

University of Florence - Faculty of Engineering
Department of Electronics and Telecommunications

PhD Course on Computer Science,
Multimedia and Telecommunications
XXII Cycle
SSD: ING-INF/03 Telecomunicazioni

A dissertation for the Degree of Doctor of Philosophy

Multiple Input Multiple Output: from Signal Processing to Protocol Gain

Filippo Meucci

Supervisor: Prof. Enrico Del Re
Supervisor: Prof. Laura Pierucci

International Supervisor: Prof. Ramjee Prasad

Director of the PhD Course: Prof. Giacomo Bucci

31st December 2009, Florence

To my family

Massimo, Luana and Alberto,

their call for tenacity echoes in my ears.

Contents

Contents	i
List of Symbols	v
List of Abbreviations	ix
Abstract	xv
Introduction	xvii
Structure of this Thesis	xviii
1 The Promises of MIMO and MIMO-ARQ designs	1
1.1 Capacity in MIMO Channels	2
1.1.1 MIMO Channel Theoretical Capacity	3
1.2 From MIMO signal processing gain to MIMO protocol gain	5
1.3 Proposed MU MIMO-ARQ Protocol	7
2 PHY: MIMO Processing Gain	11
2.1 Introduction	11
2.2 Space-Time Block Codes	13
2.2.1 Alamouti scheme	15
2.2.2 LSTBC	16
2.3 Spatial Multiplexing	17
2.4 MIMO support in WiMAX	19
2.4.1 Adaptive Modulation and Coding	24
2.5 Link Adaptation with Dynamic MIMO	25
2.5.1 MIMO switching based on Euclidean Distance	26
2.5.2 Adaptive MIMO with upper-bounded BER	29
2.5.3 Pilot Structure and Channel Estimation	30
2.5.4 Performance Analysis	31
2.6 Dynamic STFBC	37
2.6.1 Space-Time-Frequency Block Codes (STFBC)	37
2.6.2 System Model	38
2.6.3 The Time-Frequency Correlation in the Frequency Domain	39
2.6.4 Diversity Schemes in Uncorrelated Channels	43
2.6.5 Performance Analysis	45
2.7 Conclusions	48
3 MAC: ARQ – DSA protocols	51
3.1 Introduction	51

3.2	Cross-Layer ARQ	52
3.2.1	State of the Art on Cross-Layer ARQ	53
3.2.2	ACL ARQ System Model	53
3.2.3	Adaptive Cross Layer (ACL) Strategy	54
3.2.4	Performance Analysis	57
3.3	DSA Protocol	63
3.3.1	System Model	64
3.3.2	DSA Signalling Analysis	65
3.3.3	DSA Analysis Validation	68
3.3.4	Performance Analysis in Mobility	70
3.4	Conclusions	78
4	MIMO Protocol Gain	81
4.1	Introduction	81
4.2	MU MIMO State of the Art	82
4.3	SU MIMO-ARQ	84
4.3.1	Single User (SU) Protocol Operation	85
4.4	MU MIMO-ARQ	87
4.5	System Model	88
4.5.1	MT Receiver	93
4.5.2	Retransmission Strategy S1	94
4.5.3	Retransmission Strategy S2	95
4.5.4	Retransmission Strategy MU	96
4.5.5	SIC principle	99
4.5.6	New User Selection	99
4.6	PO - Packet Overhearing	100
4.6.1	Received Signals at the Second Time Slot	101
4.6.2	Generalisation of PO for any time slot	102
4.7	Results	103
4.7.1	Cross-layer MIMO-ARQ versus single layer adaptivity	103
4.7.2	S1 - S2 - MU - MU PO Comparison	107
4.7.3	Performance with Log-normal SNR Distribution	108
4.7.4	Performance for Short Channel Coherence Times	109
4.8	Conclusion	111
5	MIMO Protocol Gain System Design	115
5.1	PMP 802.16 Architecture	117
5.1.1	Packet Processing Workflow	117
5.1.2	MAC Layer	118
5.1.3	ARQ	121
5.2	Scheduling in IEEE802.16	124
5.2.1	BS and MT Schedulers	125
5.2.2	QoS Classes	127
5.3	MIMO-ARQ protocol stack	128
	Conclusions	131
	Bibliography	137
A	Closed-form SEP Calculation	151
A.1	SEP Calculation using MGF	151
A.1.1	SEP for AWGN Channel	151
A.1.2	SEP with Generic Channel Fading	153
A.1.3	Rayleigh Fading Channel Statistic	154
B	IMT-A Requirements	157
B.1	4G Technical Requirements and Challenges	159

B.2	IMT-Advanced Performance Requirements	160
B.3	WiMAX Mobile IMT-A Candidate	164
B.3.1	Functional Requirements	164
C	MIMO SVD Capacity	167
C.1	MIMO Channel Rank	168
D	Publications	171
	List of Figures	173
	List of Tables	176
	Index	177

List of Symbols

	Definition
$()^\dagger$	Pseudo-Inverse operator
$()^H$	Hermitian operator
$()^T$	Transpose operator
$\arg \min_x()$	arg min operator respect to x
A	Alamouti coding matrix
B	LSTBC coding matrix
B_C	Coherence Bandwidth
β	Instantaneous channel fading amplitude
β_i	Single Path channel fading amplitude in multipath channel
$\mathcal{CN}(0,1)$	Complex gaussian random variable with unitary variance
C_{SW}	DIV / SM switching condition
C	VBLAST coding matrix
\mathbb{C}	Complex Numbers ($a + jb$)
ΔF	Frequency subcarriers separation
d	Diversity order
d^*	Optimal diversity order
\bar{D}	Average packet delay
\bar{D}_{MAX}	Average Maximum packet delay
e^{MMSE}	Estimation error for biased MMSE receiver
E_H	Expectation over Channel Realisation
$\mathcal{F}\{\}$	Fourier transform
f_d	Doppler frequency
γ	Signal to Noise Ratio (SNR)
$\bar{\gamma}$	Average SNR

Γ	Log-normal distributed SNR
h_{ij}	Channel coefficient (complex) from tx antenna j^{th} to receiving antenna i^{th}
h_j^n	Channel coefficient from tx antenna j^{th} at signalling time (n)
$h_{ij(n)}$	Channel coefficient from tx antenna j^{th} to receiving antenna i^{th} at signalling time (n)
I	Identity Matrix
$j = \sqrt{-1}$	Imaginary unit
H	MIMO channel matrix
$\mathbf{H}_{eq}^{(u)}$	Equivalent MIMO channel for user (u)
$\ \mathbf{H}\ _F$	Frobenius norm
\mathbf{H}^\dagger	Moore-Penrose Pseudo Inverse
$\mathbf{H}^{(j)}(t)$	2×2 MIMO channel for user j at time t
$\mathbf{H}_i^{(j)}(t)$	$N_R \times 1$ channel column vector from transmit antenna i to user j at time t
$\tilde{\mathbf{H}}_{k_i}$	Channel matrix with zeroed columns $\{k_1, k_2, \dots, k_i\}$
$I^{(1)}(a, \bar{\gamma})$	Integral of $erfc(x)$ with MFG
$I^{(2)}(a, \bar{\gamma})$	Integral of $erfc^2(x)$ with MFG
$I_Q^{(Ray)}(a, \bar{\gamma})$	Integral of $erfc(x)$ with MFG over Rayleigh fading
$I_{Q^2}^{(Ray)}(a, \bar{\gamma})$	Integral of $erfc^2(x)$ with MFG over Rayleigh fading
K	ARQ retransmissions index
K_{opt}	Optimal ARQ retransmissions
K_{max}	Maximum ARQ retransmissions
K_D	Demmel condition number
L	Latency of DSA protocol
L	MIMO-ARQ retransmission index
λ	Packet arrival rate
λ_i	i^{th} Singular value of the MIMO channel
λ_{min}	Minimum singular value
λ_{max}	Maximum singular value
$M_\gamma(s)$	Moment Generating Function of $p_\gamma(\gamma)$
μ	waterfill level
n	Gaussian Noise
η	Cell spectral efficiency
ν	Doppler shift
n_R	DSA-REQ retransmissions
n_S	DSA-RSP retransmissions

\mathcal{N}	Gaussian noise vector, $\mathcal{N}\{i\} = \mathcal{CN}(0, 1)$
N_{data}	Data subcarriers
N_{fft}	DFT/IDFT points
N_{guards}	Guard subcarriers
N_{ofdm}	Number of OFDM symbols per frame
N_{pilots}	Pilot Subcarriers
N_T	Number of transmitting antenna
N_R	Number of receiving antenna
$p_\gamma(\gamma)$	PDF of SNR
p_R, p_S, p_A	DSA-REQ, DSA-RSP and DSA-ACK messages error probability
P_e	Probability of error
P_i	PHY PDU error rate
P_T	Total transmit power
$Q(x)$	Slicing operator
r	Multiplexing Gain
R_{stbc}	STBC rate
ρ	Correlation coefficient
\mathbf{R}	Covariance matrix of input symbols vector
σ	Log-normal st. deviation for SNR
σ^2	Noise variance
Ω	Mean-square value of channel fading
s	Modulation symbol
\hat{s}	Estimated symbol
S_i	PHY profile spectral efficiency
SNR_i^{MMSE}	Post-processing SNR at the output of the MMSE receiver
τ	Expected duration of service flow
τ_i	Single path delay in multipath channel
τ_{rms}	Root-mean-square delay spread
Θ	Channel impulse response length
Θ_i	Single path phase in multipath channel
T_{ofdm}	OFDM symbol duration
T_{ofdm}^{CP}	OFDM symbol duration with CP
T_f	OFDM/OFDMA frame duration
T_x	DSA timeouts
$Tr()$	Trace Operator
$(x)^+$	$\max(0, x)$
$x_K^{(j)}$	m^{th} modulation symbol for user j inside K^{th} packet

$X_K^{(j)}$	K^{th} packet for user j
X	Expected service time
\bar{X}	First moment of X
\bar{X}^2	Second moment of X
v	velocity of the MT
$\dagger_u \tilde{t}(t)$	Cancelled received signal for user (u) at time t
$\mathcal{Y}_u(t+k)$	Received signal for user (u) at the k^{th} ARQ re-transmission
$\tilde{\mathcal{Y}}_u(t)$	Cancelled received signal for user (u) at the time t
ζ	Queue utilisation factor

List of Abbreviations

1G	First Generation	xvii
1xEV-DO	1x Evolution-Data Only	158
2G	Second Generation	xvii
3G	Third Generation	xviii
3GPP	3rd Generation Partnership Project	157
4G	Fourth Generation	xvii
AAS	Adaptive Antenna Systems	20
AC	Admission Control	55
ACK	Acknowledgement	xix
ACL	Adaptive Cross Layer	53
AMC	Adaptive Modulation and Coding	13
AN	Analytically	68
ARQ	Automatic Repeat Request	xv
BE	Best Effort	25
BER	Bit Error Rate	xv
BLAST	Bell Laboratories lAyered Space-Time Architecture	4
BPSK	Binary Phase Shifting Key	13
BS	Base Station	xix
BSN	Block Sequence Number	122
CDF	Cumulative Distribution Function	162
CDI	Channel Distribution Information	3
CDMA	Code Division Multiple Access	158
CID	Connection IDentifier	21
COID	Co-Ordinated Interleaved Orthogonal Designs	37
CP	Cyclic Prefix	22
CQI	Channel Quality Indicator	53
CRC	Cyclic Redundancy Check	84
CSI	Channel State Information	xv
CSIR	CSI at the Receiver	2
CSIT	CSI at the Transmitter	2
DCD	Downlink Channel Descriptor	21
DFT	Discrete Fourier Transform	20

DHCP	Dynamic Host Control Protocol	120
DL	DownLink	xv
DLC	Data Link Control	116
DLFP	DownLink Frame Prefix	21
DPC	Dirty Paper Coding	82
DSA	Dynamic Service Addition	xv
ED	Euclidean Distance	27
ertPS	extended rtPS	128
FD	Frequency Domain	40
FDD	Frequency Division Duplex	165
FDMA	Frequency Division Multiple Access	158
FSMC	Finite State Markov Chain	53
FTP	File Transfer Protocol	120
FUSC	Fully Used SubCarriers	20
GMH	General Mac Header	120
HARQ	Hybrid Automatic Repeat Request	157
HOL	Head Of Line delay	93
HSXPA	High-Speed Downlink/Uplink Packet Access	158
ICI	InterCarrier Interference	47
IDFT	Inverse Discrete Fourier Transform	22
IEEE802.16	Institute of Electrical & Electronics Engineers 802.16	xviii
IFDMA	Interleaved Frequency Division Multiple Access	158
iid	identically independent distributed	26
IMT-A	International Mobile Telecommunications Advanced	xviii
IP	Internet Protocol	xvii
IS-2000	Information System 2000	158
ISI	InterSymbol Interference	29
ITU	International Communication Union	xvii
LLC	Link Layer Control	51
LoS	Line of Sight	47
LSTBC	Layered STBC	xix
LTE-A	Long Term Evolution Advanced	xviii
LTV	Linear Time Variant	39
MAC	Medium Access Control	xv
MC-CDMA	Multi Carrier Code Division Multiple Access	158
MCS	Modulation and Coding Schemes	xv
MGF	Moment Generating Function	xix
MIMO	Multiple Input Multiple Output	xv
MISO	Multiple Input Single Output	160
ML	Maximum-Likelihood	14
MMSE	Minimum Mean Square Error	7
MT	Mobile Terminal	xvii
MU	Multiple User	xv

NACK	Not ACK.....	xix
NET	Networking.....	2
nrtPS	non-real-time Polling Service.....	127
NU	New User.....	100
OFDM	Orthogonal Frequency Division Multiplexing.....	xviii
OFDMA	Orthogonal Frequency Division Multiple Access.....	20
OSIC	Ordered Successive Interference Cancellation.....	12
PAM	Pulse Amplitude Modulation.....	14
PAPR	Peak to Average Power Ratio.....	158
PC	Packet Combining.....	7
PDF	Probability Density Function.....	153
PDP	Power Delay Profile.....	38
PDU	Protocol Data Unit.....	54
PER	Packet Error Rate.....	86
PHY	Physical.....	xv
PMP	Point to MultiPoint.....	20
PO	Packet Overhearing.....	84
PSK	Phase Shifting Key.....	15
PtP	Point to Point.....	31
PUSC	Partially Used Subcarriers.....	20
QAM	Quadrature Amplitude Modulation.....	15
QoS	Quality of Service.....	9
QOSTBC	Quasi-Orthogonal Space-Time Block Code.....	16
RAT	Radio Access Technology.....	37
RMS	Root Mean Square.....	41
RR	Round Robin.....	xv
RRM	Radio Resource Management.....	159
rtPS	real-time Polling Service.....	127
SC	Single Carrier.....	20
SCa	Single Carrier advanced.....	20
SC-FDMA	Single Carrier Frequency Division Multiplexing.....	158
SDMA	Spatial Division Multiple Access.....	116
SDU	Service Data Unit.....	118
SE	Spectral Efficiency.....	xviii
SEP	Symbol Error Probability.....	xix
SER	Symbol Error Rate.....	46
SFBC	Space-Frequency Block Coding.....	xv
SFR	Single Frequency Reuse.....	159
SIC	Successive Interference Cancellation.....	7
SIM	Simulations.....	68
SINR	Signal to Interference Noise Ratio.....	11
SIR	Signal to Interference Ratio.....	20
SISO	Single Input Single Output.....	2

SM	Spatial Multiplexing	xv
SNMP	Simple Network Management Protocol	120
SNR	Signal-to-Noise Ratio	2
SON	Self Organising Networks	159
ST	Space-Time	37
STBC	Space Time Block Code	xv
STC	Space Time Code	160
STF	Space-Time-Frequency	37
STFBC	Space Time Frequency Block Code	37
STTC	Space Time Trellis Code	11
STTD	Space Time Transmit Diversity	15
SU	Single User	xviii
SVD	Single Value Decomposition	5
TD-CDMA	Time Division CDMA	158
TDD	Time Division Duplex	63
TDM	Time Division Multiplexing	8
TDMA	Time Division Multiple Access	84
TD-SCDMA	Time Division Synchronous CDMA	158
UCD	Uplink Channel Descriptor	21
UGS	Unsolicited Grant Service	54
UL	UpLink	3
UMTS	Universal Mobile Telecommunications System	157
VBLAST	Vertical BLAST	xix
VoIP	Voice over IP	163
WARC92	World Administrative Radio Conference of 1992	165
WiMAX	Worldwide Interoperability for Microwave Access	xviii
WRC07	World Radio Conference of 2007	165
WRC2000	World Radio Conference of 2000	165
WSSUS	Wide Sense Stationary - Uncorrelated Scattering	xix
ZF	Zero Forcing	45

Acknowledgements

This work is the result of the research conducted in two years at the University of Florence and during a visiting period of one year at Center for TeleInfrastruktur (CTIF) at the University of Aalborg, Denmark. This thesis would not have been possible without the support of several people during this very intense period: I would like to thank all of them.

I am grateful to Prof. Enrico Del Re for supporting me during this three year period. I owe my deepest gratitude to Prof. Laura Pierucci: she have always provided a nice and relaxed environment where research could grow freely. Her encouragement and support enabled me to develop an understanding of the subject and overcome all the difficulties that I encountered.

I am very thankful to Professor Ramjee Prasad for his unconditioned support during my visiting period at CTIF, Aalborg University. His approach to research and to discover of truth has fascinated me. His visionary attitude and deep commitment will be a indelible mark in my personal and professional life.

I would like to acknowledge my debt to Albena Mihovska and Neeli Prasad: their support and advises during my year at CTIF were precious. They trusted me and offered the opportunity to participate in several European actions. An experience I will not forget.

I would like to express my gratitude to Isabella Cerutti, Elisabeth De Carvalho and Petar Popovski for the vibrant and deep technical discussions that improved much of the work presented in this thesis.

Last, but not least, I wish to thank all the colleagues with whom I have spent pleasant time in various offices: Luca Simone Ronga, Simone Morosi, Rosalba Suffritti, Bayu Anggorojati, Satya Ardhy Wardana, Antonietta Stango, Bilge Kartal, Snaris Blindā, Nuno Pratas and many more.

Abstract

Multiple Input Multiple Output (MIMO) techniques for wireless networks are investigated; adaptive strategies are analysed separately at PHY and MAC layers. After, cross-layer MIMO-ARQ is proposed to further push the resource exploitation in Multiple User (MU) MIMO networks.

Multiple antennas are used for enhanced signalling at the Physical (PHY) layer, providing adaptivity to the channel status. Automatic Repeat Request (ARQ) and Dynamic Service Addition (DSA) protocols are studied at the Medium Access Control (MAC) layer. Then, MIMO is used with cross-layer strategies spanning over PHY and MAC layers. The cross-layer approach allows to increase resource exploitation with per-antenna ACKs, shifting

from *MIMO Signal Processing Gain* to *MIMO Protocol Gain*.

MIMO protocol gain can impact with a decisive role in achieving the aggressive requirements of 4G wireless networks. Five original contributions are presented.

Firstly, two strategies at the PHY layer are proposed and analysed: one provides variable spectral efficiency via adaptive MIMO scheme selection based on upper-bounded Bit Error Rate (BER). MIMO transmission is varied among full spatial multiplexing and full diversity schemes also including a trade-off between the former two. A second adaptive strategy selects among Space Time Block Code (STBC) and Space-Frequency Block Coding (SFBC) depending on the channel delay spread and Doppler shift.

Secondly, the MAC layer is studied and ARQ and DSA protocols are analysed. The length of ARQ retransmissions is optimised jointly with adaptive Modulation and Coding Schemes (MCS). Then, the probability of signalling and admission blocking of the DSA protocol are expressed in closed-form based on the parameters of the three-way handshake protocol.

The last contribution refers to a set of designs where adaptive MIMO schemes (namely Spatial Multiplexing (SM) and STBC Alamouti) are used with ARQ. A MU network is served in DownLink (DL) with a Round Robin (RR) scheduler; the design is ready to include more advanced schedulers. The ARQ state machine at the MAC layer is aware of per-antenna ARQ. The interaction between the ARQ and the PHY layer, with a per-antenna ACK, opens the floor for several re-transmission strategies which are shown to improve the network throughput with no need of Channel State Information (CSI) feedback. The absence of CSI feedback at PHY layer is an important characteristic of the proposed MIMO-ARQ cross-layer designs: MIMO CSI feedback (when feasible) drastically reduces the network efficiency.

The added degrees of freedom offered by MIMO transmissions can make the difference if correctly exploited both at the physical and medium access layers, in particular for overcoming the problem of low MIMO channel ranks. The advantages of the paradigm shift from signal processing to protocol gain - together with the modifications to be applied at the classical protocol stack - are discussed in the final chapter of the thesis.

With no loss of generality, the designs proposed in the thesis are evaluated inside the IEEE802.16 framework: OFDMA is used at the PHY layer with a connection-oriented MAC.

Introduction

Currently, the traffic demand in wireless communication networks is growing with very fast pace: data transfers, voice calls, video streaming, interactive Web 2.0 browsing are requiring high capacity links with very tight delay requirements. The scarcity of spectrum resources is forcing wireless communication engineers to push for high spectral efficiencies which is even more challenging when the Mobile Terminal (MT)* has an high degree of mobility, as required in future wireless networks.

The research and the industry communities are facing the ambitious goal of providing the so-called "Mobile Internet", i.e. an access to the network on hand-held devices, data traffic and multimedia streaming, achieving the same user experience available on a cable-connected computer, but on the move. The achievement of a good performance in such harsh setting requires several improvements; modifications are needed in all the layers of the communication stack: if the physical layer spectral efficiency must be increased, the network architecture is also to be simplified with backward compatibility with the actual legacy systems already deployed.

Currently, the standardisation bodies are working on building the specifications of Fourth Generation (4G) wireless networks. Looking back at the history of wireless systems, each migration among the cellular generations was driven by a specific market force. From the First Generation (1G) to the Second Generation (2G), the main goal was to extend the access to mobile communications to a big mass of users, as the 1G network was providing small room to accommodate an increasing amount of terminals. One of the main technology upgrades was the passage from analog to digital voice transmission with a subsequent smaller system bandwidth. From 2G to 3G, the technological upgrades provided a good data connection both in terms of capacity and costs, supporting packet switching networking on a network specifically designed for the voice calls. From a native circuit switched network, packet based services have been introduced.

4G systems are expected to provide a secure Internet Protocol (IP)-based connection where dishomogeneous services are provided to users in high mobility with much enhanced performance compared with the previous generations. Broadband and high mobility support are thus the must-to-be-fulfilled requirements and high spectral efficiencies have to be achieved with practical system implementation.

The technical performance expected by the 4G systems are defined globally by the International Communication Union (ITU) which published a

*Varied terms are used in literature to indicate the mobile equipment in a cellular network. The IEEE standards use Subscriber Station (SS) or Mobile Station (MS). LTE-A prefers to use User Terminal (UT) or User Equipment (UE). In order to keep the nomenclature as general as possible and to highlight a fully mobile equipment, in this thesis the acronym Mobile Terminal (MT) is used.

set of requirements collected in a series of recommendations which define the International Mobile Telecommunications Advanced (IMT-A). Though 4G is a broader term and could include standards outside IMT-A, 4G is here associated with IMT-A specification as published in [ITU08]. The requirement for IMT-A candidates are broadly detailed in Appendix B.

The data rate targets for research have been set to 100 Mbit/s for high and 1 Gbit/s for low mobility and MIMO has a crucial role and it has been intensively proposed for the PHY layers of 4G systems.

As well defined the technical specifications of the upcoming 4G networks are, as many doubts lay on the ground regarding which will be the standard that will lead the commercial deployment of 4G infrastructures and devices. Two big standardisation bodies are working to submit a candidate for IMT-A evaluation process: the data-centric Worldwide Interoperability for Microwave Access (WiMAX) and the voice-centric Long Term Evolution Advanced (LTE-A). Although the different vision, the latter evolves from his Third Generation (3G) predecessors, instead WiMAX, which was started before LTE-A, adopts a clean-slate approach. The two standards have profound technology similarities and are based on Orthogonal Frequency Division Multiplexing (OFDM) and heavily exploit MIMO antenna technology to attain very high spectral efficiency at the PHY layer.

On the opposite side, each standard has its peculiar focus and hence they can be seen as non competing but complementary technologies. The commercial success or predominance of one against the other is not to be forecasted on their technical specifications but considering the combination of technology performance, supported services and applications, deployment coverage and requested initial investments.

The performance evaluations have been carried out in the framework of WiMAX, following the specifications of Institute of Electrical & Electronics Engineers 802.16 (IEEE802.16) standard and referring in particular to IEEE802.16e and IEEE802.16-2009: the versions including high mobility. Nonetheless, what is proposed here is applicable also in LTE-A. This is possible since the two PHY layers share the same space-time-frequency-antenna resource allocation mechanisms and both MACs are connection-oriented.

Structure of this Thesis

The structure of this thesis is briefly outlined in this section. The thesis is composed by five chapters.

Chapter 1 presents an introduction to the topic of this thesis. Single User (SU) MIMO theoretical capacity is presented. While numerous information theoretic results are available for SU MIMO, research is currently on-going for several open problems in MU MIMO. After initial enthusiasms during the first experiments more than one decade ago, MIMO has found several practical difficulties to achieve the theoretical promised gains on a single-radio link. The cross-layer design of multiple antenna PHY together with MAC can be the key technical approach to push for better resource exploitation with no (or limited) CSI feedback. The MU MIMO-ARQ system proposed in this thesis is briefly outlined.

In *Chapter 2*, adaptive strategies at the PHY layer are analysed. Dynamic MIMO scheme selection based on channel status is presented for SU radio-links. The result is a PHY layer with variable Spectral Efficiency (SE) driven by the average SNR. At low SNR, diversity schemes are used for lower bit error rates. When the channel is in a good state, spatial multiplexing can be used, transmitting multiple data streams over the same time-frequency PHY layer resource. A 4×2 MIMO system is studied which

permits to use a Layered STBC (LSTBC) where both spatial multiplexing and diversity gains are simultaneously achieved.

Then, the dynamic selection of STBC/SFBC over OFDMA is studied. A performance loss is experienced in STBCs if the channel is not constant over the length of the code. The constancy of the channel is usually not verified in the case of high velocities of the MTs. Based on Wide Sense Stationary - Uncorrelated Scattering (WSSUS) channel model and on the measures of delay spread and Doppler shift, the proposed system can adapt the symbols allocation over adjacent subcarriers or adjacent OFDMA symbols. This mitigates the detriment effect of time varying channel for STBC. A closed-form analysis for the Symbol Error Probability (SEP) is also proposed, where the error rate is derived using the Moment Generating Function (MGF) of the instantaneous SNR at the output of the Alamouti combiner for non static channels.

Chapter 3 analyses two protocols at the MAC layer: namely the ARQ and the DSA. The length of ARQ retransmissions is optimised jointly with adaptive MCS at the PHY layer. This results in bandwidth savings with time varying channels which are modelled including a SNR log-normal distribution (shadowing). The DSA is a three way handshake protocol used to admit service-flows in the network. Signalling blocking probability and admission control blocking probability are derived in closed-form for a block fading channel. The protocol is evaluated with fast varying channels providing insights over the DSA parameters impact on service-flow activation for fast MTs.

In *Chapter 4*, a set of cross-layer designs is proposed for MU MIMO network. The system employs open-loop MIMO schemes - Alamouti and Vertical BLAST (VBLAST) - together with ARQ. The MAC layer is aware of per-antenna Acknowledgement (ACK)s and/or Not ACK (NACK)s. Several re-transmission strategies are proposed, in both cases of SU and MU MIMO. The system uses a typical RR MAC scheduler. More advanced schedulers can be readily integrated in the current design. The performance are compared with fixed MIMO and adaptive MIMO schemes based on upper-bounded BERs, as presented in Chap. 2.

The MU MIMO-ARQ state machine is at the Base Station (BS). The MU MIMO-ARQ system does not require any CSI feedback but it is still able to capture a degree of MU diversity via a random user selection. The absence of CSI feedback is an important characteristic of the proposed design, as CSI feedback in MU MIMO has proven to be very capacity demanding and generally practically unfeasible with high mobility MTs.

At the end, *Chapter 5* presents the protocol stack of an IEEE802.16 network. The protocol stack is described firstly in its classical form. Then, the issues related to the implementation of cross-layer MU MIMO-ARQ strategies are discussed: an enriched PHY/MAC interface is discussed. Several schedulers proposed in literature are also reviewed: the MU MIMO-ARQ protocol proposed in Chapter 4 is ready for the integration of more advanced scheduling techniques, finally resorting in a complete PHY/MAC system.

Chapter 1

The Promises of MIMO and MIMO-ARQ designs

*La libertà è un fatto
dell'intelligenza: ed è quella che
dipende da questa, non
l'intelligenza dalla libertà.*

*Freedom is a fact of the intelligence: it
depends on that and not the
intelligence on liberty.*

KURT ERICH SUCKERT
CURZIO MALAPARTE
Maledetti Toscani (1898 - 1957)

Recently, channel coding has reached impressive results, almost reaching the Shannon capacity: in [Bou02] a BER of 10^{-5} is reached at $SNR = -0.46dB^*$ Until today, the Shannon capacity barrier has never been broken and the potent mathematics underneath Shannon formulation seem to assure that this will never happen.

The raising traffic demand calls for lightweight implementation of powerful coding but also requires new techniques for achieving higher spectral efficiencies. Space for further improvement in channel capacity was surely offered by the use of multiple antenna: MIMO have sparked a vibrant research area since a decade ago. The added dimension of signalling coming from the multiple antenna elements, offered the floor for a further leaps in performance.

MIMO transmissions have played a crucial role in the last decade in wireless communication research. Nowadays, MIMO is a fundamental component included in all the standardisation activities for the current and future wireless networks.

The first results with MIMO systems were showing impressive capacity gains. The first MIMO testbed, [Fos96], showed that capacity could be improved linearly with the minimum number of antennas installed at the

*The BER is reached as closely as $0.03dB$ from Shannon capacity with a considerably simple form of turbo-coding.

transmitting or receiving sides of the radio link. A SE of 32bit/s/Hz was achieved compared to almost unity efficiency for a SISO link over the same channel.

MIMO can support not only capacity improvements but also be crucial in several other directions: the energy efficiency can be leveraged with an improved Signal-to-Noise Ratio (SNR) coming from array gain, larger coverage areas can be achieved. MIMO also mitigates channel fading, reducing the error rates compared to Single Input Single Output (SISO) with equal link budget. MIMO can be used to partition various users in the cell, thus reducing interferences and improving the frequency re-use factor.

Beside all these opportunities, numerous practical limitations have been discovered lying between the SU MIMO theoretical results and practical implementations.

MU MIMO, e.g. the scheduling of multiple user sharing the spatial dimensions of the channel, provides again new scenarios for improvements. While SU MIMO techniques are relatively well understood, new investigations are needed to fully understand the relations among MIMO and MU networks. The interaction of multiple MIMO links impacts over PHY, MAC and Networking (NET) layers: they need to be addressed with a wide comprehensive approach. During these last few years, the information theory community has achieved important breakthroughs in the understanding of MU MIMO systems and these systems will play a crucial role in achieving the performance requirements in terms of better spectral efficiency, improved coverage and decreased latency.

In the following, first, the main results regarding SU MIMO capacity are briefly presented. The gains obtainable when MIMO is used in such a way are referred to as *MIMO processing gains*: the added degrees of signalling are used uniquely for detection and processing of the radio signals. These operations are fully contained at the PHY layer and the upper layers cannot interact with the dynamics of MIMO transmissions. Then, the *MIMO protocol gain* concept is presented, where the layers are interacting to grab all the opportunities which arise in MU MIMO networks.

1.1 Capacity in MIMO Channels

The capacity improvements promised by MIMO technology have encountered several difficulties to be achieved. The main reasons of this implosion of throughput gain come from:

- the difficulties of having an accurate and updated knowledge of the CSI
- the propagation characteristics of the radio channel when low scattering is encountered
- the correlation among antenna elements, mainly at the MT.

Classical SU MIMO channel capacity has been derived when the perfect CSI at the Transmitter (CSIT)/CSI at the Receiver (CSIR) knowledge is usually assumed, [GJJV03]. While this is usually considered a strong assumption regarding CSIT, there are cases where a very accurate knowledge of CSIR is also impracticable, especially when MIMO schemes are used with very fast MTs. The problem gets worse as the number of antennas grows: an high number of antennas is desired (beside high terminal cost), since the capacity gain is increasing linearly with the minimum number of antennas available at the transmitter and receiver sides. An high number

of antenna can exacerbates the problem of a performing CSI feedback and requires high-complexity receivers.

Some capacity results have been found for SU MIMO when only some statistical information about the channel are known: Channel Distribution Information (CDI), but several problems are still open in the field of MU MIMO.

In general, even if CSI is perfectly known, correlation among antennas and poor scattering in the radio propagation environment, cause a big loss on the channel capacity.

1.1.1 MIMO Channel Theoretical Capacity

The Shannon capacity is defined as the maximum mutual information between the channel input and the channel output. The mutual information is the maximum data rate that the channel can sustain with an arbitrarily small error probability.

In the case of static channel, the previous definition is complete. When the channel is time-varying, several capacity definitions can be expressed since the capacity has a statistical distribution. The ergodic (Shannon) capacity is the maximum mutual information averaged over all the channel states[†].

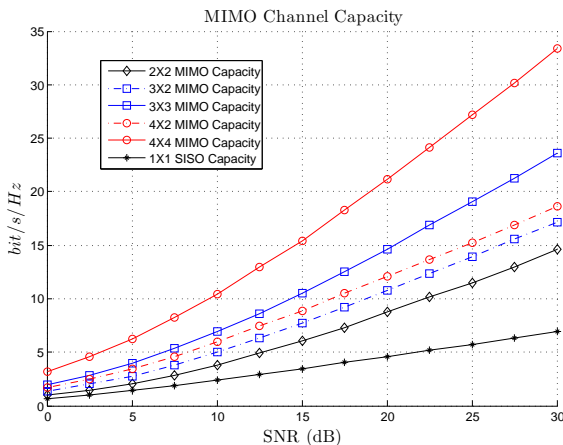


Figure 1.1: MIMO Capacity with perfectly known CSIT and CSIR.

In Fig. 1.1, the ergodic capacity for a SISO and several MIMO systems is reported with perfect CSIT and CSIR knowledge and perfectly uncorrelated antenna elements. The capacity improvement grows linearly with $\min(N_R, N_T)$.

The DL and UpLink (UL) CSI are represented by a single $N_T \times N_R$ matrix \mathbf{H} where $\mathbf{H} = \{h_{ij}\} \in \mathbb{C}$ with complex entries. Each entry represents the channel baseband equivalent h_{ij} from transmitting antenna j^{th} to the

[†]Other common capacity definitions are: the outage capacity and the minimum-rate capacity.

receiving antenna i^{th} . In the case where the channel reciprocity does not hold, DL and UL links can have substantially different capacities.

The channel capacity (which is deterministic, since the channel is static) is expressed as:

$$C = \max_{\mathbf{R}; \text{Tr}(\mathbf{R}) \leq P_T} \log |\mathbf{I}_{N_R} + \mathbf{H}\mathbf{R}\mathbf{H}^\dagger| \quad (1.1)$$

where \mathbf{I} is the identity matrix, $\text{Tr}(\cdot)$ is the trace operator, \dagger is the transpose conjugate operator, P_T is the total maximum transmit power and \mathbf{H} is the MIMO channel matrix (CSI). The capacity depends on the covariance matrix of the input symbols: $\mathbf{R} = E[x(\cdot)^\dagger]$, where x is the $N_T \times 1$ symbol vector at the input of the transmitting antenna system[‡]. The maximum capacity is achieved when the inputs are Gaussian independent streams[§].

In the case of fading (not constant channel), the result when CSIT and CSIR are perfectly known is straightforward and it is expressed as the expectation of the capacity over all the channel instantaneous realisations:

$$C = E_{\mathbf{H}} \left[\max_{\mathbf{R}; \text{Tr}(\mathbf{R}) \leq P_T} \log |\mathbf{I}_{N_R} + \mathbf{H}\mathbf{R}\mathbf{H}^\dagger| \right] \quad (1.2)$$

When perfect CSIR is available, but CSIT is not available, the capacity has been evaluated by Foschini in [FG98] and Telatar in [Emr99] assuming that the channel is Zero-Mean Spatially White (ZMSW), i.e. MIMO matrix \mathbf{H} is assumed to have uncorrelated i.i.d. complex Gaussian entries. This assumption is very strong since it does not consider the followings effects usually encountered in practical systems:

- Antenna elements correlation
- Coloured Spatial Channel: radio channel with low scattering and \mathbf{H} with low rank

In the ZMSW model without CSIT, the problem is still to find the correct input covariance matrix to maximise the channel capacity, i.e. to chose \mathbf{R} such that the following is maximum:

$$C(\mathbf{R}) = \max_{\mathbf{R}; \text{Tr}(\mathbf{R}) \leq P_T} C(\mathbf{R}) \quad (1.3)$$

where $C(\mathbf{R})$ is:

$$C(\mathbf{R}) = E_{\mathbf{H}} \left[\log |\mathbf{I}_{N_R} + \mathbf{H}\mathbf{R}\mathbf{H}^\dagger| \right] \quad (1.4)$$

In this case, it is found that the input covariance matrix which maximise the capacity is the scaled identity matrix, i.e. the power is equally divided among all the eigenmodes[¶] and the ergodic capacity results as:

$$C(\mathbf{R}) = E_{\mathbf{H}} \left[\log |\mathbf{I}_{N_R} + \frac{P_T}{N_T} \mathbf{H}\mathbf{H}^\dagger| \right] \quad (1.5)$$

Foschini and Gans proposed a scalar code to achieve the predicted $C(\mathbf{R})$ capacity: the Vertical Bell Laboratories lAyered Space-Time Architecture

[‡]The input covariance matrix includes eventual precoding at the transmitter and antenna power allocation.

[§]In practise, the input streams cannot be Gaussian variables, but source data is modulated, i.e. represented with discrete symbols taken from digital modulations. This is also a source of capacity degradation for MIMO channels when low-order modulations are used.

[¶]Intuitively, since no information is known at the transmitter, the best uninformed choice is to divide the power equally over each antenna.

(BLAST) (VBLAST) vector coding scheme, [Fos96]. At 1% outage, at $SNR = 12dB$, they shown, with a 12×12 antenna system, a spectral efficiency of $32bit/s/Hz$ compared to $1bit/s/Hz$ for the SISO channel with perfectly uncorrelated antennas and flat-fading channels^{||}.

When also CSIR is not available, which is the case for highly mobile terminals, the theoretical capacity achievable is even lower. In the case that only CDIT and CDIR are known, the MIMO channel capacity has been demonstrated to grow only with the double log of the SNR with the number of antennas as a constant additive term, [ZT00].

MIMO Channel Rank

So far, the problem of CSIT and CSIR availability has been discussed. However, even if CSIT and CSIR are perfectly known, the propagation characteristics of the MIMO channel can cause severe losses in capacity. In particular, this depends on the MIMO channel matrix rank. High ranks indicates a radio channel with rich scattering which leads to several independent spatial channels. If the channel rank is low, this means that only a principal propagation direction is available and that the spatial paths are highly correlated. In App. C at page 167, MIMO capacity is analysed from the perspective of the Single Value Decomposition (SVD) of the channel matrix to highlight the impact of MIMO channel rank on the achievable capacity.

As briefly outlined, theoretical MIMO capacity gain assumes perfect CSIR/CSIT and Spatially White Channels which are very difficult to achieve in practical system designs. In the following section, a MIMO ARQ system design is proposed for exploitation of MU diversity aiming to a performance improvement over the SU MIMO systems.

1.2 From MIMO signal processing gain to MIMO protocol gain

Historically, MIMO has been employed as a PHY layer booster, where added degree of signalling is used to reduce error probability or to communicate at higher data rates.

For several reasons, MIMO has failed to meet the promises. This is mainly due to the difficulty in having decorrelated antenna elements (mainly at the MTs) and to the fact that the radio channel may present not-suitable conditions for spatial transmissions. Shifting from SU to MU MIMO transmissions can be a key to recover some of performance losses with a cross-layer design.

Cross-layer designs and layered-layerless architectures have been long debated in the research community, with evolutionary and revolutionary attitudes, [Pra08]. Without any doubt, layering plays an important role in the development of wireless standards and sustains commercial viable solutions, guaranteeing a stable setting for introducing technical ameliorations in a coordinated way. However, it is also widely admitted that layering imposes performance losses due to the sub-optimality of the current stack subdivision for multiple-access channels. In [Pop06], an interesting overview of the layered-layerless topic is proposed discussing the fundamental features of a MU communication network.

^{||}In a more realistic system, as Gans reported in [GAY⁺02], the real capacity improvements were in the order of the 30% of the idealised setting. In [GAY⁺02], the results were still measured with independent antenna elements and totally avoiding any kind of interference.

SU MIMO is acting as a performance booster on the single radio link and higher layers benefits from PHY greater Spectral Efficiency but are not aware of what transmission is being used and cannot exploit multiple access mechanisms at the antenna level. The recent developments of cross-layer techniques have made widely recognised that separating channel access, scheduling and MIMO schemes is limiting the total capacity attainable. MIMO is not anymore regarded as a pure PHY technology, [GKH⁺07].

MU-MIMO is based on the spatial sharing of the channel, where several users can be allocated along the spatial dimensions with an upgrade respect to SM schemes in SU links where all the spatial directions were used for a single user. MU-MIMO requires additional hardware cost and more complex scheduling and resource allocation techniques but can achieve very high spectral efficiencies without bandwidth expansion. MU-MIMO can provide gains in several directions:

- MU-MIMO is more robust to channel rank-loss coming from low scattering and antenna elements correlation at the MT
- MU-MIMO can use SM schemes with single antenna MTs (via uplink collaborative schemes) reducing the cost of terminals and concentrating the wide antenna set at the BS
- MU-MIMO offers a capacity improvement exploiting multiuser diversity

On the opposite side, some drawbacks limit also the practicality of MU-MIMO schemes. From an information theoretic point of view, in MU-MIMO DL Broadcast Channel the sum-rate capacity is achieved with Dirty Paper Coding (DPC), [Cos83], which is an interference cancellation technique combined with user scheduling and power control.

Firstly, the complexity of signal processing and the feedback needed for DPC are too heavy to be implemented in a practical system and a plethora of solutions have been proposed to approach, as close as possible, the capacity upper-bound represented by DPC.

Secondly, user selection requires to solve an optimisation problem with a search space proportional to the number of users. Brute force search can be unfeasible and several trade-off solutions have been proposed, as in [WML02].

Thirdly, the CSI feedback overhead can deeply affect the performance of the network in terms of throughput. A lower feedback MU-MIMO scheduling with linear and non-linear precoding is presented in [GKH⁺07]. CSI feedback is challenging for two main reasons: feedback channel capacity is proportional to the number of users and antennas and, for highly mobile MTs, the information quickly becomes outdated. The heavy feedback poses serious limitations on the applicability of DPC and other closed-loop MIMO techniques even for slowly varying channels. For these reasons, partial (or no) feedback schemes have been proposed, [SU09, XZ09, DL08, KK08], just to cite some recent works.

A simple MU-MIMO system design with low feedback was proposed in [VTL02]: Opportunistic Random BeamForming (ORBF). The system proposes a reduced feedback system assuming that the channel fading is not too fast and the number of users in the system is not too low. The antenna system is a dumb set of isotropic radiators where each one transmits a signal multiplied with a time-varying random complex coefficient (one or each antenna) communicated by the MTs to the BS. The aim is to artificially create large and fast channel fluctuations. Large fluctuations allow to schedule each user on the peak of its induced channel state and to maximize the system throughput. Fast changing random coefficients are

able to follow fast fading peaks but require high feedback rates: to maintain the amount of feedback tolerable the coefficients cannot be chosen arbitrarily fast. However, if the period of the coefficients is longer than the correlation time of the fast fading process, ORBF will result in poor/no gain.

ORBF needs also an high number of users in order to have an effective system throughput gain. In [KHGT05], the issue of a low/realistic number of MTs has been addressed using more random coefficients inside a single transmission slot but exacerbating the rate of the feedback. In [KG05], a memory-based approach is proposed to exploit the time correlation of the channels when the number of users is low.

So far, the MIMO schemes presented needed CSIT: the feedback problem can be solved resorting to open-loop MIMO schemes, such as STBC or SM, which do not require CSIT. These schemes can be exploited to achieve diversity and/or rate gain; the diversity/spatial multiplexing trade-off** has been settled down in [ZT03]. The optimal scheme can be dynamically selected based on the spatial characteristics of the MIMO channel.

A switching system based on the Demmel condition number^{††} of the channel matrix was presented in [HP05]. The system in [HP05] is able to select among MIMO scheme with equal spectral efficiency: a 16-QAM Alamouti and a 4-QAM SM schemes, for instance, can be optimally chosen based on constellation distances expected at the receiver, as described in Sec. 2.5.1. The decision is taken at the receiver based on locally estimated CSI and need to be communicated back to the BS. However, adaptation to the channel state can be challenging in very fast channels and, as a rule of thumb, the time to elapse the adaptive procedure needs to be at least one-tenth of the channel coherence time.

A variable spectral efficiency at the PHY layer can be obtained when adaptive MIMO schemes are chosen based on CSI maintaining the BER under a target BER threshold, as in [MMPR07] and as reported in Sec. 2.5 at page 25.

1.3 Proposed MU MIMO-ARQ Protocol

A new approach for the exploitation of MIMO advantages is proposed in this thesis which is based on the integration of MIMO, ARQ and scheduling without CSI feedback. With no loss of generality, a 2×2 antenna system is studied for DL transmission in a MU network.

A simpler MIMO-ARQ protocol was firstly proposed in [dCP08] for SU transmission. The proposed retransmission strategy is summarised here. The MT had a Minimum Mean Square Error (MMSE) with Successive Interference Cancellation (SIC) and Packet Combining (PC). The protocol for single user operates as in the following. In the first time slot, two packets are sent in SM, then the system evolves depending on the ACKs of the previously transmitted packets as:

- when there are two packets in error, the packets are retransmitted with STBC (Alamouti)
- when there is only one packet in error, the packet is retransmitted in SM with a new packet. The packet is retransmitted from the weaker

**The formal definitions of diversity and rate gain are discussed in Chap. 2 in Sections 2.2 and 2.5.

^{††}The Demmel Condition Number is a particular form of the condition number. The condition number, in general, expresses how reversible a matrix is. The condition number is related to the product of the norm of the matrix times the norm of the inverse. The Demmel condition number is defined in Sec. 2.5.1 at page 26.

antenna such that the combined signal does not get excessive SNR. The weaker antenna is the one which gives the lowest post-processing SNR at the output of the receiver

- when the two packets are received correctly, two new packets are sent to the user in SM

As deeply analysed in Chap. 4, the system proposed in [dCP08] offers improvements over the adaptive MIMO systems presented in Chap. 2. However, the antenna allocation needs one bit feedback and the ARQ state machine cannot be realised as a free-running state machine^{††}.

SU MIMO-ARQ could be straightforwardly applied in MU systems in a Time Division Multiplexing (TDM) fashion. When the user is scheduled, the system retransmits the packets until they are received or until the maximum number of ARQ retransmissions are achieved prior moving to the next user in the scheduler queue.

With the SU strategy applied in TDM, in the case of dishomogeneous channel states, a MT in bad channel state forces the system to wait until the ARQ protocol elapses the full retransmission time reducing the total system throughput and serving the users with higher average delays.

A MU MIMO-ARQ protocol is proposed which integrates the ARQ process and open-loop MIMO schemes in a MU network with an implicit Round Robin (RR) scheduler with any need of CSI feedback.

Several retransmission strategies are proposed. Firstly, two strategies are studied. Assume that in the first time slot two packets are transmitted in SM. In this case, one strategy is to retransmit the wrong packet from both antennas to the same user; a second strategy is to retransmit the wrong packet with double power from a single antenna while the other is switched off.

A third strategy is proposed, which shows the best performance, and observes the following: when there is one packet in error for user (1), the packet is assumed to have been transmitted from the weaker antenna. The antenna allocation is thus switched (with no need of feedback signals) and the packet is retransmitted from the stronger antenna. A packet for a new user ($u + 1$) is transmitted on the other antenna. The retransmission on the stronger antenna permits to have the lowest service time for the user (1) thus reducing the average delay in the system and maximising the throughput. The antenna allocation follows the opposite principle with respect to [dCP08].

