left(X_s;h_{rs}X_s+Z_r,\ \mathbf{h}_s X_s+Z_d|X_r,S_r=0\right)+(1-\gamma) I\left(X_s;Z_r,\ \mathbf{h}_s X_s+Z_d|X_r,S_r=1\right)\Big\} \\
&\overset{(c)}{=} \max_{\gamma\in[0,1],\ |\rho|\leq 1,\ \gamma P_{s|0}+(1-\gamma)P_{s|1}\leq 1,\ \gamma P_{r|0}+(1-\gamma)P_{r|1}\leq 1} \min\Big\{ \\
&\quad \gamma\log\left(1+\|\mathbf{h}_s\|^2 P_{s|0}\right)+(1-\gamma)\log\left|\mathbf{I}_{n_d}+\begin{bmatrix}\mathbf{h}_s & \mathbf{h}_r\end{bmatrix}\mathbf{K}_1\begin{bmatrix}\mathbf{h}_s & \mathbf{h}_r\end{bmatrix}^{\dagger}\right|+\mathscr{H}\left(\gamma\right), \\
&\quad \gamma\log\left(1+\left(\|\mathbf{h}_s\|^2+|h_{rs}|^2\right)P_{s|0}\right)+(1-\gamma)\log\left(1+\|\mathbf{h}_s\|^2 P_{s|1}\right)\Big\} \\
&\overset{(d)}{\leq} \max_{\gamma\in[0,1],\ \alpha\in[0,1],\ \beta\in[0,1]} 2\mathscr{H}(\gamma)-(1-\gamma)\log(1-\gamma)+\min\Big\{ \\
&\quad \gamma\log\left(1+\|\mathbf{h}_s\|^2\alpha\right)+(1-\gamma)\log\left(1+x+y+2\sqrt{xy}|v|+xy(1-|v|^2)\right)\Big|_{\substack{x=\|\mathbf{h}_s\|^2(1-\alpha)\\ y=\|\mathbf{h}_r\|^2(1-\beta)\\ v=\frac{\mathbf{h}_s^{\dagger}\mathbf{h}_r}{\|\mathbf{h}_s\|\|\mathbf{h}_r\|}}}, \\
&\quad \gamma\log\left(1+\left(\|\mathbf{h}_s\|^2+|h_{rs}|^2\right)\alpha\right)+(1-\gamma)\log\left(1+\|\mathbf{h}_s\|^2(1-\alpha)\right)\Big\} \\
&\leq \max_{\gamma\in[0,1]} 2.51\log(2)+\min\Big\{ \\
&\quad \gamma\log\left(1+\|\boldsymbol{h}_s\|^2\right)+(1-\gamma)\log\left(1+\|\boldsymbol{h}_s\|^2+\|\boldsymbol{h}_r\|^2+2\|\boldsymbol{h}_s\|\|\boldsymbol{h}_r\||v|+\|\boldsymbol{h}_s\|^2\|\boldsymbol{h}_r\|^2(1-|v|^2)\right), \\
&\quad \gamma\log\left(1+\|\boldsymbol{h}_s\|^2+|h_{rs}|^2\right)+(1-\gamma)\log\left(1+\|\boldsymbol{h}_s\|^2\right)\Big\}
\end{aligned}
$$

$$\stackrel{(e)}{=} 2.51\log(2)+\log\left(1+\|\boldsymbol{h}_s\|^2\right)+\max_{\gamma\in[0,1]}\min\left\{(1-\gamma)\,\overline{a},\gamma\overline{b}\right\}$$
$$=2.51\log(2)+\log\left(1+\|\boldsymbol{h}_s\|^2\right)+\frac{\overline{a}\,\overline{b}}{\overline{a}+\overline{b}}, \tag{B.1}$$

where the different (in)equalities are explained in what follows.

(a): By using the chain rule for the mutual information and because $I(S_r;\cdot)\leq H(S_r)$, since S_r is a discrete random variable.

(b): By defining $\gamma:=\mathbb{P}[S_r=0]\in[0,1]$ and $H(S_r)=\mathscr{H}(\gamma)=-\gamma\log(\gamma)-(1-\gamma)\log(1-\gamma)\leq\log(2)$.