The new user is chosen randomly. The Hybrid SM transmission – the data streams are sent to different users – attains higher throughput exploiting the multi-user diversity. The performance is compared to the case where the best user is chosen in order to maximise the throughput (i.e. the user with the highest post-processing SNR at the output of the receiver).

The proposed MU MIMO-ARQ protocol principal features are:

- No feedback at the BS for antenna allocation
- No feedback at the BS for user selection
- Open-loop STBC and SM are used (no CSIT needed) with seamless diversity/rate gain selection
- Integration of the ARQ state machine with the packet scheduler

^{††}i.e. an ARQ state machine which does not any feedback with the exception of the ACKs/NACKs.

An optimised SIC-MMSE receiver is proposed in Sec. 4.6 at page 100. The MT selected for Hybrid SM transmission stores and exploits unintended packets sent in previous time slots to other users. The MT removes the contribution of the "overheard" packets to maximise the probability of the good reception of his packet.

The system design allows for integration of any kind of scheduler accounting for Quality of Service (QoS) attributes of the traffic flows. More advanced schedulers that can be combined inside the MIMO-ARQ system will be discussed in Chap. 5.

Chapter 2

PHY Layer MIMO Processing Gain

*I have not failed, I've just found
10.000 ways that won't work.*

THOMAS ALVA EDISON
(1847-1931)

2.1 Introduction

In this chapter, MIMO is analysed as a performance enhancer for SU PHY layer, this gain is referred to as the *MIMO processing gain*. When MIMO is employed at the PHY layer, there is a performance enhancement which is originated from the processing of the signals with the added degrees of freedom coming from multiple antennas.

The gains of multiple antennas can be divided into three main categories. MIMO can provide *diversity* gain exploiting the independent fading in the multiple antenna links to transmit several coded version of a data stream to enhance signal diversity with STBCs and Space Time Trellis Code (STTC)s. The first category is based on orthogonal designs, [TSC98, Rad22] while STTC distributes a trellis code over the antennas.

MIMO can also increase the rate (especially at high SNR) via *spatial* gain – Spatial Multiplexing – when multiple data streams are transmitted from different antennas and arrives at the receiver with sufficient different spatial signatures. SM can be implemented with or without CSIT.

A third way is to use MIMO for beamforming or precoding. In closed-loop, the MT sends back the CSI information at the transmitter and the transmitter "precodes" the signal, for instance, in order to maximise the Signal to Interference Noise Ratio (SINR) at the MT. The precoding is done via multiplication of each antenna signal with a complex coefficient, thus modulating the amplitude and phase of the antenna signals individually

such that when added together they add constructively after being passed through the channel.

The diversity gain can be obtained exploiting every domain of the wireless transmission. Typical examples are delay diversity and frequency diversity where signals are repeated over time or over different frequencies*.

For multiple antennas, an effective way of providing diversity is through STBC; usually multiple transmissions are done over the two dimensions of space (antennas) and time (successive time slots), thus the name Space-Time Block Coding. STBCs assume that the channel is constant over time for the whole duration of the STBC block. The performance is degraded if the channel is varying over the successive time slots[†].

The simplest STBC is the Alamouti code, [Ala98]. In its base version, it employs two antennas at the BS and one antenna at the MT. The scheme transmits two symbols from two antennas in two successive time slots, thus a total of four symbols per Alamouti-block are transmitted. Alamouti has been widely adopted because it is simple and requires low signal processing load at the receiver, which is a mobile and energy-limited portable device. A simple receiver, with a linear combiner and a single antenna, can extract an order of diversity equal to 2 if the assumption of constant channel over two time slots is met. Alamouti code with two transmitting antennas and a single receiving antenna provides a diversity gain of two; if two receiving antennas are used, the diversity gain is four. No spatial gain is achieved with STBCs[‡].

Block Coding is also possible over the two domains of antenna and frequency with Space-Frequency Block Coding. In a Multi Carrier system, such as OFDM, this means that the SFBC coded blocks are transmitted during the same symbol time but from different antennas and over various subcarriers.

Higher spectral efficiencies can be achieved with SM schemes which transmit multiple data streams in the same time slot over the same frequency allocation resource. SM needs independently faded spatial channels. This condition is easily achievable in channels with a lot of scatterers and several spatial paths between the transmitter and the receiver. The capacity supported by the MIMO channel in SM is related to the MIMO channel matrix rank. In particular, channel matrix with low condition number have highly correlated spatial paths and the throughput achieved by SM is low. Regarding MIMO channel rank impact on capacity, see also App. C. At the receiver, a BLAST [Fos96] or other Ordered Successive Interference Cancellation (OSIC) algorithms are used to extract the spatially multiplexed data streams. Pure SM schemes gives no diversity gain but a capacity gain depending on the number of parallel streams simultaneously transmitted. The number of antennas at the receiver has to be major or equal to the number of multiplexed streams.

The choice of MIMO schemes must be traded-off between total diversity STBC (robustness against fading) and SM (high rate). In [TVZ04], a rigorous definition of diversity rate is given and the tradeoff between diversity and spatial multiplexing is assessed; the impact on capacity is analysed from an information theory point of view. This tradeoff on the usage of the antenna system gives the floor for research about adaptive MIMO schemes. For instance, based on the instantaneous SNR, the best MIMO scheme (STBC or SM) can be chosen dynamically to maximise the

*See [Pro95] in Chap.14 at pag.777.

[†]Such situation is analysed in Sec. 2.6 for Alamouti scheme over uncorrelated Rayleigh channels.

[‡]Spatial and Diversity gains are formally defined in the following Sec. 2.2 and Sec. 2.3 respectively.

capacity while having the BER lower than a design threshold as presented in Sec. 2.5 at page 25.

Hybrid schemes have been presented as an alternative to MIMO switching for achieving diversity and spatial multiplexing tradeoff enhancement since they combine together diversity and rate enhancement in the same transmission. In the case of four transmit antennas and two receiving antennas, it is possible to transmit two parallel Alamouti blocks, getting a spatial multiplexing of 2 and diversity gain of 2. The scheme is referred to as LSTBC and it is described in Sec. 2.2.2.

The schemes studied in this chapter are: Alamouti code with one receiving antenna (2×1) and with two receiving antennas (2×2), VBLAST (2×2) and LSTBC (4×2).

The reference system is WiMAX IEEE802.16e in OFDM/OFDMA configuration. In Sec. 2.4, the MIMO support in the IEEE802.16 standard is described. The Adaptive Modulation and Coding (AMC) support in 802.16 is also presented, as the modulation order will be dynamically tweaked jointly with the MIMO schemes before-mentioned.

Two solutions for MIMO adaptivity at the PHY layer are proposed. In Sec. 2.5, a dynamic PHY layer – switching among Alamouti, LSTBC and SM – is studied for having the best Spectral Efficiency (SE) guaranteeing an upper-bounded BER. In Sec. 2.6, the allocation of STBC over the time-frequency grid of the OFDMA-PHY layer is analysed. The impact of the frequency correlation among subcarriers and the time variability due to mobility of the MT is studied. A general criterion for choosing between Space-Time or Space-Frequency Block Codes is assessed under the WSSUS channel assumption. The closed form of the Symbol Error Probability (SEP) is derived for the Alamouti scheme based on the Laplace transform of the channel fading distribution, as detailed in App. A.

2.2 Space-Time Block Codes

Diversity techniques are employed to counterfight fading and are widely used in wireless systems. The basic idea behind this approach is to transmit more copies of a signal which permits to be more resistant to *deep fading*: if the receiver is supplied with more than one replica of the signal, transmitted over independent fading channels, the probability that all the copies are in deep fade is greatly reduced.

For example, let's assume d frequency-non-selective Rayleigh independent channels and Binary Phase Shifting Key (BPSK) modulation: the slope of the BER curve over SNR depends on the number of replicas sent over the channels, [Pro95]. When SNR is sufficiently large (say greater than $10dB$) the bit error probability P_e can be approximated as:

$$P_e(SNR) = \left(\frac{1}{4 SNR} \right)^d \binom{2d-1}{d} \quad (2.1)$$

and for single or double transmitting antenna:

$$\begin{aligned} P_e(SNR)|_{d=1} &= \frac{1}{4} SNR^{-1} \\ P_e(SNR)|_{d=2} &= \frac{3}{16} SNR^{-2} \end{aligned} \quad (2.2)$$

The diversity gain (or order) is related to the number of parallel channels used. Since at high SNR the BER curve is characterised by its exponent,

we can define the *diversity gain* d for a transmission scheme as:

$$d = - \lim_{SNR \rightarrow +\infty} \frac{\log P_e(SNR)}{\log(SNR)} \quad (2.3)$$

where P_e is the bit error probability at fixed average SNR as expressed in (2.1).

The hypothesis of independent channels is not easily attainable and can be strengthened with several techniques. A first method is to exploit *frequency diversity*. The same signal is transmitted over d frequency carriers with a separation greater than the coherence bandwidth of the channel. This frequency separation ensures statistical independence between signal replicas. A second method provides diversity transmitting in successive time slots: *time diversity*. The signals are to be transmitted with a time separation greater than the channel coherence time.

A third interesting approach provides diversity using wideband signals. If a signal with bandwidth greater than the channel coherence bandwidth (B_C) is sent over the channel; the receiver can resolve all the multipath components. The receiver then exploits all the multipath replicas via a RAKE receiver to achieve diversity, see [Che97] and reference therein.

Other diversity techniques regard *polarisation* and *angle-of-arrival* diversity. With MIMO, diversity can be achieved over the antenna domain via STBCs or STTCs.

STBC term was first reported in [TSC98] when Tarokh, Jafarkhani, and Calderbank proposed a new scheme for transmission over wireless Rayleigh-fading channels using multiple transmit antennas. Actually, the work in [TSC98, TJC99, TJC00] can be thought as a generalisation for an arbitrary number of transmitting antennas of the Alamouti code previously published in 1998, [Ala98]. Alamouti code is the simplest STBC construction possible and it was proposed as a simple diversity technique to achieve double diversity at an MT equipped with a single antenna with a BS equipped with two antennas: Alamouti code is able to "transfer" the double diversity from the BS to the MT maintaining the more complex antenna system at the BS, given the assumption that the channel is constant over two symbols periods.

Design of STBC in [TSC98] was based on the intuition that orthogonal designs[§] could be used for the design of wireless transmission schemes; with an underlying orthogonal structure, STBCs possess full diversity order and a decoupled Maximum-Likelihood (ML) decoding algorithm avoids the exponential complexity of the ML receivers.

Orthogonal designs have particular properties; according to Radon's results on real orthogonal designs, they only exist if they have dimension of $\{1, 2, 4, 8\}$. The real orthogonal design can be used for code construction with real symbols (for instance for Pulse Amplitude Modulation (PAM) or BPSK modulations). In [TJC99] Radon's results have been extended to both rectangular and complex orthogonal designs (code construction for complex modulations).

An important characteristic of STBCs is the design rate: the rate of an STBC is defined as the ratio of the number of transmitted information symbols and the decoding delay of these symbols at the receiver (for any given number of transmit antennas).

A STBC design with a unitary rate ($R_{stbc} = 1$) means that it can send, for instance, two symbols over two transmission time slots without loss

[§]The orthogonal designs have been studied in the field of pure mathematics by J. Radon in [Rad22].

in bandwidth. In this case, the antenna resource is used to add diversity without sacrificing the capacity, [TSC98] Corollary 3.3.1, p.756.

For real orthogonal designs, it was shown that it exists a full-rate design for any given number of transmit antennas. With a real modulation, STBCs can achieve the maximum possible rate for any given number of transmit antennas.

However, regarding complex signal constellations, such as an arbitrary order Phase Shifting Key (PSK) (i.e. order different from two) and Quadrature Amplitude Modulation (QAM), the authors were not able to prove the existence of orthogonal design for any number of transmit antennas even if hypothesised. In [TSC98], the complex orthogonal design with a dimension of 2 (Alamouti code) was pointed out as an example of rate-one complex design which was probably part of a bigger family of complex orthogonal designs with unitary rate. Contrary to this logical intuition, in [LX03], it was proved that there does not exist any member of the rate-one complex designs except of that with dimension equal to two even if an infinite decoding delay is tolerated at the receiver. This result is in sharp contrast to the existence of rate-one generalised real orthogonal designs for any given number of transmit antennas.

2.2.1 Alamouti scheme

The Alamouti scheme [Ala98], also referred to as Space Time Transmit Diversity (STTD), needs two transmit antennas at the BS and provides a transmit diversity of 2 when the MT is equipped with a single antenna. In this configuration, the Alamouti is able to provide a diversity gain of two using the double antenna system at the BS instead of requiring 2 antennas at the MT. This is very beneficial, since the added costs of the multiple transceivers and antennas are concentrated in a central point and not replicated over each MT. The Alamouti scheme can offer a diversity gain of 4 in 2×2 MIMO system.

In the case of complex modulation, Alamouti scheme is the only scheme for complex modulations known (and possible) to be fully orthogonal, with full diversity and rate $R_{stbc} = 1$. The Alamouti coding matrix can be expressed as:

$$\mathbf{A} = \begin{pmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{pmatrix} \quad (2.4)$$

where symbol $s = a_{ij}$ is transmitted at the i^{th} time slot over j^{th} transmit antenna. The receiver is a linear combiner [Ala98] where symbols can be reconstructed by using orthogonal properties of space-time coding matrix \mathbf{A} .

Channel variations that occur during two symbol time intervals are a severe source of performance degradation. In [TJC99, Ala98], the channel is assumed to be constant over the time length of the code. However, this assumption is usually unverified especially when the code length is quite long and MTs are moving at high speed. In this case, the channel value at time t can be expressed in relation to the preceding one ($t - 1$) with a correlation coefficient $\rho < 1$: the SEP was derived for the Alamouti scheme as function of ρ . In [VLB04], the BER for Alamouti scheme over non perfectly correlated channels has been derived in closed form for BPSK in a narrow band Rayleigh channel. In Sec. 2.6.4, the BER will be extended in closed form to QAM modulation of any order and used as a benchmark for the performance evaluation of the dynamic allocation of STBC proposed in Sec. 2.6.5. App. A at page 151 reports the closed-form calculation in detail for a M-QAM modulation of any order with coherent receiver.

2.2.2 LSTBC

STBC from orthogonal designs from [TSC98] has attracted considerable attention thanks to their easy ML decoding and full diversity. A drawback of such designs is that the orthogonality cannot be achieved with an high number of antennas without losing the unitary rate of the scheme. For increasing the symbol rates, the orthogonal STBC have been generalised by relaxing the orthogonality among the columns of the design, firstly presented in [Jaf01].

LSTBC is a Quasi-Orthogonal Space-Time Block Code (QOSTBC): the scheme is operating over a 4×2 MIMO system: two Alamouti blocks are "layered", i.e. transmitted from two set, composed of two antenna each. Each Alamouti block is transmitted in Spatial Multiplexing. In LSTBC, orthogonality is lost but the full diversity gain and the receiver linearity are preserved. Of course, this scheme is a tradeoff between a full-diversity (Alamouti) and a SM approach, which offers maximum capacity gain but no diversity gain, as shown in Sec. 2.5.

Since layered schemes are subject to interference among symbols transmitted from different antennas, this results in BER degradation with respect to the orthogonal case. The receiver needs a number of receiving antennas $N_R > N_{blocks}$, where N_{blocks} are the orthogonal blocks which have been spatially multiplexed (or layered). In the case of a 4-antenna BS, the STBC matrix is:

$$\mathbf{B} = \begin{pmatrix} s_1 & s_2 & s_5 & s_6 \\ -s_1^* & s_2^* & -s_6^* & s_5^* \\ s_3 & s_4 & s_7 & s_8 \\ -s_4^* & s_3^* & -s_8^* & s_7^* \end{pmatrix} \quad (2.5)$$

where the rows of matrix \mathbf{B} represent the antennas and the columns the time slots. Two Alamouti blocks are transmitted at the same time (over two symbol times) from a subset of the available BS antennas. This solution achieves the same transmit diversity as in the classic Alamouti scheme and a spatial gain which is upper limited to 2. A spatial gain equal to two is achieved when the channel matrix is orthogonal, see also C at page 167. In this case, no detection loss is experienced since cross-stream interference is null. However, if the MIMO channel rank is low, the receivers can have difficulties to extract the multiplex streams.

The receiver operates as follows, [Jaf01]. Let us consider a system with a number of transmitting antennas $N_T = 4$ and a number of receiving antennas $N_R = 2$. The channel coefficient between the i^{th} transmitting element and j^{th} receiving one at the n^{th} signalling time is $h_{ij(n)}$ and $\mathbf{s} = [s_1, s_2, s_3, s_4]^T$.

At the receiver, the following composition of the two $N_R \times 1$ received vectors \mathbf{y} at time n and $n + 1$ is considered:

$$\begin{aligned} \begin{bmatrix} \mathbf{y}^{(n)} \\ \mathbf{y}^{*(n+1)} \end{bmatrix} &= & (2.6) \\ &= \begin{bmatrix} h_{11(n)} & h_{21(n)} & h_{31(n)} & h_{41(n)} \\ h_{12(n)} & h_{22(n)} & h_{32(n)} & h_{42(n)} \\ h_{21(n+1)}^* & -h_{11(n+1)}^* & h_{41(n+1)}^* & -h_{31(n+1)}^* \\ h_{22(n+1)}^* & -h_{12(n+1)}^* & h_{42(n+1)}^* & -h_{32(n+1)}^* \end{bmatrix} \mathbf{s} + \begin{bmatrix} \mathbf{n}^{(n)} \\ \mathbf{n}^{(n+1)} \end{bmatrix} \end{aligned}$$

where $\mathbf{n}^{(n)}$ is the $N_R \times 1$ noise vector at time n .

Channel matrix \mathbf{H} has peculiar orthogonality properties that can be exploited in order to separate the spatially multiplexed Alamouti blocks. Let us partition the channel matrix as $\mathbf{H} = [\mathbf{H}_1 | \mathbf{H}_2]$.

The Least-Square (LS) receiver is obtained from the Moore-Penrose pseudo-inverse \mathbf{H}^+ , as follows:

$$\hat{\mathbf{s}} = \mathbf{H}^+ \mathbf{y} = \mathbf{U} \mathbf{V} \mathbf{y} = \begin{bmatrix} \mathbf{H}_1^H \mathbf{H}_1 & \mathbf{H}_1^H \mathbf{H}_2 \\ \mathbf{H}_2^H \mathbf{H}_1 & \mathbf{H}_2^H \mathbf{H}_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{H}_1^H \\ \mathbf{H}_2^H \end{bmatrix} \mathbf{y}. \quad (2.7)$$

The matrix \mathbf{U} contains the auto-products of Alamouti blocks \mathbf{H}_1 and \mathbf{H}_2 on the principal diagonal and it has two cross-products representing the interferences due the spatial multiplexing. Let $(\cdot)^H$ and $(\cdot)^T$ be the hermitian and real transpose operators, respectively. Due to Alamouti orthogonality, the auto-products of Alamouti blocks are:

$$\mathbf{H}_1^H \mathbf{H}_1 = k_1 \mathbf{I} \quad \mathbf{H}_2^H \mathbf{H}_2 = k_2 \mathbf{I}$$

where

$$k_1 = \|\mathbf{h}_1\|^2 + \|\mathbf{h}_2\|^2 \quad k_2 = \|\mathbf{h}_3\|^2 + \|\mathbf{h}_4\|^2.$$

The vector \mathbf{h}_i contains $N_R \times 1$ channel coefficients from the i^{th} transmitting antenna. The anti-diagonal elements of matrix \mathbf{U} satisfy:

$$\mathbf{H}_1^H \mathbf{H}_2 \mathbf{H}_2^H \mathbf{H}_1 = (\epsilon + \chi) \mathbf{I}$$

where $\epsilon = |\mathbf{h}_1^H \mathbf{h}_4 - \mathbf{h}_2^T \mathbf{h}_3^*|^2$ and $\chi = |\mathbf{h}_1^H \mathbf{h}_3 + \mathbf{h}_2^T \mathbf{h}_4^*|^2$.

It can be shown, [ZVRLG03], that a Moore-Penrose channel inversion can be attained using linear processing exploiting orthogonal properties of the LSTBC and the estimated transmit vector $\hat{\mathbf{s}}$ can be written as:

$$\hat{\mathbf{s}} = \frac{1}{k_1 k_2 - (\epsilon + \chi)} \overbrace{\begin{bmatrix} k_2 \mathbf{I} & -\mathbf{H}_1^H \mathbf{H}_2 \\ -\mathbf{H}_2^H \mathbf{H}_1 & k_1 \mathbf{I} \end{bmatrix}^{-1}}^{\mathbf{F}_2} \overbrace{\begin{bmatrix} \mathbf{H}_1^H \\ \mathbf{H}_2^H \end{bmatrix}}^{\mathbf{F}_1} \mathbf{y} \quad (2.8)$$

First, the receiver vertically processes two spatially multiplexed streams with matrix \mathbf{F}_1 . Then, with matrix \mathbf{F}_2 , it cancels cross-products (interferences). According to (2.8), SNR undergoes a reduction that is proportional to $(\epsilon + \chi)$. The equivalent SNR values for the estimation of the received symbols are:

$$SNR_{12} = \frac{k_1 - (\epsilon + \chi)}{k_2} \quad \text{and} \quad SNR_{34} = \frac{k_2 - (\epsilon + \chi)}{k_1}. \quad (2.9)$$

LSTBC offers the ability to transmit simultaneously to two different users at the cost of a BER degradation due to the spatial multiplexing of Alamouti blocks. This can be very important when multiple connections are active in a cell and strict QoS requirements are to be met, in particular regarding delay constrains, and a multi-user scheduling on the same time-frequency resource is desirable.

2.3 Spatial Multiplexing

Multiple antennas can be exploited to leverage the rate of the system, when the channel is in a good status, instead of being used for strengthening the radio signal against deep fading. The diversity gain has been defined in the previous section at page 14. The definition of multiplexing gain is similarly expressed as:

$$r = \lim_{SNR \rightarrow +\infty} \frac{R(SNR)}{\log(SNR)} \quad (2.10)$$

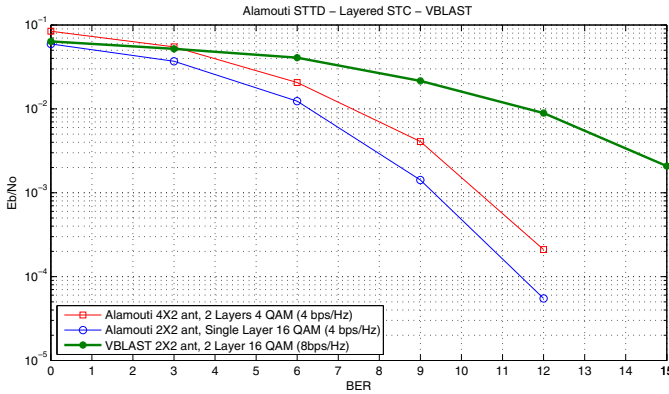


Figure 2.1: STTD and Layered STBC comparison.

where $R(SNR)$ is the rate achievable for a specified SNR.

As already discussed in Chap. 1, the ergodic capacity of the MIMO channel in SM scales with the minimum of the number of receiving and transmitting antennas as:

$$R(SNR) \approx \min(N_T, N_R) \log(SNR) \quad (\text{bps/Hz}) \quad (2.11)$$

The first and most famous SM implementation is the Vertical BLAST. Independent data streams are spatially multiplexed. Each data stream is referred to as a layer. Other forms have been proposed in literature such as Diagonal BLAST (DBLAST) and as in [JPWY08, Mou09, ZBDS06].

Since VBLAST coding is operating over a single transmission time slot, it is a space-only coding and the coding matrix is reduced to a vector (VBLAST is sometimes called Vector Modulation):

$$\mathbf{C} = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \quad (2.12)$$

Layers can be separated at the receiver if the number of receiving antennas is $N_R > N_{streams}$. While Alamouti and its hybrid version (two layered Alamouti blocks) do not exploit all the degree of freedoms of the MIMO channel, VBLAST can extract the complete spatial gain. On the other hand, VBLAST does not achieve any diversity gain.

Optimum decoding order has been demonstrated in [WFGV98] and is obtained detecting the highest SNR layer at each OSIC step. At each stage, linear combinatorial nulling is applied to remove interference caused by undetected layers. VBLAST receiver based on Zero Forcing (ZF) linear detection technique can be briefly summarised with the following steps. The initialisation phase is composed by the followings phases: initialisation, cancellation and update.

Initialisation

I1 $\mathbf{G}_1 = \mathbf{H}^+$, the channel matrix is inverted

I2 $k_1 = \arg \min_j \|(\mathbf{G}_1)_j\|^2$, the index of the column of \mathbf{G}_1 with minimum squared norm is found

Then a loop is performed for cancellation – the number of times the loop is executed is equal to the number of receiving antennas N_R :

Cancellation

C1 find the highest SNR received signal out of N_R signals available

C2 $\mathbf{w}_{k_i} = (\mathbf{G}_i)_{k_i}$, compute the weight vector \mathbf{w}_{k_i} selecting the k_i^{th} column of the inverted channel matrix \mathbf{G}_i

C3 perform signal detection (weighting and slicing)

$$\begin{aligned} y_{k_i} &= \mathbf{w}_{k_i}^T \mathbf{r}_i \\ \hat{a}_{k_i} &= Q(y_{k_i}) \end{aligned}$$

C4 operate interference cancellation

$$\mathbf{r}_{i+1} = \mathbf{r}_i - \hat{a}_{k_i} (\mathbf{H})_{k_i}$$

Update

U1 update channel matrix before next OSIC iteration

$$\mathbf{G}_{i+1} = \tilde{\mathbf{H}}_{k_i}^+$$

U2 $k_{i+1} = \arg \min_{j \notin \{k_1, \dots, k_i\}} \|(\mathbf{G}_{i+1})_j\|^2$, the index for the next weighting vector is found

U3 $i + 1 \rightarrow i$, update of the loop index for the next iteration

where $Q()$ is the detection operator, $\mathbf{G}_i = \mathbf{H}^+$ is the channel matrix pseudo-inverse at the i^{th} step of the detection algorithm, $(\mathbf{G}_i)_j$ is the j^{th} row of \mathbf{G}_i , $\tilde{\mathbf{H}}_{k_i}$ denotes the matrix obtained by zeroing columns k_1, k_2, \dots, k_i and $()^+$ denotes the Moore-Penrose Pseudo-inverse. The process is repeated until the last interference-free layer has been detected.

As in the former hybrid transmission scheme, the receiver here described can be used in a multiuser context and data layers transmitted with a SM approach can belong to a single or multiple users as specified in the frame map broadcasted from the BS, according to 802.16e standard: Hybrid SM transmission. Hybrid SM will be studied in conjunction with ARQ in Chap. 4. Hybrid SM will form the base for the MU MIMO-ARQ protocol proposed to achieve better opportunistic scheduling and network capacity.

2.4 MIMO support in WiMAX

IEEE 802.16 standards [IEE04, IEE06] are ready to support multiple antenna systems. In particular, 802.16e standard includes mobile support and adopts all the three solutions above: Alamouti, LSTBC and SM. Mainly due to their cost, antenna elements are principally deployed at the BS and receivers generally benefit from no more than one or two antennas. The Multi Antenna support in the standard falls inside two main classes:

the Adaptive Antenna Systems (AAS) and the MIMO schemes applied to OFDM or OFDMA PHY layer.

IEEE 802.16-2004 (also known as IEEE 802.16d) standard [IEE04] introduced multiantenna support. The “Amendment 2 and Corrigendum 1 to IEEE std. 802.16-2004”, namely 802.16e version [IEE06], added several details about AAS and MIMO [Pig06].

The support of multiple antenna systems by IEEE802.16d/e is summarised in Tab. 2.1. The columns report the 802.16 PHY layer configurations: Single Carrier (SC), Single Carrier advanced (SCa), OFDM and Orthogonal Frequency Division Multiple Access (OFDMA). The first two configuration operate in Single Carrier over a narrowband flat-fading channel. SC is design to work on the 10-60 GHz band, while the second configuration in the lower 2-11 GHz band. The difference between OFDM and OFDMA is mainly related to the allocation mechanism. In OFDM, a single OFDM symbol is to be directed to a single user, while in OFDMA the transmissions are based on frames. Each frame is divided in smaller resource units: time-frequency bins (or chunks), which can be assigned to different users. AAS is briefly described in the next section.

	SC	SCa	OFDM	OFDMA
802.16d	no	STTD, AAS	STTD, AAS	2/4 BS antennas and SM, AAS
802.16e	no	STTD, AAS	STTD, AAS	2/3/4 BS antennas, grouping, selection, STTD, LSTBC, SM, Collab. UL SM, AAS with AMC perm.

Table 2.1: IEEE802.16d/e MIMO support

Adaptive Antenna Systems (AAS)

The Adaptive Antenna Systems (AAS) scheme works as a beamforming technique. In AAS, the signals from several antenna elements (not necessarily a linear array) are weighted (both in amplitude and in phase) and combined to maximise the Signal to Interference Ratio (SIR) at the receiver side. AAS is used only in the OFDMA PHY layer configuration. The smart antennas beams are not fixed and can place nulls in the radiation pattern to cancel interferences and mitigate fading, with the objective of increasing SIR and the SE of the system.

The IEEE 802.16d standard defines a signalling structure that enables the use of adaptive antenna system. A Point to MultiPoint (PMP) frame structure is defined for the transmission on downlink and uplink using directional beams, each one covering one or more MTs. The BS forms a beam based on channel quality reported by the MTs.

The part of the frame with adaptive antenna transmission is sent in a dedicated AAS zone, which spans over all the subchannels until the end of the frame or until the next permutation zone, as shown in Fig. 2.2. AAS support for AMC permutation has been added in IEEE 802.16e version of the standard.

When a Discrete Fourier Transform (DFT) size greater than or equal to 512 is used, the BS can decide to allocate an AAS Diversity-Map Zone. The Diversity-Map Zone is placed accordingly to the permutation used. In Partially Used Subcarriers (PUSC), Fully Used SubCarriers (FUSC), and

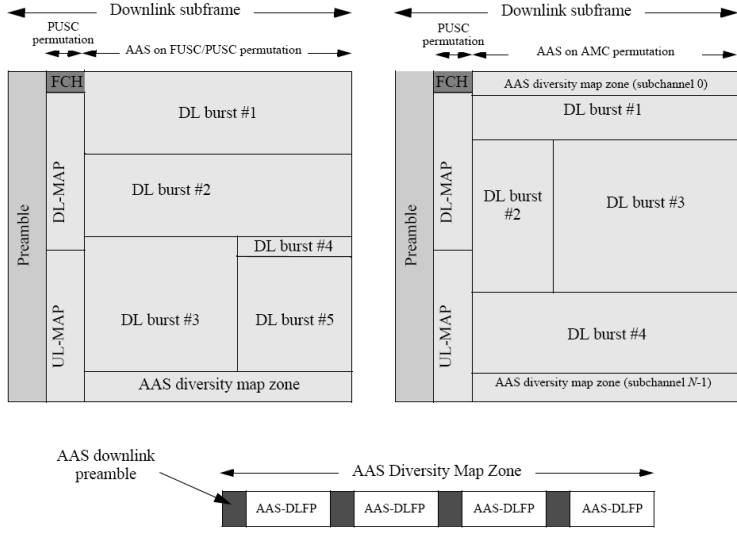


Figure 2.2: AAS Frame Structure in IEEE802.16

optional FUSC permutation, it is positioned on the two highest numbered subchannels of the DL frame, while in AMC it is positioned on the first and last subchannels of the AAS Zone, [IEE06].

These subchannels will be used to transmit the AAS-DownLink Frame Prefix (DLFP) whose purpose is to provide a robust transmission of the required BS parameters to enable MT initial ranging, as well as the beam selection via beam indexes and the resource assigned in the AAS zone.

In order to enter the network involving the DLFP, an AAS-MT follows a specific procedure. The Downlink Channel Descriptor (DCD) and Uplink Channel Descriptor (UCD) offer information useful for decoding and demodulation. A channel with broadcast Connection Identifier (CID) is readable by every MT. The channels are allocated accordingly to the DL/UL maps transmitted after the frame header. An MT entering the network in AAS mode follows these steps:

- the AAS-MT synchronises time and frequency by using the DL preamble
- after synchronisation, AAS-MT can receive necessary messages to identify: used modulation and coding (as the DCD and UCD pointed to by the allocation specified in the AAS-DLFP using the broadcast CID)
- then, after the AAS-MT has decoded the DCD and UCD, the AAS-MT should perform ranging
- the AAS-MT can receive a ranging response message through a DL-MAP allocation pointed to by an AAS-DLFP with the broadcast CID;
- then AAS-MT can receive initial downlink allocations through a DL-MAP allocation pointed to by the AAS-DLFP with broadcast or specific CID

- other allocations can be managed by private DL-MAP and UL-MAP allocations.

Multiple Input Multiple Output System (MIMO)

Multiple Input Multiple Output (MIMO) support has been introduced in WiMAX since the IEEE 802.16d version has been released. In IEEE 802.16e, several features have been added for OFDM and OFDMA PHY layers. These systems can be used to achieve diversity gain in a fading environment or to increase capacity.

When transmit diversity is desired, multiple copies of the same data stream are transmitted over independent spatial channels which are created by employing multiple antennas. Transmission is more robust to wireless channel fluctuations, since it is unlikely that all the channels will fade simultaneously.

When higher capacity is needed, spatial multiplexing is utilised to transmit various streams of data simultaneously over different antennas in the same time slot, over the same frequencies. If statistical decorrelation among antenna elements is available, multiple transmit and receive antennas can create independent parallel channels and the transmitted symbol can be correctly reconstructed at the receiver. Decorrelation condition can be satisfied by using antennas well separated (by more than $\lambda/2$) and/or with different polarisation.

IEEE 802.16e standard defines three MIMO operation modes: Alamouti Code, referred to as STTD, as in Sec. 2.2.1 and in [Ala98], LSTBC, as described in Sec. 2.2.2, and VBLAST SM, see Sec. 2.3 and [Fos96, EBS07].

Adaptive selection among the previous MIMO schemes, jointly with adaptive modulation and coding techniques, can offer high flexibility at the PHY layer and allows to maximise data throughput and coverage.

Alamouti STTD is standardised for SCa, OFDM, and OFDMA PHY Layers. The receiver is a linear combiner [Ala98] where symbols can be reconstructed by using orthogonal properties of space-time coding matrix \mathbf{A} as in (2.4).

For SCa (Single Carrier for NLoS operation in the frequency bands below 11 GHz), STTD is used at a burst level. A burst p , composed by QAM symbols, has length F as in Fig. 2.3. The bursts are arranged on a time basis. The STTD requires the processing of a pair of time bursts. A Cyclic Prefix (CP) is pre-appended to absorb the effect of signal dispersion due to the multipath: the length of the CP must be longer than the root-mean-square of the channel delay spread.

First antenna transmits two sequence of F symbols while the second antenna not only reverses the order in which burst are transmitted but also conjugates the transmitted complex symbols and time-reverses the sequence of data within each burst. The index n is the running position inside the burst. Due to the DFT/Inverse Discrete Fourier Transform (IDFT) operations, the bursts need also a time reversing operation which is realised by the $(F - n)_{mod(F)}$ operation, as in Fig. 2.3, where the transmission for a pair of bursts (p_0 and p_1) is reported. A portion of the bursts (U symbols) is copied to form the CP. The receiver for the STTD scheme in SCa PHY layer can be found in [IEE04] in Section 8.2.1.4.3.1.

In OFDM and OFDMA PHY layers, STTD transmission operates differently. OFDM and OFDMA are differentiated by the minimum data unit that can be manipulated. In OFDM PHY layer, Alamouti operates on two subsequent OFDM symbols composed by all the data subcarriers available in the system, while in OFDMA, STTD can operate on a single group of subcarrier in the time-frequency allocation grid (a resource chunk).

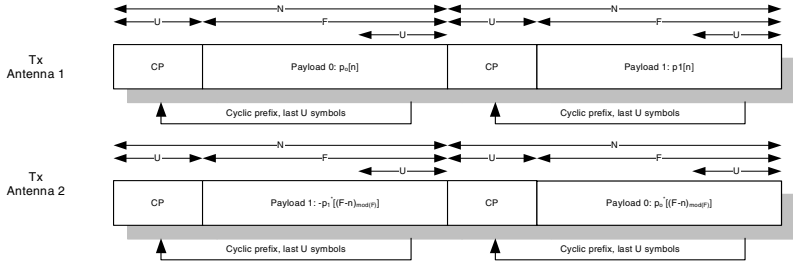


Figure 2.3: Alamouti transmission for SCa PHY layer profile.

In both cases, the symbols are coded accordingly to the matrix in (2.4): rows represent the transmitting antennas and columns represent OFDM symbols or subcarriers for OFDM/OFDMA PHY layer respectively. Compared to the SCa PHY layer case, no time-reverse operation is needed, since STBC coding is done on the single subcarrier and not on a full burst.

When the BS has three or four antennas, it is not possible to achieve a full-diversity approach, since it has been demonstrated that a full-rate, fully orthogonal Space Time Code only exists for two antennas, [HTW03]. When a full rate transmission is desired, LSTBC schemes have to be used. Data rate is increased at the expense of one of the followings: diversity gain, linearity or orthogonality. In IEEE 802.16e, the proposed scheme for four antennas is expressed by the coding matrix \mathbf{B} , as in(2.5).

LSTBC offers the ability to transmit simultaneously to two different users with the drawback of a BER degradation due to the spatial multiplexing of the Alamouti blocks.

As a third MIMO option, the system can switch on a VBLAST transmission [Fos96]. Independent data streams are spatially multiplexed, i.e., they are transmitted from different antennas in the same OFDM symbol time. VBLAST follows the coding matrix \mathbf{C} as in (2.12).

While Alamouti and its hybrid version (two layered Alamouti blocks) do not exploit all the freedom degrees of the MIMO channel, VBLAST can extract the complete spatial gain.

VBLAST receiver decodes the streams in successive steps as described in Sec. 2.3. Receiver algorithm starts from the detection of the signal with the highest SIR. The detected symbol is then cancelled from each one of the remaining received signals, and so on. It is evident the analogy with the multi-user interference cancellation. The drawback of every SIC method is that an error in the first stages of processing can propagate destructively to the successive steps of detection. The maximum SNR signal is decoded first, lowering the probability of error propagation. For a VBLAST performance evaluation refer to [ZBDS06].

In uplink transmission, an additional scheme is available: the *collaborative SM* scheme. In a multiuser collaborative SM context, two users may share the same subchannel in the uplink. In this case, an MT can request an allocation in the uplink to be used in coordination with a second MT in order to perform a Collaborative Spatial Multiplex. MT participating at the Collaborative SM can be equipped with a single antenna. The BS will separate the transmitted streams by each MT.

2.4.1 Adaptive Modulation and Coding

In wireless systems, the signal transmitted to and by a station can be modified to counter-react to the signal quality variations through a process commonly referred to as *link adaptation*. This allows to improve system capacity, peak data rate, and coverage reliability. Traditionally, wireless systems use fast *power control* as the preferred method for link adaptation. In a system with power control, the power of the transmitted signal is adjusted in order to meet a target carrier-to-interference-plus-noise ratio at the receiver. Therefore, typically, the transmit power is low when a user is close to the BS and it increases when the user moves away from the BS.

Adaptive Modulation and Coding (AMC) is offering an alternative link adaptation method that promises the increase of the overall system capacity. In a system with AMC, the power of the transmitted signal is held constant, but the modulation and coding formats are changed to match the current received signal quality. AMC provides the flexibility to match the modulation-coding scheme to the average channel conditions of each station. Users close to the BS are typically assigned higher-order modulations and high code rates. The modulation-order and/or the code rate may decrease as the distance from the BS increases.

In particular, in an OFDM/OFDMA wireless system, the inherent multi-carrier nature of OFDM allows the use of AMC according to the behaviour of each narrow-band channel (subcarrier) and different subcarriers can be allocated to different users to provide a flexible multiuser access scheme, that exploits multiuser diversity.

In this section, the Adaptive Modulation and Coding mechanisms of IEEE802.16 standard family are briefly explained since adaptive MIMO and AMC are jointly used in the dynamic MIMO *link adaptation* in Sec. 2.5 at page 25.

The key metric for link adaptation is usually the BER and is related to the SE of the scheme used. In this sense, MIMO schemes will be adaptively changed combined with different modulation orders to vary the spectral efficiency of the transmission based on the channel quality.

AMC is twofold supported by IEEE 802.16 standard [IEE04]. First, a large selection of modulation and channel coding is available at BS and MT. All the possible combinations between the modulations QPSK, 16-QAM, and 64-QAM and the coding rates $1/2$ and $3/4$ are allowed. In addition, the 64-QAM modulation can be also combined with a coding rate $2/3$. Then, a special AMC permutation scheme is defined, where subchannels are composed by groups of contiguous subcarriers.

In AMC permutation, the smallest allocation unit is called *bin* and consists of 9 contiguous subcarriers (eight data subcarriers and one pilot subcarrier) on a single OFDM symbol. Let N be the number of contiguous bins and M be the number of contiguous symbols. A slot in AMC is defined as a collection of bins of type $(N \times M)$, with $(N \times M) = 6$. Thus, the allowed combinations are: (6 bins, 1 symbol), (3 bins, 2 symbols), (2 bins, 3 symbols), and (1 bin, 6 symbols).

In addition to power control and AMC, *Dynamic Subcarrier Assignment* schemes can be used in OFDM/OFDMA systems. In particular, due to the contiguous subcarrier allocation, each user can experience highly-variable channel conditions and may benefit from multiuser diversity by choosing the subchannel with the best frequency response. High spectrum efficiency can be achieved with adaptive subchannel allocation strategies.

In conclusion, there is a wide degree of flexibility for radio resource management in the context of OFDMA. Since channel frequency responses are different at different frequencies for different users, adaptive power allocation, AMC, and dynamic subcarriers assignment, can significantly improve

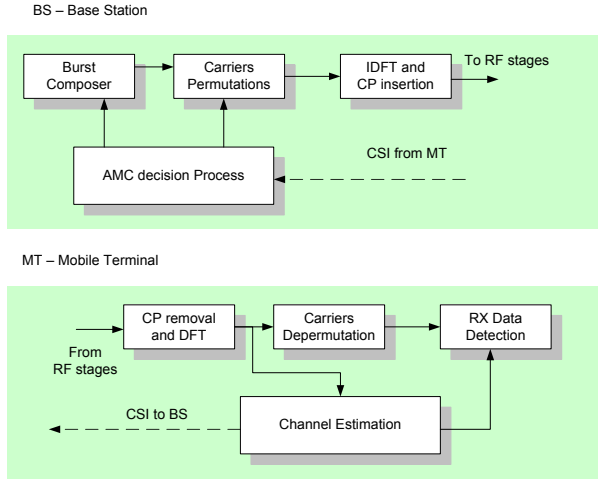


Figure 2.4: IEEE802.16 AMC System Diagram.

the performance of OFDMA systems. The link adaptation algorithms can be designed to maximise the overall network throughput or to achieve target error performance. The former objective may be appropriate for Best Effort (BE) services but does not meet QoS requirements in terms of error performance. The latter one may ensure the QoS requirements, at the expense of network throughput.

Dynamic link adaptation in the DL is obtained through MT feedback, by providing the transmitting BS with CSI estimates, as illustrated in Fig. 2.4. In the uplink the MT can change its transmission parameters based on its own channel estimates, if a good channel reciprocity can be assumed.

The accuracy of the channel estimates and the latency of the feedback affect the AMC algorithms performance. AMC strategies have to define how quickly the transmitter must change its constellation size. Since the constellation size is adapted to an estimate of the channel fade level, several symbols may be required to obtain a good estimate. These practical considerations are relevant in a mobile context with fast varying channels and require accurate solutions such as prediction models for the channels.

The MU MIMO-ARQ system proposed in Chap. 4 will address the problem of fast feedback of the CSI which is challenging for power-control and AMC. A cross-layer ARQ state machine will be designed which do not need any CSI feedback. An opportunistic multiuser scheduling will be shown to attain performance improvements comparable with an informed user selection method based on the selection of the user with the best post-processing SNR.

2.5 Link Adaptation with Dynamic MIMO

An adaptive selection of MIMO schemes is proposed considering the performance of STTD, LSTBC and SM. A selection of different transmission schemes based on the CSI can increase the Spectral Efficiency (SE) while maintaining the BER under a specified threshold. Link adaptation

is achieved via dynamic modulation (AMC strategy) and switching MIMO schemes (DIV/SM tradeoff).

This tradeoff has been analysed in [TVZ04] from an information theoretic point of view and with an analysis at high SNR. It has been demonstrated that spatial and diversity gains can be achieved simultaneously. The work quantifies the amount of diversity and multiplexing gains with ideal coding and the tradeoff between them depending on the number of antennas of the MIMO system.

With some hypothesis, this tradeoff can be simply expressed. Consider a $N_t \times N_R$ MIMO system transmitting over an identically independent distributed (iid) spatially white Rayleigh-flat-fading channel; consider the channel constant over the coding block length l ($l \geq N_T + N_R - 1$) and a coding scheme with a multiplexing gain equal to r . The optimal diversity gain achievable $d^*(r)$ can be expressed as:

$$d^*(r) = (N_T - r)(N_R - r) \quad (2.13)$$

This means that the antenna resource, in the case of ideal coding and constant channel over the block code, can be independently exploited for multiplexing or diversity gains. Out of the total $N_T \times N_R$ antenna, r are used for spatial multiplexing while the remaining $(N_T - r)$ at the transmitter and $(N_R - r)$ at the receiver are used for diversity. The multiplexing-diversity tradeoff represents the tradeoff between data rate and error probability of a MIMO system.

In the preceding case, an optimal tradeoff curve can be plotted varying the multiplexing gain from zero towards its maximum: $0 \mapsto \min(N_T, N_R)$. The correspondent diversity gain varies from $0 \mapsto N_T \times N_R$, as illustrated in Fig. 2.5.

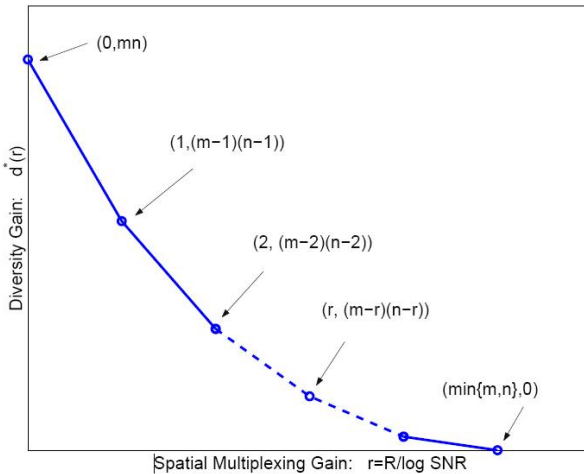


Figure 2.5: Diversity – Spatial Multiplexing Tradeoff, $m = N_T$ and $n = N_R$, from [TVZ04].

2.5.1 MIMO switching based on Euclidean Distance

The previous analysis with an ideal coding of infinite length and for $\text{SNR} \rightarrow \infty$ furnishes a theoretical tool useful to compare existing MIMO coding

schemes but has no direct implication in the design of adaptive MIMO systems.

Several designs for MIMO adaptivity have been proposed in literature. In [MCFH07] and in [MMPR07], switching algorithms between diversity and spatial schemes were proposed based on theoretical BER; in [XZG04] the effect of partial channel estimation for dynamic MIMO systems was considered.

A practical system for DIV/SM dynamic switching among equal SE MIMO schemes was proposed in [HP05] which considered a narrowband system with a Rayleigh channel. The adaptive criterion is based on the *Demmel condition number* of the channel matrix. When equal SE schemes are used, it is meaningful to assess a criterion based on the Euclidean Distance (ED) of the codebook (or constellation).

The minimum codebook ED metric can be expressed for MIMO multiplexing and diversity codes; based on instantaneous channel condition, a decision can be performed in order to adapt to the requirements in terms of symbol and error rates. Let's define the following codebook distances:

- $d_{min,SM}^2$ the squared ED for SM at the source
- $d_{min,MD}^2$ the squared ED for DIV at the source
- $d_{min,sm}^2$ the squared ED for SM at the receiver
- $d_{min,md}^2$ the squared ED for DIV at the receiver

For the SM schemes, the codebook ED at the receiver is related to the channel eigenvalue distribution, where each eigenvector corresponds to the energy of a spatial path in the channel. It can be demonstrated that for SM, the ED at the receiver is bounded as:

$$\lambda_{min}^2(\mathbf{H}) \frac{d_{min,sm}^2}{N_t} \leq d_{min,SM}^2(\mathbf{H}) \leq \lambda_{max}^2(\mathbf{H}) \frac{d_{min,sm}^2}{N_t} \quad (2.14)$$

where λ_{min} and λ_{max} are the minimum and maximum eigenvalues of the MIMO channel matrix \mathbf{H} and N_T is the number of transmitting antennas.

For the diversity schemes, the distance at receiver is directly proportional to the number of transmit antennas and inversely proportional to Frobenius norm of the matrix: an higher codebook (or constellation) ED is measured at the receiver when the eigenvalues λ_i are uniformly distributed. The following relation holds:

$$d_{min,MD}^2 \leq \frac{1}{N_t} \|\mathbf{H}\|_F^2 d_{min,md}^2 \quad (2.15)$$

where $\|\mathbf{H}\|_F^2$ is the squared Frobenius norm of the MIMO channel matrix[¶].

The Frobenius norm (or the Hilbert-Schmidt norm) can be defined in various ways:

$$\begin{aligned} \|\mathbf{H}\|_F &= \sqrt{\sum_{i=1}^{N_t} \sum_{j=1}^{N_r} |h_{ij}|^2} \\ \|\mathbf{H}\|_F &= \sqrt{\text{Tr}(\mathbf{H}^\dagger \mathbf{H})} \\ \|\mathbf{H}\|_F &= \sqrt{\sum_{i=1}^{\min(N_T, N_R)} \lambda_i^2} \end{aligned} \quad (2.16)$$

[¶]For Alamouti (2.15) holds with equality

From (2.14) and (2.15), it is clear that a criterion can be derived for MIMO scheme selection. Fixing the same transmit codebook ED at the receiver and referring to the lower bound in (2.14), a conservative criterion for choosing SM can be derived. SM is chosen when the following relation holds:

$$\lambda_{\min}^2(\mathbf{H}) \frac{d_{\min,sm}^2}{N_T} \geq \frac{1}{N_T} \|\mathbf{H}\|_F^2 d_{\min,md}^2 \quad (2.17)$$

The previous can be rewritten using the definition of *Demmel Condition Number*^{||} of the matrix \mathbf{H} . The Demmel Condition Number is defined as:

$$K_D = \|\mathbf{H}\|_F / \lambda_{\min}(\mathbf{H}) \quad (2.18)$$

Equation (2.17) can be expressed as:

$$\|\mathbf{H}\|_F^2 d_{\min,md}^2 \leq \lambda_{\min}^2(\mathbf{H}) d_{\min,sm}^2 \quad (2.19)$$

and using the definition of the Demmel condition number:

$$K_D \leq \frac{d_{\min,sm}}{d_{\min,md}} \quad (2.20)$$

The DIV/SM criterion can be summarised as the following: *the SM scheme is chosen if the ratio on the right hand side of (2.20) is greater than the MIMO channel Demmel condition number K_D .*

If small ratios of $d_{\min,md}/d_{\min,sm}$ are accepted by design, the criterion allows to exploit SM with lower scattering channels with the penalty of higher BERs.

In Fig. 2.6, the Demmel condition number measured over time is reported along with the SM-STBC switching condition calculated for 4-QAM SM (VBLAST) and 16-QAM STBC (Alamouti). The channel simulation is based on WINNER II channel model. The time sampling period is 20ms and the link is simulated in Urban Macro Cell (scenario C2) for a distance between BS and MS of 1200m and a velocity of 120 Km/h. The channel simulator is available at [Ky08]. The C2 channel model is described in [K⁺07]. It is evident that the switching condition can change several times per second: this requires a very fast adaptive stack able to switch from diversity to spatial multiplexing schemes with very low delay and hysteresis.

The switching condition C_{SW} is calculated considering the minimum distance among symbol in M-QAM modulation, [Pro95], which is:

$$d_{\min,M}^2 = \frac{12}{2^M - 1} \quad (2.21)$$

Evaluating for M=2 (4-QAM) and M=4 (16-QAM), the following is obtained:

$$C_{sw} \leq \sqrt{\frac{d_{\min,2}^2}{d_{\min,4}^2}} = \sqrt{\frac{4}{4/5}} = \sqrt{5} \quad (2.22)$$

If $C_{SW} \geq K_D$, the SM can be used, or, equivalently, SM can be used if the condition number is lower than $\sqrt{5}$.

^{||}The "standard" condition number can be calculated as: $\|\mathbf{H}\| \|\mathbf{H}^{-1}\|$ for any order of the matrix norms; in the case of Demmel condition number, it is calculated as: $K_D = \|\mathbf{H}\|_F \|\mathbf{H}^{-1}\|_2$.

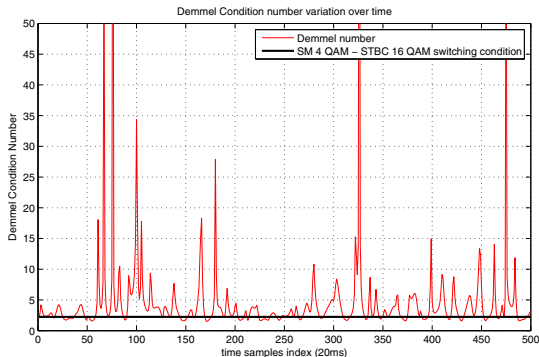


Figure 2.6: Demmel Condition Number simulated for WINNER II channel and SM-STBC Switching Condition as in [HP05].

2.5.2 Adaptive MIMO with upper-bounded BER

A selection of different multiantenna systems is analyzed in this section with the aim of keeping an upper-bounded BER over a wide range of SNRs while maximising the SE, as reported in [MMPR07]. An adaptive selection is performed based on BER of the multiantenna schemes proposed in the IEEE802.16e standard, also considering channel estimation issues. Sensitivity to channel estimation is analysed using a channel estimator operating with linear interpolation only in the frequency domain.

The channel estimator is based on pilots transmitted according to the adjacent data-pilot permutation. Channel estimator feeds the multiantenna receivers and provides a CSI measure to select the correct AMC profile and MIMO scheme.

The adjacent data-pilot permutation, or AMC permutation, is implemented as defined in [IEE04, IEE06]. The choice of this permutation permits to modulate and code the chunks independently on the basis of CSI measured by the MT.

In Figure 2.7 a 4×2 system is shown. The BS receives from the MT a control signal to operate the best selection of MIMO transmission scheme. The return link is supposed to be delay-less and error-free. The MIMO encoded streams are allocated on the OFDM symbols which are composed of encoded data, pilots, left and right guard subbands and a central DC tone. Data and pilots subcarriers are ordered according to AMC permutation. More details about IEEE802.16e AMC permutation are reported in Sec. 2.5.3. A CP is added to avoid InterSymbol Interference (ISI). The length of CP is $L_{CP} > \Theta - 1$ where Θ is the maximum length of the channel impulse response.

At MT, after CP removal and FFT, pilots are extracted to obtain channel estimation. Channel coefficients and data subcarriers are fed into MIMO receivers. The correct receiver is selected on the basis of the indication provided by the BS in the downlink map. Channel estimator operates a linear interpolation over the frequency domain, as detailed in section 2.5.3.

Received signal at time n for each subcarrier can be expressed as:

$$\mathbf{y}_n = \mathbf{H}_n \mathbf{s}_n + \mathbf{n}_n \quad (2.23)$$

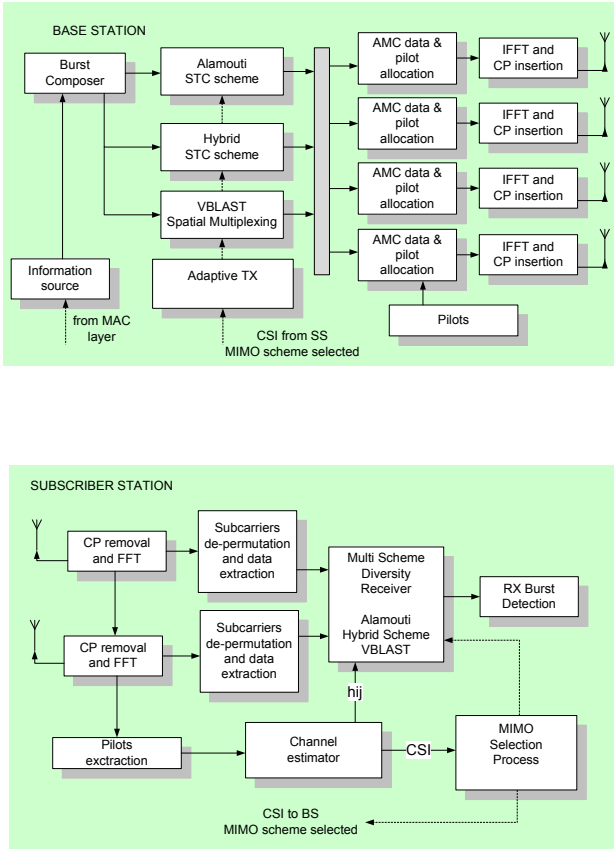


Figure 2.7: Functional blocks of the considered BS and MT

where $\mathbf{H}_n \in \mathbb{M}_{N_T \times N_R}$ is the channel matrix, $\mathbf{s} = [s_1, s_2, s_3, s_4]^T$ is transmitted symbol vector and vector \mathbf{n} is additive gaussian noise with zero mean and variance σ^2 .

2.5.3 Pilot Structure and Channel Estimation

IEEE802.16 PHY layer considers different data and pilots permutations. In PUSC and FUSC modes data positions are scrambled on the whole OFDM symbol, thus channel state is averaged in the frequency domain and upper layers cannot exploit heterogeneous link qualities with Dynamic Subcarriers Assignment. AMC permutation, instead, groups data subcarriers in adjacent blocks with heterogeneous channel qualities.

AMC permutation allows the MAC layer to chose independent profiles for each burst in order to match QoS and upper layers requests.

According to Sec. 2.4, in this work the $6 \cdot 1$ configuration is used which is allocated only in the frequency domain.

Pilots are positioned in linear ascending fashion on a regular grid, as in Figure 2.8, where three bins are shown. The processing of regular spaced

grids can be performed as indicated in [MS83].

The adopted channel estimator operates a linear interpolation over the frequency dimension to extract channel coefficients for each data subcarriers, no time-domain correlation is exploited and a new channel estimation is performed for each OFDM symbol. Channel estimation has a two-fold role: first it feeds the equalizer with frequency domain channel coefficients for decoding process, then provides a CSI measure helpful to select the correct AMC profile and MIMO scheme. Pilots from multiple antennas are multiplexed; when an antenna is transmitting its pilot tone, the other three antennas are switched off thus allowing a simple reconstruction of MIMO channel coefficients, as visible in Fig. 2.8. Feedback to the transmitter is herein considered delay-less and with perfect quantization.

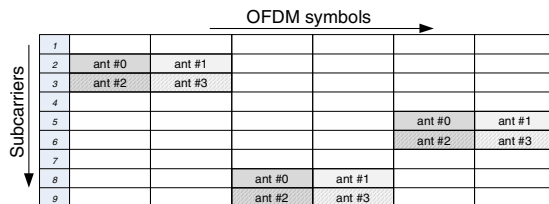


Figure 2.8: AMC permutation pilots position in 4×4 multi-antenna system.

The position of the pilots inside the symbol, in the case of $N_T = 4$, is ruled by the following equations:

$$\begin{aligned}
 \text{Pilot Location}|_{ant0} &= 9k + 3 m_{mod3} + 1 & m \text{ even} \\
 \text{Pilot Location}|_{ant1} &= 9k + 3 (m - 1)_{mod3} + 1 & m \text{ even} \\
 \text{Pilot Location}|_{ant2} &= 9k + 3 m_{mod3} + 2 & m \text{ odd} \\
 \text{Pilot Location}|_{ant3} &= 9k + 3 (m - 1)_{mod3} + 2 & m \text{ odd}
 \end{aligned} \tag{2.24}$$

where m is the time index and k is the subcarrier index.

2.5.4 Performance Analysis

Computer simulation results for the different STBC, LSTBC and SM schemes considered in the IEEE802.16e standard are presented. A spatially white MIMO Rayleigh fading channel with iid coefficients is considered. The simulated system selects a MIMO transmitting scheme accordingly to the estimated SNR. The communication is in Point to Point (PtP) from the BS to the MT.

Simulation parameters have been chosen accordingly to IEEE 802.16e standard. Tab. 2.2 reports the main parameters used in simulation. Channel is modelled accordingly to ITU Pedestrian channel A, as in Tab. 2.3. The transmission modes available (as described in the previous sections of this chapter) are summarised in Tab. 2.4, and consist in all the combination of the three scheme with the three order of the QAM modulations. The diversity d and spatial r gains of the schemes are reported.

In Figure 2.9, the Alamouti scheme with $N_R = 2$ and the hybrid LSTBC scheme with $N_T = 4$ and $N_R = 2$ are compared. Both schemes obtain the same bandwidth efficiency of 4 bit/s/Hz . In the LSTBC, the loss of orthogonality is paid with a reduction of E_b/N_o values, at fixed BER, due to interference from spatially multiplexed Alamouti blocks. The E_b/N_o

Parameter	Value
OFDM useful time	89.6 μs
CP time	11.2 μs
CP Ratio	1/8
FFT size	1024
Subcarrier frequency spacing	11.16071429 kHz
Guard subcarriers	159
Used subcarriers	865
Data subcarriers	768
Pilot subcarriers	96
Frame duration	2 μs
Per frame OFDM symbols	19 (1 preamble)

Table 2.2: Upper-Bounded BER MIMO switching parameters.

Delay		Gain	
0	ns	0	dB
110	ns	-9.7	dB
190	ns	-19.2	dB
410	ns	-22.8	dB

Table 2.3: PDP for ITU Pedestrian Type A Multipath Channel

Parameter Value					
MIMO modes	Mod	d	r	Ant.	SE bit/s/Hz
A	4 QAM	4	1	2×2	2
A	16 QAM	4	1	2×2	4
A	64 QAM	4	1	2×2	6
B	4 QAM	2	2	4×2	4
B	16 QAM	2	2	4×2	8
B	64 QAM	2	2	4×2	12
C	4 QAM	1	2	2×2	4
C	16 QAM	1	2	2×2	8
C	64 QAM	1	2	2×2	12

Table 2.4: Upper-bounded BER MIMO switching – MIMO transmission modes.

degradation is around 2 dB in the range from 2 to 12 dB. Despite of this performance loss, this scheme allows to allocate multiple user on the same time-frequency resource**.

A first adaptive system has been simulated with the constrain of a minimum spatial gain of 2 which realises a selection of the adopted MIMO architectures allowing an increase in bandwidth efficiency. The scheme which provides a diversity gain of 2 are LSTBC and SM. The selection is performed in order to maintain the uncoded BER within the operational

**Multi-User spatial multiplexing will be the key of the MIMO-ARQ protocol proposed in Chapter 4.

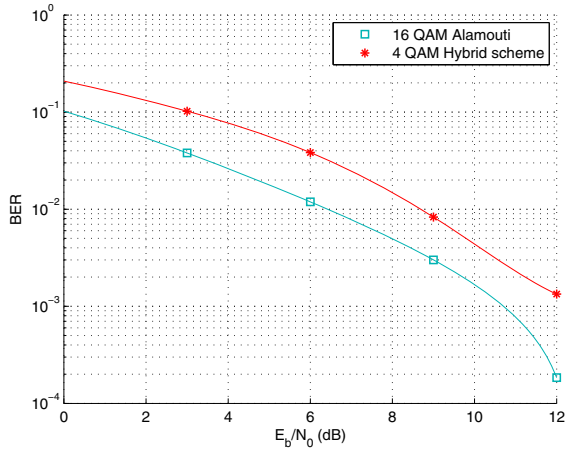


Figure 2.9: Alamouti and hybrid STC schemes comparison at low E_b/N_0 .

range of $8 \cdot 10^{-4} < BER < 10^{-2}$, as in Fig. 2.10. The adaptive scheme aims at the highest spectral efficiency and will not use LSTBC. In the high SNR range, a SE of 8 bit/s/Hz is achieved; at 8.3 dB the system doubles efficiency, from 4 to 8 bit/s/Hz

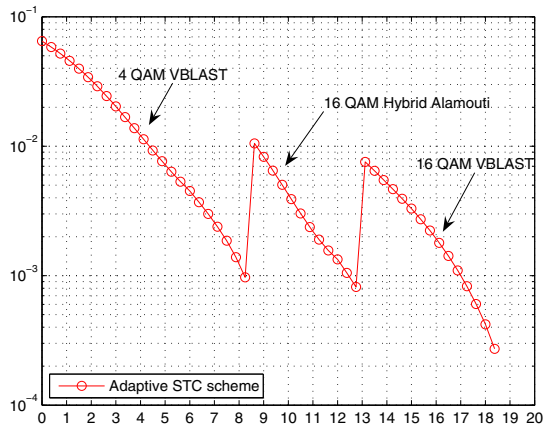


Figure 2.10: BER over SNR (dB) of a constrained Adaptive system with Multiplexing Gain of 2, ITU Vehicular Channel A.

If no diversity constrains are applied and the unique adaptation criterion is the maximisation of the system SE with a constrained BER, LSTBC will not be chosen. This situation is illustrated in the following results which report uncoded bit error rates for Alamouti and SM MIMO schemes. Modulations used are 4 – 16 – 64 QAM.

Figs.2.11-2.12 show uncoded BER for Alamouti 2x2 and VBLAST 2x2 schemes with respect to the SNR. The SE is reported. The target BER is

set to 0.01.

As shown in Fig.2.11, it is evident that in the case of flat fading channel, diversity schemes are chosen for any order of QAM modulation before resorting to spatially multiplexed transmissions and VBLAST 4QAM is never used. VBLAST SM has very poor performance at low SNR since the scheme needs a multipath environment to ensure multiple independent paths from the transmitter to the receiver. If a non-selective channel is encountered, the BLAST receiver tends to propagate errors in the decoding stages, leading to high bit error rates.

In the case of ITU multipath channel, Fig. 2.12, BER performance of 16-QAM Alamouti and 4-QAM VBLAST schemes (same SE) are equal. In this specific case, that ranges from 6 dB to 10 dB, the selection may depend on other requirements, such as multi-user allocation on the same radio resource or independent ARQ acknowledgements, as later illustrated in Chap. 4.

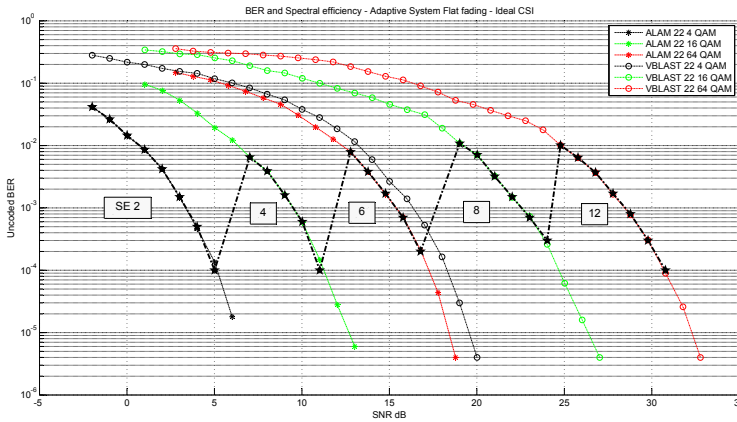


Figure 2.11: Adaptive MIMO Link Adaptation for upper-bounded BER, Flat Fading Rayleigh Channel.

A general remark is to use SM for higher data throughput at high SNR regions as long as the BER is below a desired threshold. When the channel quality is detected to be poorer, data transmission should be switched to exploit spatial diversity so that the reliability can be increased at low SNR regions: operative SNR thresholds need to be determined to switch among AMC/MIMO schemes.

However, when high throughput is requested and the channel quality is sufficient to support it with a specified BER, data transmission can switch on SMSchemes with high order modulations increasing the overall data rate.

Throughput graphs have been also produced. In Figure 2.13, the throughput for an upper-bounded BER of 0.01 is shown in the case of ideal channel estimation. As expected, it can be seen that LSTBC has always a lower throughput respect to the equal efficiency (doubled modulation order) of the Alamouti scheme due to the imperfect orthogonality of the code. Be-

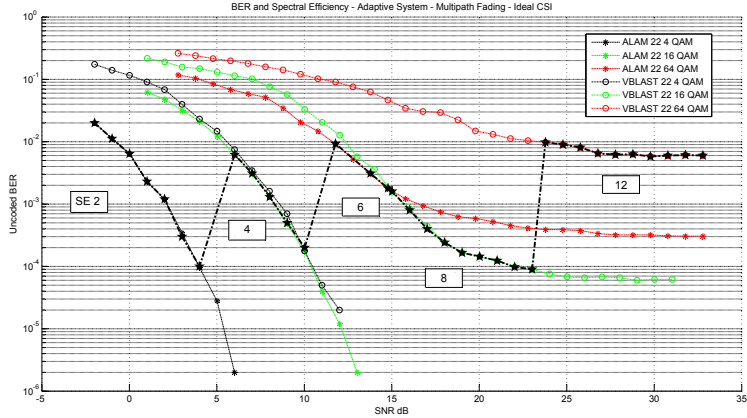


Figure 2.12: Adaptive MIMO Link Adaptation for upper-bounded BER, ITU Veh channel A.

tween 16-QAM Alamouti and 4-QAM LSTBC, there is around a 2 dB loss and more than 1 dB between 64-QAM Alamouti and 16-QAM LSTBC. The VBLAST schemes start having an acceptable performance at higher SNRs reaching the target BER for 64-QAM only at around 23 dB.

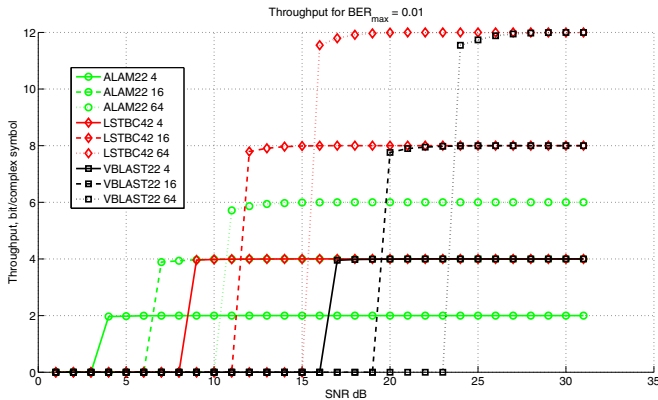


Figure 2.13: Throughput for ideal channel estimation for a $BER_{max} = 0.01$.

In Figure 2.14, the throughput of Alamouti and VBLAST schemes is compared in the case of a channel estimators operating as detailed in Sec. 2.5.3. The reference channel is ITU multipath channel (specified in Tab. 2.3 with a MT velocity set at 5 Km/h. Coloured (red and green) curves report Alamouti scheme performance, while black curves are re-

lated to the SM scheme. It can be noted that the channel estimation needs to be accurate in order to have good throughput performance. The simple interpolator in the frequency domain seems to be too far not enough robust to support the MIMO schemes analysed; the 4-QAM Alamouti scheme has around 4 dB loss. At 16-QAM the dB loss is around 7 dB.

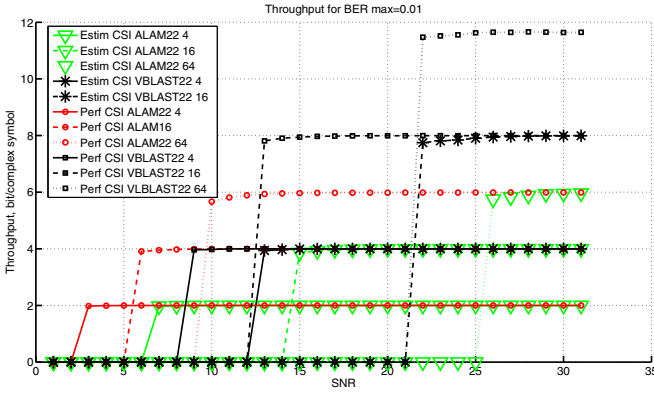


Figure 2.14: Influence of channel estimation over throughput for Alamouti and Spatial Multiplexing schemes.

Based on their characteristics, the MIMO schemes adopted in IEEE 802.16e standard are selected in order to maximise the system capacity while having a BER under a fixed threshold. If no constraints are posed on the minimum multiplexing gain, the system maximises the throughput. The system is proven to obtain a constrained BER over a wide SNR range. The sensitivity of MIMO schemes regarding channel estimation has also been reported for a linear estimator in the frequency domain.

2.6 Dynamic Space-Time-Frequency Block Codes

OFDMA and MIMO are the key technologies for the Radio Access Technology (RAT) in 4G networks. They are so suitable to be applied together that the keyword MIMO-OFDM has been around in these last years as if their respective advantages were naturally born to inter-work together. A panorama over the challenges and technical solutions proposed in the field of MIMO-OFDM can be found in [SBM⁺04].

In OFDMA, the division of the wide-band channel in several subcarriers allows to allocate data in time and/or frequency domains with the opportunity for optimisation of STBCs. The three dimensions of time, frequency and antennas can be exploited together for allocation of Space Time Frequency Block Code (STFBC)s.

In the next sections, the STFBCs proposed in the literature are briefly reviewed and then the proposed approach is described: a dynamic allocation of STBC and SFBC is studied. The method is based on a single bit feedback and allows to counterfight the performance degradation due to the MT mobility and channel frequency selectivity. The dynamic allocation criterion is assessed for WSSUS channels.

2.6.1 Space-Time-Frequency Block Codes (STFBC)

In order to combine the advantages of the MIMO systems and OFDM, STFBC coded MIMO-OFDM systems have been designed in which coding is done across space, time and frequency to achieve maximum possible diversity gain. It has been proved that MIMO transmissions over a frequency-selective fading channels can provide a diversity gain which is, at maximum, the product of the number of transmit antennas, number of receive antennas and the channel length. It has been shown, in [BBP03], that space-frequency-only coded OFDM systems cannot achieve full diversity: coding over the three Space-Time-Frequency (STF) dimensions is needed. In [JRSKR04] a full-diversity STFBC is proposed using the concept proposed in [KR02] of Co-Ordinated Interleaved Orthogonal Designs (COID). Three dimension coding has very complex implementation and simpler STF solutions have been proposed in literature.

In [LXG02], constellation precoding and Space-Time (ST) are used. The OFDM subcarriers frequency decorrelation can be exploited in order to preserve (and enhance) the diversity when the channel spatial correlation is high (thus the channel matrix rank is low). This can be accomplished using subcarriers grouping as in [MWW02, LXG02] or multiple channel fading blocks as proposed in [ZXC07]. In [MWW02] the diversity is improved with subcarrier grouping: subcarriers with independent fading are chosen achieving the highest frequency separation via a specific subcarriers permutation. A permutation is needed since the subcarrier frequency separation ΔF is designed in order to have high correlation among adjacent subcarriers, for channel estimation purposes, [vNP00].

Designs in [LXG02] and in [KR02] are no rate-one orthogonal design for more than two transmit antennas. A rate-one STFBC for four transmit antenna MIMO-OFDM systems with double symbol decoding is presented in [GR05].

The solution proposed here, [MPDP09], relies on a single bit feedback for choosing between STBC and SFBC. The dynamic allocation in OFDM can be exploited to optimise the reception of MIMO streams for mobile MTs. The effect of the frequency selectivity and time variation of the wireless channel over STBCs in an OFDM system for a mobile MT is studied.

In order to mitigate the influence of the channel variability, a dynamic allocation over the time-frequency OFDM grid is proposed. The criterion for dynamic allocation direction is derived for a WSSUS channel.

It is shown that medium to high velocity terminals (50-300 Km/h) have comparable correlation coefficients along time and frequency directions and results show that mobile users can greatly improve BER performance with dynamically allocated STFBC symbols. The SEP gain is analytically derived and computer simulations are provided for ITU Vehicular channel for a DL connection in a IEEE802.16e system.

When the MTs are mobile, the channel has high variability. The movement of the MT causes variations in the frequency domain due to varying multipath delay spread and in the time domain due to the Doppler effect. The time variability is worsen in the higher spectrum bands^{††}.

A dynamic allocation of the STBC data symbols along the time or frequency direction with a single-bit feedback is proposed aiming to maximise the correlation between successive STBC symbols. It is shown that STBCs can maintain the diversity if dynamically allocated. The effect of the channel multipath fading and Doppler spectra over STBCs in an OFDM system is evaluated. The dynamic allocation of STBC proposed has low impact on the legacy OFDM systems already developed and can enhance upper-layer performance leveraging the throughput at the interface between PHY and MAC layer, [MMA08].

Based on WSSUS channel model [Bel63], the correlation coefficients are derived among subcarriers – $\rho_f(\Delta f)$ – and between different signalling times – $\rho_t(\Delta t)$. The correlation coefficients are used to analytically calculate the SEP expression for the Alamouti-coded symbols for any order QAM modulations. The STBC symbols are allocated based on whether correlation coefficient is the highest between ρ_t and ρ_f , tracking the variation in channel Power Delay Profile (PDP) and Doppler variations. Antennas are considered uncorrelated both at the MT and at the BS.

In Section 2.6.2 the OFDM system model is explained. In Section 2.6.3, the channel model is presented: correlation coefficients along time and frequency domains are derived based on PDP and terminal velocity. After, in Section 2.6.4, the method for calculating closed-form SEP for correlated channel is reported for the case of Alamouti scheme using the MGF of the channel fading statistic. The averageSEP calculation based on MGF is reported in details in App. A. The proposed dynamic allocation with feedback is presented in Section 2.6.5 and then results are shown in terms of BER gains.

2.6.2 System Model

OFDM is an excellent transmission technique for broadband wireless links since it divides the frequency-selective (time-dispersive) channel into a number of narrow subchannels that can be considered as flat-fading via an easy implementation with DFT. In OFDMA systems, the data to be transmitted is organised in frames which can be represented as a grid or a matrix with $N_{data} \times N_{ofdm}$ entries, [vNP00], where N_{data} is the number of data subcarrier per OFDM symbol and N_{ofdm} is the number of OFDM symbols in one frame.

The number of data subchannels, N_{data} in which the spectrum is divided depends on the number of the IDFT/DFT points N_{fft} . The number

^{††}In order to have large bandwidths, next generation systems will need to exploit higher bands thus worsening the Doppler effect which is directly proportional to the carrier frequency.

of IDFT/DFT points are usually bounded to the powers of 2 for easy implementation with the radix-2 and radix-4 algorithms, [DR97].

The total IDFT/DFT points are equal to sum of data subcarriers plus guards and pilots: $N_{fft} = N_{data} + N_{guards} + N_{pilots} + 1$ where 1 is the central DC tone.

If we consider a system with N_T transmitting, N_T frames are composed, one for each antenna. The available subcarriers for data transmission are less than the total number of DFT points because part of the tones are used for guard intervals and then set to zero and some of them are assigned to pilots or training sequence transmission. The collection of all the subcarriers along frequency is called an OFDM symbol. T_{ofdm} , time duration of the OFDM symbol, is related to the inverse of the frequency subcarrier separation, ΔF . T_{ofdm} and ΔF are defined in the first phase of the system design [vNP00]. The OFDM symbol duration T_{ofdm} and the subcarrier spacing in frequency ΔF are fundamental for calculating the STBC BER performance since they represent the minimum distance, thus maximum channel correlation, that the symbols can experience once allocated on the OFDMA grid, [vNP00].

Fig. 2.15 shows a diagram of a MIMO-OFDM system with a generic STBC encoder. The symbols from the source are first modulated and then enter the STBC encoder. The number of the outputs of the encoder is equal to N_T , the number of transmitting antennas. For each of these outputs, a mapping function places the coded symbols on the frame based on a single-bit feedback from the MT.

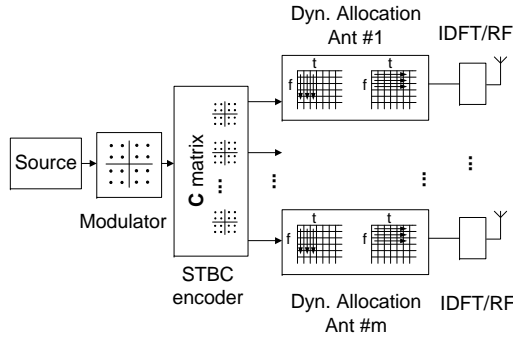


Figure 2.15: MIMO-OFDM system model with STBC encoder.

2.6.3 The Time-Frequency Correlation in the Frequency Domain

A model of the correlation coefficients ρ_t and ρ_f between subcarriers both along time and frequency direction respectively is needed in order to asset a generalised criterion to decide in which direction the STBC symbols should be allocated. In Sec. 2.6.4, the SEP of the Alamouti scheme is computed for different correlation coefficients which are obtained based on the derivation of this section.

Almost in all the practical cases, the radio channel behaviour can be effectively described as a Linear Time Variant (LTV) system. Stemming

from the nomenclature used in [Bel63], where a set of dual frequency-time functions were defined for LTV systems, the performance of diversity schemes based on the *spaced-frequency, spaced-time correlation function*: $\phi_H(\Delta t, \Delta f)$ is studied.

In a multipath channel the *time-variant channel impulse response* $h(t, \tau)$ can be expressed as a sum of various paths:

$$h(\tau, t) = \sum_i \beta_i(t) e^{-j\theta_i(t)} \delta(\tau - \tau_i(t)) \quad (2.25)$$

where the index i represent the propagation paths, $\beta_i(t)$ are the time varying path gains and $\theta_i(t)$ are the path phases. Each path is delayed, respect to a first reference path, of $\tau_i(t)$.

From the *time-variant channel impulse response*, $h(t, \tau)$, the *time-variant channel transfer function* in the Frequency Domain (FD), $H(f, t)$, can be expressed via a the Fourier transform as:

$$H(f, t) = \mathcal{F}\{h(\tau, t)\} = \int_{-\infty}^{+\infty} h(\tau, t) e^{-j2\pi f\tau} d\tau = \sum_i \beta_i(t) e^{-j(2\pi f\tau + \theta_i(t))} \quad (2.26)$$

The baseband equivalent of the channel matrix $\mathbf{H}(i, j)$ ($h_{ij} \in \mathbb{C}$) is obtained sampling the value of $H(f, t)$ at $f = f_0 + \Delta F \cdot i$ where f_0 is the lowest frequency subcarrier and at $t = T_0 + T_{ofdm} \cdot j$. In Fig. 2.16 the channel magnitude ($\|\mathbf{H}_{i,j}\| - [dB]$) of one frame in the frequency domain is reported. The channel model is the ITU Vehicular Channel Type A with a terminal velocity of 100 Km/h. The plot shows the channel frequency profile and its variation over time due to MT mobility. When MIMO is used, a two-dimensional channel transfer function is measured for each couple of transmitting-receiving antennas.

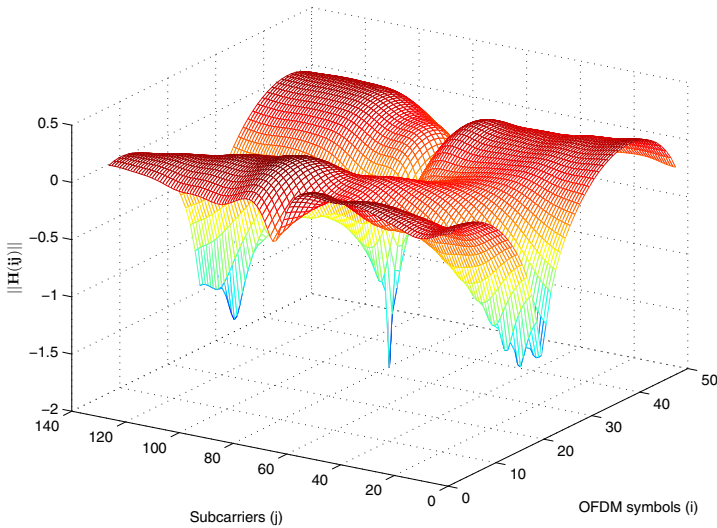


Figure 2.16: OFDM Transfer Function Vehicular Channel, 128 subcarriers and 48 OFDM symbols, 150 km/h.

The general *channel correlation function* in the FD is:

$$\phi_H(f_1, f_2, t_1, t_2) = E\{H^\dagger(f_1, t_1)H(f_2, t_2)\} \quad (2.27)$$

Under the assumption that the channel is WSSUS, the correlation function is depending only on time and frequency differences, thus (2.27) can be written as:

$$\phi_H(\Delta t, \Delta f) = E\{H^\dagger(f, t)H(f + \Delta f, t + \Delta t)\} \quad (2.28)$$

which is the *spaced-frequency, spaced-time correlation function* in the frequency domain.

The real value of $\phi_H(\Delta t, \Delta f)$, sampled at multiplies of ΔF and T_{ofdm} , represents the correlation between various \mathbf{h}_{ij} : $\phi_H(i, j)$. The correlation along time or frequency direction on the OFDMA grid can be calculated as:

$$\begin{aligned} \rho_f(i) &= \phi_H(0, \Delta f)|_{\Delta f=f_0+i\cdot\Delta F} \\ \rho_t(j) &= \phi_H(\Delta t, 0)|_{\Delta t=T_0+j\cdot T_{ofdm}} \end{aligned} \quad (2.29)$$

The minimum time and frequency separations are lower bounded by the OFDM system design parameters, respectively T_{ofdm} and ΔF .

The functions $\phi_H(\Delta f)$ and $\phi_H(\Delta t)$ can be obtained by evaluating the spaced-time spaced-frequency correlation function as in (2.29) by nulling the domain which is not of interest or also by evaluating the Fourier transforms of the *PDP* $\phi_h(\tau)$ and of the *Doppler Spectrum* $S_H(\nu)$ (where ν is the Doppler shift variable) respectively. The proposed dynamic allocation can minimise the BER for the STBCs through the estimation of the PDP and of the Doppler Spectrum. In OFDM systems, the Root Mean Square (RMS) delay spread τ_{rms} and the Doppler frequency can be easily estimated and are usually calculated at the receiver. Various methods have been proposed in literature: [RB01, WLLC08, WKP01, YA08]. When only the τ_{rms} and f_d are available, an assumption about the shape of the PDP and of the Doppler Spectrum is needed.

If an exponential decaying PDP model is assumed, the related spaced-frequency correlation function can be expressed in function of the τ_{rms} as:

$$\phi_H(\Delta f) = \frac{1}{1 + j2\pi\tau_{rms}\Delta f} \quad (2.30)$$

The frequency correlation (ρ_f) is illustrated in Fig. 2.17. The PDP is assumed as in (2.30) and various correlation are reported for different channel delay spreads τ_{rms} and subcarriers spacing ΔF .

The correlation among successive OFDM symbols on the same subcarrier index is related to the MT velocity and the Doppler Spectrum shape. The Doppler spectrum shape depends on the geometrical assumption made on the directions of the multipath signals impinging on the receiving antenna. The most common model assumed is the Jakes model [DBC93] which is between the two extreme cases of two rays and uniform distribution of the arriving rays at the receiving antennas.

In Tab. 2.5, the three most common cases are reported with the respective space-time correlation functions.

The time correlation is shown for Jakes spectrum in Fig. 2.18. Figs. 2.19 and 2.20 plot the theoretic correlation among subcarriers for the other two Doppler models. The Doppler frequencies shown are corresponding to MT velocities of 60 – 125 – 250 Km/h at a carrier frequency of 6GHz. In the case of Jakes model, when $T_{ofdm} = 100\mu s$, the correlation coefficient ranges from 1 to 0.8. The Jakes model represent an of average situation for the

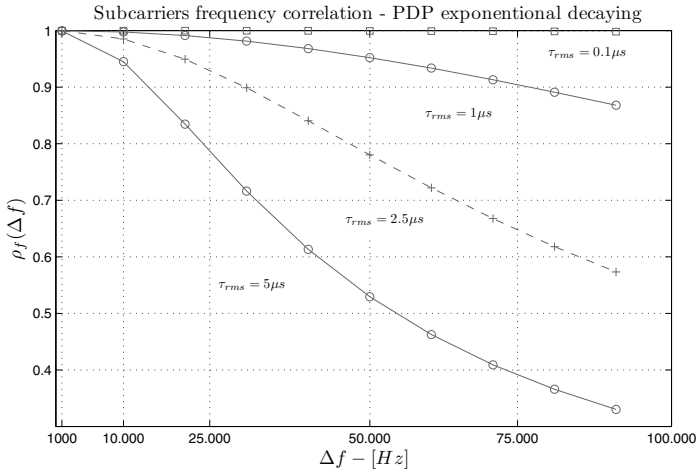


Figure 2.17: Exponential Decaying PDP - subcarriers correlation over frequency separation for various τ_{rms} delays.

Doppler Spectra	$S_H(\nu)$	$\phi_H(\Delta t)$
Jakes Model	$\begin{cases} \frac{1}{\pi f_d} \frac{1}{\sqrt{1 - (\frac{\nu}{f_d})^2}}, & \text{if } \nu < f_d \\ 0, & \text{otherwise} \end{cases}$	$J_0(2\pi f_d \tau)$
Two Rays	$\begin{cases} \frac{1}{2f_d}, & \text{if } \nu < f_d \\ 0, & \text{otherwise} \end{cases}$	$\text{sinc}(f_d \tau)$
Uniform	$\frac{1}{2} [\delta(f + f_d) + \delta(f - f_d)]$	$\cos(2\pi f_d \tau)$

Table 2.5: PDP models and $\phi_h(\Delta f)$ correlation function

impinging signals on the antenna that can be encountered in typical urban scenario. However, when the MT is in a worse scenario and the channel has principal multipath rays, meaning that energy is received at the MT antenna concentrated from certain directions, the correlation can be much lower than in the Jakes case. In the case of two rays model it ranges from 1 to almost 0.6, as visible in Fig. 2.19. The uniform model is providing the highest correlation values but represents the ideal situation of a very high scattering channel where the antenna at the MT sees signals coming from all the direction in the space. Further, high scattering channels are usually encountered indoor which is not the typical environments for medium and high velocity MTs.

Comparing the values of ρ_t and ρ_f from Figs. 2.17- 2.18, we note that the correlation coefficient can be lower in the time or frequency domain, depending on the relative ratio of terminal velocity and delay spread. Thus, the BER performance of STBCs can be enhanced if a dynamic allocation direction is implemented as shown in Sec. 2.6.5.

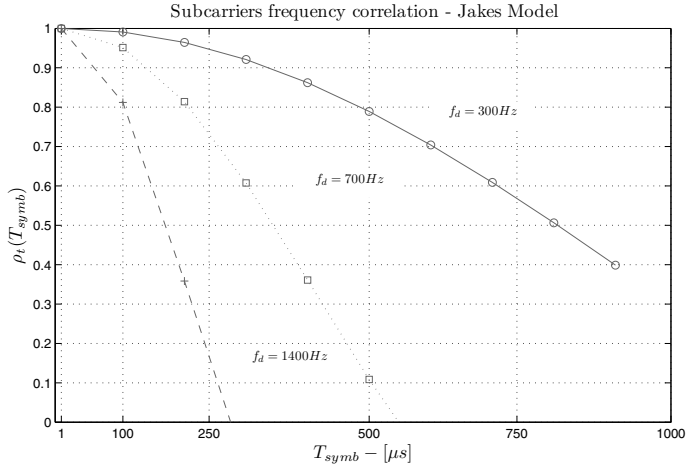


Figure 2.18: Jakes model - subcarriers correlation over time for various Doppler frequencies.

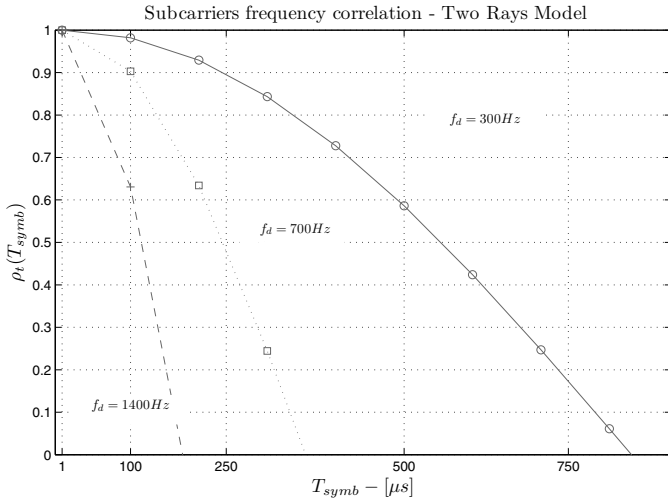


Figure 2.19: Two Rays model - subcarriers correlation over time for various Doppler frequencies.

2.6.4 Diversity Schemes in Uncorrelated Channels

A space-time block code is defined by a coding matrix which maps constellation symbols over the N_T transmitting antennas in successive T time slots. The coding matrix \mathbf{A} has to be orthogonal since a non-orthogonal \mathbf{A} will result in ISI interference at the receiver, as explained in Sec. 2.2.1. Matrix $\mathbf{A} \in \mathbb{R}^{L \times N_T}$ maps $L \times N_T$ data symbols over successive L time slots over the transmitting antennas. The decoding process for the ST codes is simple and it is based on linear combination of the transmitted

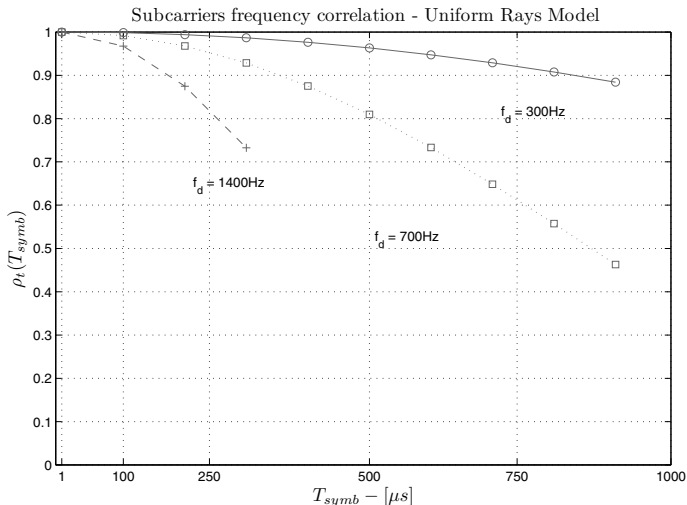


Figure 2.20: Uniform model - subcarriers correlation over time for various Doppler frequencies.

signals from each of the N_T transmitting antennas.

The first and simplest STBC was proposed in [Ala98], where $N_T = 2$ both for the case of $N_R = 1$ and $N_R = 2$. The coding matrix \mathbf{A} can be written as, see Sec. 2.2.1:

$$\mathbf{A} = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix} \quad (2.31)$$

STBCs assume that the channel remains constant during the time duration of the code (L signalling time slots). Alamouti code requires that the channel is constant over 2 signalling period. The constant channel assumption is easily not verified particularly in quick time variant channels.

When channel coefficients are not correlated ($\rho < 1$), the BER performance of the STBCs is influenced by the correlation coefficients between two consecutive symbols: ρ_f if symbols are allocated along frequency, ρ_t otherwise.

The SEP performance for the Alamouti transmit diversity scheme with a time-varying channel has been derived for narrow-band partially correlated Rayleigh channels for BPSK, in [VLB04]. We can express the channel coefficients accounting for the correlation as:

$$h_1^1 = \rho h_1^2 + \sqrt{1 - \rho^2} \epsilon_1 \quad (2.32)$$

$$h_2^2 = \rho h_2^1 + \sqrt{1 - \rho^2} \epsilon_2 \quad (2.33)$$

where h_j^i is the complex channel coefficient, j is the transmitting antenna and i is the time slot of the Alamouti block, $i|j \in \{1, 2\}$. The random variables ϵ_i are distributed with Rayleigh distribution, as the channel coefficients. If the correlation $\rho = 1$, the transmit diversity scheme is not suffering any performance losses. The other extreme case is represented by $\rho = 0$, when no correlation exists between two successive channel coefficients.