(c): By letting n_d be the number of antennas at the destination (in our system model in (5.1) we have $n_d=2$), since "Gaussian maximizes entropy" and by defining

$$\mathbf{K}_\ell:=\mathrm{Cov}\begin{bmatrix}X_s\\X_r\end{bmatrix}\Big|_{S_r=\ell}=\begin{bmatrix}P_{s|\ell} & \rho_\ell\sqrt{P_{s|\ell}P_{r|\ell}}\\ \rho_\ell^*\sqrt{P_{s|\ell}P_{r|\ell}} & P_{r|\ell}\end{bmatrix}:|\rho_\ell|\leq 1,$$

for some $(P_{s|0},P_{s|1},P_{r|0},P_{r|1})\in\mathbb{R}_+^4$ satisfying the average power constraint

$$\gamma P_{u|0}+(1-\gamma)P_{u|1}\leq 1,\ u\in\{s,r\}. \tag{B.2}$$

It can be implicitly seen that $|\rho_0|=0$ is optimal, since the relay is listening. For notation convenience we let $\rho:=\rho_1$ in what follows.

(d): By defining $v:=\frac{\boldsymbol{h}_s^\dagger\boldsymbol{h}_r}{\|\boldsymbol{h}_s\|\|\boldsymbol{h}_r\|}:|v|\leq 1$, $t_s:=\|\boldsymbol{h}_s\|^2P_{s|1}$, $t_r:=\|\boldsymbol{h}_r\|^2P_{r|1}$, we obtain

$$\begin{aligned}
T_1&:=\log\left|\mathbf{I}_{n_d}+\begin{bmatrix}\boldsymbol{h}_s & \boldsymbol{h}_r\end{bmatrix}\mathbf{K}_1\begin{bmatrix}\boldsymbol{h}_s & \boldsymbol{h}_r\end{bmatrix}^\dagger\right|\\
&\stackrel{(f)}{=}\log\left|\mathbf{I}_2+\begin{bmatrix}\boldsymbol{h}_s & \boldsymbol{h}_r\end{bmatrix}^\dagger\begin{bmatrix}\boldsymbol{h}_s & \boldsymbol{h}_r\end{bmatrix}\mathbf{K}_1\right|=\log\left|\mathbf{I}_2+\begin{bmatrix}\|\boldsymbol{h}_s\|^2 & \boldsymbol{h}_s^\dagger\boldsymbol{h}_r\\ \boldsymbol{h}_r^\dagger\boldsymbol{h}_s & \|\boldsymbol{h}_r\|^2\end{bmatrix}\mathbf{K}_1\right|\\
&=\log\left(1+t_s+t_r+t_st_r\left(1-|v|^2\right)\left(1-|\rho|^2\right)+2\sqrt{t_rt_s}\Re\{v\rho^*\}\right)\\
&\stackrel{(g)}{\leq}\log\left(1+t_s+t_r+2\sqrt{t_st_r}|v||\rho|+t_st_r(1-|v|^2)(1-|\rho|^2)\right)\\
&\stackrel{(h)}{\leq}\log\left(1+t_s+t_r+2\sqrt{t_st_r}|v|+t_st_r(1-|v|^2)\right)\\
&\stackrel{(i)}{\leq}-2\log(1-\gamma)+\log\left(1+x+y+2\sqrt{xy}|v|+xy(1-|v|^2)\right),
\end{aligned}$$

where: (i) the equality in (f) follows from the Sylvester's determinant identity; (ii) the inequality in (g) follows since $\Re\{v\rho^*\}\leq|\rho||v|$; (iii) the inequality in (h) follows by us-

ing $|\rho|=1$ in the term increasing in $|\rho|$ and $|\rho|=0$ in the term decreasing in $|\rho|$; (iv) the inequality in (i) follows by parameterizing the power constraints in (B.2) as $\gamma P_{s|0}=\alpha$, $(1-\gamma)P_{s|1}=1-\alpha$, $\gamma P_{r|0}=\beta$, $(1-\gamma)P_{r|1}=1-\beta$, for some $\alpha\in[0,1]$, $\beta\in[0,1]$, and by introducing $x:=\|\boldsymbol{h}_s\|^2(1-\alpha)$, $y:=\|\boldsymbol{h}_r\|^2(1-\beta)$. By means of similar steps, we have

$$\begin{aligned}T_2 &:= \log\left(1+\|\mathbf{h}_s\|^2 P_{s|0}\right)=\log\left(1+\|\mathbf{h}_s\|^2\frac{\alpha}{\gamma}\right)\\ &=\log\left(\gamma+\|\mathbf{h}_s\|^2\alpha\right)-\log(\gamma)\le\log\left(1+\|\mathbf{h}_s\|^2\alpha\right)-\log(\gamma).\end{aligned}$$