When ρ is comprised between these two extremes, the SEP for a Zero Forcing (ZF) detector can be expressed based on the SNR distribution at the output of the STBC combiner. In [VLB04], the closed-form SEP for BPSK has been derived; in [MPDP09], the closed-form SEP is derived for any order QAM modulation, more details about the calculations can be found in App. A.

If the channel coefficient h_j^i are assumed to be Rayleigh distributed, the SNR distribution at the output of the receiver is:

$$p_\gamma(\gamma) = \frac{2(1-\rho^2)}{\bar{\gamma}} e^{-2\gamma/\bar{\gamma}} + 4\rho^2 \frac{\gamma}{\bar{\gamma}} e^{-2\gamma/\bar{\gamma}} \quad (2.34)$$

To obtain the averaged SEP over the fading channel, the SEP conditioned on the instantaneous SNR needs to be integrated over the SNR distribution.

$$SEP_M(\bar{\gamma}) = \int_0^\infty SEP(M, \gamma) p_\gamma(\gamma) d\gamma \quad (2.35)$$

where $\bar{\gamma}$ is the average SNR. Following the approach in [SA05], the SEP for coherent detection of Alamouti symbol with QAM modulation of any order is derived as:

$$SEP_M(\bar{\gamma}) = 4KI^{(1)}(a(M), \gamma) - 4K^2I^{(2)}(a(M), \gamma) \quad (2.36)$$

where $K = (\sqrt{M} - 1)/\sqrt{M}$, M is the modulation order and

$$a(M) = \sqrt{(3\log_2(M))/(M-1)}$$

$$I^{(1)}(a(M), \gamma) = 1/\pi \int_0^{\pi/2} M_\gamma(-s(\theta)) d\theta$$

$$I^{(2)}(a(M), \gamma) = 1/\pi \int_0^{\pi/4} M_\gamma(-s(\theta)) d\theta$$

In the integrand functions: $s = \frac{a^2}{2s \sin^2(\theta)}$ and $M_\gamma(s)$ is the Moment Generating Function (MGF) function of the SNR distribution at the output of the ZF receiver:

$$M_\gamma(s) = \frac{(1-\rho^2)}{(s\bar{\gamma}/2+1)} + \frac{\rho^2}{(s\bar{\gamma}/2+1)^2} \quad (2.37)$$

In Fig. 2.21, analytical SEP curves are reported for 4-16-64 QAM modulations with correlation coefficients $\rho \in [1, 0.9, 08]$. For SNR levels above 4dB, the Alamouti scheme is very sensitive to channel variability and the performance is highly degraded. For a 4-QAM modulation at $SNR = 20dB$, the SEP shifts from around 10^{-4} to more than $3 \cdot 10^{-3}$.

2.6.5 Performance Analysis

In the previous Sec. 2.6.3 we have analysed the correlation coefficient along time and frequency allocation direction based on the *spaced-time, spaced-frequency correlation function* in the frequency domain for a WSSUS channel. Then, based on general correlation ρ , SEP has been expressed in closed-form for the Alamouti scheme in Sec. 2.6.4.

In this section we propose the adaptive allocation on the OFDM grid based on the estimated delay spread τ_{rms} and Doppler frequency f_d . In Tab. 2.6 we report the OFDM system settings, referring to the standard IEEE802.16e.

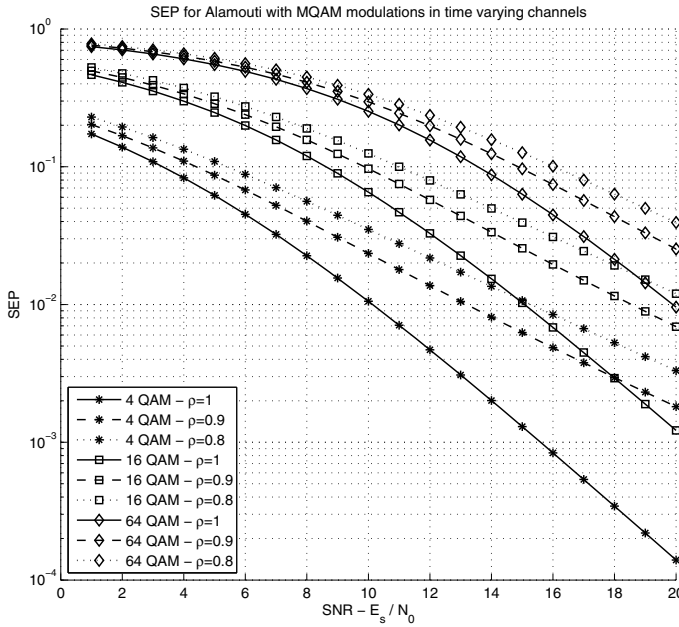


Figure 2.21: Symbol Error Probability for Alamouti scheme for 4-16-64 QAM modulations over narrow-band partially correlated Rayleigh channel.

Param.	Description	Value in 802.16e
c_f	Carrier frequency	2.5 GHz
BW	Total bandwidth	10 MHz
N_{fft}	Number of FFT points	1024
ΔF	Subcarrier spacing	10.9375 kHz
T_{ofdm}	OFDM symbol duration without CP	91.43 μ s
CP	Cyclic prefix length	$1/8 T_{ofdm}$
T_{ofdm}^{CP}	symbol duration with CP	102.86 μ s
F	Frame length	5 ms
N	Number of OFDM symbols in frame	47

Table 2.6: STFBC: system parameters.

The proposed system is illustrated in Fig. 2.22. The system calculates ρ_t and ρ_f . Then it decides the allocation direction where the correlation is higher. These two coefficients can be exactly calculated if the PDP and the Doppler Spectrum of the channel are known. However, assumptions can be made on the Doppler Spectrum and PDP shapes. In that case, the system needs only to estimate the channel delay spread and the Doppler frequency.

The computer simulations show the Symbol Error Rate (SER) performance of the OFDM system with the Alamouti scheme in a DL connection

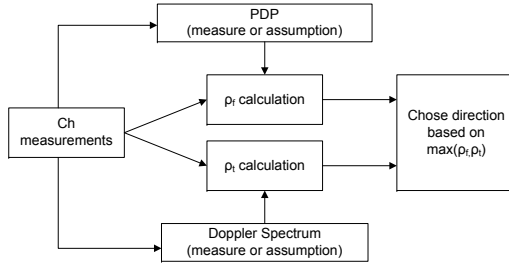


Figure 2.22: The dynamic allocation direction of STBC in an OFDM system.

to a high velocity MT moving at 150 Km/h. The results are averaged over a simulated real-time of 10 seconds. The dynamic allocation operates with the assumption of Jakes Doppler spectrum and Exponential decaying PDP. The ITU vehicular channel model A has a $\tau_{rms} = 3.32\mu s$. In the case of the exponential decaying PDP, we can calculate the expected correlation between adjacent subcarriers along the frequency direction obtaining $\rho_f = \phi_H(\Delta F) = 0.98$, $\Delta F = 10.9375kHz$.

In Fig. 2.23, we can see that the dynamic allocation direction preserves the diversity gain of the Alamouti scheme. The time allocated STBC is losing performance due to low correlation between symbols allocated on the same subcarrier but on different time slots. The dynamic allocation is able to track the best correlation and dynamically change the allocation direction of STBC. This reduces the BER and preserves the system throughput.

Since the MT velocity is high (150 Km/h), the best correlation is over the frequency domain ($\rho_t < \rho_f$). The SER of the dynamic allocation tends to the theoretical one calculated in closed-form with the frequency correlation coefficient $\rho_f = \phi_H(\Delta F)$. The difference between the theoretically calculated Alamouti BER performance over correlated channels and the simulated one derives from the fact that the closed-form calculation is valid for narrow-band transmissions. The simulation results reported are the average of STBC transmitted symbols over all the subcarriers of the frame and include the effect of InterCarrier Interference (ICI) coming from Doppler shift.

In the case of receivers with very low Doppler frequencies^{‡‡}, the symbols are likely to be mainly allocated along the time direction, especially in scenarios with high delay spreads, providing better SER performance respect to a fixed allocation along the frequency direction.

The multipath channel has been simulated with Rayleigh statistic distribution. Other fading statistic model can be inserted in the analysis. The presence of various degrees of Line of Sight (LoS) can be studied using Nakagami- m fading statistics. Correlated space-time coding over correlated Nakagami- m channel fading has been studied in [MA06, SL04]. The Nakagami index m has an important role being directly related to the channel correlation in the frequency domain.

Based on these derivations, multivariate Gamma distribution can be used to described the SNR at the receiver thus obtaining SEP for the cor-

^{‡‡}This can be the case of static MTs which generally show an f_d of the order of some units of Hertz, due to moving object in the radio environment.

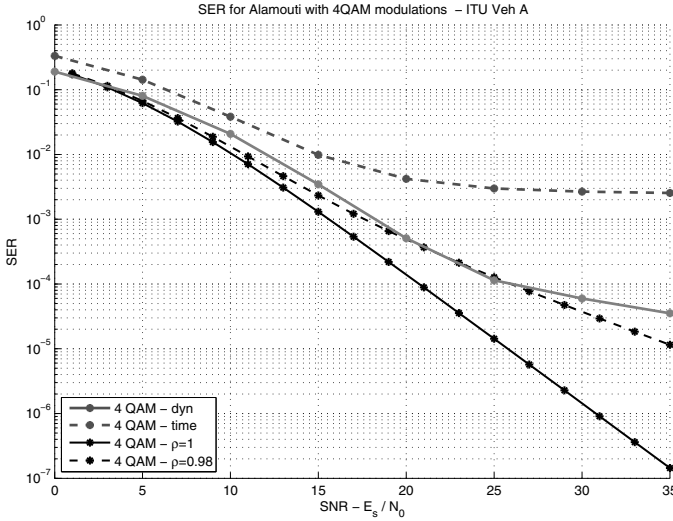


Figure 2.23: BER improvement for ITU Vehicular Channel model A with MT at 150 Km/h

related STCs over Nakagami- m channel. When the fading severity factor of the channel is decreasing (when $m > 1$ in the Nakagami- m channel statistic), the channel will show a reduced frequency selectivity, raising the correlation between adjacent subcarriers: when LoS component is sufficiently high, SFBC would perform better also with medium to high speed MTs.

2.7 Conclusions

In this chapter, MIMO has been studied aiming at PHY layer enhancements. The performance gains are coming from the added degrees of freedom of the radio signals which can be exploited with suitable receiving techniques. However, this approach can be fragile and the performance are very sensitive to several physical characteristics of the channel and of the devices, among them:

- Antenna Correlation at transmitter and receiver sides
- MIMO Channel Matrix rank
- Channel variability in STBC

While the antenna correlation has not been addressed in this chapter, MIMO channel matrix rank and channel variability have been studied and two solution proposed.

To react to low channel ranks, an adaptive MIMO link adaptation strategy has been presented. The system spectral efficiency is maximised given an upper-bound on the BER. Hybrid LSTBC scheme have been analysed jointly with SM for a system constrained with a multiplexing gain of 2. If the diversity constrain is not required, Alamouti scheme has been used

instead of LSTBC in order not to suffer from BER degradation due to the quasi-orthogonal LSTBC structure.

A solution with a light feedback channel has been proposed for fast channel variation in the presence of high velocity MTs via a dynamic STFBC allocation over the OFDMA frame. The system proposed is able to adapt the allocation direction on the OFDMA grid in order to best match to the channel characteristics of the MT. Current standards (as IEEE802.16e) are using fixed allocation direction and they may lose the diversity enhancement of the STBC. The proposed dynamic allocation requires low modifications in the air-interface design and preserves the STBC performance which otherwise would be lost due to the MT mobility. The feedback needs only a single-bit feedback.

The strategy has been simulated with the Alamouti scheme, without a loss of generality. The results can be extended to a general L length STBC where the improvements will show even greater enhancements, considering that the assumption of quasi-static channel is more difficult to be verified when the STBC length increases.

Chapter 3

MAC Layer ARQ and DSA analysis

*We know the truth, not only by
the reason, but by the heart.*

BLAISE PASCAL
(1623-1662)

3.1 Introduction

This chapter is focused on some of the mechanisms at the MAC layer of the IEEE802.16 stack. The 802.16 networks have three sublayers: the security sublayer, the Link Layer Control (LLC) sublayer and the MAC sublayer. The MAC security sublayer will not be discussed in this thesis. However a work by the author was published in [MWP09b], where subcarrier permutation at PHY layer is used to encapsulate higher layers security mechanisms providing enhanced network security.

At the LLC, the communications are organised in flows and are managed with a connection-oriented approach, each service flow has a connection-ID. The reference standard for LLC is the IEEE802.2, [IEE89].

The MAC sublayer manages the ARQ retransmissions and the multiple access to the radio resource. In 802.16 the multiple access to the channel is supported with dynamic resource allocation. Depending on the PHY layer configuration – SC,SCa,OFDM,OFDMA – the dynamic resource allocation follows different approaches. In OFDMA, the DL and UL frames are divided in chunks: a set of subcarriers that can be assigned to a service flow connection. The resources are allocated both in the time and frequency domain via a two-dimensional partition of the OFDMA grid, which is visible in Fig. 2.16 at page 40.

The resource assignment mechanism in OFDMA is based on statistical multiplexing which is a multiplexing technique where the communication channel (in this case the RF bandwidth of the cell) is partitioned into a number of variable capacity digital channels. The resources are shared and

the allocation is adapted to the instantaneous traffic demands of the service flows that are admitted in the system. Statistical multiplexing is opposed to the fixed multiplexing where the sharing of the resource is static, as in TDM.

Statistical multiplexing provides a system performance gain due to the opportunistic scheduling of the admitted service flow. This gain is called *statistical multiplexing gain* and permits to achieve higher link utilisation factor: the scarce resources of the wireless networks (namely time and spectrum) are thus better utilised. Statistical multiplexing accounts for QoS requirements of each service flow when calculating the priorities for the access to the radio resources, for instance calculating the priorities based on rate and delay measurements at the scheduler queues as in [MMAP08].

Statistical multiplexing with QoS awareness and aggressive resource exploitation is facilitated by a packet-based design. The packet segmentation of the traffic flows permits better flexibility and network performance.

The more complicated prediction of the network load and the more complex admission control mechanisms are the main drawbacks of packet-based statistical multiplexing compared to the circuit-switched designs. In circuit switched networks, once the line has been leased for a communication, QoS is surely met and the load is readily calculable.

In this chapter the followings will be analysed:

- MAC sublayer: PHY/MAC cross-layer optimisation for a Stop and Wait ARQ protocol
- LLC sublayer: Dynamic Service Addition blocking analysis

The work on ARQ and DSA protocols has been carried out in collaboration with Isabella Cerutti and Scoula Superiore Sant’Anna Pisa, Italy.

3.2 ARQ: a Cross-Layer Strategy for QoS-guaranteed Links

Future wireless networks are going to offer high transmission rates to stationary and mobile users supporting the convergence of data, voice, and video transmissions. Such convergence imposes strict requirements on the QoS that must be ensured to users’ transmission. The problem of ensuring QoS to 4G network users is especially hard due to the high data rate involved and the high mobility of the MTs. QoS guarantees are extremely difficult to maintain in wireless networks, especially due to the intermittent quality degradation that radio channel may experience. Degradations and fluctuations of the channel quality can be caused by various impairments and are further exacerbated by high user mobility and high data rates. In order to achieve the requested QoS, strategies that counteract the channel quality variations in an *adaptive* way are, thus, necessary.

Aggressive QoS strategies can be greatly enhanced via cross-layering design. The classical layered communication stack has been shown to be very useful for a coordinated technological development of data networks. However, dividing the problem of communication in strictly separated layers has reduced the overall networks efficiency. It is true that some operations of the data networks are fundamentally separated in their nature and cannot be solved with a single approach. In this sense, physical communications (radio, optic, cable, etc.) are a widely separated research fields from those studying networking capabilities with data flowing from multiple sources to multiple destinations. However, strictly layered stack, even if represents a good base for standardised and well-organised development, greatly reduces the performance and the flexibility of a wireless network.

For a good overview of the fundamental capabilities of a data network and the limitations in the current layered stack the reader can refer to [Pop06].

An Adaptive Cross Layer (ACL) strategy, that jointly optimises the parameters of PHY and MAC layers using queueing theory, is proposed and applied to a mobile user communication in an IEEE802.16 wireless network, supporting different modulations and the ARQ protocol. PHY and MAC layer parameters are here optimally combined by ACL strategy to meet the QoS requirements.

The evaluation of the impact of mobility on ACL strategy performance is assessed by using an accurate channel model that accounts for fast fading and shadowing effects. The simulation results show that the ACL strategy outperforms non-adaptive or single-layer strategies, in terms of bandwidth savings. The ACL strategy is able to guarantee the requested QoS in the different mobility scenarios.

3.2.1 State of the Art on Cross-Layer ARQ

Adaptive strategies for guaranteeing QoS may exploit the various degrees of flexibility and redundancy that are offered by the different layers of a 4G wireless network. It is well-known that the choice of the layer at which the adaptive strategy is implemented affects the overall performance of the wireless network. In particular, strategies implemented at the upper layers (e.g., IP layer) could be slowly counteract the channel variations. In addition, if different disjoint strategies are implemented at the different layers, lower layers (e.g., PHY layer) may tend to promptly under or over-compensate the channel quality degradation and, thus, the upper layer may strive to achieve the QoS requirements. Therefore, the adaptive strategy must feature a cross-layer approach to optimally achieve the QoS requirements and a prompt adaptivity to compensate channel quality variations.

A number of previous works considered cross-layer strategies suitable for 4G networks (e.g., [LZG05, WLG06]). The performance of the proposed solutions has been mainly evaluated for stationary users. The effects of slowly varying Channel Quality Indicator (CQI) on ARQ protocol are evaluated in [AB00]. In such paper, channel state information is derived as long-term BER. The estimation is computed on a sliding windows and four simple adaptation algorithms are provided. These strategies can react only after a certain number of errors, thus with a delay that is not tolerable by real time traffic.

A more accurate channel model that accounts for CQI variations is considered in [LWG06, RLM02, LZG05, WLG06]. The model is based on a Finite State Markov Chain (FSMC) that alternates good and bad channel quality states, according to simulation parameters. Transition probabilities and state holding times are inferred from statistical measurements of the wireless channel and lack of a proper model for the user mobility inside a wireless network cell.

3.2.2 ACL ARQ System Model

The proposed ACL strategy jointly and promptly optimises the PHY and MAC parameters, [CMCP08]. The objective of ACL strategy is two-fold, i.e., to meet QoS requirements and to allocate the minimum amount of bandwidth necessary for transmissions. The proposed adaptive strategy is applied to a IEEE802.16 network based on an OFDM PHY layer; system parameters are set according to [IEE04, CVM⁺07, vNP00]. PHY and MAC layer of IEEE802.16 support, respectively, different modulation profiles and

a stop and wait ARQ protocol. In the work proposed herein, they are optimally combined by the ACL strategy. The optimisation is based on the analytical evaluation of the QoS performance in terms of packet loss and expected delay, derived using queueing theory.

The ACL strategy is applied to a mobile user scenario, in which the transmission channel is affected by both fast and slow fading. The slow fading inter-frame process is originated by the shadowing due to buildings or obstacles, encountered along the mobile user trajectories. The performance of the ACL strategy is evaluated assuming perfect and immediate knowledge of CQI. Simulation results assess the adaptivity and the optimality of the ACL strategy under different user mobility scenarios and quantify the impact of user mobility on QoS.

The ACL strategy is tested over the radio link between the BS and a single MT. BS and MT communicates using IEEE 802.16 protocols operating in point-to-multipoint mode.

IEEE 802.16 is a connection-oriented protocol and transmissions are time-division multiplexed. In each time frame, the BS schedules the transmissions and indicates the DL and UL bandwidth allocation in a map along with the PHY layer parameters. Data packets, i.e., MAC Protocol Data Unit (PDU), are transmitted during the scheduled time allocation.

Without loss of generality, a single connection, referred to as "service flow", is established on the radio link from BS to MT. The following assumptions are made on the derivation of the cross-layer optimisation scheme.

For further details regarding IEEE802.16 packet processing refer to Chap. 5 at page 115.

It is assumed that the transmitted power at BS is kept constant. The BS has the perfect and immediate knowledge of the CQI. The service flow is an Unsolicited Grant Service (UGS) type. The PDU packets of UGS have fixed size and are assumed to be periodically issued (e.g., UGS supporting voice traffic streams). Issued PDUs are stored in a buffer with unlimited capacity. In each time frame, the bandwidth required for PDU transmission is dynamically and optimally allocated upon ACL strategy computation.

3.2.3 Adaptive Cross Layer (ACL) Strategy

The main objective of the cross-layer optimisation is to minimise the amount of bandwidth necessary for transmitting the service flow PDUs, while ensuring the requested QoS. The bandwidth required to transmit a single PDU is defined as the bandwidth (e.g., allocated time) for the first time transmission of the PDU and for the subsequent retransmission of the same PDU, until either the PDU is successfully received or the maximum number of retransmissions is reached.

QoS requirements include maximum PDU loss rate, L_{max} , and may include a set of other QoS requirements (such as delay and jitter); the QoS set is indicated with Q .

To minimise the bandwidth while ensuring QoS requirements, the PHY and MAC layer parameters are optimally and jointly decided. The PHY layer parameters (e.g., modulation order, code type, and rate) allows to define the PHY profile to select. The MAC layer parameters include whether ARQ protocol is required, and if so the maximum number of times, K , the transmission of the same PDU should be attempted, before dropping the PDU.

Let S be the set containing the SE of each PHY profile. Let λ be the expected arrival rate of PDUs. Let P_i be the expected probability that a fixed-size PDU, transmitted using PHY profile i , is received incorrectly, for

a given CQI measured at the PHY layer. The expected P_i can be calculated measuring the channel status and using look-up table when complex MCS are used and no closed-form expressions are available.

Under the assumption of unlimited buffer capacity, PDU losses occur only when the maximum number of transmission attempts (i.e., K) is reached. Thus, let $f(L_{max}, P_i)$ be the function that optimally selects the MAC parameter K to meet L_{max} values. Also, assume to know the set of functions $g_j(K, \lambda, P_i)$ that evaluate each QoS metric j .

The cross-layer optimisation works as follows. The various PHY profiles are considered, starting from those at maximum spectral efficiency; a set of physical layer profile with the same SEs (which is the maximum possible) is found as:

$$\xi = \max_i S[i] \quad (3.1)$$

Among the PHY profiles having the same spectral efficiency, the PHY profile at minimum PDU error rate is selected:

$$\rho = \min_{i:S[i]=\xi} P_i \quad (3.2)$$

Then, the optimal number of transmission attempts, K_{opt} , for the given value of ρ and L_{max} can be derived as:

$$K_{opt} = f(L_{max}, \rho) \quad (3.3)$$

If constraint $g_j(K_{opt}, \lambda, \rho) < Q[j]$ holds for each QoS metric j , then the optimum MAC parameter K_{opt} and PHY profile $s_{opt} = s$ are found. Otherwise, the QoS for the other PHY burst options at the same spectral efficiency ξ (if any) is evaluated. If QoS is not met, the other PHY profile options at decreasing SEs are evaluated.

Such cross-layer strategy is performed each time the CQI deviates significantly. A flow-chart of the strategy is sketched in Fig. 3.1. Notice that the PHY profile at maximum spectral efficiency is selected when QoS requirements are met. Also, for the same spectral efficiency, PHY profile with the best performance in terms of PDU error rate is selected.

A derivation of functions $f(\cdot)$ and $g_j(\cdot)$ is presented next.

The above presented ACL strategy is applied to the case in which the QoS requirements are given in terms of maximum PDU loss rate, L_{max} , and expected PDU delay, \bar{D}_{max} . The estimation of PDU loss rate, L , and the expected PDU delay, \bar{D} , are analytically derived in this section using a queueing model based on the following assumptions.

Realisation of the statistical process representing the BS-MT channel behaviour during a time frame are assumed to be independent on a time frame basis, i.e., fast fading is assumed to be block fading. BS can schedule a single PDU per time frame. Fragmentation or packing of PDU is not considered. ARQ acknowledgements are promptly received. Therefore, the ARQ protocol is based on stop-and-wait mechanism, with up to K transmissions of the same PDU, as shown in Fig. 3.2.

Under the above mentioned assumptions, the performance of stop-and-wait protocol can be evaluated using a discrete time queueing model with Bernoulli arrivals at rate λ . Traffic model is Poisson distribution which is well-accepted for modelling voice calls and UGS service flows*. Let X be the expected service time; with Bernoulli arrival rates, first moment (\bar{X}) and second moment ($\overline{X^2}$) of X are as follows:

*More accurate traffic models can be used for (non)real-time traffic: they may affect Admission Control (AC) blocking, but not the signalling blocking or delay.

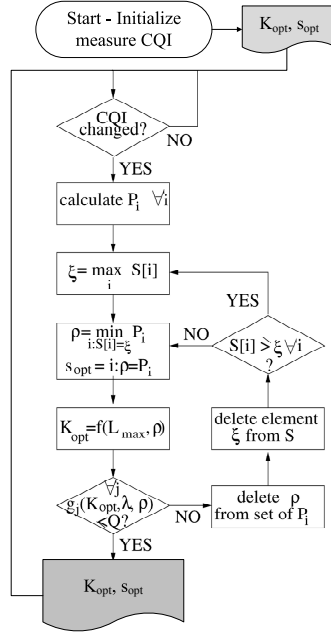


Figure 3.1: ARQ Cross-layer optimisation algorithm.

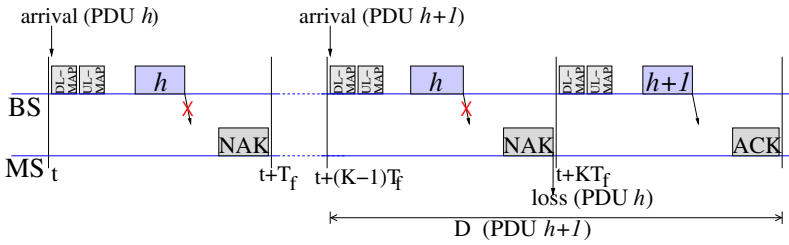


Figure 3.2: Considered Stop-and-Wait ARQ protocol.

$$\bar{X} = T_f \left(\sum_{h=1}^K h P_i^{h-1} (1 - P_i) + K \cdot P_i^K \right) \quad (3.4)$$

$$= T_f \left(\frac{(1 - P_i^{K+1})}{(1 - P_i)} - (K + 1) P_i^K \right) + T_f \cdot K \cdot P_i^K,$$

$$\begin{aligned} \bar{X}^2 &= \left(\sum_{h=1}^K h^2 P_i^{h-1} (1 - P_i) + K^2 \cdot P_i^K \right) \\ &= \left(\frac{(1 + P_i - P_i^K (K + 1))^2 + P_i^{K+1} (2K^2 + 2K - 1)}{(1 - P_i)^2} \right. \\ &\quad \left. + \frac{P_i^{K+2} \cdot (-K^2)}{(1 - P_i)^2} \right) T_f^2 + T_f^2 \cdot K \cdot P_i^K, \end{aligned} \quad (3.5)$$

where T_f is the time frame duration.

The expected delay experienced by correctly received PDUs is given by the expected time spent in the buffer ($\overline{D_w}$) and the expected time spent for retransmissions (i.e., \overline{X}). The expected value of the first term, $\overline{D_w}$, is derived using P-K formula [Kle75]

$$\overline{D_w} = \frac{\lambda(\overline{X^2} - \overline{X})}{2(1 - \lambda\overline{X})}, \quad \overline{D} = \overline{D_w} + \frac{\overline{X} - K \cdot L}{1 - L}. \quad (3.6)$$

PDU loss rate is given by $P_{i,n}^K$. In order to meet the PDU loss rate requirement, it is possible to derive:

$$f(L_{max}, P_i) = \left\lceil \frac{\log L_{max}}{\log P_i} \right\rceil, \quad (3.7)$$

and find K_{opt} as indicated in (3.3). Notice that such value of K_{opt} is the minimum value that allows to meet PDU loss rate requirement. Therefore, the bandwidth required to transmit a PDU is minimal. The QoS metric function concerning the expected delay is $g_j(K_{opt}, \lambda, P_i) = \overline{D}$, as in (3.6).

Finally, the expected amount of bandwidth necessary in each time frame for supporting the service flow is equal to $\lambda \cdot \overline{X}$ times the bandwidth required for transmitting one PDU.

3.2.4 Performance Analysis

In this section, after a detailed description of the channel model for mobile users and of the system parameters is provided. The performance of the proposed ACL strategy is evaluated in different mobility scenarios. The ACL strategy is compared against two strategies. The first, a single-layer (SL) strategy, optimises the MAC parameters only, according to the updated CQI, when PHY parameters are fixed and given. A second a cross-layer (CL) strategy, that is not aware of CQI variations over time, optimises jointly PHY and MAC parameters, according to the average average channel quality (i.e., SNR).

The channel model accounts for mobility. The MT is assumed to move inside a network cell, along a circular path centred at the BS. The channel model for the MT is based on ITU Vehicular type A model [SMF05]. MT mobility induces shadowing effects, that make the received average signal-to-noise, Γ , fluctuate according to a log-normal distribution. Path loss variations are neglected in this work as the changes are slower than shadowing, especially in urban mobility scenario. The ACL strategy which is able to react to the shadowing, can easily adjust the MCS to the slower variations of the link budget due to the long term variations of the path loss.

For comparison purposes, a frequency flat fading model (LoS transmissions without Doppler effects) is considered. In the LoS case, power of the delayed multipath replicas is very low and no frequency selectivity is appreciated.

System Parameters

The performance of the proposed cross-layer approach is evaluated for an OFDM-based PHY layer, according to the parameters indicated in Table 3.1. Parameters are defined according to IEEE 802.16 [IEE04] standard. PHY profiles consist of three different types of modulation, i.e.,

4-QAM, 16-QAM, and 64-QAM, applied to uncoded data. Carrier frequency is centered at 3.5 GHz and the wavelength is 8.6 cm. MT speed is 50 Km/h, which leads to a Doppler frequency of about 160 Hz for the specified carrier frequency. Shadowing is supposed to have a spatial coherence of 20 wavelengths. For the considered MT speed and time frame duration, shadowing causes a log-normal Γ deviation every 120 ms or 24 time frames, [GGF94, Gud91]. It is assumed that only 2 dB variations of Γ are detected (i.e., the log-normal distribution is quantized with a minimum step of 2 dB step over a given Γ range). The standard deviation of Γ depends on the mobility scenario. Shadowing parameters for the different mobility scenarios are selected according to [FWC⁺06] and reported in Table 3.2.

Upon detection of a channel quality variation, BS promptly performs the cross-layer optimisation, presented in Sec. 3.2.3, and schedules PDU transmissions according to the computed MAC and PHY parameters. Size of PDU packets, including MAC header and CRC, is fixed to 288 bytes, i.e., 1 (2 or 3) OFDM symbol(s) at 64-QAM (16 or 4 QAM).

Parameter	Value
Data subcarriers	384
FFT size	512
Subcarrier freq. separation	31.25 KHz
OFDM symbols in frame	48
Frame duration (T_f)	5 ms
Symbol duration	102.86 μ s
Useful Symbol duration	91.43 μ s
Cyclic Prefix Ratio	1/8

Table 3.1: ARQ ACL strategy, system parameters.

Numerical Results

In the simulation results, unless otherwise indicated, the ITU vehicular A channel model with log-normal deviation $\sigma = 2.3$ dB and Γ range in [26-40] dB is used. Unless otherwise stated, the QoS requirements to be met for any value of Γ are: $L_{max} = 10^{-4}$ and $D_{max} = 25$ ms. PDU arrival rate is $\lambda = 0.1$ PDUs per time frame. In the simulation, retransmissions of the same PDU use the same MAC and PHY parameters computed for the first transmission.

Fig. 3.3 plots the percentage of bandwidth saved for transmitting PDUs of a single service flow (with QoS requirements) when bandwidth is dynamically allocated, instead of a fixed bandwidth allocation of 3 OFDM symbols per time frame. Dynamic bandwidth allocation is achieved using ACL strategy or an adaptive MAC (SL) strategy that optimises MAC parameter K for a given modulation. Results are shown for different values of Γ , when $\sigma = 0$. In the figure, the curves are obtained using the queueing model, while the points with the error bars indicate the simulation results. Notice that at low values of Γ the modulations at higher spectral efficiency are unable to meet QoS requirements. The plot in Fig. 3.3 shows the perfect matching between theoretical results and simulation results and the optimality of ACL approach. Moreover, the results quantify the bandwidth savings achieved by strategies with dynamic bandwidth allocation. Also, the results clearly indicate the best performance of the cross-layer strategy (i.e., ACL) with respect to a single-layer strategy (i.e., SL) in terms of both

Scenario	Log-normal σ
Urban Micro Cell	2.3 dB
LOS fixed station	3.4 dB
Urban Macro Cell	8.0 dB
Indoor small office	12 dB

Table 3.2: ARQ ACL strategy, mobility scenarios.

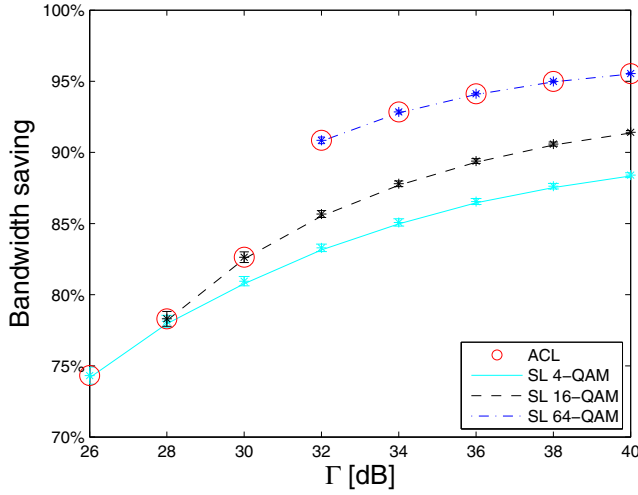


Figure 3.3: ITU vehicular A channel: percentage of bandwidth saving vs. Γ , for ACL and SL strategies, $\sigma = 0$ (no shadowing).

bandwidth savings (for a given Γ) and wider Γ range over which QoS can be guaranteed.

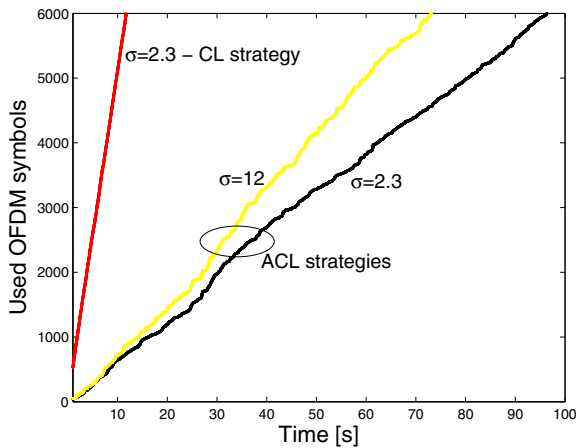


Figure 3.4: ITU vehicular A channel: overall number of OFDM symbols used vs. time, for different values of σ .

Fig. 3.4 plots the overall number of OFDM symbols used for PDU transmissions versus the time, when the ACL strategy ($\sigma = 2.3$ dB and $\sigma = 12$ dB) and when a (non-adaptive) cross-layer (CL) strategy ($\sigma = 2.3$ dB) is used. The CL strategy optimises both the PHY and MAC layer parameters for a mean Γ value (i.e., $\Gamma = 33$ dB) at the MT, without tracking the shadowing variations. The gap between the curves of ACL and CL strategy indicates the amount of bandwidth (in number of OFDM symbols) saved thanks to the adaptive approach. For higher values of σ (e.g., user in office), a higher amount of bandwidth is required, because there is a higher probability that the channel quality degrades and thus lower spectral efficiency modulations must be selected.

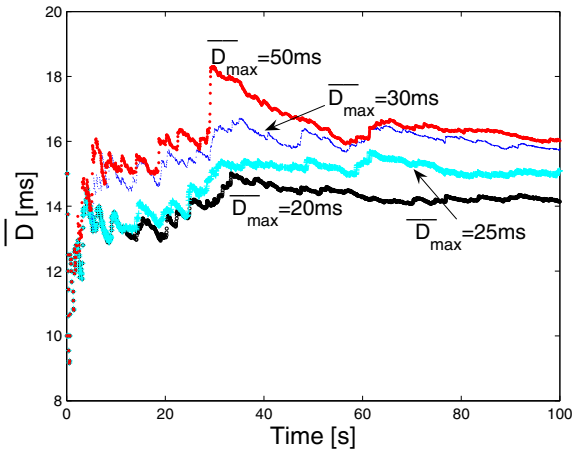


Figure 3.5: ITU vehicular A channel: \bar{D} vs. time, for different QoS requirements of \bar{D}_{max} , $L_{max} = 10^{-4}$.

Fig. 3.5 plots the expected delay, \bar{D} , computed from time instant at 0 s, versus time, for different requirements of \bar{D}_{max} . Results show that the QoS is met even for the most stringent \bar{D}_{max} requirements and that \bar{D} is not significantly affected by channel quality fluctuations. The gap between \bar{D} and \bar{D}_{max} indicates that it is possible to propose other ACL strategies that trade the delay for PDU loss rate (i.e., by increasing the value of K). However, such strategies may require a larger bandwidth utilisation with respect to the proposed one.

It is interesting to compare the performance of Fig. 3.5 against the value of \bar{D} achieved in a flat fading channel, as shown in Fig. 3.6. The comparison shows that user terminal mobility cause wider fluctuations in the MAC layer performance. Note that the performance of flat fading channel are achieved on a wider Γ range (i.e., 12-40 dB) and for a log-normal distribution centered at 26 dB, due to the very low channel frequency selectivity.

Fig. 3.7 plots \bar{D} versus time, for different mobility scenarios, i.e., for different values of σ . Although, the value of \bar{D} tends to stabilize in an interval of few ms, it is interesting to notice that a channel with high Γ variance, due to shadowing, leads to better performance in terms of expected delay. This counter-intuitive result is due to the combination of shadowing effects and ACL strategy performance.

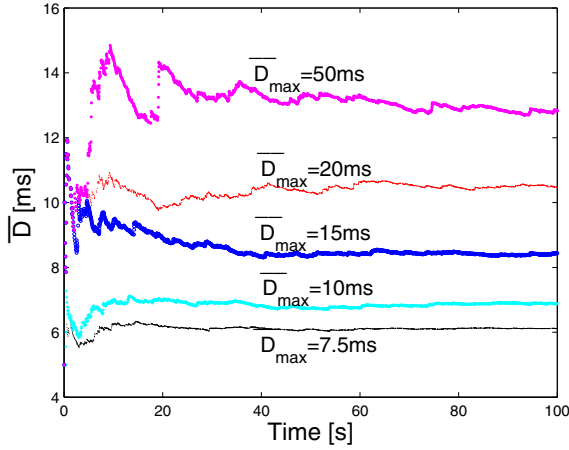


Figure 3.6: Flat fading channel: \bar{D} vs. time, for different QoS requirements of \bar{D}_{max} , $L_{max} = 10^{-4}$.

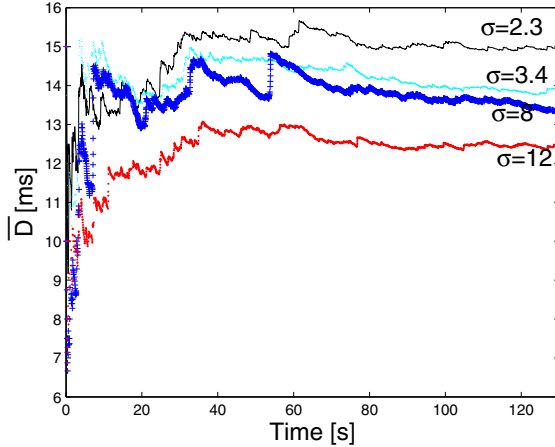


Figure 3.7: ITU vehicular A channel: \bar{D} vs. time, for different values of σ .

Fig. 3.8 plots \bar{D} versus time, for different service flow bandwidth requirements, i.e., for different values of λ , when $L_{max} = 0.001$. The increase of λ leads to an increase of \bar{D} , making it difficult to ensure requested QoS. For values of $\lambda > 1.5$ PDU per time frame, QoS requirements cannot be guaranteed anymore on the Γ range. In such case, Γ range on which QoS should be guaranteed, service flow bandwidth (e.g., to allow more than one PDU transmission per time frame) or service flow QoS (e.g., PDU loss rate and expected delay) should be renegotiated with the BS.

Simulation results show that the ACL strategy is able to promptly adapt to channel quality variations and to meet QoS requirements using the minimum amount of bandwidth.

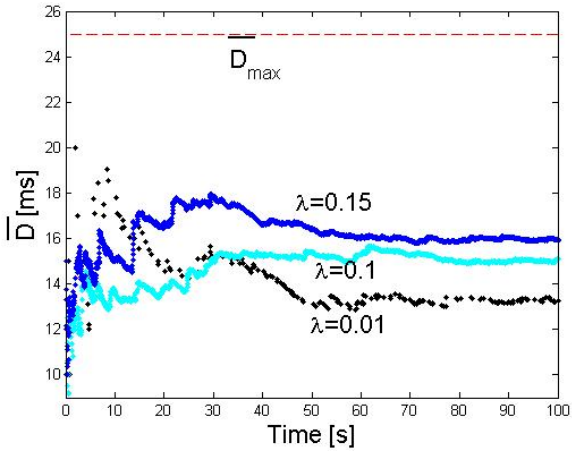


Figure 3.8: ITU vehicular A channel: \bar{D} vs. time, for different values of λ .

3.3 DSA: Dynamic Service Flow Addition

The family of IEEE 802.16 standards permits to easily manage and guarantee the requested QoS to each connection, referred to as *service flows*, established between the BS and the MTs [CLME06].

At the LLC sublayer of the 802.16 networks, admission control is operated via a three-way handshake procedure, referred to as DSA and service flows can be dynamically established upon request with specific QoS requirements.

The decision whether admitting or not a new service flow to the network has to guarantee that the QoS levels are maintained for the service flows already active and for the one to be added.

The DSA protocol is based on a request message, a response message, and an acknowledgement message, exchanged between BS and MT. Service flow is activated and the requested QoS is guaranteed, when the protocol terminates successfully (i.e., both the response and the acknowledgement are positive and the corresponding messages are received).

However, the unreliable nature of the radio medium may seriously compromise DSA message exchange. Thus, to ensure a successful reception of DSA messages, DSA protocol is provided with a robust retransmission strategy. When a reply to a transmitted DSA message (e.g., the request or the response message) is not received correctly within a timeout, a retransmission of the same DSA message can be attempted. Maximum number of message transmission attempts and timeout duration can be flexibility selected within a wide range of values defined by the standard. Although robust, DSA signalling can still be blocked if all the multiple attempts of sending a request (or a response) message fail. Thus, service flow requests can be blocked due to unavailable resources to meet QoS (i.e., admission control blocking) as well as erroneous termination of DSA protocol (i.e., signalling blocking). Blocking and latency of DSA protocol are both affected by the selection of timeout values and the maximum number of attempts.

A number of works addressed the performance of data transmissions over wireless channels [TDN08] and ARQ protocol [CMCP08] in WiMAX networks. Also, strategies to ensure QoS of activated service flows [LXG02, BSY08], cross-layer approaches that adapt to the channel conditions [EF08, CMCP08] and service flow scheduling [SIJT09, BdF08, LMP07, MMAP08] have been proposed for WiMAX networks. Although relevant, these works are unable to capture the peculiar behaviour of DSA protocol and its performance, especially in a mobile environment. The problem is that a degradation of the radio channel, for instance due to mobility, could lead to an increase of DSA protocol blocking and in turn to a reduction of the overall network performance. Similar problems occur also during the message exchange for handover procedure as investigated in [YHK08].

In this section, the performance of DSA protocol are thoroughly evaluated in a mobile networks based on OFDM physical layer in Time Division Duplex (TDD). DSA is investigated in terms of signalling blocking, admission control blocking, and latency for a variety of scenarios, i.e., for different channel quality conditions, degrees of reciprocity between DL and UL radio channels, MT speeds, OFDM configurations and DSA protocol parameters.

Analytical expressions for admission and signalling blocking probabilities and for service activation latency will be derived. The probability of service activation without MT knowledge will be also derived.

A comprehensive study of the impact of the different parameters on DSA protocol performance is assessed. Simulation results aim at quantifying

the impact of mobile user speeds and PHY layer parameters on MAC performance. This helps to derive useful considerations on the strategies to use to counter-react the performance degradation due to high speeds, long channel coherence time at low (pedestrian) speeds, and absence of reciprocity between DL and UL channels. The work has been submitted with the work in [MPCC10].

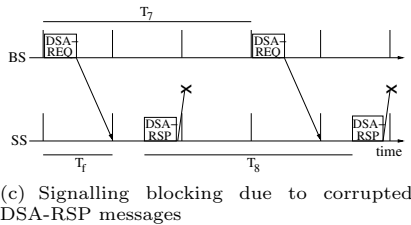
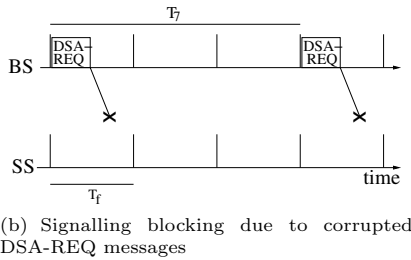
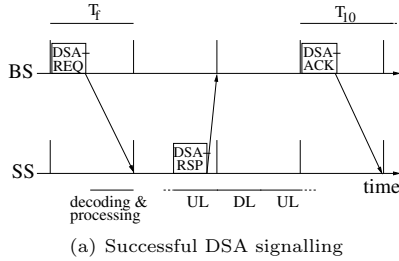


Figure 3.9: Time chart of DSA signalling.

3.3.1 System Model

The considered physical layer is based on OFDM as standardised in [IEE06]. In the OFDM physical layer, each time frame is composed by N_{ofdm} OFDM symbols. Each OFDM symbol consists of N_{fft} subcarriers. The frame duration is $T_f = N_{ofdm} \cdot T_{ofdm}$, where T_{ofdm} is the OFDM symbol duration and $T_{ofdm} = \frac{1}{\Delta F}$ where ΔF is the subcarrier frequency separation. Given the maximum MT velocity and the maximum level of tolerable ICI, a minimum ΔF (i.e., a maximum duration of the OFDM symbol) is derived according to [LC01, vNP00]. The frequency separation is kept fixed to a value that usually ensures an ICI below $-27dB$. Typical values of T_{ofdm} are in the order of tens of μs and time frame duration in the range of $[1, 20]ms$.

An OFDM symbol is considered incorrectly received if the symbol carried by any subcarrier is detected incorrectly (i.e., no FEC is assumed). The probability that an OFDM symbol is received incorrect depends on the SNR distribution over the frequency domain. The instantaneous SNR of the i^{th} subcarrier for the j^{th} OFDM symbol of the frame is defined as: $\gamma_{ij} = \frac{E_s}{N_0} \|\mathbf{H}(i, j)\|^2$, where $\frac{E_s}{N_0}$ is the average SNR and $\mathbf{H}(i, j)$ is the channel response in the frequency domain for the i^{th} subcarrier for the j^{th} OFDM symbol, Sec. 2.6.3. The distribution γ_{ij} along i depends on the channel PDP. The PDPs are chosen accordingly to the ITU channel model for pedestrian and vehicular scenarios [SMF05]. Each path of the multipath channel is Rayleigh distributed, thus the distribution of the SNR for the subcarriers follows a Rayleigh distribution too [KYL03]. The PDP determines the channel frequency profile; mobility and the Doppler Spectrum affect the channel response reducing the correlation of the subcarrier SNR over time, according to the speed, see Sec. 2.6.3.

3.3.2 DSA Signalling Analysis

Consider a wireless network based on IEEE 802.16 operating in point-to-multipoint mode, i.e., the BS communicates with a number of MTs. DL and UL transmissions are scheduled by the BS in TDD mode, within each time frame. Time frame duration is fixed and indicated as T_f .

When a DL or UL service flow is requested, admission control operations are executed. The admission control blocking is derived under the assumption that a fixed number m of service flow requests can be accommodated, i.e., each service flow requests $1/m$ of the available bandwidth in the network. Under the above assumption and assuming that the signalling latency is negligible with respect to the service flow duration, the network can be modelled as a $M/M/m/m$ queue. The probability that an arriving customer finds $M/M/m/m$ queue full is given by the well-known Erlang B formula [GH85].

If the service flow is admitted, it can be activated using a three-way handshake between the BS and the MT. In the followings, the BS-triggered signalling[†] is considered for establishing DL or UL service flows. For simplicity and without loss of generality, a single MT is assumed.

Upon request of a service flow, the BS performs admission control operations. If the service flow is admitted, the service flow can be activated using a three-way handshake between the BS and the MT. Signalling is triggered by the BS by sending a DSA-REQ message (i.e., dynamic service flow activation request) to the MT. Upon correct reception of the DSA-REQ message, the MT replies with a DSA-RSP (i.e., dynamic service flow activation response). Upon correct reception of the DSA-RSP message, the BS replies with a DSA-ACK (i.e., dynamic service flow activation acknowledgement). Service flow is activated at the BS after the expiration of timeout T_{10} , as described in Sec.6.3.14.8 in [IEEE09a]. This case of successful signalling is sketched in Fig. 3.9(a).

The messages are sent on the primary management connection, using the most robust physical layer profile. However, due to the unreliability of the radio channel, one or more of the messages may be lost or incorrectly received. For this purpose, DSA-REQ and DSA-RSP messages may be retransmitted multiple times, upon expiration of a timeout.

[†]According to IEEE 802.16 standard, support of BS-triggered signalling is a mandatory requirement, while support of MT-triggered signalling is optional.

Upon expiration of timeout T_7 , BS can retransmit a copy of the DSA-REQ message. BS can transmit up to n_R copies of a DSA-REQ message. This case is represented in Fig. 3.9(b), for $n_R = 3$.

Upon expiration of timeout T_8 , the MT can retransmit a copy of the DSA-RSP message. MT can transmit up to n_S copies of a DSA-RSP message. This case is represented in Fig. 3.9(c), for $n_S = 3$.

A service flow request can be blocked due to:

- admission control: BS may decide to not admit a service flow when the available bandwidth or QoS cannot be guaranteed;
- erroneous termination of signalling: when the BS does not receive any DSA-RSP from the MT, the requested service flow is not activated, Fig. 3.9(b);
- negative responses: for various reasons, MT (BS) can reply with a negative DSA-RSP (negative DSA-ACK). In such cases, the requested service flow is not activated.

In addition to the above cases, a service flow can be successfully activated at the BS, without that MT is informed. This happens when the MT does not received any positive DSA-ACK sent from the BS (Fig. 3.9(c)) and may create burdens at the MAC and upper layers.

The performance of the DSA signalling is analytically quantified in terms of blocking probability and latency, under the assumption of absence of negative responses from BS and MT: the blocking analysed here is due to the radio transmission unreliability.

The analysis is based on the following parameters:

- p_R, p_S, p_A : expected probability that DSA-REQ, DSA-RSP or DSA-ACK is received incorrectly, respectively;
- n_R, n_S : maximum number of copies of a DSA-REQ or DSA-RSP message, respectively;
- T_f : time frame duration;
- T_7 : timeout that triggers a DSA-REQ retransmission;
- T_8 : timeout that triggers a DSA-RSP retransmission;
- T_{10} : timeout after which a service flow is considered blocked or ready for data transmission;
- λ : expected arrival rate of service flow requests;
- τ : expected duration of a service flow.

DSA performance is analytically derived assuming that timeouts satisfy (see Fig. 3.9):

$$T_7 \leq T_8 \qquad T_{10} > (n_S - 1) \cdot T_8. \qquad (3.8)$$

The first assumption ensures that for a requested service flow up to n_R copies of DSA-REQ can be transmitted. In other words, two consecutive DSA-REQ for the same service flow cannot trigger a new sequence of n_S copies of the DSA-RSP message. The second assumption ensures that the BS takes the final decision about the activation or blocking of a service flow only when all the copies of DSA-RSP should have been sent.

The expected blocking due to erroneous termination of signalling, or *signalling blocking*, the expected admission control blocking and the expected probability that a service flow is activated without MT knowledge are derived. The analysis is performed assuming that the radio channel in different time frames is uncorrelated.

Signalling Blocking

Let B_R be the expected probability that no DSA-REQ is correctly received by the MT, for a service flow request, i.e.,

$$B_R = p_R^{n_R}. \quad (3.9)$$

Let B_S be the expected probability that no DSA-RSP is correctly received by the BS, under the assumption that MT correctly received one or more copies of DSA-REQ for a service flow request, i.e.,

$$B_S = p_S^{n_S}. \quad (3.10)$$

Thus, signalling blocking (B_{DSA}) can be evaluated as

$$B_{DSA} = B_R + (1 - B_R) \cdot B_S. \quad (3.11)$$

Admission Control Blocking

Admission control blocking is derived under the assumption that a fixed number m of service flow requests can be accommodated, i.e., each service flow requests $1/m$ of the available bandwidth in the IEEE 802.16 network. Assuming that the signalling latency is negligible with respect to the service flow duration, the network can be modelled as a M/M/ m/m queue. In such queue, customers are the service flows, the service time is the duration of a service flow, and the queue length is finite and equal to the maximum number of service flows that can be accommodated, i.e., m .

Admission control blocking, B_{adm} , can be computed as the probability that an arriving customer finds M/M/ m/m queue full, i.e., m customers waiting in queue. Such probability is given by the well-known Erlang B formula [GH85]:

$$B_{adm} = \frac{(1 - \zeta^m)}{m! \sum_{h=0}^m \zeta^h}, \quad (3.12)$$

where ζ is the queue utilisation, i.e.,

$$\zeta = \lambda \cdot \tau \cdot (1 - B_{DSA}) \quad (3.13)$$

When removing the assumption of negligible signalling latency with respect to the service flow duration, Eq. (3.12) is a lower bound of the admission control blocking.

Activation without MT Knowledge

The expected probability that a service flow is activated without MT knowledge (B_{awk}) is given by the probability that MT is not able to decode correctly any of the copies of a DSA-ACK message sent by the BS, i.e.,

$$\begin{aligned} B_{awk} &= \sum_{i=1}^{n_S} \text{prob}\{\text{first correct DSARSP is } i^{\text{th}}\} \\ &\cdot \sum_{m=1}^{n_S-i} \text{prob}\{m \text{ correct DSA-RSP} | (n_S-i) \text{ DSA-RSP sent}\} \\ &\cdot \text{prob}\{m+1 \text{ incorrect DSA-ACK}\} \\ &= \sum_{i=1}^{n_S} \frac{1-p_S}{1-p_S^{n_S}} p_S^{i-1} \sum_{m=0}^{n_S-i} (1-p_S)^m p_A^{m+1} p_S^{n_S-i-m} \binom{n_S-i}{m} \\ &= \frac{p_A(1-p_S)}{p_S(1-p_S^{n_S})} \left[s^{n_S} \left(\frac{1 - (\frac{p_S}{s})^{n_S+1}}{1 - (\frac{p_S}{s})} - 1 \right) \right] \end{aligned}$$

where $s = (p_S + (1 - p_S) \cdot p_A)$.

Signalling Latency Analysis

Signalling latency is derived under the assumption that the primary management connection has always bandwidth sufficient to transmit any DSA messages, i.e., no buffering delay of DSA messages is considered.

Signalling latency for a successfully activated service flow is defined as the time interval from the arrival of a service flow request to the activation of the service flow at the BS. Signalling latency can be decomposed into three terms: the time interval from the arrival of a service flow request to the start of the transmission of the first DSA-REQ message (L_{arr}), the time interval from the first DSA-REQ to the first correctly received DSA-RSP (L_{DSA}), and time interval from the activation of the service flow to when data transmission can take place (i.e, T_{10}), see Fig. 3.9. Thus, the expected signalling latency (L) is equal to

$$L = L_{arr} + L_{DSA} + T_{10} \quad (3.14)$$

When DSA requests follow a Poisson process, the first term can be approximated[‡] as:

$$L_{arr} \approx \frac{T_f}{2} \quad (3.15)$$

where T_f is the time frame duration. The second term, L_{DSA} , is given by

$$\begin{aligned} L_{DSA} = & \sum_{i=1}^{n_R} p_R^{i-1} \frac{1 - p_R}{1 - p_R^{n_R}} (i - 1) T_7 \left[\frac{T_7}{T_f} \right] \\ & + \sum_{i=1}^{n_S} p_S^{i-1} \frac{1 - p_S}{1 - p_S^{n_S}} (i - 1) T_t + T_s \end{aligned} \quad (3.16)$$

where T_t is the expected latency between the transmission of two DSA-RSP copies and T_s is the latency between the transmission of a DSA-REQ and the corresponding DSA-RSP, according to transmission scheduling. The first term of L_{DSA} indicates the latency from the first transmitted DSA-RSP to the first correctly received DSA-REQ. The second term is the latency required to correctly receive a DSA-RSP once a DSA-REQ has been correctly received. T_t depends on the number of DSA-RSP attempts remaining to transmit, on the timeout values for T_7 and T_8 and their ratio. An approximation is given in Sec. 3.17 in the following section.

3.3.3 DSA Analysis Validation

Blocking and latency experienced by DSA signalling are evaluated and compared under different values of n_R and n_S . Results are obtained Analytically (AN) and by Simulations (SIM). The theoretical analysis is verified with computer simulations for a block fading channel: the channel is uncorrelated among successive frames.

The considered IEEE 802.16 physical layer is as follows: $N_{sub} = 1024$ [§], $\Delta f = 11.16071429$ kHz, $T_{ofdm} = 102.86\mu$ s, $N_{ofdm} = 48$, $T_f = 5$ ms [Yag04]. Signal modulation is uncoded BPSK and carrier frequency is 3.5 GHz. A single OFDM symbol is considered incorrectly received if

[‡] When $B_{adm} = 0$, $L_{arr} = \frac{T_f}{2}$.

[§] For simplicity, the number of gap and null subcarriers is 0.

one or more (or all) subcarriers contain an error. For such physical layer, a single OFDM symbol is sufficient to support a DSA message.

Inter-arrival times ($1/\lambda$) and duration (τ) of the service flow requests are exponentially distributed. The timeouts used for DSA are: $T_7 = 3 \text{ ms}$, $T_8 = 8 \text{ ms}$, $T_{10} = 20 \text{ ms}$. With such selection of timeout values, it is possible to approximate T_t in (3.16) as

$$T_t = \left\lceil \frac{T_8}{T_f} \right\rceil p_R + \left\lceil \frac{T_7}{T_f} \right\rceil (1 - p_R), \quad (3.17)$$

i.e., retransmission of a DSA-RSP message is triggered either by a DSA-REQ message when the copy of a DSA-REQ message (transmitted upon expiration of timeout T_7) is received correctly or by the expiration of timeout T_8 for DSA-RSP message. Scheduling latency in Eq. (3.16) is set to $T_s = 0 \text{ ms}$.

To compare the analytical results with the simulation results, an OFDM channel characterised by independent and uncorrelated and uniformly distributed OFDM symbol errors is considered. MAC layer parameters are set as follows: $n_R = n_S$, $n_R \in \{1, 2, 3\}$, $\lambda = 1 \text{ s}^{-1}$ and $\tau = 1 \text{ s}$.

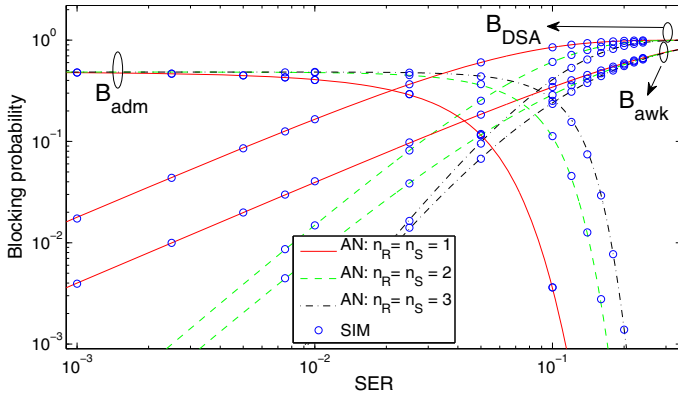


Figure 3.10: Simplified channel model: blocking probability vs. SER.

Fig. 3.10 shows the different components of blocking probability versus the average OFDM SER. Analytical results closely match simulation results. This is true also for the approximation given for the admission control blocking (B_{adm}) in (3.12).

The figure shows that signalling blocking (B_{DSA}) and admission without MT knowledge (B_{awk}) increases with SER. On the other hand, B_{adm} decreases with SER. This behaviour is due to the fact that a lower number of service flows can be activated due to the higher signalling blocking, resulting in a higher probability that resources are available for an incoming service flow request.

Increasing the number of retransmissions of control messages (i.e., n_R and n_S) helps to significantly reduce B_{DSA} and B_{awk} . On the other hand, B_{adm} increases with n_R and n_S , due to the decrease in B_{DSA} . Finally, it is important to note that B_{awk} is non-negligible, especially at high SER. This result calls for protocol improvement and strategies, in order to inform the MT of the activated service flows.

The performance improvement of B_{DSA} with n_R and n_S comes at the price of an increased signalling latency (L), as shown in Fig. 3.11. Analyt-

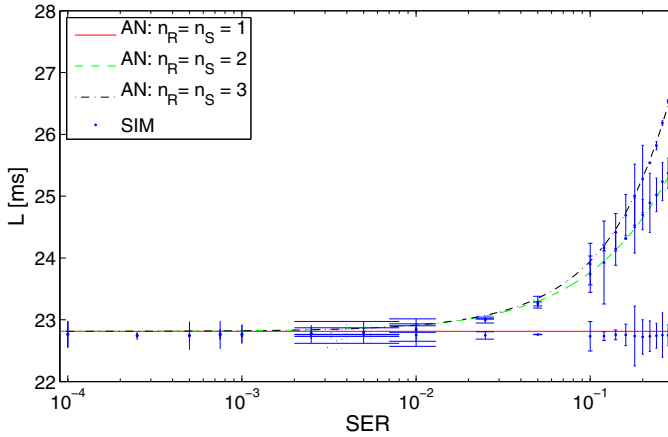


Figure 3.11: Simplified channel model: signalling latency vs. SER.

ical results well follow the simulation curves. Signalling latency is constant with SER when $n_R = n_S = 1$, i.e., only one DSA-REQ and DSA-RSP is allowed to activate a service flow. When increasing n_R and n_S , latency arises with SER, as multiple retransmissions of the same control message are allowed and may be required.

3.3.4 Performance Analysis in Mobility

In this section, DSA protocol performance is evaluated in terms of:

- signalling blocking probability: it is the probability that no DSA-RSP is correctly received by the BS within T_{10} timeout in response to DSA-REQ message.
- admission control blocking probability: it is the probability that a service flow request cannot be admitted due to lack of available resources.
- signalling latency: it is the time interval from the arrival of a service flow request to the instant in which data transmission has taken place

A simplified admission control mechanism it is derived under the assumption that a fixed number m of service flow requests can be accommodated, i.e., each service flow requests $1/m$ of the available bandwidth in the IEEE 802.16 network.

To quantify the performance, MAC and PHY layer are implemented as described in Section 3.3. in a IEEE802.16 C/C++ custom-made event-driven simulator. Simulation results are collected using the following configuration, unless otherwise indicated. Parameters of the OFDM physical layer are: $N_{fft} = 1024$, $\Delta F = 11.16071429$ kHz, $T_{ofdm} = 102.86$ μ s, $N_{OFDM} = 48$, $T_f = 5$ ms [Yag04]. With this selection of N_{fft} , each DSA message can be accommodated on a single OFDM symbol. Signal modulation is uncoded BPSK and carrier frequency is 3.5 GHz.

At the MAC layer, the timeouts are set to: $T_7 = 3 \cdot T_f$, $T_8 = 3 \cdot T_f$, $T_{10} = 4 \cdot T_f$. Scheduling latency for transmitting DSA messages (i.e., buffering delay) is considered negligible. Latency due to signal and information

processing at the PHY layer is accounted as indicated in Fig. 3.9, i.e., half time frame is required for decoding the message and generating a reply. Maximum number of copies of DSA-REQ and DSA-RSP messages are $n = n_R = n_S \in \{1, 3\}$. Inter-arrival times ($1/\lambda$) and duration of the service flow requests are exponentially distributed. Inter-arrival rate of DSA requests is set to $\lambda = 20 \text{ s}^{-1}$ and expected duration of service flows is set to 1 s. Each service flow requests 20% of the available bandwidth in the IEEE 802.16 network. Effects of MT mobility are taken into account by implementing a channel model with Jakes Doppler spectrum. The ground speed v of MT with respect to the BS is selected in the range $[5, 300]$ km/h. For comparison purposes, a single-path block-fading channel without mobility is also considered and indicated as $v = 0$. DL and UL channels are assumed to be independent.

Impact of Mobility

Figs. 3.12-3.14 show the impact of MT mobility in terms of signalling blocking, admission control blocking and latency, respectively, when $n = 1, 2$, and 3. Mobility deteriorates the channel performance and degrades the DSA performance in terms of signalling blocking and latency. However, when MT speed exceed 50 km/h (i.e., from 50 km/h to 300 km/h), DSA performance do not degrade any further. Signalling blocking can be improved by resorting to higher number of transmission attempts (n), to detriment of an increased latency. Admission control blocking is reduced with mobility. This happens because the number of service flow requests blocked by erroneous termination of signalling increases with MT speed and, thus, the network load decreases lowering the probability of finding a full admission queue.

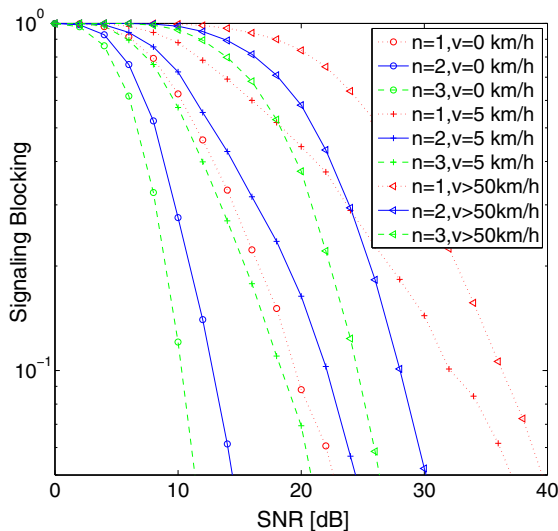


Figure 3.12: Signalling blocking vs. SNR for different values of v and n .

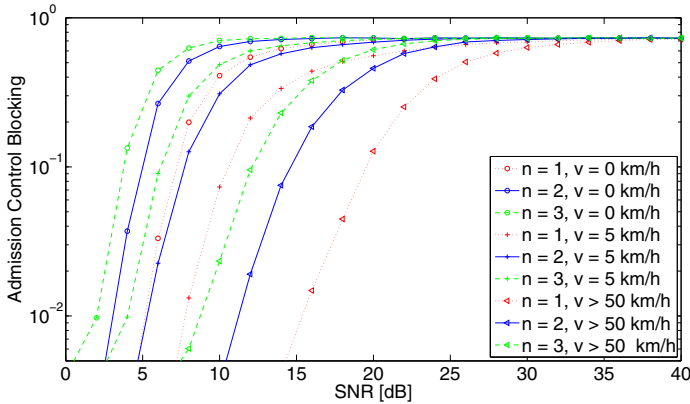


Figure 3.13: Admission control blocking vs. SNR for different values of v and n .

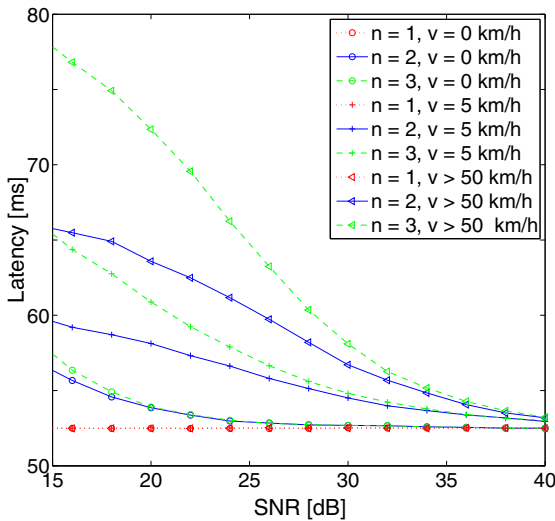


Figure 3.14: Latency vs. SNR for different values of v and n .

Impact of Correlation between UL and DL Channels

Figs. 3.15 and 3.16 compare the signalling blocking and the latency, respectively, experienced at pedestrian speed ($v = 5$ km/h), when the DL and UL channels are reciprocal (i.e., DL and UL are considered as a single fading channel) or independent fading processes. In Fig. 3.15, independent block-fading channels (with no time correlation) are also included for comparison. The results indicate that the reciprocity is beneficial when $n = 1$ (no retransmissions) for any value of SNR. When $n > 1$, reciprocity benefits are limited to low SNRs and may compensate the negative effects of channel correlation which has worse performance of the case of block-fading channels (no time correlation).

Fig. 3.16 shows that the correlation between the DL and UL channels allows to greatly lower the latency of the admitted service flows, especially at low SNR.

In the presence of time correlation, the reciprocity of DL-UL channels is beneficial. This is due to a lower coherence time of the error process when DSA messages are sent over a single fading channel: the coherence time of the independent channels is lower bounded by the coherence time of the reciprocal DL/UL channels. Loss of reciprocity between DL and UL channels can be recovered with calibration techniques [GSK05] which could be available in future MTs.

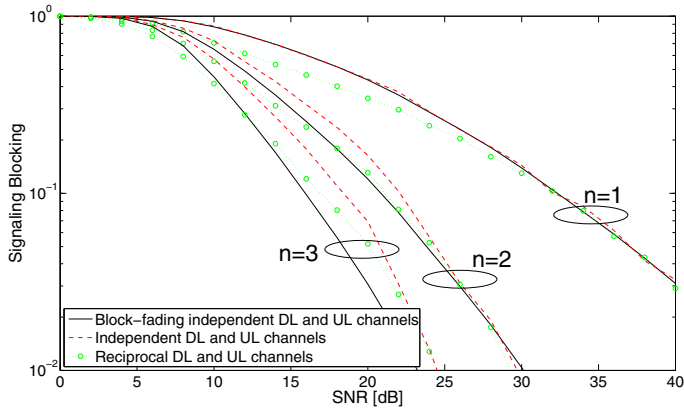


Figure 3.15: Signalling blocking vs. SNR for $v = 5$ km/h, when UL and DL channels are either independents or fully reciprocal.

Impact of the Number of Subcarriers per OFDM Symbol

In IEEE802.16 scalable OFDM [Yag04], ΔF is fixed and different bandwidths can be exploited varying the number of subcarriers (N_{fft}), thus, the impact of N_{fft} over DSA is analysed. For $N_{fft} < 1024$, the DSA messages are spread over two (or more) consecutive OFDM symbols. A DSA message is received incorrectly when at least one of OFDM symbols is in error. In Figs. 3.17 and 3.18, the signalling blocking is evaluated as a function of SNR for various value of N_{fft} , when $v = 50$ km/h and $v = 300$ km/h, respectively. At 50 km/h, the correlation of the channel is favourable: if an OFDM symbol is received correctly, with high probability the consecutive OFDM symbols are received correctly too, leading to a correct reception of a DSA message. Thus, the channel correlation, combined with the small values of N_{fft} , lowers the signalling blocking. In Fig. 3.18, at 300 km/h, the difference among the curves at different N_{fft} is greatly narrowed because, at high velocities, the channel varies sensibly during consecutive OFDM symbols and the probability that two or more consecutive OFDM symbols are received incorrect is lower.

Impact of Time Frame Duration and DSA Timeouts

In scalable OFDM, the frame duration, T_f , is also configurable; in particular, short T_f with low N_{fft} have high overheads and should be avoided, [Yag04]. Fig. 3.19 shows the impact of T_f on the signalling blocking probability as

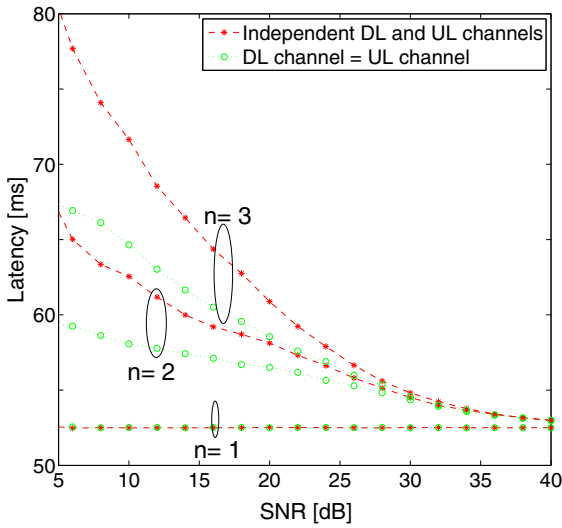


Figure 3.16: Latency vs. SNR for $v = 5$ km/h, when UL and DL channels are either independents or fully reciprocal.

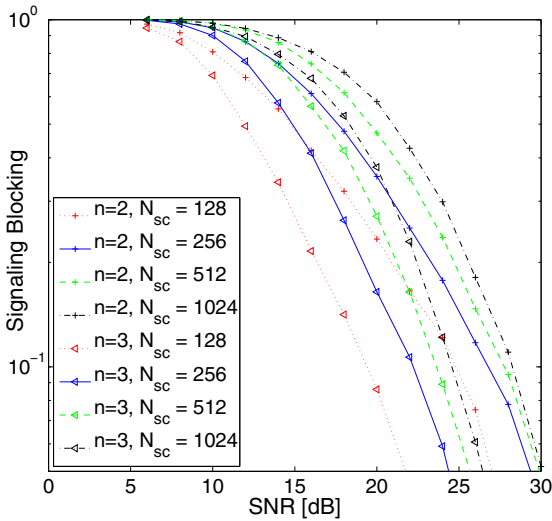


Figure 3.17: Signalling blocking vs. SNR = 20 dB, for $v = 50$ km/h and different values of n and N_{fft} .

a function of the SNR, for $v = 5$ km/h. The signal blocking probability decreases not only with an higher number of retransmissions but also with longer frame duration. Indeed, by using long time frames, the channel correlation can be reduced, with a beneficial effect on the signalling blocking. The positive effect of T_f duration is particularly evident at high SNR -

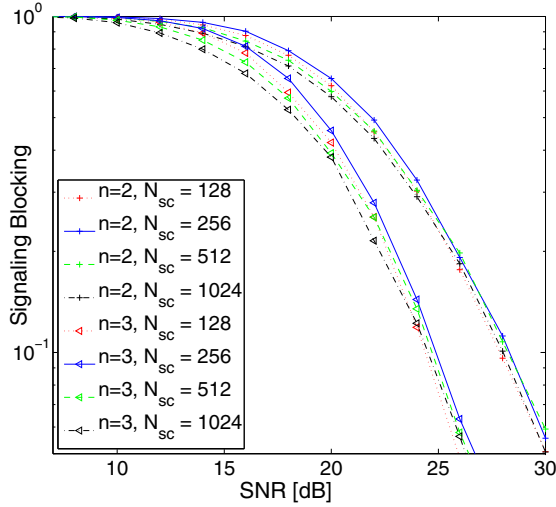


Figure 3.18: Signalling blocking vs. SNR = 20 dB, for $v = 300$ km/h and different values of n and N_{fft} .

where the channel performance is dominated by the fading coherence time with respect to the additive white noise.

Fig. 3.20 shows the latency normalised to number of time frames, for $v = 5$ km/h and various value of T_f . Correlation experienced at low T_f plays a positive effect on the latency, i.e., if a message is received correctly so do the following ones with high probability. Thus, low values of T_f are preferable for a fast activation of service flows. However, in a practical case, small values of T_f may lead to a loss of throughput due to overheads (e.g., for synchronization and data maps). Since the normalized latency for various T_f is comparable while T_f may up to 8 times longer, it is evident that low values of T_f are preferable for a fast activation of service flows. However, in a practical case, small values of T_f may lead to a loss of throughput due to overheads (e.g., for synchronisation and data maps).

Viceversa, the use of long time frames allows to exploit the time diversity and, thus, increase the probability of a successful reception to detriment of a higher latency.

Impact of Timeout Duration

In addition to time frame duration, the channel correlation experienced by successive transmission of DSA messages can be controlled by varying the duration of DSA timeouts. Figs. 3.21 and 3.22 show the impact of timeout duration on DSA performance, for increasing values of SNR and $v = 5$ km/h. Consistently with the results in Fig. 3.19, the lower signalling blocking is achieved for longer timeout duration (i.e., $T_7 = T_8 = 4$ time frames in the figures) as time diversity can be exploited. Even in this case, the price to pay for the improved blocking is an increase of the latency.

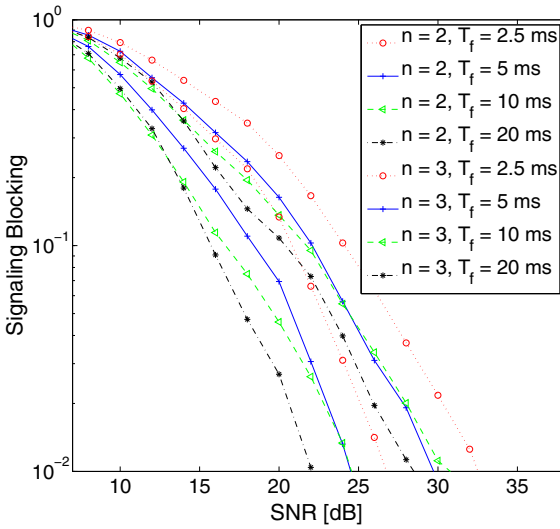


Figure 3.19: Signalling blocking vs. SNR for $v = 5$ km/h and different values T_f .

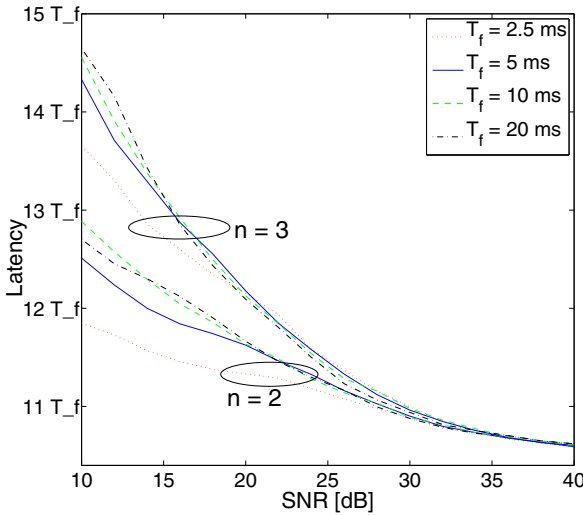


Figure 3.20: Latency vs. SNR for $v = 5$ km/h and different values T_f .

Impact of Network Load

Fig. 3.23 shows the admission control blocking experienced by a MT moving at $v = 5$ and 300 km/h, respectively, and with a channel SNR of 20 dB, for increasing inter-arrival rate λ . The admission control blocking increases with λ , but also with n . The reason is that, for larger values of n , a

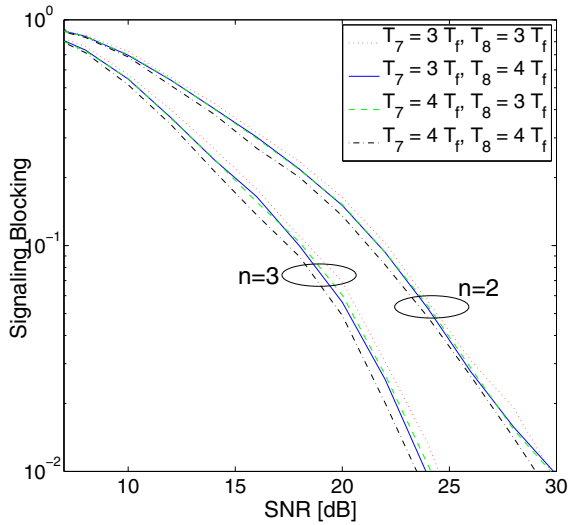


Figure 3.21: Signalling blocking vs. SNR for different values of timeouts and $v = 5$ km/h.

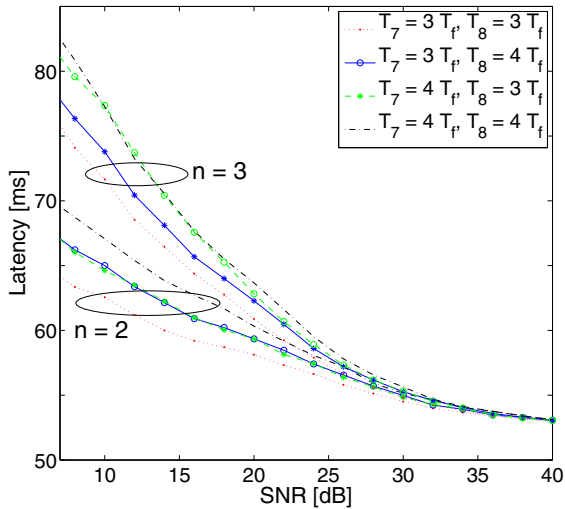


Figure 3.22: Latency vs. SNR for different values of timeouts and $v = 5$ km/h.

higher number of service flows are admitted, leading to an increase of the effective network load. The increase of admission control blocking with n is particularly evident at MT speed of $v = 300$ km/h.

Results indicate that robustness of DSA protocol can be achieved by increasing the number of message transmission attempts, at the expenses

of an increased latency. This is especially important for fast moving MT. Indeed, mobility has been shown to have a detrimental effect on the DSA performance. Contrary to fast moving MT, slowly moving MT (e.g., pedestrian speed) suffers from the effect of the long channel coherence times and from the independence of DL and UL channels. To compensate the negative impact of the long channel coherence time on the signalling blocking, time diversity could be exploited, equivalently attainable by selecting longer time frames or longer duration of timeouts, at the expenses of the latency. In addition, a careful selection of the number of OFDM subcarriers (e.g., as in scalable OFDM), that are used to carry the DSA message, may help to reduce signalling blocking and latency. Finally, these results help to properly configure DSA parameters (i.e., timeouts and number of retransmissions) to achieve a requested level of DSA blocking and latency.

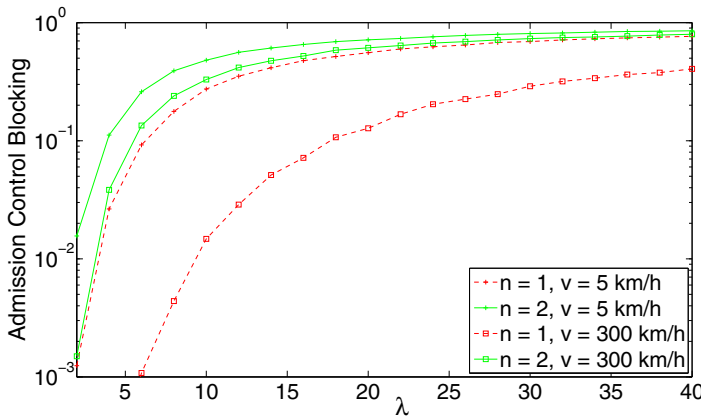


Figure 3.23: Admission control blocking vs. λ for SNR = 20 dB and $v = 5$ and 300 km/h

3.4 Conclusions

This chapter has been focused on the analysis at the MAC Layer of a IEEE802.16e wireless network[¶]. In particular, firstly an ARQ Adaptive Cross Layer strategy that jointly optimises PHY and MAC parameters to meet users' QoS was presented. ARQ ACL strategy is based on a queueing theory model. The theoretical results were shown to perfectly match the simulation results. ACL strategy was tested in different scenarios of user mobility.

Secondly, the performance of the Dynamic Service Addition (DSA) protocol has been analysed. An extensive performance evaluation of the DSA protocol has been carried out through theoretical analysis. Theoretical results have been verified with computer simulations for block fading channel. Theoretic results perfectly matched the simulations. Performance evaluation have been performed also for a variety of scenarios. Signalling blocking, admission control blocking, and latency experienced by DSA protocol

[¶]These conclusions can also refer to a generic wireless network with an OFDM PHY layer and a connection-oriented MAC layer, such as LTE-A.

have been quantified for different channel and mobility conditions and various PHY and MAC parameters. DSA protocol parameters thus can be trimmed to meet QoS requirements especially for highly moving MTs to react to detrimental signal shadowing.

The ARQ protocol will be used with dynamic MIMO schemes in Chapter 4 and the MIMO signal processing gain will be jointly exploited with the protocol gain coming from the ARQ process.

The integration of DSA and ARQ with MIMO schemes will be discussed in Chapter 5 where the impact of introducing MIMO protocol gain in current network architectures will be studied.

Chapter 4

MIMO Protocol Gain

By gaining Paramatma, everything will become automatic. If you leave Paramatma and try to gain everything else, you will never be able to gain anything. Whatever is gained, will look so little, that you will not feel happy.

BRAHMANANDA SARASWATI
(1870-1953)

4.1 Introduction

During the last decade, the research community has witnessed the transition from MIMO as a theoretical concept to a practical system component for enhancing performance of wireless networks. Point-to-point MIMO over a single radio link promised large gains for both channel capacity and reliability, essentially via the use of spatial multiplexing and space diversity respectively; the extra spatial degrees of freedom coming from MIMO signalling are exploited for expanding the dimensions available for signal processing and detection, acting mainly as a pure PHY layer performance booster.

However, several difficulties have been encountered to achieved very high SEs. This is mainly due to channel propagation characteristics (loss of MIMO channel ranks) and high correlation of the antenna elements mainly at the MT side, where the small dimensions pose challenges in obtaining independent fading statistics on each antenna. Further, in single user MIMO, the link layer protocol cuts the performance benefits of MIMO antenna because it does not have full awareness of the MIMO capabilities, even if allowing for greater per-user data rate or more reliable communications.

The recent development of cross-layer techniques, aiming at the joint optimisation of the PHY and MAC layers, has shown that noticeable per-

formance improvement is at hand. The research community is taking off from single-link MIMO to Multi-User designs to attain a protocol gain.

An important benefit of this approach comes from the opportunistic MU scheduling. As opposed to point to point communication, multi-user communications introduces several new features like cooperation, superposition of information, use of side information, joint source and channel coding, interference management and competition techniques, [BBCS06, ME09, TD08, LJ08].

Fundamental recent results in this area have pointed that PHY layer (modulation/coding), link layer, resource allocation and scheduling are deeply connected. A remarkable improvement can be achieved when all these parameters are considered with the objectives of capacity maximisation and best resource exploitation.

4.2 MU MIMO State of the Art

Multiple antenna communications play an important role for achieving the high data rates required by the increasing traffic demand to be supported in future wireless networks.

When, from a processing gain exploited on a single link, MIMO is used jointly with higher layers protocols, the SU MIMO limitations discussed in Chap. 1 can be mitigated and information theory has shown that consistent gains can be achieved when MIMO is applied in MU scenarios, [EGC80, ZLH04, WK09, RH07, PW07, HCK09].

MU-MIMO can achieve capacity gains not only via multiuser multiplexing schemes but also transmitting with multiple antennas at the BS towards different users simultaneously (MU-MISO systems) and with transmission from multiple MTs (even with a single antenna) towards the BS (MU-SIMO systems). IEEE802.16 Uplink Collaborative scheme falls inside the latter of the beforementioned classes.

The scheduling of the set of the best users in the cell maximise the sum-rate, exploiting the *multiuser diversity*. Hybrid SM transmission can be used to counterfight the problem of low MIMO channel rank for a certain MT jointly scheduling different users: the equivalent MIMO matrix rank can be preserved due to the statistical independence of the channels of widely separated MTs.

However, from an information theoretic point of view, the sum-rate capacity in MU-MIMO DL Channel is achieved with Dirty Paper Coding (DPC), [Cos83], which is an interference cancellation technique combined with user scheduling and power control. The complexity of signal processing and the heavy feedback needed for DPC are too high to be implemented in a real system and a plethora of solutions have been proposed to approach, as close as possible, the capacity upper-bound represented by DPC.

Another limitation comes from the CSI feedback. When CSI feedback is feasible, its overhead can affect the performance of the network in terms of throughput. The CSI feedback is challenging for two reasons: feedback channel capacity is proportional to the number of users and antennas and, for mobile MTs, the information quickly becomes outdated. Partial (or no) feedback schemes have been proposed, [SU09, XZ09, DL08, KK08], just to cite some recent works.

User selection mechanisms are also required. This calls for a problem solver with a search space proportional to the number of users currently admitted in the network. Brute force search can be not feasible and trade-off solutions have been proposed, as in [WML02].

A simple MU-MIMO system design with low feedback was proposed in [VTL02]: Opportunistic Random BeamForming (ORBF). ORBF needs

an high number of users in order to have an effective system throughput gain. In [KHGT05], the issue of a low/realistic number of MTs has been addressed using multiple random coefficients inside a single time slot but exacerbating the problem of the feedback rate. In [KG05], a memory-based approach is proposed to exploit the time correlation of the channels when the number of users in the cell is low.

The feedback problem can be solved resorting to open-loop MIMO schemes, such as STBC or SM, which do not require CSIT. These schemes can be exploited to achieve diversity and/or rate gain. The optimal scheme can be selected dynamically, based on the channel status and on the spatial characteristics of the MIMO channel. A switching system based on the Demmel condition number of the channel matrix was presented in [HP05] and discussed in Sec. 2.5.1 at page 26.

Equal SE can be optimally chosen based on constellation distances expected at the receiver. The decision is taken at the receiver based on locally estimated CSI and then sent back to the BS. Adaptation to the channel state can be challenging in very fast channels since the Demmel condition number can be subjected to quick variations, as visible in Fig. 2.6 at page 29. When variable spectral efficiency MIMO schemes are available, the selection can be based on a minimum BER to be guaranteed at the PHY layer, for instance as in [MMPR07] presented at page 25.

This chapter presents the analysis of a system with ARQ and MIMO schemes jointly integrated and optimised over a narrowband Rayleigh channel. Several retransmission strategies are proposed for MU networks, as detailed in Sec. 4.4.

The reference MIMO system is chosen as a 2×2 system as it constitutes the baseline for the next generation wireless networks. When more antennas are available, more than two users can be served balancing the trade-off between SM and STBC, e.g. using a Layered STBC [HP05, ITU08, AI-05], see Sec. 2.2.2 at page 16.

Retransmission strategies can be designed for larger antenna sets or the antenna elements can be partitioned in sets of two antennas and MIMO-ARQ protocol applied independently on each set.

The work proposed in this chapter stems from the results obtained in [dCP06, dCP08] where an optimal ARQ retransmission strategy has been defined for a single user using SM and STBC MIMO schemes.

In MU scenario, the MIMO-ARQ transmission protocol addresses the previous drawbacks of MU-MIMO exploiting multiuser diversity via random user selection without any need of CSI feedback. Open-loop MIMO schemes are used – with the added Hybrid SM scheme – jointly with Multi-User ARQ; the strategy designed in [dCP08] has been extended for a Multi-User scenario.

The SM-STBC switching allows for preserving diversity order in the low SNR region in a smoother way than the strategies presented in [HP05, MMPR07]. The system proposed allows for sending data streams to different users in SM: Hybrid SM. When transmitting to two different users, the users to be coupled together can be chosen randomly or informed by the feedback of the post-processing SNR from the MTs to the BS. Hybrid SM transmissions gives the opportunity to send a packet to a new user while retransmitting the wrongly decoded packet to a previous user: the users are served with a lower average delay and the total system throughput is leveraged.

Multi-User ARQ has been investigated for optimising multicast/unicast flows in broadcast transmissions, [LJ06a, Lar07a].

Antenna Allocation during retransmissions is also analysed. An optimal strategy for single and multi user cases is defined. The packet scheduling

is done in RR fashion. The system serves a pool of users in DL. The users are served with an implicit RR scheduler with Time Division Multiple Access (TDMA); fairness among users is thus guaranteed (as in [VTL02]). However, the system design easily permits to introduce QoS aware scheduling, taking advantage of the increasing capacity achieved via the MIMO-ARQ transmissions, as visible in Fig. 4.1 at page 91.

The MIMO-ARQ protocol decides the packet scheduling after the BS has decided the AMC profile based on the long term SNR of each MT and the analysis is conducted for homogeneous AMC profiles among all the MTs.

In Sec. 4.3, the work from [dCP08] is presented, as it is the basis for the further developments in MU scenario. Then, in Sec. 4.4, the optimal antenna allocation and MU retransmission strategies are analysed. In this case, it is possible to enhance the performance of the system with an optimised receiver which stores the unintended packets (the packets that were sent to other users in the previous transmission slots). This receiver is later referred to as: Packet Overhearing (PO) and detailed in Sec. 4.6. In Chapter 5, the architectural issues related to the introduction of the MIMO-ARQ protocol proposed in this chapter will be discussed.

4.3 ARQ Strategies for Single User MIMO

This section presents the work of [dCP08]. The design proposed herein is then further generalised for the Multiple User case in [MDCP09].

The channels between the transmitting and receiving antennas are modelled with flat fading Rayleigh statistics and are considered constant over the retransmission period. The packets are composed by a number of data symbols plus an appended Cyclic Redundancy Check (CRC) to enable error detection at the receiver. The data is 4-QAM modulated and the duration of the transmission slot is fixed. During each slot two packets are sent, one from each transmitting antenna. The MIMO system is a 2×2 . On each antenna, the packet is coded/modulated independently and it can be independently retransmitted. The retransmission strategy has the following characteristics:

- retransmitted packets have the same symbol content of the original message
- sign and conjugation operations are allowed

The retransmission strategy is decided at the MT; 1 additional bit of feedback is required to indicate if the antenna allocation is to be switched, together with another bit (for each packet) indicating if the packet was received or not (ACK/NACK). The feedback is received error-free and without delay. The CSI at the receiver is assumed perfect.

The SU strategy is summarised as follows:

- when at time t only 1 packet is in error, the erroneous packet is retransmitted at $t + 1$ from one of the antenna and a new packet is transmitted from the other antenna
- when both packets are erroneously decoded, both are retransmitted using Alamouti coding, [Ala98]
- when both packets are decoded correctly, in the next time slot two new packets are sent.

The receiver is a SIC-MMSE receiver with Packet Combining (PC): the receiver stores the signals containing the packets decoded with errors until the current time slot and removes the contribution of the correctly decoded packets. Those past signals ensure a higher probability for correct decoding of the transmitted packets.

The work in [dCP08] assumes the same AMC level. The aim of the design is to maximise the throughput. Two designs are proposed: one design based on post-processing SNRs at the output of the SIC-MMSE receivers at the MTs and one based on an approximation of the throughput.

4.3.1 Single User (SU) Protocol Operation

The notation used in the following presentation is reported:

- X_A is a packet (data plus CRC symbols), in the first time slot, X_A is transmitted from antenna one and X_B from antenna two
- x_a is the single modulation symbol inside packet X_A
- $\mathbf{H}(t)$ is the 2×2 MIMO channel matrix
- $h_{ij} \in \mathbb{C}$ is the channel coefficient from transmit antenna j^{th} to receiver antenna i^{th}
- \mathbf{H}_A or \mathbf{H}_B are the column vectors defined respectively as the channel coefficients from the transmitting antenna for packet X_A or X_B to the two antennas at the receiver.

The BS always optimistically starts with SM transmission. If two packets are decoded, the receiver sends 2 ACKs and the system remains in SM mode.

Assume that X_A and X_B have been transmitted in the first time slot. The first case applies when X_A and X_B are decoded with errors. At time $t + 1$, Alamouti coding starts, antenna allocation is switched – X_B is retransmitted from antenna 1 and X_A from antenna 2 – and sign and conjugation operations are applied, as in [Ala98]. The system design permits to seamless switch between SM and STBC modes. This appears more flexible than the solution proposed in [HP05] or in [MMPR07].