Hence, the first term within the min in the equality in (c) can be upper bounded as

$$\begin{aligned}\gamma T_2+(1-\gamma)T_1+\mathscr{H}(\gamma)\le{}&2\mathscr{H}(\gamma)-(1-\gamma)\log(1-\gamma)+\gamma\log\left(1+\|\mathbf{h}_s\|^2\alpha\right)\\&+(1-\gamma)\log\left(1+x+y+2\sqrt{xy}|v|+xy(1-|v|^2)\right).\end{aligned}$$

Similarly, the second term is upper bounded within the min in the equality in (c) as

$$\begin{aligned}T_3 :=&\gamma\log\left(1+\left(\|\mathbf{h}_s\|^2+|h_{rs}|^2\right)P_{s|0}\right)+(1-\gamma)\log\left(1+\|\mathbf{h}_s\|^2P_{s|1}\right)\\=&\gamma\log\left(1+\left(\|\mathbf{h}_s\|^2+|h_{rs}|^2\right)\frac{\alpha}{\gamma}\right)+(1-\gamma)\log\left(1+\|\mathbf{h}_s\|^2\frac{1-\alpha}{1-\gamma}\right)\\=&\gamma\log\left(\gamma+\left(\|\mathbf{h}_s\|^2+|h_{rs}|^2\right)\alpha\right)+(1-\gamma)\log\left(1-\gamma+\|\mathbf{h}_s\|^2(1-\alpha)\right)+\mathscr{H}(\gamma)\\\le&\gamma\log\left(1+\left(\|\mathbf{h}_s\|^2+|h_{rs}|^2\right)\alpha\right)+(1-\gamma)\log\left(1+\|\mathbf{h}_s\|^2(1-\alpha)\right)+\mathscr{H}(\gamma).\end{aligned}$$

Therefore, the expression in the equality in (c) can be upper bounded as

$$\begin{aligned}&\min\left\{\gamma T_2+(1-\gamma)T_1+\mathscr{H}(\gamma),T_3\right\}\\&\le\min\Big\{\gamma\log\left(1+\|\mathbf{h}_s\|^2\alpha\right)+(1-\gamma)\log\left(1+x+y+2\sqrt{xy}|v|+xy(1-|v|^2)\right)\\&\quad+2\mathscr{H}(\gamma)-(1-\gamma)\log(1-\gamma),\gamma\log\left(1+\left(\|\mathbf{h}_s\|^2+|h_{rs}|^2\right)\alpha\right)\\&\quad+(1-\gamma)\log\left(1+\|\mathbf{h}_s\|^2(1-\alpha)\right)+\mathscr{H}(\gamma)\Big\}\\&\le 2\mathscr{H}(\gamma)-(1-\gamma)\log(1-\gamma)+\min\Big\{\gamma\log\left(1+\|\mathbf{h}_s\|^2\alpha\right)\\&\quad+(1-\gamma)\log\left(1+x+y+2\sqrt{xy}|v|+xy(1-|v|^2)\right),\gamma\log\left(1+\left(\|\mathbf{h}_s\|^2+|h_{rs}|^2\right)\alpha\right)\\&\quad+(1-\gamma)\log\left(1+\|\mathbf{h}_s\|^2(1-\alpha)\right)\Big\},\end{aligned}$$

which is precisely the expression in the inequality in (d).

(e): By defining

$$\bar{a}:=\log\left(1+\frac{\|\boldsymbol{h}_r\|^2+2\|\boldsymbol{h}_s\|\|\boldsymbol{h}_r\||v|+\|\boldsymbol{h}_s\|^2\|\boldsymbol{h}_r\|^2(1-|v|^2)}{1+\|\boldsymbol{h}_s\|^2}\right)\le\underline{a}+\log(2),\tag{B.3}$$

$$\overline{b} := \log\left(1+\frac{|h_{rs}|^2}{1+\|\boldsymbol{h}_s\|^2}\right) \leq \underline{b}+\log(2), \tag{B.4}$$

where $\underline{a}$ and $\underline{b}$ are defined in (5.2b) and (5.2c), respectively.

Notice that the above analysis holds for the case when the destination is equipped with a general number n_d of antennas.

It is shown that the upper bound in (B.1) and the lower bound in (5.2) are within a constant gap of one another. By taking the difference between (B.1) and (5.2) we obtain

$$\mathsf{gap} \leq \frac{\overline{a}\,\overline{b}}{\overline{a}+\overline{b}} - \frac{\underline{a}\;[\underline{b}]^+}{\underline{a}+[\underline{b}]^+} + 2.51\log(2) \leq 3.51\log(2).$$

Thus, the two-phase three-part-message scheme designed in Section 5.3 is at most $3.51\log(2)$ bits/dim from capacity, irrespective of the number of antennas at the destination and of the channel gains.