The second case applies when only 1 packet, say X_A , is decoded with errors: X_A is retransmitted from the weakest antenna because if it were retransmitted from the strongest antenna, the total information available for decoding the packet after two transmissions could have been unnecessarily large.

The first 2 schemes are providing gain if the channel is highly correlated between two successive time slots (see Sec. 4.7.4 for an analysis in fast varying channels for SU and MU networks). The MIMO channel $\mathbf{H}(t)$ is a 2×2 matrix and it is assumed constant over 2 transmission slots. If a more complex MIMO system is available, the retransmission format could be applied at each couple of antennas independently; it could happen that two antenna are swapped while the other transmit other packets in SM mode or LSTBC schemes could be applied as in Sec. 2.6.

Those schemes are optimal (according to the design criterion) when X_A and X_B are new transmitted packets at time t . When either X_A or X_B are (re)transmitted before time t , the schemes are not always optimal. The design aims at maximising the system throughput. In the case of retransmissions, the throughput cannot be expressed easily in closed-form. Two retransmission designs are proposed with using an approximation of the actual throughput: one based on throughput and one based on post-processing SNR at the output of SIC-MMSE receiver.

Throughput based design

Since $\mathbf{H}^P(t+1) = \mathbf{H}(t)$ (correlated channel over two time slots assumption), the receiver calculates the predicted SNRs for both streams at time $t+1$: $\gamma_K^P(t+1)$, where K stands for the particular stream and $K \in A, B$. The predicted throughput at time $t+1$ can be computed as:

$$Thr^P(t+1) = Thr_A^P(t+1) + Thr_B^P(t+1) \quad (4.1)$$

where

$$Thr_K^P(t+1) = r_K(t+1)[1 - PER_K(\gamma_K^P(t+1))]$$

and $r_K(t+1)$ and $PER_K(t+1)$ are the nominal bit rate and Packet Error Rate (PER) for stream K at $t+1$. The MT decides to swap antenna allocation in order to obtain the maximum throughput evaluating (4.1) for the two allocations possible. The reliability of the decision decreases as the channel correlation decrease. Equation (4.1) do not express the real throughput of the system since it does not take into account the whole retransmission. This criterion is also valid when both streams have different AMC levels.

When the two streams have the same modulation and coding profile, maximizing (4.1) is equivalent to minimising the sum of the PER_A and PER_B . When closed-form PER expressions are not available for the PHY profile in use, look-up tables can be used.

SNR based design

In the region where $Thr^P(\gamma)$ is a concave function of γ , maximising (4.1) is equivalent to maximising (4.2) and antenna switching is decided for maximisation of the sum of the SNRs of both streams at the SIC-MMSE receiver output:

$$\gamma^P(t+1) = \gamma_A^P(t+1) + \gamma_B^P(t+1) \quad (4.2)$$

Maximising (4.2) is equivalent to minimising the difference, over the retransmission slots, of the sums of the channel coefficients from the transmitting antenna to both the receiver antennas:

$$\min \left(|\chi_A^P(t+1) + \chi_B^P(t+1)| \right) \quad (4.3)$$

where

$$\chi_X^P(t+1) = \sum_{s \geq 0}^L \|\mathbf{H}_X(t-s)\|^2 + \|\mathbf{H}_X(t+1)\|^2 \quad (4.4)$$

and L is the number of retransmissions elapsed. The previous relations mean that if packet X_A is decoded with errors, retransmission is done from the weakest antenna. When packets X_A and X_B are decoded with errors, antenna allocation is switched or not, depending on the relative value of $\chi_A^P(t+1)$ and $\chi_B^P(t+1)$.

After antenna allocation has been decided, the retransmission has to decide if sign or conjugate operators are to be used. Contrary to the case where both packets are decoded with errors and are retransmitted following the structure of a STBC using sign and conjugation operators, if only 1 packet is in error, the packet is retransmitted unchanged. In any case, the format is unique and is known a-priori by the transmitter: the transmitter only needs to know the antenna allocation.

4.4 MU MIMO-ARQ

This work proposes a MIMO-ARQ cross-layer protocol without feedback from the MTs to the BS which integrates the characteristics of SM and STBC MIMO schemes and ARQ in a MU networks.

The system proposed in [dCP08] could be straightforwardly applied in MU systems in a TDM fashion. For the current user scheduled, the system retransmits the packets until they are received or until the maximum number of retransmissions are achieved, then the next user in the scheduler queue is served with the same strategy. However, a MT in bad channel state forces the system to wait until the ARQ protocol elapses the full retransmission time, which is not efficient.

The MU MIMO-ARQ protocol integrates the ARQ process and open-loop MIMO schemes in a MU system with an implicit RR scheduler with the ability to send packets in Hybrid SM, so that other users can be served while the wrong packet directed to an user in bad channel condition is being retransmitted. The system design allows for integration of any kind of scheduler considering QoS attributes of the traffic flows.

Several retransmission strategies are proposed. Firstly, two strategies are studied. Assume that in the first time slot two packets are transmitted in SM. A first strategy is to retransmit the wrong packet with double power from the best antenna (strategy S1) while the other antenna is switched off. The other strategy is to retransmit the wrong packet from both antenna to the same user in spatial repetition coding, (strategy S2).

A more comprehensive strategy is proposed in order to exploit the MU diversity (strategy MU) which observes the following: when there is one packet in error for user (1), the packet is assumed to have been transmitted from the weaker antenna. The antenna allocation is thus deterministically switched and the packet retransmitted from the (assumed) stronger antenna. A packet for a new user ($u + 1$) is transmitted on the other antenna. The antenna allocation follows the opposite principle respect to [dCP08] and aims to give the maximum SNR of the recombined signal in order to get rid of the bad user as quickly as possible. This strategy permits to have the lowest service time for the user (1), reducing the average delay in the system.

The Hybrid SM transmission (data streams sent to different users) attains higher throughput exploiting the multi-user diversity: indeed, while the wrong packet is being retransmitted, other users have the chance to receive a packet. The new user is chosen randomly. The performance of the strategy is compared to the case where the best user is chosen in order to maximise the throughput (i.e. the user with the highest post-processing SNR at the output of the receiver for the antenna to be used).

During Hybrid SM, ARQ is used with two different users; this constitutes a Multi-User ARQ. Multi-User ARQ was previously introduced at MAC layer for multiple unicast flows in [LJ06b, Lar07b] where each MT stores not only the packets directed to the local MT but also the packets directed to the other users (unintended packets). ACKs are sent back by each user for both types of packets (intended and unintended). The aggregate throughput efficiency is enhanced via wise selection of the users and joint encoding of retransmitted information. A joint network and channel coding with Multi-User ARQ has been proposed in [Tho08] where all the received packets (intended and unintended) are exploited for iterative decoding.

The benefits of the MU MIMO-ARQ protocol can be recognised in:

- No feedback at the BS for antenna allocation

- No feedback at the BS for user selection
- Open-loop STBC and SM used seamlessly providing higher diversity order at low SNR regions
- Robustness in the presence of very fast varying channels
- Good down-scalability: as the number of users decreases the performance is lower-bounded by the SU protocol, as in [dCP08].

An optimized SIC-MMSE receiver is also proposed. The MT selected for Hybrid SM transmission stores and exploits unintended packets sent in previous time slots to other users. The MT removes the contribution of the "overheard" packets to maximize the probability of the good reception of his packet. This receiver is later referred to as PO.

The following Sec. 4.5 presents in details the MU MIMO-ARQ state machine and the retransmission strategies proposed. Sec. 4.5.6 describes the user selection methods: the random user selection and the best user selection used to upper bound the performance achievable with MU diversity exploitation. The optimised PO receiver is explained in Sec. 4.6. Then, Sec. 4.7 reports the results before final conclusions in Sec. 4.8.

4.5 System Model

The DL transmission from a BS to a set of MTs - in number of N_{users} - is studied. The users are served with a RR scheduler with TDMA and fairness among users is guaranteed, as in [VTL02]. The transmission is in broadcasting to all the users. Without loss of generality, the MIMO system is assumed with $N_T = N_R = 2$, i.e. two transmitting antenna at the BS and two receiving antenna for each MT. If more antennas are available, more than two users can be served balancing the trade-off between SM and STBC, e.g. using a layered STBC [HP05, ITU08, Al-05]. Open-loop MIMO schemes are used: STBC, SM and Hybrid SM. In Hybrid SM, two data streams are sent to different users simultaneously. An ARQ protocol is used with same symbol content and with independently coded packets. Such ARQ configuration allows to retransmit packets independently to different users in Hybrid SM mode.

The packets are composed by N_{data} bits plus an appended CRC consisting of N_{crc} bits. The notation used in the design of MU MIMO-ARQ protocol is reported here:

- $X_K^{(j)}$ is the K^{th} packet for user j , the upper case letter X indicates the whole packet
- $x_K^{(j)}$ is the m^{th} symbol for user j , the lower case indicates one modulation symbol inside K^{th} packet
- $\mathbf{H}^{(j)}(t)$ is the 2×2 MIMO channel matrix from the transmitter to user j at time t
- $\mathbf{H}_i^{(j)}(t)$ is the $N_R \times 1$ channel column vector from transmit antenna i to user j at time t

The MU MIMO-ARQ protocol transmission is started in SM mode to the first user (1) in the RR scheduler queue. If the two packets are received correctly, the BS serves the second user in the scheduler queue again in SM mode. If the two packets are received correctly, the BS transmits to the

third user until the maximum number of users is reached and then it starts again from the first user. The BS schedules the packets in RR fashion.

If both packets are decoded with errors, they are retransmitted in Alamouti mode: the two transmissions (first and second time slots) are equivalent to one Alamouti block. The Alamouti transmission is carried out until both packets are received or until the maximum number of retransmissions K_{max} is reached. In this status, the two packets are directed to the same user and we have SU STBC transmission.

When one only packet is in error, the erroneous one is retransmitted coupled with a new packet. In this case, three retransmission strategies are proposed. The first strategy (S1), retransmits the packet from the best antenna with double power. This strategy needs a single bit feedback from the MT to the BS (for choosing one antenna out of two). Then, a second strategy is proposed (S2) where the packet is retransmitted from both antennas. In these first two strategies, Hybrid SM is never used.

A third retransmission strategy (MU) uses Hybrid SM to exploit multi-user diversity gain. It operates as follows. Assume that only one out of two packets for user (1) is received correctly. The packet in error is assumed to have been transmitted from the weakest antenna. On the first retransmission, the antenna allocation is switched. The erroneous packet is transmitted from what it is supposed to be the best antenna for communicating with user (1) and a new packet is transmitted on the other antenna for a new user (u). In RAND user selection, the new user is selected choosing randomly among all the users in the system. The performance is upper-bounded by the BEST user selection, in this case the new user is selected as the best user for throughput maximisation, the user selection is detailed in Sec. 4.5.6. The BEST user selection needs the feedback from all the MTs to the BS of the post-processing SNRs at the two outputs of the receiver. The feedback link is assumed error-free and delay-less.

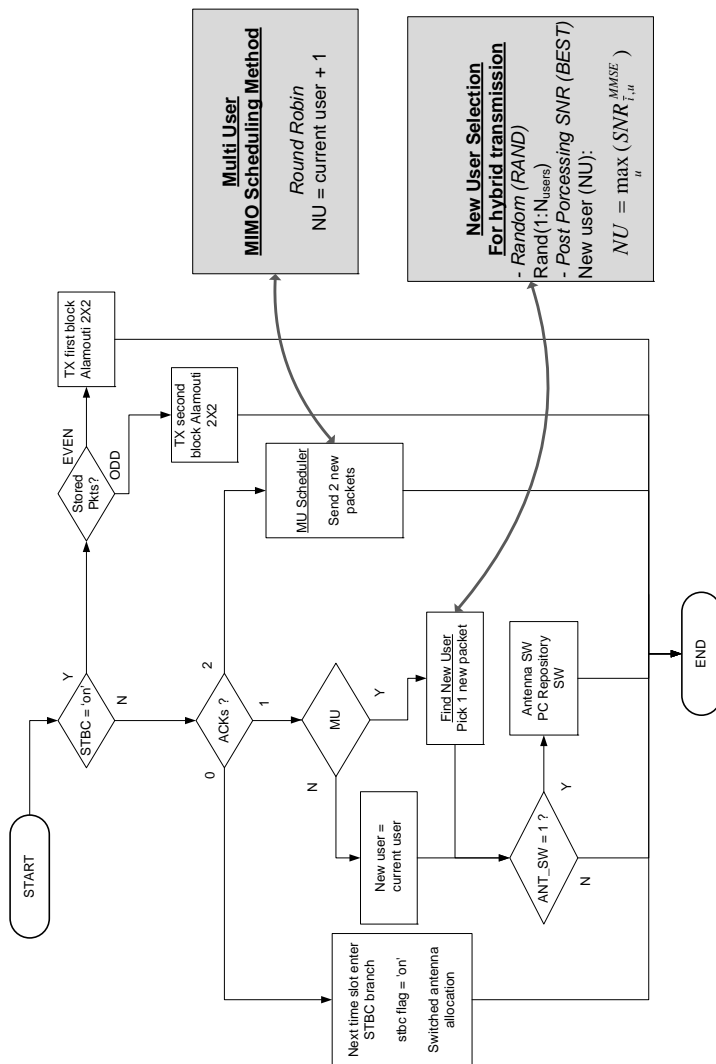


Figure 4.1: MU MIMO-ARQ Tx State Machine Diagram.

The MIMO ARQ protocol can be described by a state machine located at the BS which is reported in Fig. 4.1. The state machine includes SU and MU systems. The main block of the state machine is the central ACKs check block. Three output conditions are possible: no packets correctly decoded (0), only one (1) or (2) correctly decoded. If both packets are not decoded (Acks=0), the antenna are switched and the STBC flag is set to "on"; in the next time slot, the transmission state machine will enter the STBC branch. In the STBC branch, if the number of Alamouti coded blocks transmitted is odd, the packets are transmitted with sign and conjugate operation as in [Ala98]. If an even number of packets have been transmitted – i.e. an integer number of Alamouti transmissions have been completed – the packets are retransmitted unaltered to form the basis for the next Alamouti block.

If only one packet has been decoded (Acks=1), in the case of a SU system, a new packet is sent for the same user. Antenna switching criteria is evaluated at the MT and the allocation is sent back to the BS, as in [dCP08]. In MU systems, a new user is selected (RAND / BEST). For the first retransmission, the antenna allocation is switched, while it remains fixed in the following time slots if Hybrid SM transmission is still used. Hybrid SM transmission continues to be used if these two conditions are both met: the retransmitted packet has not been received and both packets are not in error. If both packets are received in error during Hybrid SM mode, both packets are retransmitted in STBC. Differently from the previous case, the STBC packets are not directed to the same user and we have a MU STBC. MU STBC is terminated if the packet retransmitted for the user (1) – the scheduled user in the first time slot – is received or K_{max} retransmissions have been tried. In MU STBC, if the packet for the new user (u) is received, the system continues to retransmit both packets in STBC. This is inefficient from the user (u) prospective, but permits to preserve the information stored in the previous time slots via PC for user (1) and to preserve the diversity order of the already sent Alamouti blocks. Finally, if both packets are decoded correctly (Acks=2), the scheduler transmits two new packets for the next user.

Fig. 4.1 highlights where other schedulers and user selection strategies can be inserted in the proposed MIMO-ARQ protocol architecture. In particular, the RR scheduler can be enhanced considering QoS parameters at higher layers for the determination of the users to be scheduled. Further, the user selection process is also highlighted and other selection strategies can be inserted in the system design. In particular, a joint design of scheduling and user selection methods seems very promising for exploiting at its best the MU network resources. The methods can be designed when variable degrees of information about the MTs CSI are available at the BS. User selection method can be dynamically adapted based on the queue status – e.g. when an user has a delay approaching its maximum latency time, this user can be selected for the next Hybrid SM transmission, thus giving the chance to reset his Head Of Line delay (HOL) delay.

The system seamlessly switches between STBC and SM furnishing a variable diversity order depending on the channel status. At low SNR regions, a diversity order of two is available. In the high SNR region, SM is usually used, while with moderate SNRs (from 5 to 20 dB) both schemes are used.

4.5.1 MT Receiver

The MT receiver applies Packet Combining (PC) which combines the signals received during all the retransmissions in order to have the highest

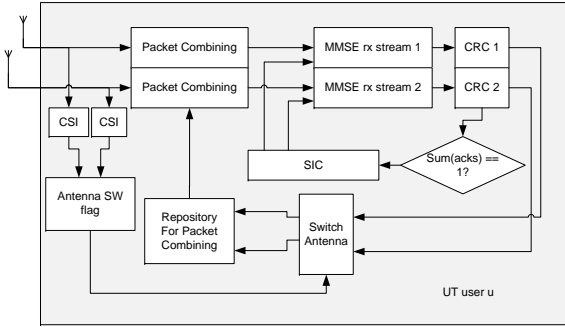


Figure 4.2: MIMO-ARQ Receiver Diagram.

probability of correct decoding the packets. The PC exploits only the packets directed to the intended user. In Sec. 4.6, an optimized receiver will be described where the user not only exploits the packets directed to itself, but stores also the packets sent previously to other users.

In Fig. 4.2, a receiver diagram is reported. The MT has a SIC-MMSE receiver which inverts the equivalent channel matrix $\mathbf{H}_{eq}^{(u)}$, where (u) is the user index. The data symbols are estimated via a pseudo-inversion as:

$$\begin{aligned} \hat{\mathcal{X}} &= \mathcal{F}_{MMSE}^{(u)} \cdot \mathcal{Y}_u(t+k) \\ &= \left(\mathbf{H}_{eq}^{\dagger(u)}(t+k) \mathbf{H}_{eq}^{(u)}(t+k) + \frac{\sigma_n^2}{P} \right)^{-1} \mathbf{H}_{eq}^{\dagger(u)}(t) \cdot \mathcal{Y}_u(t+k) \end{aligned}$$

where \dagger is the conjugate transpose operator, $\mathcal{Y}_u(t+k)$ is the vector of all the received signals until the current time slot and has dimension of $2 \times (k+1)$ and k is the number of retransmissions elapsed until the current time slot. The construction of $\mathbf{H}_{eq}^{(u)}$ depends on the protocol state evolution and it is explained in the followings for each retransmission strategy.

After the MMSE channel inversion, the CRC of both streams is checked. If only one packet is decoded, the SIC part is activated: the correct packet is subtracted from the received signals and decoding is performed again. After this cancellation, the final ACKs/NACKs are communicated to the BS and the MIMO-ARQ protocol proceeds to the next transmission slot.

4.5.2 Retransmission Strategy S1

Assume that user (1) is the actual user scheduled for transmission and assume that in the first time slot only packet $X_B^{(1)}$ was correctly received and $X_A^{(1)}$ has to be retransmitted. In S1 strategy, the packet in error is retransmitted from the antenna which gives the best post-SNR at the receiver, while the other antenna is kept silent. The transmission power over the best antenna is doubled. This strategy needs one-bit feedback at the BS in order to select one antenna out of two available at the BS. It is evident that if more than two antenna are used at the BS, the feedback would need a number of bits equal to $N_b = \log_2(N_T)$.

The correctly received packet at the first time slot is cancelled and the received signals at MT₁ at time $t+1$ for the m^{th} symbol of the packet can

be written as:

$$\begin{aligned} \mathcal{Y}_1 &= \begin{bmatrix} \tilde{y}_1(t) \\ y_1(t+1) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_1^{(1)}(t) \\ \mathbf{H}_B^{(1)}(t+1) \end{bmatrix} \cdot x_A^{(1)} + \mathcal{N} \end{aligned}$$

where $\mathbf{H}_B^{(1)}(t+1)$ is the channel column vector for the antenna with the best post-processing SNR for user (1), $\mathcal{N} = (n_A(t) \ n_B(t))^T$ is the Gaussian noise term, $n_i(t)$ is $\mathcal{CN}(0, 1)$ and the cancelled signals is:

$$\hat{y}_1(t) = y_1 - \mathbf{H}_1^{(1)}(t) \cdot x_B^{(1)} \quad (4.5)$$

If the packet is still not received, at the second time slot, the received signal is expressed as:

$$\begin{aligned} \mathcal{Y}_1 &= \begin{bmatrix} \tilde{y}_1(t) \\ y_1(t+1) \\ y_1(t+2) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_1^{(1)}(t) \\ \mathbf{H}_B^{(1)}(t+1) \\ \mathbf{H}_B^{(1)}(t+2) \end{bmatrix} \cdot x_A^{(1)} + \mathcal{N} \end{aligned}$$

This process holds until the packet is received or the maximum number of retransmission K_{max} has not been reached.

4.5.3 Retransmission Strategy S2

Assume that user (1) is the actual user scheduled for transmission and assume that in the first time slot only packet $X_B^{(1)}$ was correctly received and $X_A^{(1)}$ has to be retransmitted. In S2 strategy, the packet in error is retransmitted from both antenna in spatial repetition coding. The MT does not need to send the feedback of the preferred antenna to the BS, since both antenna are used.

In strategy S2, after packet $X_B^{(1)}$ has been cancelled in the first time slot, the received signal at the second time slot for the m^{th} symbol of the packet is expressed as:

$$\begin{aligned} \mathcal{Y}_1 &= \begin{bmatrix} \tilde{y}_1(t) \\ y_1(t+1) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_1^{(1)}(t) \\ \mathbf{H}_1^{(1)}(t+1) + \mathbf{H}_2^{(1)}(t+1) \end{bmatrix} \cdot x_A^{(1)} + \mathcal{N} \end{aligned}$$

The retransmission strategy S2 transmits the same signal from the two antenna and do not doubles the power at the antenna port. The power doubling required by strategy S1 is correct from an information theoretical perspective but practical problems would be encountered if S1 is applied in real system. The operating point of the RF transmitter chain will not tolerate a power totally unbalanced on a single antenna element. For this reason, S2 strategy seems more appropriate for a practical system design. The act of transmitting from both antennas can be regarded as a form of repetition coding: *spatial repetition coding*.

Time Slot	Tx Ant 1	Dec	Tx Ant2	Dec
t	$X_A^{(1)}$	✓	$X_B^{(1)}$	×
$t + 1$	$X_B^{(1)}$	×	$X_A^{(u+1)}$	✓
$t + 2$	$X_B^{(1)}$	×	$X_A^{(u+2)}$	×
$t + 3$	$-X_A^{\dagger(u+2)}$	×	$X_B^{\dagger(1)}$	×
$t + 4$	$X_B^{(1)}$	✓	$X_A^{(u+2)}$	✓

Table 4.1: MU MIMO-ARQ: case study

4.5.4 Retransmission Strategy MU

The third proposed strategy MU uses hybrid SM transmission in order to exploit multiuser diversity. This strategy leads to a more complex MIMO-ARQ state machine and several possible states are possible.

Assume that user (1) is scheduled in the first time slot, strategy MU is illustrated for the following time evolution which summarises all the configurations possible:

- First time slot t : transmission of packets $X_A^{(1)}$ and $X_B^{(1)}$ to user (1), only packet $X_A^{(1)}$ is decoded;
- $t + 1$: retransmission of $X_B^{(1)}$ from switched antenna; user $(u + 1)$ is selected for hybrid SM transmission of packet $X_A^{(u+1)}$. Packet $X_B^{(1)}$ still not decoded, while packet $X_A^{(u+1)}$ is correctly decoded at MT_{u+1} ;
- $t + 2$: retransmission of $X_B^{(1)}$ (antenna is not anymore switched); user $(u + 2)$ is selected for hybrid SM transmission of packet $X_A^{(u+2)}$. Packet $X_B^{(1)}$ is not decoded at MT_1 and $X_A^{(u+2)}$ is not decoded at MT_{u+2} ;
- $t + 3$: retransmission in MU STBC with Alamouti 2×2 ; both packets are not decoded;
- $t + 4$: retransmission of the first half of a new Alamouti block; both packets are decoded correctly; the scheduler is reset and packets $X_A^{(2)}$ and $X_B^{(2)}$ are transmitted for the second user in the scheduler queue.

Tab. 4.1 reports the packets sent on each antenna at the BS with the relative CRC checking outcome at the intended MTs for each time slot.

In the following, the received signals at the MTs are expressed along with the construction of the equivalent channel matrix: $\mathbf{H}_{eq}^{(u)}$. In the first time slot, user (1) receives packets $X_A^{(1)}$ and $X_B^{(1)}$ in SM at time t . The corresponding received signal for the m^{th} symbol of the packet is:

$$y_1(t) = \mathbf{H}^{(1)}(t) \cdot \begin{pmatrix} x_A^{(1)} \\ x_B^{(1)} \end{pmatrix} + \mathcal{N}(t)$$

where $\mathbf{H}^{(1)}(t) = [\mathbf{H}_1^{(1)}(t) | \mathbf{H}_2^{(1)}(t)]$.

Assume that packet $X_A^{(1)}$ is decoded. The SIC is performed and the contribution from packet $X_A^{(1)}$ is removed: the interference-cancelled signal is:

$$\tilde{y}_1(t) = y_1 - \mathbf{H}_1^{(1)}(t) \cdot x_A^{(1)} \quad (4.6)$$

After the cancellation, packet $X_B^{(1)}$ is decoded again, if not successful in the next time slot the packet X_B^1 is transmitted from switched antenna (i.e. one) and the packet X_A^{u+1} for a new user is sent over antenna two. The received signal at the MT₁ can be written including the previous cancelled received signal \hat{y}_1 and the new time slot transmission:

$$\begin{aligned} \mathcal{Y}_1 &= \begin{bmatrix} \tilde{y}_1(t) \\ y_1(t+1) \end{bmatrix} = \\ &= \begin{bmatrix} \mathbf{H}_2^{(1)}(t) & \mathbf{O}_{2 \times 1} \\ \mathbf{H}_1^{(1)}(t+1) & \mathbf{H}_2^{(1)}(t+1) \end{bmatrix} \begin{bmatrix} x_B^{(1)} \\ x_A^{(u+1)} \end{bmatrix} + \mathcal{N}(t) \end{aligned}$$

where $\mathcal{N}(t)$ is a four term additive Gaussian noise for \mathcal{Y}_1 .

The received signal at the MT _{$u+1$} is composed by the two terms, one for each antenna for the current $t+1$ time slot:

$$y_{u+1}(t+1) = \mathbf{H}^{(u+1)}(t+1) \cdot \begin{pmatrix} x_B^{(1)} \\ x_A^{(u+1)} \end{pmatrix} + \mathcal{N}(t) \quad (4.7)$$

where $\mathbf{H}^{(u+1)}(t+1) = [\mathbf{H}_1^{(u+1)}(t+1) \mid \mathbf{H}_2^{(u+1)}(t+1)]$.

Now suppose that packet $X_B^{(1)}$ is not correctly decoded and the new packet in hybrid transmission $X_A^{(u+1)}$ is decoded at MT _{$u+1$} . In the third time slot, packet $X_B^{(1)}$ is retransmitted without antenna switching and a new packet $X_A^{(u+2)}$ is sent on the second antenna for a new user ($u+2$). At MT₁ the signal is thus composed as:

$$\begin{aligned} \mathcal{Y}_1 &= \begin{bmatrix} \tilde{y}_1(t) \\ y_1(t+1) \\ y_1(t+2) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_2^{(1)}(t) & \mathbf{O}_{2 \times 1} & \mathbf{O}_{2 \times 1} \\ \mathbf{H}_1^{(1)}(t+1) & \mathbf{H}_2^{(1)}(t+1) & \mathbf{O}_{2 \times 1} \\ \mathbf{H}_1^{(1)}(t+2) & \mathbf{O}_{2 \times 1} & \mathbf{H}_2^{(1)}(t+2) \end{bmatrix} \cdot \\ &\quad \begin{bmatrix} x_B^{(1)} \\ x_A^{(u+1)} \\ x_A^{(u+2)} \end{bmatrix} + \mathcal{N} \end{aligned}$$

The received signal at MT _{$u+2$} is:

$$y_{u+2}(t+2) = \mathbf{H}^{(u+2)}(t+2) \cdot \begin{pmatrix} x_B^{(1)} \\ x_A^{(u+2)} \end{pmatrix} + \mathcal{N}(t) \quad (4.8)$$

where $\mathbf{H}^{(u+2)}(t+2) = [\mathbf{H}_1^{(u+2)}(t+2) \mid \mathbf{H}_2^{(u+2)}(t+2)]$. Now assume that both packets at MT₁ and MT _{$u+2$} are decoded with errors. A MU-STBC transmission is started with a 2×2 Alamouti scheme. The STBC transmission ends only when both packets are decoded correctly at the MT₁ or K_{max} retransmission are achieved. In the next time slot, at MT₁,

the received signal is expressed as:

$$\begin{aligned}
\mathcal{Y}_1 &= \begin{bmatrix} \tilde{y}_1(t) \\ y_1(t+1) \\ y_1(t+2) \\ y_1^\dagger(t+3) \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{H}_2^{(1)}(t) & \mathbf{O}_{2 \times 1} & \mathbf{O}_{2 \times 1} \\ \mathbf{H}_1^{(1)}(t+1) & \mathbf{H}_2^{(1)}(t+1) & \mathbf{O}_{2 \times 1} \\ \mathbf{H}_1^{(1)}(t+2) & \mathbf{O}_{2 \times 1} & \mathbf{H}_2^{(1)}(t+2) \\ \mathbf{H}_2^{\dagger(1)}(t+3) & \mathbf{O}_{2 \times 1} & -\mathbf{H}_1^{\dagger(1)}(t+3) \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} x_B^{(1)} \\ x_A^{(u+1)} \\ x_A^{(u+2)} \end{bmatrix} + \mathcal{N} \tag{4.9}
\end{aligned}$$

At MT_{u+2} the signal is composed by a full STBC block as:

$$\begin{aligned}
\mathcal{Y}_{u+2} &= \begin{bmatrix} y_{u+2}(t+2) \\ y_{u+2}^\dagger(t+3) \end{bmatrix} \tag{4.10} \\
&= \begin{bmatrix} \mathbf{H}_1^{(u+2)}(t+2) & \mathbf{H}_2^{(u+2)}(t+2) \\ \mathbf{H}_2^{\dagger(u+2)}(t+3) & -\mathbf{H}_1^{\dagger(u+2)}(t+3) \end{bmatrix} \cdot \begin{pmatrix} x_B^{(1)} \\ x_A^{(u+2)} \end{pmatrix}
\end{aligned}$$

The expression of the equivalent channel matrix to be inverted by the MMSE requires that the signal received for the second half of the STBC block is conjugated, as expressed with $y_1^\dagger(t+3)$ and $y_{u+2}^\dagger(t+3)$ in (4.9) and (4.10) respectively. If the packets are not correctly decoded, a new first half of STBC block is transmitted. For brevity, the received signal is reported only for MT_1 (note that the received signal for the first half of the new STBC block is not conjugated):

$$\begin{aligned}
\mathcal{Y}_1 &= \begin{bmatrix} \tilde{y}_1(t) \\ y_1(t+1) \\ y_1(t+2) \\ y_1^\dagger(t+3) \\ y_1(t+4) \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{H}_2^{(1)}(t) & \mathbf{O}_{2 \times 1} & \mathbf{O}_{2 \times 1} \\ \mathbf{H}_1^{(1)}(t+1) & \mathbf{H}_2^{(1)}(t+1) & \mathbf{O}_{2 \times 1} \\ \mathbf{H}_1^{(1)}(t+2) & \mathbf{O}_{2 \times 1} & \mathbf{H}_2^{(1)}(t+2) \\ \mathbf{H}_2^{\dagger(1)}(t+3) & \mathbf{O}_{2 \times 1} & -\mathbf{H}_1^{\dagger(1)}(t+3) \\ \mathbf{H}_1^{(1)}(t+4) & \mathbf{O}_{2 \times 1} & \mathbf{H}_2^{(1)}(t+4) \end{bmatrix} \cdot \\
&\quad \begin{bmatrix} x_B^{(1)} \\ x_A^{(u+1)} \\ x_A^{(u+2)} \end{bmatrix} + \mathcal{N} \tag{4.11}
\end{aligned}$$

Once the STBC transmission is started, the scheduler can proceed to the next user in the queue only when:

- the packet $X_B^{(1)}$ at MT_1 is decoded
- maximum number of retransmission has been reached (K_{max}).

4.5.5 SIC principle

To illustrate the SIC principle and how it is applied throughout the protocol operation, let's focus on MT₁. Beside the cancellation of packet X_A^1 (or X_B^1) in the first time slot, SIC can be applied at each step of the protocol evolution. The vector $\hat{\mathcal{X}}$ contains the estimates of the packet(s) to be received at the current MT and also other packets directed to other users, if hybrid SM transmission mode has been used. Suppose that the intended packet is not received. If one (or more) of the other packets is/are correctly decoded, the MT cancels their contributions and then tries to recover again the remaining packets. This can be done several times if a new packet is decoded successful at each SIC iteration.

To illustrate this mechanism, let's refer to (4.11). At the output of the receiver, three estimates are available: $\left[\hat{x}_B^{(1)} \quad \hat{x}_A^{(u+1)} \quad \hat{x}_A^{(u+2)} \right]^T$. If $X_B^{(1)}$ packet is decoded correctly the scheduler can serve the next user in the queue. Otherwise, $X_A^{(u+1)}$ and/or $X_A^{(u+2)}$ can be decoded correctly. The correctly received packets are subtracted from the received signal vector and MMSE is performed again. If, say, $X_A^{(u+1)}$ was decoded correctly, a cancelled signal is constructed as:

$$\tilde{\mathcal{Y}}_1 = \mathcal{Y}_1 - \mathbf{H}_{eq,2} \cdot X_A^{(u+1)} \quad (4.12)$$

where $\mathbf{H}_{eq,2} = \left[\mathbf{O}_{2 \times 1} \quad \mathbf{H}_2^{(1)}(t+1) \quad \mathbf{O}_{6 \times 1} \right]^T$ is the second column of the channel equivalent matrix. The estimates are then obtained via the MMSE inversion:

$$\hat{\mathcal{X}}_1 = \mathcal{F}_{MMSE} \quad \tilde{\mathcal{Y}}_1 = \left(\mathbf{H}_{eq}^\dagger \mathbf{H}_{eq} + \frac{\sigma_n^2}{P} \right)^{-1} \mathbf{H}_{eq}^\dagger \tilde{\mathcal{Y}}_1 \quad (4.13)$$

If, at this step, $X_B^{(u+2)}$ was decoded without errors, a new SIC step could have been enforced for the last chance of detecting $X_B^{(1)}$:

$$\tilde{\tilde{\mathcal{Y}}}_1 = \tilde{\mathcal{Y}}_1 - \mathbf{H}_{eq,3} \cdot X_A^{(u+2)} \quad (4.14)$$

where the third column of the equivalent channel matrix is as:

$$\mathbf{H}_{eq,3} = \left[\mathbf{O}_{4 \times 1} \quad \mathbf{H}_2^{(1)}(t+2) \quad -\mathbf{H}_1^{(1)}(t+3) \quad \mathbf{H}_2^{(1)}(t+4) \right]^T$$

4.5.6 New User Selection

With the MU strategy, Hybrid SM mode is used and a new user is coupled with the previous scheduled user which needs a single packet retransmission. On one hand, the new user selection can be performed without any knowledge at the transmitter when the BS chooses randomly the new user to be scheduled. On the other hand, if every MT reports to the BS the post-processing SNRs at the outputs of the MMSE receiver for each stream, the BS can chose the best user in order to maximise the system throughput. Since the BEST user selection needs feedback of the post-processing SNR from each MT to the BS, this strategy is here proposed as a comparison.

The BEST user selection is reliable when the channel between successive time slots are highly time-correlated. The issue of the feedback rate given the variability of the MT CSI is not in the scope of this work and the best user selection is performed assuming that $\mathbf{H}^P(t+1) = \mathbf{H}(t)$, i.e. the channels in the future time slot are assumed equal to the channels in the current time slot and the time dependency is dropped in the following.

MMSE Receiver and Post-Processing SNR

In this section, the post-processing for the MMSE is expressed: the estimation error for the m^{th} symbol for the K^{th} packet directed to the user u can be expressed as:

$$\begin{aligned} e^{MMSE} &= \min[E(x_K^{(u)} - \hat{x}_K^{(u)})(x_K^{(u)} - \hat{x}_K^{(u)})^\dagger] \\ &= \sigma_n^2 \left(\mathbf{H}^{(u)\dagger} \mathbf{H}^{(u)} + \frac{\sigma_n^2}{P} \right)^{-1} \end{aligned} \quad (4.15)$$

where $\hat{x}_K^{(u)}$ is the MMSE estimation of x_K directed to the user (u), $K \in \{A, B\}$, σ_n^2 is the noise variance, P is the power transmitted for each stream and $\mathbf{H}^{(u)}$ is the equivalent channel matrix for user (u). The post-processing SNR for each stream i thus can be expressed as:

$$SNR_i^{MMSE} = \frac{P}{e^{MMSE}} = \frac{P}{\left[\sigma_n^2 \left(\mathbf{H}^{(u)\dagger} \mathbf{H}^{(u)} + \frac{\sigma_n^2}{P} \right)^{-1} \right]} \quad (4.16)$$

Expanding the matrix notation in (4.16) for the 2×2 MIMO case, the post-processing can be expressed as:

$$SNR_i^{MMSE} = \frac{P}{\sigma_n^2} \left[\left(\|\mathbf{H}_i^{(u)}\|^2 + \frac{\sigma_n^2}{P} \right) - \frac{|\mathbf{H}_i^{(u)\dagger} \mathbf{H}_{\underline{i}}^{(u)}|^2}{\|\mathbf{H}_{\underline{i}}\|^2 + \frac{\sigma_n^2}{P}} \right] \quad (4.17)$$

where the alternative antenna to the i^{th} is indicated with:

$$\underline{i} = 3 - i \quad \text{with} \quad i = \{1, 2\}$$

. The post-processing SNR for the i^{th} stream depends on $\mathbf{H}_i^{(u)}$, which is the channel from the two transmitting antenna to the i^{th} receiving one, but it is also degraded from the non perfect orthogonality between the $\mathbf{H}_i^{(u)}$ and $\mathbf{H}_{\underline{i}}^{(u)}$ as represented in the second subtractive term.

When the same modulation is used for both streams [dCP08], the best user selection for throughput maximization is to select the user with the best post-processing SNR. Each user calculates the two post-processing SNR at the output of the receiver (SNR_1^{MMSE} and SNR_2^{MMSE}) and sends this information back to the BS. The BS has all the information needed to optimize the multi-user scheduling, the New User (NU) is decided searching the N_{users} space:

$$NU = \max_u (SNR_{\underline{i},u}^{MMSE}) \quad (4.18)$$

where $SNR_{\underline{i},u}^{MMSE}$ is the post-processing SNR for user u for the antenna \underline{i} which will transmit the new packet. The best user selection is performed after the antenna switching criterion has been evaluated.

4.6 PO - Packet Overhearing

The performance of the system can be further improved with an optimized design of the SIC-MMSE receiver. When Hybrid SM transmission is used, one antenna retransmits a previous packet which was decoded incorrectly jointly with a packet directed to a new user. The probability that the retransmitted packet will be decoded (due to the packet combining at the

receiver) can be reasonably high. The probability of decoding the new packet, instead, will be lower since it is based on a single received signal collected at the new MT.

An optimized MT receiver with the ability of recording previous packets transmitted is described here. The packet storing starts when the MIMO-ARQ protocol is reset and a new user is scheduled in the system. This condition is notified by the BS. The MIMO-ARQ protocol can be reset when the two packets directed to the scheduled user have been decoded or when the maximum number of retransmission has been reached.

The PO receiver has an impact on the system performance due to a two-fold improvement. Firstly, the joint detection at the MMSE receiver of two (or more) received signals improves the probability of decoding the new packet sent in the hybrid SM transmission. Secondly, the SIC procedure of the PO receiver has more possibilities due to the larger set of estimates.

4.6.1 Received Signals at the Second Time Slot

Let's consider a transmission to MT_1 with packets $X_A^{(u)}$ and $X_B^{(u)}$. If MT_1 has decoded correctly both packets, the BS skips to the next user in the scheduling queue and user $(u + 1)$ deletes the overheard signals from its memory. If MT_1 decodes only one packet, say $X_A^{(1)}$, $X_B^{(1)}$ is retransmitted with switched antenna allocation (as described in Sec. 4.5) and a new packet $X_A^{(u+1)}$ is sent for the new user $(u + 1)$.

The aim of user $(u + 1)$ is to detect the packet $X_A^{(u+1)}$ exploiting all the radio signals collected. In the next time slot, two cases can occur at MT_1 :

1. MT_1 has decoded $X_B^{(1)}$ after one retransmission
2. MT_1 has decoded $X_B^{(1)}$ incorrectly.

In the first case, the hybrid SM transmission is terminated and the MU scheduler will compose two new packets for the next user in the scheduler queue. The fact that MT_{u+1} decodes or not its packet is not affecting the protocol state evolution; the protocol is reset anyway. In the second case where MT_1 has decoded the packet with errors, $X_B^{(1)}$ is retransmitted from the same antenna. If MT_{u+1} has decoded its packet, a new user will be scheduled in hybrid SM, otherwise a Multi-User STBC is started for packets $X_A^{(1)}$ and $X_B^{(u+1)}$.

In either cases, at the end of the second time slot transmission, user $(u + 1)$ has, so far, collected four radio signals. In the first time slot packets $X_A^{(1)}$ and $X_B^{(1)}$; in the second time slot the packet $X_B^{(1)}$ (retransmitted to user 1) and $X_A^{(u+1)}$ transmitted for the first time. At MT_{u+1} , each signal is transmitted on channel $\mathbf{H}^{(u+1)}$ as the following:

- $X_A^{(u)}$: for user 1 from antenna 1. Corresponding (noise free) received signal for the $m - th$ symbol: $\mathbf{H}_1^{(u+1)}(t)x_A^{(1)}$
- $X_B^{(u)}$: for user 1 from antenna 2. Corresponding (noise free) received signal for the $m - th$ symbol: $\mathbf{H}_2^{(u+1)}(t)x_B^{(1)}$
- $X_B^{(u)}$: for user u retransmitted from antenna 1. Corresponding (noise free) received signal for the $m - th$ symbol: $\mathbf{H}_1^{(u+1)}(t+1)x_B^{(u+1)}$

- $X_A^{(u+1)}$: for user $u + 1$ from antenna 2. Corresponding (noise free) received signal for the $m - th$ symbol: $\mathbf{H}_2^{(u+1)}(t+1)x_A^{(u+1)}$

The received signal at MT_{u+1} for the first time slot can be expressed as:

$$y_{u+1}(t) = \mathbf{H}_1^{(u+1)}(t)x_A^{(u)} + \mathbf{H}_2^{(u+1)}(t)x_B^{(u)} + \mathcal{N}^{(u+1)}(t) \quad (4.19)$$

where $\mathbf{H}_i^{(u+1)}(t)$ is the channel from transmitting antenna i to MT_{u+1} at time t . The received signal during the retransmission ($t+1$) is:

$$y_{u+1}(t+1) = \mathbf{H}_1^{(u+1)}(t+1)x_B^{(u)} + \mathbf{H}_2^{(u+1)}(t+1)x_A^{(u+1)} + \mathcal{N}^{(u+1)}(t+1) \quad (4.20)$$

The received signals can be composed for a single MMSE channel inversion to maximise the probability of correctly decoding packet $X_A^{(u+1)}$:

$$\begin{aligned} \mathcal{Y} &= \begin{bmatrix} \mathbf{H}_1^{(u+1)}(t) \\ \mathbf{O}_{2 \times 1} \end{bmatrix} x_A^{(u)} + \begin{bmatrix} \mathbf{H}_2^{(u+1)}(t) \\ \mathbf{H}_1^{(u+1)}(t+1) \end{bmatrix} x_B^{(u)} \\ &+ \begin{bmatrix} \mathbf{O}_{2 \times 1} \\ \mathbf{H}_2^{(u+1)}(t+1) \end{bmatrix} x_A^{(u+1)} + \begin{bmatrix} \mathcal{N}^{(u)}(t) \\ \mathcal{N}^{(u)}(t+1) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_1^{(u+1)}(t) & \mathbf{H}_2^{(u+1)}(t) & \mathbf{O}_{2 \times 1} \\ \mathbf{O}_{2 \times 1} & \mathbf{H}_1^{(u+1)}(t+1) & \mathbf{H}_2^{(u+1)}(t+1) \end{bmatrix} \begin{bmatrix} x_A^{(u)} \\ x_B^{(u)} \\ x_A^{(u+1)} \end{bmatrix} \\ &= \mathbf{H}_{4 \times 3} \cdot \mathcal{X} \end{aligned} \quad (4.21)$$

Then, the MMSE estimate is obtained via multiplication with \mathcal{F}_{MMSE} which is the pseudo-inverse of $\mathbf{H}_{4 \times 3}$:

$$\hat{\mathcal{X}} = \mathcal{F}_{MMSE} \quad \mathcal{Y} = \left(\mathbf{H}_{4 \times 3}^\dagger \mathbf{H}_{4 \times 3} + \frac{\sigma_n^2}{P} \right)^{-1} \mathbf{H}_{4 \times 3}^\dagger \quad \mathcal{Y} \quad (4.22)$$

After MMSE receiver, if packet $X_A^{(u+1)}$ is not received, there is a double SIC possibility: packet X_A^1 and/or packet X_B^1 can be subtracted if correctly decoded. If PO was not used, only packet X_A^1 could have been exploited for SIC.

4.6.2 Generalisation of PO for any time slot

The PO is possible at any step of the MIMO ARQ protocol operation: i.e. for any user chosen for hybrid SM transmission at any time slot. To exemplify this case, let's present the most general case where user $(u + 2)$ is selected for hybrid SM transmission at time slot $t + 2$.

Assume that packets $X_A^{(1)}$ and packets $X_B^{(1)}$ are transmitted in the first time slot for MT_1 and assume that packet $X_A^{(1)}$ has not been decoded. Packet $X_A^{(1)}$ is retransmitted with switched antenna. A first user $(u + 1)$ is scheduled in hybrid SM together with packet $X_A^{(1)}$. MT_{u+1} decoded correctly its packet.

At the successive time slot, user $(u + 2)$ is scheduled. MT_1 and MT_{u+2} are not able to decode the packets and a multi-user STBC starts. After an Alamouti coded block, both packets are decoded and the MIMO ARQ protocol can be finally reset without excess of retransmissions.

The received signals at MT_{u+2} , after reception of a full STBC block (4^{th} time slot), can be expressed as reported in (4.23).

$$\begin{aligned}
\mathcal{Y}_{u+2}(t+3) &= \begin{bmatrix} Y_{u+2}(t) \\ Y_{u+2}(t+1) \\ Y_{u+2}(t+2) \\ Y_{u+2}^\dagger(t+3) \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{H}_1^{(u+2)}(t) & \mathbf{H}_2^{(u+2)}(t) & \mathbf{O}_{2 \times 1} & \mathbf{O}_{2 \times 1} \\ \mathbf{O}_{2 \times 1} & \mathbf{H}_1^{(u+2)}(t+1) & \mathbf{H}_2^{(u+2)}(t+1) & \mathbf{O}_{2 \times 1} \\ \mathbf{O}_{2 \times 1} & \mathbf{H}_1^{(u+2)}(t+2) & \mathbf{O}_{2 \times 1} & \mathbf{H}_2^{(u+2)}(t+2) \\ \mathbf{O}_{2 \times 1} & -\mathbf{H}_2^\dagger(t+3) & \mathbf{O}_{2 \times 1} & \mathbf{H}_1^\dagger(t+3) \end{bmatrix} \\
&= \begin{bmatrix} x_A^{(1)} \\ x_B^{(1)} \\ x_A^{(u+1)} \\ x_A^{(u+2)} \end{bmatrix} + \mathcal{N}
\end{aligned} \tag{4.23}$$

4.7 Results

In this section, results are reported for different MIMO-ARQ protocol configurations including PO receivers. The protocol configurations are listed for clarity in Table 4.2 and include SU and MU modes. The performance of the proposed strategies are compared with a system with non-adaptive MIMO schemes and in the case when the schemes are selected based on a maximum BER, as described in Chap. 2.

Protocol configuration	Con-	SU / MU	User Selection
SU		SU as in [dCP08]	–
S1		SU	–
S2		SU	–
MU RAND		Hybrid MU	Random
MU BEST		Hybrid MU	Post-proc SNR
MU RAND PO		Hybrid MU PO	Random
MU BEST PO		Hybrid MU PO	Post-proc SNR

Table 4.2: MU MIMO-ARQ: configurations

The PHY layer is configured as reported in Table 4.3. The channel is simulated in block fading. The channel power is normalized to one. Where not otherwise specified, each fading block contains 250 packets, which means a channel coherence time of $\approx 108ms$. The total number of packets simulated is 25000.

Results are reported in terms of system throughput and HOL delays. The RR scheduler provides fairness among all the users and the results are presented averaged over the users.

4.7.1 Cross-layer MIMO-ARQ versus single layer adaptivity

In this section, the average rate and delay of the MU strategy are compared with three systems: fixed STBC, fixed SM and with a system where the

Parameter	Value
Modulation	4 QAM
Symbol Duration	$1\mu s$
Symbols per packet	432
Packet Duration	$T_p = 432\mu s$
Length of CRC	17 bit
Data bits per packet	847
Number of Users	10
Max Number of rx (K_{max})	7
Channel	Rayleigh fading
Total packets simulated per SNR	25000
Block Fading Length in packets	250

Table 4.3: System Parameters

two former schemes are selected based on a maximum BER at the PHY layer as in Chap. 2, [MMPR07].

The fixed STBC system transmits Alamouti blocks; the CRC is checked only after an even number of transmission slots have been received (i.e. after a full Alamouti block has been received). In the fixed SM system, if two packets are in error, they are retransmitted as are. If one packet is in error the packet is retransmitted from the same antenna and a new packet is scheduled on the other antenna.

The adaptive system switches from the fixed STBC to the fixed SM transmission when the SNR is high enough to support SM transmission with a BER lower than a given threshold. From STBC to SM, the spectral efficiency is doubled. Fig. 4.3 plots the BER curve for 4QAM Alamouti and 4QAM VBLAST schemes which is used for the determination of the cross-over point between SM and STBC. Two BER thresholds are chosen in the following: $BER \leq 0.01$ and $BER \leq 0.03$. In all the three systems (adaptive, fixed STBC and fixed SM), the receiver is a SIC-MMSE with Packet Combining.

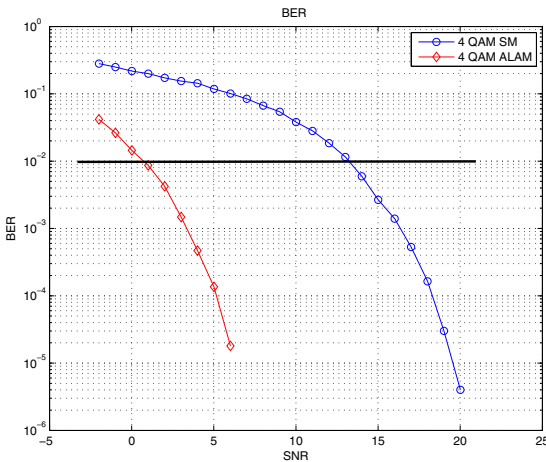


Figure 4.3: 4-QAM Alamouti and VBLAST BERs over SNR.

Fig. 4.4 reports the rate averaged over all the users for a fixed STBC and a fixed SM system, the rate for an adaptive scheme (with two BER thresholds) and the rate achieved by the MU strategy with PO in the case of BEST and RAND user selection. The adaptive system chooses the best scheme depending on the channel status and clearly outperform the fixed MIMO systems.

The adaptive system is outperformed by the proposed MU strategy: the faster adaptivity of the MIMO-ARQ state machine offers higher data rates and lower average delays due to added flexibility in the receiver and exploitation of the multi-user diversity.

As evident in Fig. 4.4, the low SNR region is dominated by the diversity gain of the STBC. The fixed STBC system data rate is converging to almost $200Kbps$ when $SNR \rightarrow \infty$, while fixed SM reaches $400Kbps$: as expected due to the double spectral efficiency. The different slopes of STBC and SM in the range of $[0, 5]dB$ represent the double and single diversity of the SM and STBC schemes respectively. When the BER is constrained to be lower than 0.01, the system switches to SM after $12.5dB$. In the case of 0.03 BER threshold, the system switches after $10dB$.

The MU strategy offers a better performance over all the SNR range. In the low SNR region, the gain of the MU in respect to the fixed STBC system comes from the possibility of receiving also halves of the STBC transmission blocks (as explained in Sec. 4.5.4) and from the fact that sometimes, in the case of particularly good CSI, SM can be used with an overall capacity gain. Around a $1dB$ gain can be observed when $SNR = 0dB$. MU strategy and fixed STBC offer same data rate when $5 \leq SNR \leq 7.5dB$: in this SNR region the higher probability of having a single packet in error during the first time slot pushes the MU strategy in longer retransmission states, compared to the fixed STBC system. After $7.5dB$, the MU strategy offers almost $5dB$ increase over the fixed SM scheme. The gain comes from exploitation of multiuser diversity and from the seamlessly usage of STBC and SM: in the high SNR region, still STBC can be seldom used when the users are occasionally in bad fading.

In Fig. 4.5 and Fig. 4.6, the average delays and the average delay jitters are plotted. In fixed STBC, the delay is converging to around $8ms$ against the $4ms$ of the fixed SM scheme because a full Alamouti block has to be received before decoding. The adaptive schemes present a delay which stabilises at $8ms$ for $SNR = 10dB$. After, the delay lowers to the level of the fixed SM system*.

The MU strategy sensibly improves the delay performance due to the Hybrid SM transmissions: when one packet is in error, it is still possible to chose another user and to send a packet to him. In the case of correct reception the new user chosen for Hybrid SM has a reduction of his average delay. This results in the fact that MU strategy approaches the $4ms$ average delay already at around $10dB$ with a considerably lower delay in the low SNR region. At $5dB$ the average delay is $2ms$ lower than in the STBC system, i.e. the users are served in average four time slots earlier than in the other three systems.

The delay jitter of the MU strategy is higher than in the fixed STBC strategy but it is lower-bounding the one of the fixed SM scheme. The higher delay jitter is due to the larger set of states available in the MIMO-

*The average delay is converging to $4.32ms$ when $SNR \rightarrow \infty$: indeed all the packets are transmitted in SM and successfully decoded. With RR scheduler, each MT receives 2 packets every $\frac{1}{N_{users} \cdot T_p}$ and the asymptotic average delay depends on the number of users in the scheduler queue. If other schedulers are integrated in the system, the asymptotic delay for each user may be different.

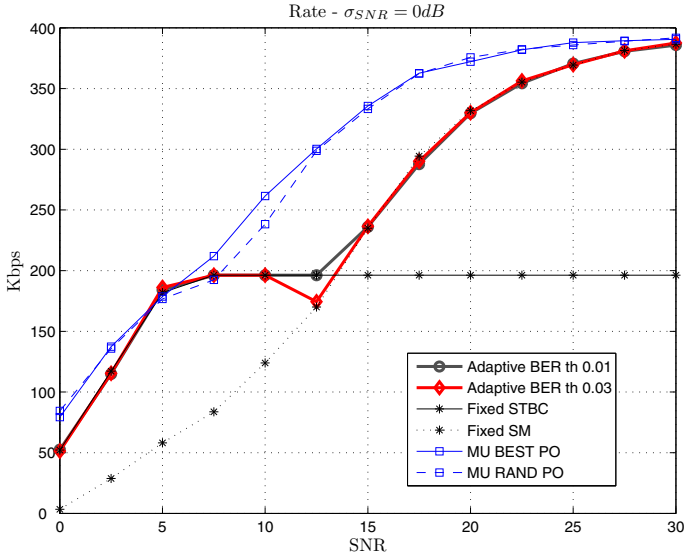


Figure 4.4: Average Rate for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems.

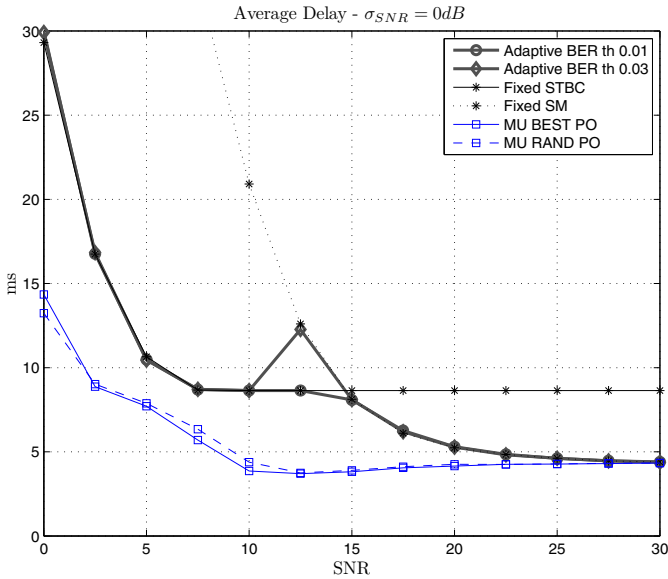


Figure 4.5: Average Delay for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems.

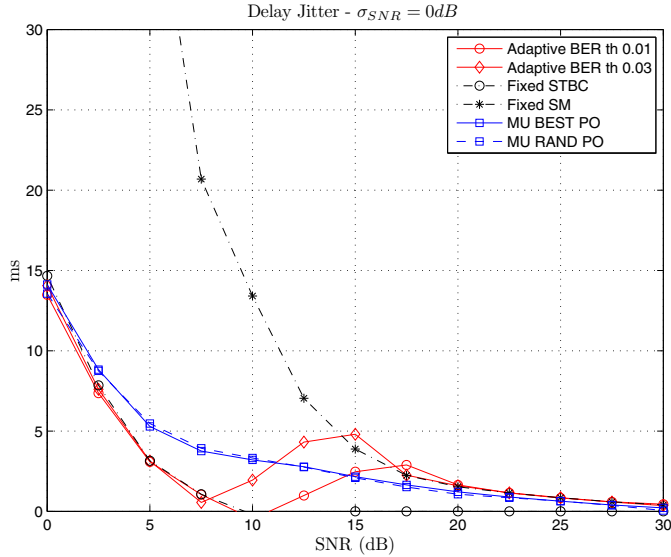


Figure 4.6: Delay Jitter for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems.

ARQ protocol compared to the unique retransmission possibility for the fixed STBC.

4.7.2 S1 - S2 - MU - MU PO Comparison

In this section, the various retransmission strategies, namely S1, S2 and MU, are compared when all the users have the same SNR.

Fig. 4.7 shows that SU strategy (as in [dCP08]) has a better performance in the low SNR region. S1,S2 and MU strategies outperforms the SU strategy in the medium to high SNR range where the SU strategy lacks to exploit the multi-user diversity.

SU is the best choice in the low SNR region because retransmitting to the same user implies an higher probability of packets in error and thus higher probability of using double diversity transmission (STBC). When STBC is used, the errors are quickly recovered and the system can send two new packets. When the STBC is less used and its contribution is vanishing ($SNR > 10dB$), SU performance quickly drops and S1,S2 and MU strategies show around $3dB$ gain[†]. SU strategy also shows an higher average delay as reported in Fig. 4.8.

The three strategies (S1-S2-MU) and the PO and non-PO receivers have comparable performance when all the user have the same SNR, with slightly better results for the MU strategy. When the user have different SNRs, as reported in Sec. 4.7.3, S1 and S2 will show sensible performance loss.

[†]The system studied in this work is 4-QAM uncoded; at $10db$ is $BER \approx 0.03$. A proper channel coding can move the cross-over point between SU and MU strategies at lower SNRs comparable with the typical operating point of the wireless systems.

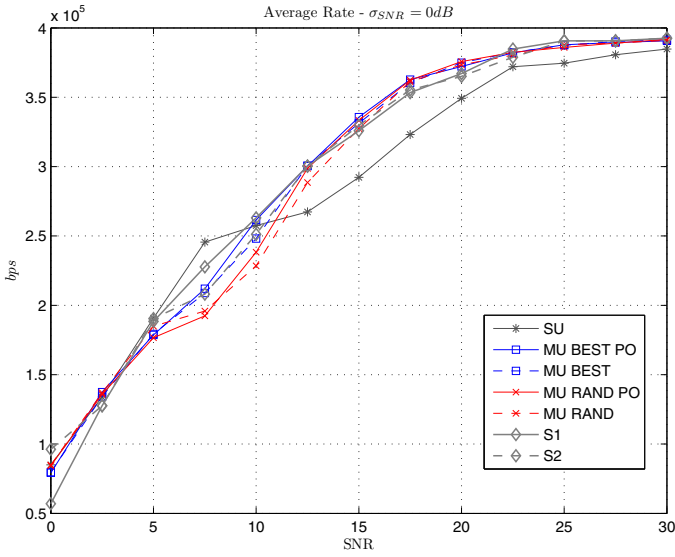


Figure 4.7: Average Rate for S1, S2 and MU retransmission strategies.

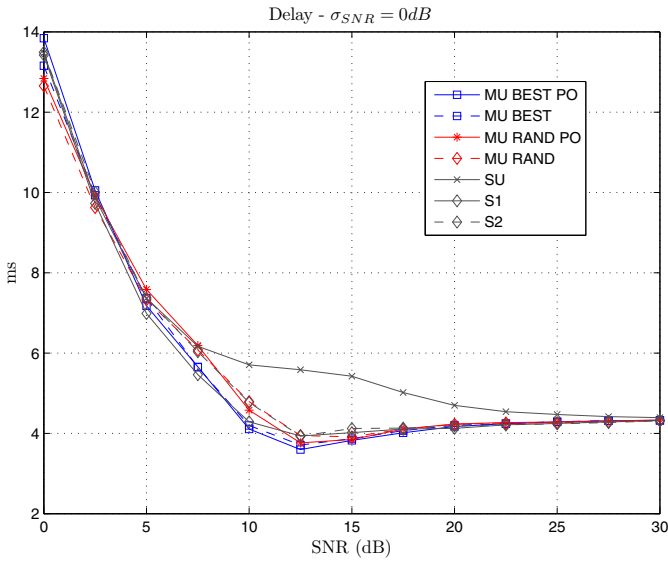


Figure 4.8: Average Delay for S1, S2 and MU retransmission strategies.

4.7.3 Performance with Log-normal SNR Distribution

The MIMO-ARQ protocol proposed is able to cope with heterogeneous SNR distribution of the users in the system. In Fig. 4.9 the average rate

of the MU MIMO-ARQ with random user selection is compared with fixed and adaptive systems when the SNR follows a lognormal distribution with a standard deviation of $\sigma_{SNR} = 2dB$. The adaptive system chooses the MIMO scheme based on the average SNR calculated over all the users. In Fig. 4.10 a tangible delay reduction can be observed compared to the adaptive systems. This comes from the Hybrid SM transmissions the and exploitation of MU diversity which has a bigger impact compared to the case of $\sigma_{SNR} = 0dB$. When $5 \leq SNR \leq 10dB$, a $5ms$ improvement is available, which means that in average users are served 10 packets earlier and the $4ms$ average delay is already reached at $SNR \approx 10dB$. The delay jitter is also improved, as shown in Fig. 4.11. Fixed STBC has, the lowest jitter but the jitter of MU strategy is still low compared to the fixed SM and the adaptive systems.

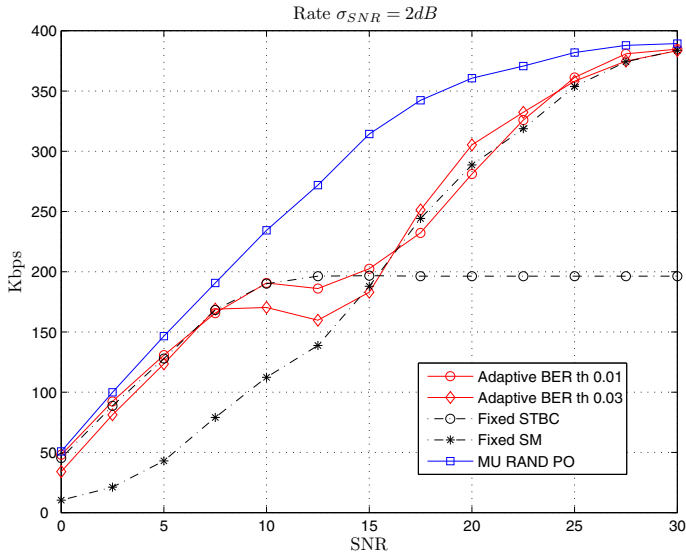


Figure 4.9: Average Rate for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems with 2dB lognormal SNR distribution.

Fig. 4.12 reports the average rate for all the strategies when $\sigma_{SNR} = 2dB$. MU strategy shows the best performance while S1 and S2 show a considerable reduction of performance of around $4dB$, starting from $12dB$. The MU with PO and BEST user selection also achieves slightly better rate compared to the RAND user selection due to optimal user selection for throughput maximisation.

4.7.4 Performance for Short Channel Coherence Times

The MU MIMO-ARQ protocol has been evaluated also for short channel coherence times. Two coherence periods of 250 packets ($\approx 100ms$) and of 5 packets ($\approx 4ms$) are compared. The performance of MU strategy with random user selection and SU (as in [dCP08]) are compared. The results is that MU strategy is more robust the fast changing channels. The average rate is reported in Fig. 4.13, where at $15dB$ for $4ms$ coherence time, the SU shows a loss of $2.5dB$ against the $0.5dB$ of the MU strategy. The reduction

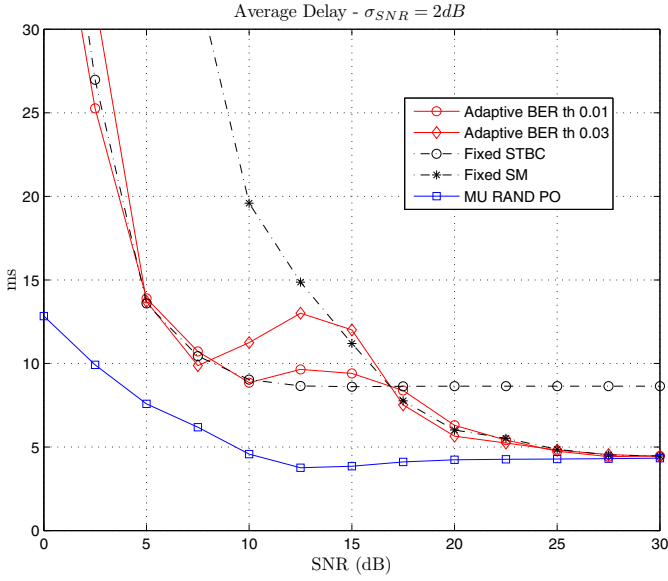


Figure 4.10: Average Delay for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems with 2dB lognormal SNR distribution.

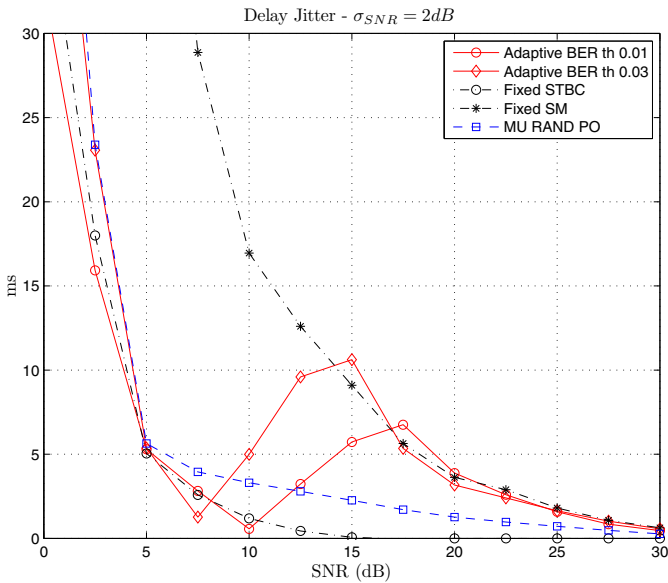


Figure 4.11: Delay Jitter for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems with 2dB lognormal SNR distribution.

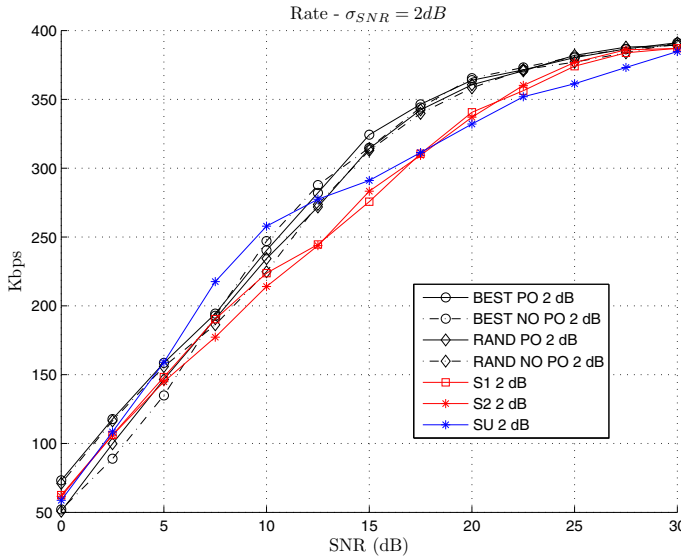


Figure 4.12: Average Rate for S1, S2 and MU retransmission strategies with 2 dB lognormal SNR distribution.

of average delay is maintain also with log-normal shadowing, as visible in Fig. 4.14.

4.8 Conclusion

A MIMO-ARQ protocol has been developed for a MU network with three retransmission strategies. An optimised receiver, able to store previous transmitted packets, has been developed (Packet Overhearing). The users are scheduled with a RR scheduler; the system proposed can be directly extended including any kind of known scheduler.

The work extends the results achieved in [dCP08], where retransmission strategy was designed for a SU link. The antenna allocation criterion was extended for MU networks. The MU MIMO-ARQ protocol can operate without any feedback (with the exception of the ACKs/NACKs) and avoids the one-bit feedback needed in [dCP08]. Three strategies were proposed: retransmission from single best antenna (which still needs one bit-feedback), retransmission from boths antenna in spatial repetition coding and MU strategy, which uses Hybrid SM transmissions.

In the last strategy, Multi-User diversity is exploited with a random (RAND) user selection and no feedback is needed at the BS. The performance is compared with the BEST user selection for throughput maximisation based on post-processing SNR. The performance degradation of the random user selection is small compared to the best user selection.

The MIMO-ARQ strategies for SU and MU are compared to wireless systems where fixed STBC/SM transmissions are used and to an adaptive system where the former two schemes are adaptively chosen to maintain the BER under a specified target. Better throughput and lower average delay are achieved over all the SNR range with a peak gain of 7dB.

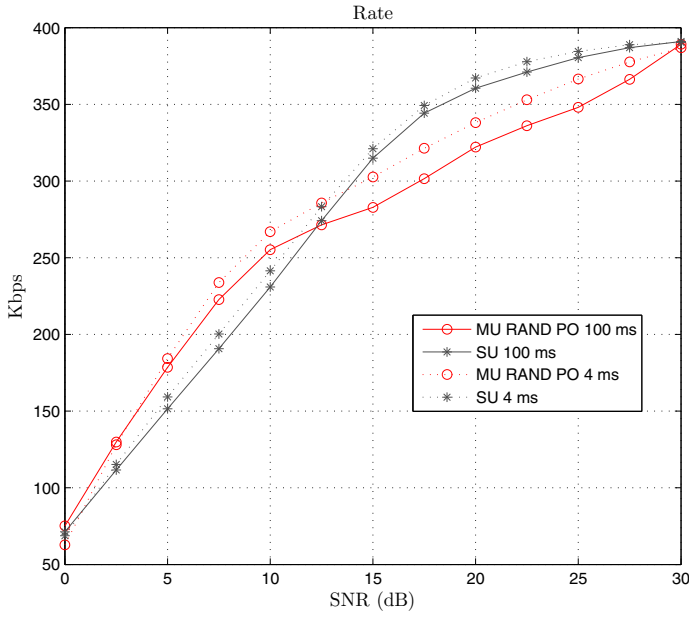


Figure 4.13: Average Rate for MIMO-ARQ in MU and SU modes for short channel coherence times.

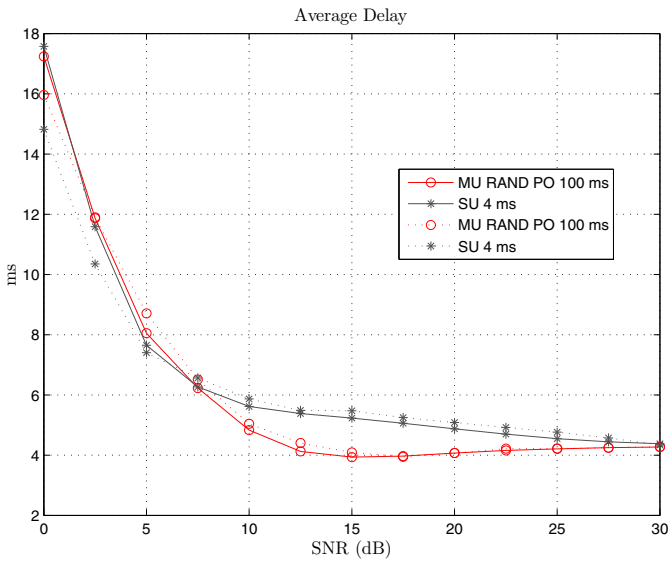


Figure 4.14: Average Delay for MIMO-ARQ in MU and SU modes for short channel coherence times.

The protocol proposed outperforms the one in [dCP08] for medium to high SNRs. At low SNR, strategy in [dCP08] achieves better throughput. A cross-over point between MU and SU strategies is found at $10dB$ for a 4-QAM uncoded modulation.

MU strategy outperforms the other strategy proposed (S1-S2), especially in the case of users with heterogeneous SNR. In log-normal SNR distribution, the MU gain over the SU strategy is more tangible due to the increasing impact of multi-user diversity.

The protocol proposed has been designed to maintain a good performance with fast time varying channels (open-loop MIMO schemes are seamlessly used with no CSI feedback). The MIMO-ARQ protocol proposed in this work can adapt rapidly to the channel status with no need of CSI feedback while exploiting Multi-User diversity.

Chapter 5

MIMO Protocol Gain System Design Issues

The mental imagery involved with pianistic tactilia is not related to the striking of individual keys but rather to the rites of passage between notes.

REVIEW OF PAYZANT
Glenn Gould (1932-1982)

The first communication network that was designed during the primitive years of information theory was the telephone network constituted by a set of point-to-point links with fixed allocated resource. When a call is originated, a path is reserved and the communication happens with no restrictions inside the assigned bandwidth.

The origin of packet-switched networks and the concept of packet payload itself was born with the need of interconnecting computing machines. The burstiness of voice calls travelling over analog connection was, by far, much more reduced compared to that of the messages originated by the machines. Kleinrock was among the firsts to address this problem, formulating the concept of the backbone elements of a packet-switched networking, which originated the well-known queueing theory, later published in [Kle75].

During the late 1960s and early 1970s, several computer scientists and engineers proposed the formulation of the seven-layer Open System Interconnection (OSI) framework. At that time, the separation between the problem of physical data transmission and link management was very useful. The idea of layers is even found in the work by Shannon, [Sha48]. Shannon clearly stated that the digital channel that he has been studying was a layer above the underlying physical channel where the data was transmitted in analog form. Shannon also considered the process of channel coding as residing a layer below the process of data compression.

However, since some years, there is a strong revisionist feeling with respect to the notion of layering. The research community has fully re-

alised that the strict layering of a packet-switched networks leads to under-exploitation of system resource. Nonetheless, the layered framework offers a stable basis for technology development and standardisation.

The constraints introduced by the protocol stack are sometime clear and less evident in some other cases. It is evident that the OSI layering was not designed for multiaccess mediums, such as the wireless mediums. The optimal exploitation of multiaccess networks requires a joint design of physical, link and network mechanisms. When multiple antenna or sectorized cells are available, the spatial dimension of the channel can be exploited and Spatial Division Multiple Access (SDMA) can offer high gains when designed with a cross-layer approach.

The field of multiuser detection theory represents a link between multiaccess information theory and network multiaccess and there has been considerable activity on signal modelling, power-control, antenna patterns, channel interference models, and receiver structures.

The provisioning of a channel for end to end transmission can be divided in several subtasks and traditionally each one of this tasks is accomplished in a layer of the stack. The PHY layer provides a bit tunnel adapted to the physical medium used for data transmission. The organisation of bits in packets is in charge of the Data Link Control (DLC) layer which provides reliable transmission service to the NET layer. The route from a source to a destination is provided by the NET layer, which manages a set of ordered links, each one with its own DLC and PHY processing.

The OSI layering was designed for wired-networks, and the stack needed to be modified to add the MAC layer to support communication over multiaccess channel, as ALOHA protocol in [Abr70] just to cite a famous example. The MAC resides between the DLC and the PHY and basically provides an intermittent data connection to the DLC layer.

The MAC layer can be totally unaware of the underlying PHY or some information can be shared. This is the example of CQI signalling between PHY and MAC for channel-aware schedulers. Other example is the joint design of ARQ parameters based on CSI, as described in Sec. 3.2 at page 52.

An interaction among PHY and MAC layers has been proposed in Chap. 4 when multiple antenna are available. In particular, the MAC is aware of a per-antenna ACK and several retransmission strategies have been designed for aggressive resource exploitation in multi-user networks with fast mobile terminals without the need for feedback.

Introducing control of the ARQ process at the antenna elements level has some impact on the network architecture. This impact will not require substantial modification to the current approach, however, the key architectural ideas are outlined in this Chapter. In the current PHY/MAC layer division, it is not possible to be aware at the MAC layer of the per-antenna channel conditions. The designs described in Chapter 4 are analysed from the signalling and architectural perspective.

The description of the current architecture refers to the last approved version of the IEEE802.16 standard, [IEE09a]. Two topologies are defined for 802.16 networks: Mesh and PMP. Only the latter is considered in this Chapter.

The PMP traffic management, basic packet processing operations and the ARQ process are described. The ARQ state machine is controlled by the ACKs sent back by the MTs. ACKs (NACKs) are calculated based on the CRC appended to the MAC PDUs, as visible in the diagram in Fig. 5.1. A modified interface between PHY and MAC is then presented (see Fig. 5.11), in order to support the SU and MU MIMO-ARQ protocols of Chapt. 4.

An overview of the schedulers presented in literature for IEEE802.16 networks is also proposed. This will clarify the possibilities offered by the integration of MU MIMO-ARQ system with more advanced schedulers compared to the RR scheduling assumed in the previous chapter.

5.1 General overview of IEEE802.16 Point-to-MultiPoint

In IEEE802.16, the BS is usually a central point with a sectorized antenna system capable of handling multiple independent sectors simultaneously. Inside a given frequency channel and antenna sector, all MTs receive the same transmissions in DL. In UL, the BS collects all the signals coming from the MTs previously associated and which are inside the same sector. Only one BS is operating in the sector and does not have to coordinate with other transmitting stations. For DL, the DL-MAP specifies the addressee (MT) for each portion of the time-frequency resources. In the case where the DL-MAP does not explicitly indicate that a portion of the DL subframe is for a specific MT, all MTs will listen. The MTs decode the connection identifiers (CIDs) in the received MAC PDUs and store only those PDUs addressed to them. Messages can be sent to a single MT (unicast) or to multiple MTs (multicast connections). The latter is used for control messages and broadcast messages to all stations in the sector/cell.

In the UL subframe, all the MTs share the UL channels on a demand basis. The right to transmit depends on the class of service flow which will access the transmission bandwidth: the MT may be issued continuing rights to transmit, or the right to transmit may be granted by the BS after receipt of a request from the user.

MTs adhere to a transmission protocol that controls contentions and enables the service scheduling to be adjusted to the QoS requirements in terms of delay and bandwidth. There are three different types of UL scheduling mechanisms:

- Unsolicited bandwidth grants
- Periodic polling
- Contention based

These three UL scheduling mechanisms allow vendors to optimise the system performance by using combinations of all the scheduling techniques, maintaining interoperability. For Unsolicited grants, the service flows has a fixed amount of bandwidth reserved independently from the load of the network and the status of the data queues. The use of polling guarantees that service flows receive resources on a deterministic basis when required. Contrary to data applications, which are delay tolerant, real-time applications, like voice and video, require service with very tightly-controlled schedule for preserving delay-related QoS (average delay and jitters) and need to have a constant check of the bandwidth requests (periodic polling). Finally, contention is used to avoid individual polling of MTs that have been inactive for a long period of time. The MT can send a request to the BS for bandwidth allocation in contention with the others MTs in the sector/cell in a specified part of the UL frame.

5.1.1 Packet Processing Workflow

The workflow for the packets to be transmitted over the wireless segment follows the scheme depicted in Fig. 5.1, where a typical IEEE802.16 protocol stack is reported with the ARQ state machine in evidence. From the

Application layer, packets are received at the air-interface. Following the IEEE802 nomenclature, each layer receives packets to be dispatched from upper layer (Service Data Unit (SDU)s) and transfer to the lower layer fixed size packets to be transmitted (PDUs).

From the Networking layer, the NET PDU are passed to the MAC layer. The MAC layer is divided into three sublayers. The MAC Convergence Sublayer divides the traffic based on QoS attributes and provide transparent access to the wireless transmission from the transport perspective.

The MAC privacy sublayer provides support for identification and association to the network and for enforcement of security protocols. The security aspect is not treated in this thesis. A work has been published proposing a PHY subcarrier permutation in [MWP09b]. The PHY layer security encapsulates upper layer protocols for an overall augmented security.

The MAC Sublayer is the central entity which covers the aspects related to the medium access. At this sublayer, NET PDU are packed or fragmented (depending on the size of the packets to be transmitted) in MAC PDUs which are later transferred to the PHY layer. The ARQ process is managed at this level of the protocol stack. Each MAC SDU, after being fragmented or packed, is appended with a CRC which is used to detect error during transmission on the radio channel. Fragmented/packed MAC SDUs are passed to the PHY layer. Each packet is then mapped in the OFDMA grid. Data randomization, modulation, subcarrier permutation and MIMO scheme are applied. Prior to be sent at the RF transmitting chains, the frame is composed along with data symbols, pilots and guard bands. The DL and UL subframes are separated by appropriate timings for permitting a correct ramp-up and ramp-down of the power amplifiers. The frame is left-padded with a Frame Preamble which is used for synchronisation and coarse channel estimation purposes.

On the receiver side, the signal is processed with the appropriate MIMO receiver. Demodulation, decoding and depermutation is performed. A single PHY SDU is passed to the upper MAC layer for the CRC check. The result of this check (ACK/NACK) is then sent back to the BS. The ARQ state machine is fully contained at the MAC layer of the BS. If a NACK is received, a retransmission of packet happens until the packet is correctly received or until the maximum number of retransmissions have been reached.

In the followings, some details of the protocol stack will be deeper analysed.

5.1.2 MAC Layer

Each air-interface of a MT (each MT can have more then one air-interface) has a 48-bit universal MAC address, as defined in the standard IEEE802, [IEE89]. The MAC address uniquely defines the air interface of the MT in the network and it is issued by the manufacturer. The MAC address is used for the initial ranging and association processes and to establish the connections between the BS and the MT. During the authentication process, the MAC address is also used to verify the identity of BS and MT.

IEEE802.16 MAC is connection-oriented: all data communications are in the context of a transport identified with a CID. Upon the registration of an MT (or the modification of the services already negotiated), higher layers of the BS and of the MT are stimulated to initiate the setup of the service flows. Each service flow is associated with a transport connection

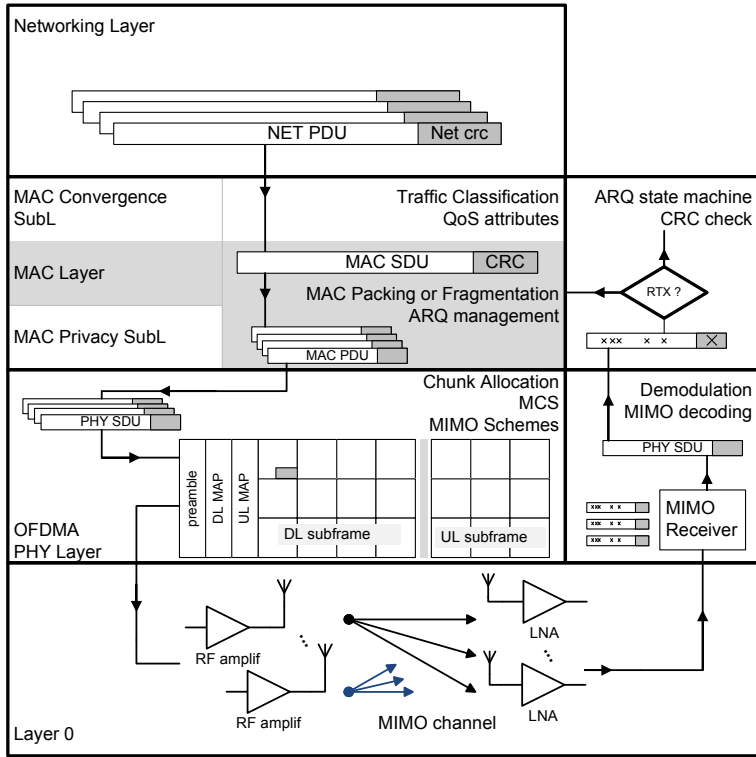


Figure 5.1: IEEE802.16 Protocol Stack for classical ARQ at MAC layer.

and each MT can have more than one connection; each service flow is identified by a unique 16-bit CID. Bandwidth requests are always referred to a single CID. New transport connections are established for new traffic flows or when a MT changes the QoS attributes of an already established service flow. Indeed, the service flow contains all the QoS attributes for the PDUs that are exchanged on the MAC connection. Bandwidth is always granted by the BS to an MT (also as an aggregate of grants) in response to per-connection requests from the MT.

In the initialisation phase, two pairs of management connections, basic connections (UL and DL) and primary management connections (UL and DL), are established between the MT and the BS*. The management connection pairs are used for three different levels of QoS of the management traffic. The basic connection is used to exchange short and urgent MAC management messages. The primary management connection is used to transport longer but more delay-tolerant MAC management messages. In [IEE09a], Table 38 specifies on which connection pair MAC messages have to be transferred, based on messages length and urgency. The table also specifies which MAC management messages are transported on the broadcast connection. An optional secondary management connection is used to transfer delay-tolerant messages of specific

*Optionally, a third pair of management connections (secondary management, DL and UL) can be generated.

standards, such as: Dynamic Host Control Protocol (DHCP), File Transfer Protocol (FTP) (only trivial messages) and Simple Network Management Protocol (SNMP). The secondary management connection allows for packet packing or fragmentation. When OFDM and OFDMA physical layers are used, all the management messages must have an appended CRC.

Once established, transport connections for variable-rate data flows may require maintenance. For example, in the case of an unchannelized T1[†] service, virtually no connection maintenance is needed, since a constant bandwidth is allocated. IP services usually require a considerable amount of maintenance due to their bursty nature and the possibility of fragmentation. Transport connections may also be terminated: the termination of a transport connection is stimulated by the BS or by the MT. The transport connection management functions are implemented through the use of dynamic addition, deletion and modification of service flows, as described in Sec. 3.3 at page 63.

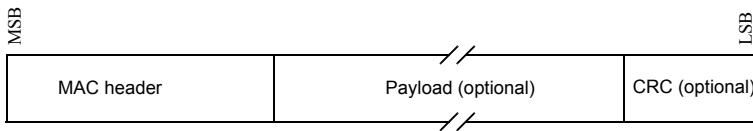


Figure 5.2: MAC PDU content in 802.16e.

The format of a MAC PDU is reported in Fig. 5.2. The first part of the PDU is the General Mac Header (GMH) which is exploded in detail in Fig. 5.3. The GMH contains several data fields, e.g.: the PDU length in bytes, the CID and the Header Check Sequence (HCS) for detecting error in the GMH. For a detailed description of all the fields of the GMH, the reader can refer to [IEE09a], Sec. 6.3.2.1.1. In the case of fragmentation,

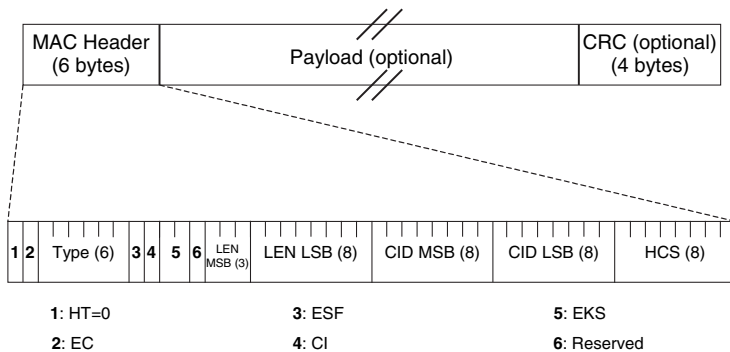


Figure 5.3: Generic MAC Header Content.

each fragment of the SDU is encapsulated into one or more PDUs, as

[†]The historical definition of T1 service is related to a digital transmission line with 1.544 Mbps data rate equivalent to 24 voice channels used in telephone network; an unchannelized T1 service defines a service flow with 1.5 Mbps bandwidth.

illustrated in Fig. 5.4. If SDUs are very small, more than one of them can also be encapsulated into a single PDU (packing), as reported in Fig. 5.5. Internal Packing Sub-Header (PSH) are inserted.

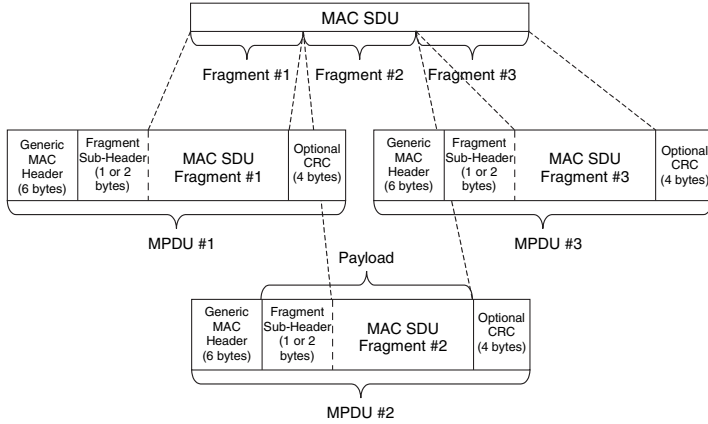


Figure 5.4: Fragmentation Procedure at the MAC layer. Reproduced from [Nua07].

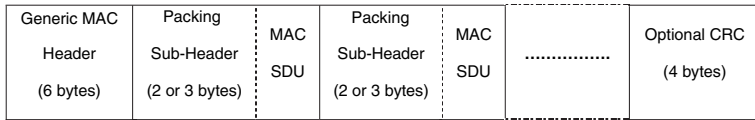


Figure 5.5: Packed SDU into a single PDU

5.1.3 ARQ

The ARQ mechanism is part of the MAC layer. ARQ is optional for implementation. If implemented, ARQ may be enabled on a each single transport connection. The per-connection ARQ is specified and negotiated during the creation of the connection and is limited to unidirectional traffic. A single connection cannot have ARQ and non-ARQ traffic mixed together.

For ARQ-enabled connections, it is possible to activate fragmentation of MAC SDUs: each SDU can be split into fragments for separate transmission. If fragmentation is not enabled, each PDU must contain all the data related to the parent SDU. The size of each MAC PDU is defined by the *ARQ_BLOCK_SIZE* parameter. An ARQ block is a distinct unit of data that is carried on an ARQ-enabled connection. The block size is a parameter negotiated during connection establishment. If the length of the SDU is not an integer multiple of the *ARQ_BLOCK_SIZE* parameter, the last fragment is padded with zero payload until the correct length has been achieved. The ARQ block size needs to be accurately set to avoid losses of network efficiency. Each PDU may contain ARQ blocks that are transmitted for the first time as well as those being retransmitted. Every ARQ

block has an assigned Block Sequence Number (BSN) and is managed as a distinct entity by the ARQ state machines.

A diagram of the ARQ state machine, which is located at the BS, is reported in Fig. 5.6. Once a MAC SDU has been scheduled for transmission and MAC PDUs have been generated after eventual fragmentation/packing, the PDUs are sent to the PHY layer and the state machine moves to the "Outstanding" state, where it waits for the first response (NACK or ACK). If an ACK is received, the state machine is reset. If NACK is received, the state machine moves to the state "Retransmission". The PDU is then transmitted again and the machine waits again in "Outstanding" for a NACK or an ACK from the MT.

When a NACK is received or a maximum Waiting-Time is elapsed the state machine retransmit the packet. The maximum Waiting-Time is specified by $ARQ_RETRY_TIMEOUT^{\ddagger}$. This parameter is fundamental to avoid starving of the state machine for two cases: 1) the PDU has not been received at the MT, 2) the packet was received at the MT and the MT transmitted the NACK/ACK which was not received at the BS. Finally, from the two intermediate states presented before, the state machine can interrupt the protocol operation when an ACK is received or when a maximum amount of time is elapsed for the ARQ process, specified by $ARQ_BLOCK_LIFETIME$. This specification is equivalent to specifying the maximum number of ARQ retransmission, as done for Stop and Wait ARQ in Sec. 3.2 at page 52 and for MU MIMO-ARQ in Sec. 4.4 at page 87. It is to be noted that an ACK signal also link the "Discarded" to the "Done" state. This is to include the case where a positive ACK is received slightly after the $ARQ_BLOCK_LIFETIME$ was over.

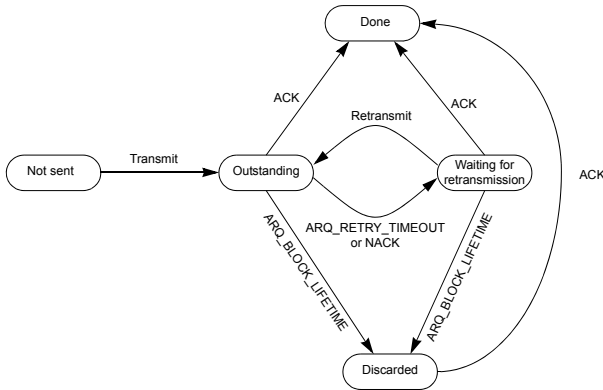


Figure 5.6: ARQ TX state machine. Reproduced from [IEE09a].

A system supporting ARQ must be able to receive and process the ARQ feedback messages. The ARQ feedback information is sent back to the BS and cannot be fragmented. ARQ feedbacks can be implemented in two forms:

- on standalone MAC management connection (in UL)
- via piggyback on an existing UL data connection

The ARQ feedbacks always refer to the ARQ BSN. Several ARQ feedback types are possible:

[‡]The definition of $ARQ_RETRY_TIMEOUT$ is specified in [?] in Sec. 6.3.4.3

An advanced features of the ARQ at the MAC layer, is the rearrangement of MAC PDUs, as shown in Fig. 5.8.

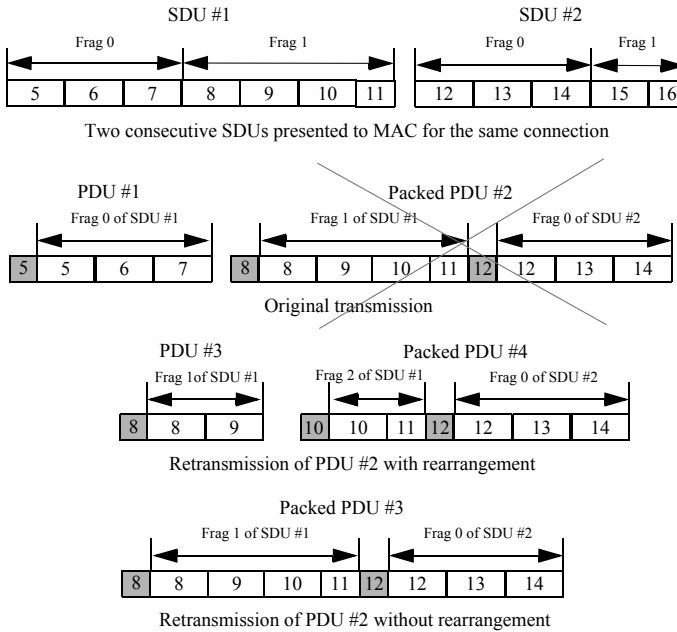


Figure 5.8: Packet rearrangement for ARQ in 802.16 2009. Reproduced from [IEE09a].

An SDU is partitioned into a set of blocks prior entering the ARQ process and the partitioning remains in effect until the SDU is discarded by the transmitter state machine (if maximum retransmission time is elapsed) or if all blocks of the SDU are successfully delivered to the receiver.

Part of ARQ blocks, selected for transmission or retransmission, are encapsulated into a PDU. When a PDU is composed, the sequence of blocks immediately between MAC subheaders must have contiguous block sequence numbers. However, if ARQ arrangement is available, the fragments of MAC SDUs can be rearranged on MAC PDUs in order to save bandwidth during retransmissions. This requires added processing at the MAC layer for both ends of the connection.

5.2 Scheduling in IEEE802.16

The Scheduling service represents the mechanisms for data packets handling on a single connection. Each connection can be associated only with a single scheduling service. QoS parameters define a scheduling service (or service class). IEEE802.16 wireless networks defined several QoS classes. These classes are used to guarantee the required rate and delay performance for heterogeneous applications, such as: data, voice and video traffic flows.

IEEE802.16 standard does not include a definition for the scheduler implementation and leaves freedom to the vendors for differentiation. Several schedulers have been proposed in literature with channel-aware and

channel-unaware techniques which needs to interwork with PHY layer to obtain CSI information. Each scheduler is characterised by the following tasks:

- guarantees of the QoS
- maximises the throughput
- optimises the resource utilisation
- minimises the average delays and delay jitters.

All these features are to be achieved with the smallest computational complexity. Also, scalability is an important characteristic of the scheduler, i.e. the scheduler has to be complexity-limited and viable when an high number of connections is to be managed. The task is far from being easy, since the radio channel has an high degree of variability (especially with fast moving MTs).

5.2.1 BS and MT Schedulers

In IEEE802.16 with PMP configuration, the BS is equipped with a sectorised antenna and can manage independently several sectors. Inside a single sector, the scheduler controls the admission of service flows and the MAC SDUs scheduling. It is also a role of the BS scheduler to control the adaptive modulation and coding of each flow and to enforce power control.

The MCS profile is decided by the BS based on several parameters:

- SIR
- Radio channel quality
- Network load
- Status of the packet queues.

The scheduler has to monitor constantly the received SIR values for each associated MT and determine the bandwidth to be allocated considering the service QoS attributes. Fig. 5.9 reports a simple diagram summarising the inputs and outputs parameters of the scheduler.

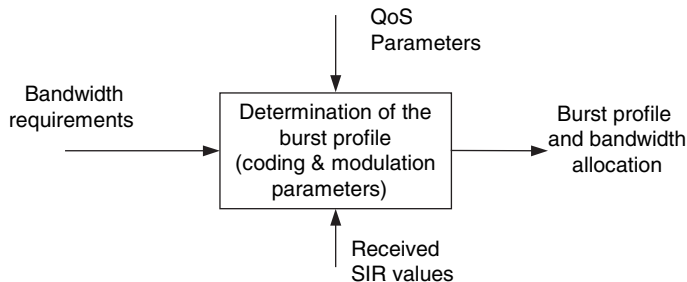


Figure 5.9: Input Parameters for the scheduler.

The BS scheduler decides for the DL and UL resource allocation. The DL allocation is broadcasted in the DL-MAP with MAC management messages. On the UL, based on the UL requests and taking into account QoS

parameters and scheduling services priorities, the BS scheduler decides an UL resource allocation which is broadcasted in the UL-MAP. There is also a scheduler in the MT which have the role to classify all the incoming packets into the MT different UL connections.

The BS and MT schedulers are reported in the diagram of Fig. 5.10. For each CID, a queue is set up in the scheduler memory space. At the BS, two schedulers manage the DL and UL connections separately. DL SDUs from the upper layers are placed in the queues after eventual packing/fragmentation and CRC appending operations. Instead, received UL PDU are directly transferred from the PHY to the MAC functions for de-fragmentation/de-packing, decoding and CRC checks. The UL scheduler at the BS is informed of the Bandwidth Request sent by the MTs in the previous UL subframes.

On the MT side, an UL scheduler is also present: UL PDUs generated at the MT are placed in the scheduler queues. After having received from the BS the UL bandwidth allocation grants in the previous DL subframes, the MT scheduler transfers the packets to be transmitted from the scheduler queues to the PHY layer. Received DL PDUs are directly transferred from the PHY to the MAC layer.

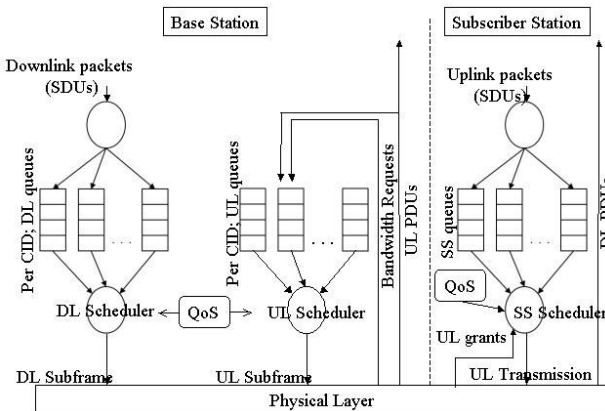


Figure 5.10: DL and UL Schedulers in IEEE802.16, from [SIJT09].

IEEE802.16 has not specified any packet scheduling or admission control mechanisms and several algorithms have been proposed in literature.

Schedulers for WiMAX can be classified into two main categories. The schedulers which do not have any information from the PHY layer regarding the CSI of the admitted users (channel-unaware schedulers) and the schedulers which accounts for the actual radio channel qualities (channel-aware schedulers). The channel-unaware schedulers do not any information of the channel state condition in making the scheduling decision and generally assume error-free channel. In wireless networks there is a high variability of radio links (signal attenuation, noise, fading, interference) and the channel-awareness is important. An overview of the impact of channel awareness on the scheduling process can be found in [BBT⁺07].

Scheduling mechanisms have a big impact on network performance and a consistent amount of research is still ongoing; some recent works are herein cited. Several scheduling algorithms have been proposed such as Fair Scheduling [XYW09, TSZS09, KSH⁺09, LBS99, LMP07], Distributed Fair

QoS Classes	Maximum traffic rate	Minimum traffic rate	Tolerated Jitter	Maximum Delay	Traffic Priority
BE	✓				✓
nrtPS	✓	✓			✓
rtPS	✓	✓		✓	
ertPS	✓	✓		✓	
UGS	✓		✓	✓	

Table 5.1: IEEE802.16 mandatory QoS classes and attributes.

Scheduling [CYK08], MaxMin Fair Scheduling [TS05, TS02], Customised Deficit Round Robin, [LAC09].

Another family of schedulers is based on traffic priorities calculated from the HOL delay and the averaged sustained data rate in respect to the QoS delay and rate requirements, as in [WL09, MMAP08].

5.2.2 QoS Classes

The IEEE802.16-2009 standard MAC Layer is able to provide QoS differentiation for the various types of applications through five types of scheduling services, or QoS classes. This classification facilitates bandwidth sharing among different MTs. According to the QoS classes and the radio channel quality, the BS scheduler allocates the necessary amount of bandwidth required for each connection: e.g. a real-time application, such as a video or voice over IP application, will be prioritised compared to FTP or email or web traffic[§].

In IEEE802.16d standard, four scheduling services were defined:

- Unsolicited Grant Service (UGS)
- real-time Polling Service (rtPS)
- non-real-time Polling Service (nrtPS)
- Best Effort (BE).

Each of these classes has a mandatory set of attributes (QoS parameters) that are included in the service flow definition. The mandatory service flow QoS parameters for each of the four scheduling services is defined in [IEE04].

The UGS QoS class is designed to support real-time data with fixed-size data packets to be transmitted at periodic intervals, e.g. Voice over IP without silence suppression. In this case, the BS provides fixed-size data grants at periodic intervals, eliminating the overhead due to MT grant requests.

The nrtPS is designed for delay-tolerant data streams with variable-size data packets. For nrtPS a minimum data rate is defined. This is the situation, for example, of an FTP connection. The BS polls nrtPS connections typically on an interval on the order of one second or less.

The rtPS scheduling class is used for delay sensitive service flows (i.e. video streaming) and it is designed to support real-time data streams with variable-sized data packets issued at periodic intervals. In this service, the BS provides periodic unicast (UL) request opportunities.

[§]This feature is not included, for example, in Wi-Fi WLAN networks where all services have the same level of QoS.

The BE service is transmitted on a best available basis. This class is designed to support data streams for which no minimum guarantees are required. When the network is congested, a long period can run without a BE connection receives any packet.

Another class, extended rtPS (ertPS), was added by the 802.16e amendment. The aim of ertPS is to reduce overhead for grant requests in rtPS: the BS provides unicast grants in an unsolicited manner like in UGS, thus saving the latency of a bandwidth request. Compared to UGS allocations, which have a fixed-allocation resource assigned, ertPS are dynamic providing variable bandwidths. The ertPS is suitable for variable rate real-time applications with rate and delay requirements, as VoIP with silence suppression. Tab. 5.1 reports the traffic attributes for each QoS service class.

5.3 MIMO-ARQ protocol stack

The MIMO-ARQ system proposed in Chap. 4 needs to be aware of a per-antenna ACK in order to implement the retransmissions strategies.

A modified stack is proposed here and illustrated in Fig. 5.11. As every cross-layer technique, an enriched interface is needed between the PHY and MAC layer. The added signalling allows the ARQ state machine to be aware of MIMO transmission with a per-antenna ARQ process. Compared to the classical stack, as in Fig. 5.1 at page 119, more data pipes are available on the PHY/MAC boundaries, one for each antenna element. The resource allocation at a chunk level is still defined at the PHY layer, but the antenna allocation is informed by the MAC layer MIMO-ARQ state machine. The MIMO coding (STBC or SM) is decided at the MAC layer. The PHY layer places the multiple MAC PDUs received to the transmitting antenna as instructed by the MAC which follows one of the strategies presented in Chap. 4. The PHY layer modulates the payloads accordingly to the MIMO-ARQ transmission strategy.

At the receiver side, the MIMO receiver at the PHY layer processes the packets and sends several PHY SDUs to the upper layer for CRC checking (one for each receiving antenna). The CRC is checked for each packet and multiple ACKs/NACKs are sent back to the BS. The MIMO-ARQ state machine resides at the MAC level at the BS but it is aware of ACK/NACK for each antenna as sent back by the MT (or MTs in the case of hybrid SM transmission).

At the MT, a Packet Combining (PC) repository is available, where all the packets previously transmitted are stored; the packets correctly received are cancelled from the received signals, as explained in Sec. 4.5.5 at page 99.

The MU MIMO-ARQ protocol proposed in the previous chapter is designed to be readily integrated with advanced packet schedulers. In particular, as shown in Fig. 4.1 at page 91, two levers are available.

The performance gains shown in the previous chapter were obtained considering a simple RR scheduler and random user selection for Hybrid SM transmission and the BEST user selection was solely analysed as a performance upper-bound of the previous un-informed user selection method.

Despite the perfect fairness provided, the RR scheduler lacks to exploit the network resource with heterogeneous traffic flows. A more advanced scheduler could enhance the average data rate via multi-user opportunistic scheduling.

An added opportunity is offered by the user selection method in Hybrid SM transmission. Instead of being chosen randomly, the user could be chosen based on some traffic parameters measured at the packet queue. For instance, if a user is approaching its HOL maximum delay, the scheduler

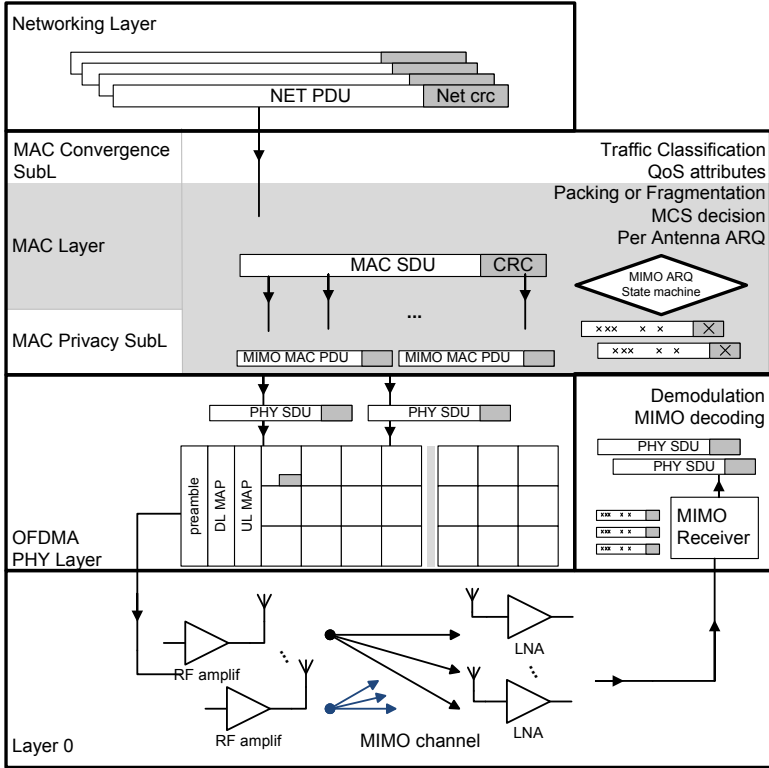


Figure 5.11: Modified Protocol Stack with per-antenna ARQ.

could promptly chose the user for the next Hybrid SM transmission thus giving an opportunity of receiving a packet without a deep impact on the current transmissions ongoing.

Various possibilities are clearly available and more then one urgency level could be defined. A "medium" urgency could be treated with the scheduling for the next hybrid transmission while a "severe" urgency could be addressed scheduling the user directly in STBC, interrupting the current transmission and thus providing the highest reception probability.

The scheduler has an enlarged set of actions compared to a classical channel-aware scheduler. Indeed, the MIMO-ARQ scheduler can decide to:

- schedule the user in SM after the next MIMO-ARQ state machine termination
- schedule the user in STBC after the next MIMO-ARQ state machine termination
- schedule the user for the next Hybrid SM transmission
- schedule the user in SM interrupting the current state machine evolution
- schedule the user in STBC interrupting the current state machine evolution

The previous decision could be taken with all or some or none of the followings: CSI, sustained data rate, average delay or average delay jitter measured at the MAC scheduler queue.

The added degree of complexity should be evaluated with memory and processing power constrains, particularly evaluating the scaling issues with the number of MTs admitted in the system.

Conclusions

The ongoing process of globalization requires interoperability of the wireless networks deployed around the world. The next generation wireless systems shall fulfil the specifications published by the International Communication Union (ITU) defining the International Mobile Telecommunications Advanced (IMT-A) system which guarantees world-wide and intra-system interoperability.

IMT-A Wireless networks shall be able to provide high Spectral Efficiency (SE) radio links towards Mobile Terminals (MTs) moving with velocities as high as 300Km/h . High transmission rates shall be offered to stationary and mobile users supporting the convergence of data, voice, and video transmissions providing the "Mobile Internet", i.e. a connection over secure IP with the same performance of a cable-connected computer but on the move. The convergence of heterogeneous services imposes strict requirements on the Quality of Service (QoS) that must be ensured to users' transmission. In wireless networks it is extremely difficult to guarantee such level of QoS requirements, especially due to the intermittent quality degradation that radio channel may experience, greatly exacerbated by the high user mobility.

In order to achieve the requested QoS, strategies that counteract the channel quality variations in an *adaptive* way are, thus, necessary. Such adaptive strategies can be designed at each layer of the stack separately. The choice of the layer at which the adaptive strategy is implemented affects the performance of the wireless network. In particular, strategies implemented at the upper layers (e.g., Internet Protocol (IP) layer) are too slow to counteract the channel variations of vehicular radio channels. In addition, if different disjoint strategies are implemented at different layers, lower layers (e.g., Physical (PHY) layer) compensate the channel quality degradation very fast and the upper layers may strive to achieve the QoS requirements and stabilise their adaptive algorithms.

The classical OSI protocol stack has been a fundamental basis for a coherent and stable technology development of practically viable communication systems. However, it is undoubtedly recognised that the strict division among Physical (PHY) and Medium Access Control (MAC) layers is unsuitable for multiple access channels which is the typical situation encountered in wireless cellular systems. The limitations are especially demanding when aggressive performance is to be achieved and multiple antennas are used. Indeed, Multiple Input Multiple Output (MIMO) enlarges the signalling dimensions and enriches the possibilities of access to the channel when the antenna elements are used for traffic directed to different users in Multiple User (MU) networks.

For the previous reasons, the adaptive strategies must feature a cross-

layer approach to optimally achieve the QoS requirements and promptly compensate the channel quality variations. In a MU MIMO network, a cross-layer approach is even more advocated since it can greatly improve the performance exploiting the gain coming from Multi-User diversity.

The work presented in this thesis stemmed from several analysis conducted separately at the PHY and MAC layers. After the characterisation of PHY signalling and MAC protocols, a set of cross-layer designs have been proposed integrating the features of each layer and advancing the state of the art of the strategies previously proposed for each layer.

In particular, a MIMO-ARQ protocol, with several retransmission strategies, have been proposed which integrates MIMO open-loop schemes jointly with ARQ process and packet scheduling. The MIMO-ARQ protocol showed increased network throughput, reduced average delay and reduced delay jitters over all the SNR range compared to fixed MIMO schemes and to adaptive MIMO schemes applied at the PHY layer, independently from upper layers.

Final Considerations

The work led to the multiple results and considerations that are summarised here.

First, at the PHY layer, a dynamic MIMO scheme selection has been designed with an upper-bounded BER. The system switches among several PHY profiles using Space Time Block Code (STBC), Layered STBC (LSTBC) and Spatial Multiplexing (SM) schemes in order to provide the highest SE maintaining the BER under a specified target. When no multiplexing gains are required, the quasi-orthogonal schemes should be avoided in order to have better throughput. Data throughput and spatial multiplexing gain can be traded-off based on channel quality. The performance of this adaptive PHY layer is able to boost the PHY layer performance but cannot address and mitigate the problem of low MIMO channel ranks due to low scattering. This adaptive strategy is later compared with the cross-layer MIMO-ARQ protocol. The impact of a dynamic PHY layer is to be accounted in the design of the protocol stack since the MAC layer has to adapt promptly to the variations of SE at the MAC/PHY interface.

Then, dynamic allocation of STBCs and Space-Frequency Block Coding (SFBC)s has been studied for an OFDMA PHY layer in order to preserve the performance of Space-Time Codes (STCs). Based on delay spread and Doppler frequency measurements, which can be easily obtained in OFDM Multicarrier systems as recently shown in literature, the diversity order of the STC scheme is preserved allocating the symbols along the direction (time or frequency) which shows the highest correlation. The correlation coefficients have been derived under the Wide Sense Stationary - Uncorrelated Scattering (WSSUS) channel assumption. It is shown that at velocities over 120Km/h , comparable correlation is measured in the time and frequency domains. This is a fundamental indication for the design of high velocity radio links with STCs, they shall adapt to the time-frequency characteristics of the MIMO channel. The proposed STBC/SFBC adaptive system needs a single bit feedback. The ratio between the spatial coherence of the channel delay spread and the MT velocity constancy period gives the required feedback rate. Such lightweight feedback is an important design feature to implement this solution with fast MTs.

At the MAC, two protocols have been analysed with the queuing theory: Automatic Repeat Request (ARQ) and Dynamic Service Addition (DSA). The number of ARQ retransmissions has been optimised with the selection of one of several Single Input Single Output (SISO) PHY profiles.

The proposed Adaptive Cross Layer (ACL) strategy jointly and promptly optimises the PHY and MAC parameters with the objective of guaranteeing the QoS requirements while allocating the minimum amount of bandwidth necessary for transmissions. Even without the use of MIMO, QoS requirements are met on a larger SNR range compared to the non-cross-layer adaptation and the ACL is able to counterfight the channel shadowing very efficiently. Eighth times less bandwidth is needed to transmit the same amount of data when the channel shows a log-normal shadowing with a standard deviation of $2dB$. A counter-intuitive result has been observed: channels with high SNR deviation show lower average delays: mobility can impact positively on the average transport delay if an ACL strategy is implemented with the ARQ protocol.

DSA protocol have been analysed. When harsh radio propagation is experienced, the DSA protocol can lead to high signalling blocking probabilities, thus preventing service flow to enter the network and reducing the resources utilisation factor. A closed-form signalling blocking, admission control blocking and latency have been derived, also for the case of the activation of the service flow with an un-informed MT. The theoretical model has been fully validated for a block fading channel. The impact of mobility, network load, DL/UL reciprocity, OFDM PHY layer parameters – such as the DFT size – have been studied over fast channels and the impact on the blocking probabilities and latency has been determined for slow and fast MTs.

Finally, an interaction among PHY and MAC layers has been designed to exploit Multi-User diversity with open-loop MIMO schemes and ARQ. In particular, the MAC is aware of a per-antenna ACK and several retransmission strategies have been proposed for aggressive network exploitation without the need of any CSI feedback. The feedback is two-fold avoided by design: 1) for the optimal antenna selection for ARQ retransmissions and 2) for the user selection in Hybrid SM transmission mode: the user is chosen randomly. The random user selection show small performance loss compared to the case where the best user is selected for network throughput maximisation.

For MIMO-ARQ protocol, both Single User (SU) and MU networks are analysed. A cross-over point for 4-QAM uncoded modulation has been found at $10dB$ between SU and MU protocols. When the SNR is higher, the MU MIMO-ARQ leads to more than $3dB$ improvement over the SU strategy previously proposed in literature. This gain increases with the log-normal deviation of the users' SNR distribution. The MU MIMO-ARQ protocol has shown very good performance even with channel coherence times as low as $4ms$: much lower than the minimum delay requirement for the data plane specified for IMT-A system in the case of fully unloaded network.

Further, the cross-layer MU MIMO-ARQ protocol had shown a throughput gain of $7dB$ over the single layer adaptive strategy based on upper-bounded BER previously discussed. Compared to single layer adaptive strategies, the average packet delay of the MU MIMO-ARQ protocol is noticeably lower over all the SNR range.

The strength of the cross-layer approach and the gains attained with the MU MIMO-ARQ protocol have been shown and compared to other adaptive strategies previously studied. The results are supported by several theoretical analysis which have been conducted for the followings:

- closed-form calculation of the average Symbol Error Probability (SEP) at the PHY layer

- ARQ and DSA signalling blocking probability
- Average Delay of the ARQ protocol
- DSA admission control blocking probability
- DSA service flow activation probability with un-informed Mobile Terminal (MT)
- DSA service flow activation latency.

The SEP calculations at the PHY layer were conducted using alternative forms of the error and squared error-functions plugged with the Moment Generating Function (MGF) of the SNR distribution at the output of the MIMO receivers.

The system design issues have been discussed inside the IEEE802.16 framework (WiMAX), referring to the last version of the standard published in 2009. However, the theoretical approach followed in the proposed designs and the several common features shared by WiMAX and LTE-A make the results achieved in this thesis valid and applicable in a very broad context.

Lady Presenter: *Well, that's the end of the film. Now, here's the meaning of life. Well, it's nothing very special. Uh, try and be nice to people, avoid eating fat, read a good book every now and then, get some walking in, and try and live together in peace and harmony with people of all creeds and nations.*

MONTY PYTHON
The Meaning of Life (1983)

Bibliography

- [AB00] A. Annamalai and V. K. Bhargava, "Simple and efficient techniques to implement a self-reconfigurable arqsystem in a slowly varying mobile radio environment," *Wirel. Pers. Commun.*, vol. 13, no. 1-2, pp. 97–117, 2000. [cited at p. 53]
- [Abr70] N. Abramson, "The aloha system: another alternative for computer communications," in *AFIPS '70 (Fall): Proceedings of the November 17-19, 1970, fall joint computer conference*. New York, NY, USA: ACM, 1970, pp. 281–285. [cited at p. 116]
- [Al-05] Al-Ghadhban, S. Buehrer, R.M. Woerner, B.D., "Outage capacity comparison of multi-layered STBC and V-BLAST systems," in *Proc. IEEE Vehicular Technology Conference, VTC Fall, 2005*. [cited at p. 83, 88]
- [Ala98] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct 1998. [cited at p. 12, 14, 15, 22, 44, 84, 85, 93]
- [BBCS06] A. Bennatan, D. Burshtein, G. Caire, and S. Shamai, "Superposition coding for side-information channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1872–1889, May 2006. [cited at p. 82]
- [BBP03] H. Bolcskei, M. Borgmann, and A. Paulraj, "Space-frequency coded MIMO-OFDM with variable multiplexing-diversity tradeoff," in *Proc. IEEE International Conference on Communications ICC '03*, vol. 4, 11–15 May 2003, pp. 2837–2841. [cited at p. 37]
- [BBT⁺07] L. Badia, A. Baiocchi, A. Todini, S. Merlin, S. Pupolin, A. Zanella, and M. Zorzi, "On the impact of physical layer awareness on scheduling and resource allocation in broadband multicellular ieee 802.16 systems [radio resource management and protocol engineering for ieee 802.16]," *Wireless Communications, IEEE*, vol. 14, no. 1, pp. 36–43, Feb. 2007. [cited at p. 126]
- [BdF08] J. Borin and N. da Fonseca, "Scheduler for IEEE 802.16 networks," *IEEE Communications Letters*, vol. 12, no. 4, pp. 274–276, April 2008. [cited at p. 63]
- [Bel63] P. Bello, "Characterization of Randomly Time-Variant Linear Channels," *IEEE Transactions on Communications Systems*, vol. 11, no. 4, pp. 360–393, December 1963. [cited at p. 38, 40]
- [Bou02] Boutros, J. and Caire, G. and Viterbo, E. and Sawaya, H. and Vialle, S., "Turbo code at 0.03 dB from capacity limit," in *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on, 2002*, pp. 56–. [cited at p. 1]

- [BSY08] X. Bai, A. Shami, and Y. Ye, "Robust QoS control for single carrier PMP mode IEEE 802.16 systems," *IEEE Transactions on Mobile Computing*, vol. 7, no. 4, pp. 416–429, April 2008. [cited at p. 63]
- [Che97] K. Cheun, "Performance of direct-sequence spread-spectrum RAKE receivers with random spreading sequences," *IEEE Trans. Commun.*, vol. 45, no. 9, pp. 1130–1143, Sep. 1997. [cited at p. 14]
- [CLME06] C. Cicconetti, L. Lenzini, E. Mingozzi, and C. Eklund, "Quality of service support in IEEE 802.16 networks," *IEEE Network*, vol. 20, no. 2, pp. 50–55, March-April 2006. [cited at p. 63]
- [CMCP08] I. Cerutti, F. Meucci, P. Castoldi, and L. Pierucci, "An adaptive Cross-Layer strategy for QoS-guaranteed links in 4G networks," in *IEEE Globecom 2008 Wireless Networking Symposium*, New Orleans, LA, USA, 11 2008. [cited at p. 53, 63, 172]
- [Cos83] M. Costa, "Writing on dirty paper (corresp.)," *Information Theory, IEEE Transactions on*, vol. 29, no. 3, pp. 439–441, May 1983. [cited at p. 6, 82]
- [Cra91] J. Craig, "A new, simple and exact result for calculating the probability of error for two-dimensional signal constellations," in *Military Communications Conference, 1991. MILCOM '91, Conference Record, 'Military Communications in a Changing World'.*, IEEE, Nov 1991, pp. 571–575 vol.2. [cited at p. 151, 152]
- [CVC⁺09] I. Cerutti, L. Valcarenghi, P. Castoldi, D. Marabissi, F. Meucci, L. Pierucci, E. Del Re, L. S. Ronga, R. Tka, and F. Kamoun, *Current Technology Developments of WiMax Systems*. Springer, 2009, ch. Enhanced operation modes in IEEE 802.16 and integration with optical MANs. [cited at p. 172]
- [CVM⁺07] I. Cerutti, L. Valcarenghi, D. Marabissi, F. Meucci, L. Pierucci, L. S. Ronga, P. Castoldi, and E. Del Re, "Enhanced Operation Modes in IEEE 802.16 and Integration with Optical MANs," in *Proceedings of the 12th Wireless Communications CNIT Thyrrrenian Symposium*, 2007. [cited at p. 53, 172]
- [CYK08] J. Choi, J. Yoo, and C. K. Kim, "A Distributed Fair Scheduling Scheme With a New Analysis Model in IEEE 802.11 Wireless LANs," *IEEE Trans. Veh. Technol.*, vol. 57, no. 5, pp. 3083–3093, Sept. 2008. [cited at p. 127]
- [DBC93] P. Dent, G. Bottomley, and T. Croft, "Jakes fading model revisited," *Electronics Letters*, vol. 29, no. 13, pp. 1162–1163, 24 June 1993. [cited at p. 41]
- [dCP06] E. de Carvalho and P. Popovski, "ARQ strategies for 2 X 2 spatially multiplexed MIMO systems," in *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, 29 2006–Nov. 1 2006, pp. 1666–1670. [cited at p. 83]
- [dCP08] —, "Strategies for ARQ in 2X2 MIMO Systems," *IEEE Commun. Lett.*, vol. 12, no. 6, pp. 441–443, Jun. 2008. [cited at p. 7, 8, 83, 84, 85, 87, 88, 93, 100, 103, 107, 109, 111, 113]
- [DL08] A. D. Dabbagh and D. J. Love, "Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1859–1868, Nov. 2008. [cited at p. 6, 82]
- [DR97] E. Del Re, *Elements of Digital Signal Processing, Elementi di elaborazione numerica dei segnali*. Pitagora Editrice, 1997, ISBN 88-371-0964-4. [cited at p. 39]

- [EBS07] B. M. E. Biglieri, A. Goldsmith and H. Sari, "MIMO link adaptation in mobile WiMAX systems," in *Proc. WCNC 2007*, Hong Kong, Mar. 2007. [cited at p. 22]
- [EF08] S.-E. Elayoubi and B. Fourestie, "Performance evaluation of admission control and adaptive modulation in OFDMA WiMax systems," *IEEE/ACM Trans. on Networking*, vol. 16, no. 5, Oct. 2008. [cited at p. 63]
- [EGC80] A. El Gamal and T. M. Cover, "Multiple user information theory," *Proc. IEEE*, vol. 68, no. 12, pp. 1466–1483, Dec. 1980. [cited at p. 82]
- [Emr99] Emre Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999. [cited at p. 4, 167]
- [FG98] G. Foschini and M. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–355, 1998. [cited at p. 4]
- [Fos96] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Laboratories Technical Journal*, pp. 41–59, 1996. [cited at p. 1, 5, 12, 22, 23]
- [FWC+06] I.-K. Fu, W. C. Wong, D. Chen, P. Wang, M. Hart, and S. Vadgama, "Path-loss and Shadow Fading Models for IEEE 802.16j Relay Task Group," IEEE 802.16 Relay Task Group, Document C80216j-06 045r1, Tech. Rep., July 2006. [cited at p. 58]
- [GAY+02] M. J. Gans, N. Amitay, Y. S. Yeh, H. Xu, T. C. Damen, R. A. Valenzuela, T. Sizer, R. Storz, D. Taylor, W. M. MacDonald, C. Tran, and A. Adamiecki, "Outdoor BLAST measurement system at 2.44 GHz: calibration and initial results," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 570–583, April 2002. [cited at p. 5]
- [GGF94] A. J. Goldsmith, L. J. Greenstein, and G. J. Foschini, "Error statistics of real-time power measurements in cellular channels with multipath and shadowing," *IEEE Trans. Veh. Technol.*, vol. 43, no. 3, pp. 439–446, Part 1–2, Aug. 1994. [cited at p. 58]
- [GH85] D. Gross and C. M. Harris, *Fundamentals of queueing theory (2nd ed.)*. New York, NY, USA: John Wiley & Sons, Inc., 1985. [cited at p. 65, 67]
- [GJJV03] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity Limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003. [cited at p. 2]
- [GKH+07] D. Gesbert, M. Kountouris, R. W. Heath, C. byoung Chae, and T. Sizer, "From Single user to Multiuser Communications: Shifting the MIMO paradigm," in *IEEE Sig. Proc. Magazine*, 2007. [cited at p. 6, 160]
- [GR05] S. Gowrisankar and B. Rajan, "A rate-one full-diversity low-complexity space-time-frequency block code (STFBC) for 4-Tx MIMO-OFDM," in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*, Sept. 2005, pp. 2090–2094. [cited at p. 37]
- [GSK05] M. Guillaud, D. Slock, and R. Knopp, "A practical method for wireless channel reciprocity exploitation through relative calibration," in *Proc. Signal Processing and Its Applications*, vol. 1, 2005, pp. 403–406. [cited at p. 73]

- [Gud91] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronics Letters*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991. [cited at p. 58]
- [HCK09] D. Hwang, B. Clerckx, and G. Kim, "Regularized channel inversion with quantized feedback in down-link multiuser channels," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 12, pp. 5785–5789, December 2009. [cited at p. 82]
- [HP05] R. Heath and A. Paulraj, "Switching between diversity and multiplexing in MIMO systems," *IEEE Transactions on Communications*, vol. 53, no. 6, pp. 962–968, June 2005. [cited at p. 7, 27, 29, 83, 85, 88]
- [HTW03] A. Hottinen, O. Tirkkonen, and R. Wichman, *Multi-antenna Transceiver Techniques for 3G and Beyond*. John Wiley & Sons, Inc., 2003. [cited at p. 23]
- [IEE89] "Information processing systems - Local area networks - Part 2: logical link control," *ISO Std 8802-2: 1998; IEEE Std 802.2-1998*, pp. –, Dec 1989. [cited at p. 51, 118]
- [IEE04] IEEE Std 802.16-2004 (Revision of IEEE Std 802.16-2001), "IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems," 2004. [cited at p. 19, 20, 22, 24, 29, 53, 57, 127]
- [IEE06] IEEE Std 802.16e-2005 and IEEE Std 802.16-2004 Cor 1-2005 (Amendment and Corrigendum to IEEE Std 802.16-2004), "IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1," 2006. [cited at p. 19, 20, 21, 29, 64]
- [IEE09a] "IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems," *IEEE Std 802.16-2009 (Revision of IEEE Std 802.16-2004)*, pp. C1–2004, 29 2009. [cited at p. 65, 116, 119, 120, 122, 124]
- [IEE09b] IEEE 802.16 Broadband Wireless Access Working Group. (2009, January) IEEE 802.16m Evaluation Methodology Document, IEEE 802.16m-08/004r5. IEEE. [Online]. Available: http://ieee802.org/16/tgm/docs/80216m-08_004r5.doc.zip [cited at p. 164]
- [IEE09c] ——. (2009, September) IEEE 802.16m System Description Document (SDD) IEEE 802.16m-09/0034r2. IEEE. [Online]. Available: http://ieee802.org/16/tgm/docs/80216m-09_0034r2.zip [cited at p. 164]
- [IEE09d] ——. (2009, September) IEEE 802.16m System Requirements Document (SRD) IEEE 802.16m-07/002r9. IEEE. [Online]. Available: http://ieee802.org/16/tgm/docs/80216m-07_002r9.pdf [cited at p. 164]
- [ITU08] ITU-R M.2134, "Requirements related to technical performance for IMT-Advanced radio interface(s)," ITU-R, Tech. Rep., 2008. [Online]. Available: http://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2134-2008-PDF-E.pdf [cited at p. xviii, 83, 88, 157, 160]
- [Jaf01] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 1–4, Jan 2001. [cited at p. 16]
- [JPWY08] Z. Jiang, W. Pan, D. Wang, and X. You, "Study on Capacity of BLAST System with Error Propagation," in *Proc. 4th IEEE International Conference on Circuits and Systems for Communications ICCSC 2008*, 26–28 May 2008, pp. 387–391. [cited at p. 18]

- [JRSKR04] D. Jagannadha Rao, V. Shashidhar, Z. Khan, and B. Rajan, "Low-complexity, full-diversity space-time-frequency block codes for MIMO-OFDM," in *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, vol. 1, Nov.-3 Dec. 2004, pp. 204–208 Vol.1. [cited at p. 37]
- [K⁺07] P. Kysti *et al.*, "WINNER II Channel Models, D1.1.2," IST-4-027756 WINNER II, Tech. Rep., 2007. [cited at p. 28]
- [KG05] M. Kountouris and D. Gesbert, "Memory-based opportunistic multi-user beamforming," in *Proc. International Symposium on Information Theory ISIT 2005*, Sep. 4–9, 2005, pp. 1426–1430. [cited at p. 7, 83]
- [KHGT05] I.-M. Kim, S.-C. Hong, S. S. Ghassemzadeh, and V. Tarokh, "Opportunistic beamforming based on multiple weighting vectors," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2683–2687, Nov. 2005. [cited at p. 7, 83]
- [KK08] T.-S. Kang and H.-M. Kim, "Optimal Beam Subset and User Selection for Orthogonal Random Beamforming," *IEEE Commun. Lett.*, vol. 12, no. 9, pp. 636–638, Sep. 2008. [cited at p. 6, 82]
- [Kle75] L. Kleinrock, *Queueing Systems*. Wiley-Interscience, 1975. [cited at p. 57, 115]
- [KR02] M. Z. A. Khan and B. S. Rajan, "Space-time block codes from co-ordinate interleaved orthogonal designs," in *Proc. IEEE International Symposium on Information Theory*, 2002, p. 275. [cited at p. 37]
- [KSH⁺09] M. Kang, Y. J. Sang, H. G. Hwang, H. Y. Lee, and K. S. Kim, "Performance Analysis of Proportional Fair Scheduling with Partial Feedback Information for Multiuser Multicarrier Systems," in *Proc. IEEE 69th Vehicular Technology Conference VTC Spring 2009*, 26–29 April 2009, pp. 1–5. [cited at p. 126]
- [Ky08] P. Kysti, *Matlab SW Documentation of WIM2 Model*, WINNER, Wireless World Initiative New Radio, www.ist-winner.org, September 2008. [cited at p. 28]
- [KYL03] Z. Kang, K. Yao, and F. Lorenzelli, "Nakagami-m fading modeling in the frequency domain for OFDM system analysis," *IEEE Communications Letters*, vol. 7, no. 10, pp. 484–486, Oct. 2003. [cited at p. 65]
- [LAC09] E. Laias, I. Awan, and P. Chan, "Fair and latency aware uplink scheduler in ieee 802.16 using customized deficit round robin," in *Advanced Information Networking and Applications Workshops, 2009. WAINA '09. International Conference on*, May 2009, pp. 425–432. [cited at p. 127]
- [Lar07a] P. Larsson, "Analysis of Multi-User ARQ with Multiple Unicast Flows Under Non-IID Reception Probabilities," in *Proc. IEEE Wireless Communications and Networking Conference*, March 2007, pp. 384–388. [cited at p. 83]
- [Lar07b] —, "Analysis of Multi-User ARQ with Multiple Unicast Flows Under Non-IID Reception Probabilities," in *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE*, March 2007, pp. 384–388. [cited at p. 87]
- [LBS99] S. Lu, V. Bharghavan, and R. Srikant, "Fair scheduling in wireless packet networks," *IEEE/ACM Trans. Netw.*, vol. 7, no. 4, pp. 473–489, Aug. 1999. [cited at p. 126]
- [LC01] Y. Li and J. Cimini, L.J., "Bounds on the interchannel interference of OFDM in time-varying impairments," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 401–404, 2001. [cited at p. 64]

- [LJ06a] P. Larsson and N. Johansson, "Multi-User ARQ," in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, vol. 4, May 2006, pp. 2052–2057. [cited at p. 83]
- [LJ06b] —, "Multi-User ARQ," in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, vol. 4, May 2006, pp. 2052–2057. [cited at p. 87]
- [LJ08] E. Larsson and E. Jorswieck, "Competition versus cooperation on the miso interference channel," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 7, pp. 1059–1069, September 2008. [cited at p. 82]
- [LMP07] A. Lera, A. Molinaro, and S. Pizzi, "Channel-Aware Scheduling for QoS and Fairness Provisioning in IEEE 802.16/WiMAX Broadband Wireless Access Systems," *IEEE Network*, vol. 21, no. 5, pp. 34–41, 2007. [cited at p. 63, 126]
- [LWG06] Q. Liu, X. Wang, and G. Giannakis, "A cross-layer scheduling algorithm with QoS support in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 3, pp. 839–847, May 2006. [cited at p. 53]
- [LX03] X.-B. Liang and X.-G. Xia, "On the nonexistence of rate-one generalized complex orthogonal designs," *Information Theory, IEEE Transactions on*, vol. 49, no. 11, pp. 2984–2988, Nov. 2003. [cited at p. 15]
- [LXG02] Z. Liu, Y. Xin, and G. Giannakis, "Space-time-frequency coded OFDM over frequency-selective fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2465–2476, Oct. 2002. [cited at p. 37, 63]
- [LZG05] Q. Liu, S. Zhou, and G. B. Giannakis, "Queuing with adaptive modulation and coding over wireless links: cross-layer analysis and design," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1142–1153, May 2005. [cited at p. 53]
- [MA06] A. Maaref and S. Aissa, "Performance analysis of orthogonal space-time block codes in spatially correlated MIMO Nakagami fading channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 807–817, April 2006. [cited at p. 47]
- [MCFH07] M. McKay, I. Collings, A. Forenza, and R. Heath, "Multiplexing/Beamforming Switching for Coded MIMO in Spatially Correlated Channels Based on Closed-Form BER Approximations," *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 5, pp. 2555–2567, Sept. 2007. [cited at p. 27]
- [MCM⁺09] F. Meucci, O. Cabral, A. Mihovska, F. Velez, and N. Prasad, "Spectrum Aggregation with Multi-Band User Allocation over Two Frequency Bands," in *IEEE Mobile WiMAX Symposium 2009*, Napa Valley, California, USA, 7 2009. [cited at p. 159, 172]
- [MDCP09] F. Meucci, E. De Carvalho, and P. Popovski, "Multi-User MIMO ARQ strategies, in submission," *IEEE Trans. Wireless Commun.*, 2009. [cited at p. 84, 172]
- [ME09] D. Marco and M. Effros, "On lossless coding with coded side information," *Information Theory, IEEE Transactions on*, vol. 55, no. 7, pp. 3284–3296, July 2009. [cited at p. 82]
- [MMAP08] F. Meucci, A. Mihovska, B. Anggorojati, and N. Prasad, "Joint resource allocation and admission control mechanism for an OFDMA-Based system," in *The 11th International Symposium on Wireless Personal Multimedia Communications*, Saariselk, Lapland, Finland, Finland, 9 2008. [cited at p. 38, 52, 63, 127, 172]

- [MMPR07] D. Marabissi, F. Meucci, L. Pierucci, and L. S. Ronga, “Adaptive selection of MIMO schemes in IEEE 802.16e,” in *IWCMC '07: Proceedings of the 2007 international conference on Wireless communications and mobile computing*. New York, NY, USA: ACM, 2007, pp. 423–428. [cited at p. 7, 27, 29, 83, 85, 104, 172]
- [Mou09] M. Moustafa, “Switched-mode BLAST technique for MIMO communications,” in *Proc. 11th International Conference on Advanced Communication Technology ICACT 2009*, vol. 03, 15–18 Feb. 2009, pp. 1499–1502. [cited at p. 18]
- [MP08] F. Meucci and L. Pierucci, “A real-time siren detector to improve safety of guide in traffic environment,” in *European Conference on Signal Processing 2008*, Lausanne, Switzerland, 8 2008. [cited at p. 172]
- [MPCC10] F. Meucci, L. Pierucci, I. Cerutti, and P. Castoldi, “Performance of dynamic service addition in mobile WiMAX networks,” in *submitted to IEEE ICC 2010 - Wireless and Mobile Networking Symposium*, Cape Town, South Africa, 2010. [cited at p. 64, 172]
- [MPDP09] F. Meucci, L. Pierucci, E. Del Re, and R. Prasad, “Dynamic allocation of STBC for OFDM based on WSSUS channel model,” in *Wireless Communication Society, Vehicular Technology, Information Theory and Aerospace & Electronics Systems Technology*, Aalborg, Denmark, 5 2009. [cited at p. 37, 45, 172]
- [MS83] R. Mersereau and T. Speake, “The processing of periodically sampled multidimensional signals,” *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, vol. 31, no. 1, pp. 188–194, Feb 1983. [cited at p. 31]
- [MVM⁺09] A. Mihovska, F. Velez, F. Meucci, O. Cabral, and N. Prasad, “Multi-Operator resource sharing scenario in the context of IMT-Advanced systems,” in *Wireless Communication Society, Vehicular Technology, Information Theory and Aerospace & Electronics Systems Technology*, Aalborg, Denmark, 2009. [cited at p. 172]
- [MWP09a] F. Meucci, S. A. Wardana, and N. R. Prasad, “Data Mules Networks Analysis with Semi-Markov-Process for Challenging Internets and Developing Regions,” in *Proceedings of the 4th Africa ICT Conference*, Kampala, Uganda, 06 - 08 May 2009. [cited at p. 172]
- [MWP09b] —, “Secure Physical Layer using Dynamic Permutations in Cognitive OFDMA Systems,” in *Proceedings of the IEEE 69th Vehicular Technology Conference: VTC2009-Spring*, Barcellona, 2009. [cited at p. 51, 118, 172]
- [MWW02] A. Molisch, M. Win, and J. Winters, “Space-time-frequency (STF) coding for MIMO-OFDM systems,” *IEEE Commun. Lett.*, vol. 6, no. 9, pp. 370–372, Sept. 2002. [cited at p. 37]
- [Nua07] L. Nuaymi, *WiMAX: Technology for Broadband Wireless Access*. Wiley InterScience, January 2007. [cited at p. 121]
- [Pig06] P. Piggin, “Emerging mobile WiMAX antenna technologies,” *IET Communications Engineer*, vol. 4, no. 5, pp. 29–33, 2006. [cited at p. 20]
- [PMZ⁺09] N. Pratas, F. Meucci, D. Zrno, N. Prasad, A. Rodrigues, and R. Prasad, “Cogito testbed: Wireless research evolved,” in *Wireless Communication Society, Vehicular Technology, Information Theory and Aerospace & Electronics Systems Technology*, Aalborg, Denmark, 5 2009. [cited at p. 172]
- [Pop06] P. Paposki. (2006) An Alternative Viewpoint on the (Cross)Layering in Wireless Networks. Unpublished Draft. [Online]. Available: <http://kom.aau.dk/~petarp/publications.html#> [cited at p. 5, 53]

- [Pra08] Prasad, Ramjee, "A Perspective of Layerless Communications," *Wirel. Pers. Commun.*, vol. 44, no. 1, pp. 95–100, 2008. [cited at p. 5]
- [Pro95] J. Proakis, *Digital Communications 4th Edition*. McGraw-Hill Book Co., 1995. [cited at p. 12, 13, 28]
- [PW07] L. Ping and P. Wang, "Multi-user gain and maximum eigenmode beamforming for mimo systems with rate constraints," in *Information Theory for Wireless Networks, 2007 IEEE Information Theory Workshop on*, July 2007, pp. 1–5. [cited at p. 82]
- [Rad22] J. Radon, *Lineare scharen orthogonaler Matrizen*. Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universitat, 1922. [cited at p. 11, 14]
- [RB01] K. Ramasubramanian and K. Baum, "An OFDM timing recovery scheme with inherent delay-spread estimation," in *Proc. IEEE Global Telecommunications Conference GLOBECOM '01*, vol. 5, 25–29 Nov. 2001, pp. 3111–3115. [cited at p. 41]
- [RH07] C. Rao and B. Hassibi, "Diversity-Multiplexing Gain Trade-off of a MIMO System with Relays," in *Information Theory for Wireless Networks, IEEE Information Theory Workshop on*, July 2007, pp. 1–5. [cited at p. 82]
- [RLM02] J. Razavilar, K. Liu, and S. Marcus, "Jointly optimized bit-rate/delay control policy for wireless packet networks with fading channels," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 484–494, Mar 2002. [cited at p. 53]
- [SA05] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels, 2nd Edition*. Wiley - IEEE Press, 2005, ISBN: 978-0-471-64953-3. [cited at p. 45, 151, 152, 154]
- [SBM⁺04] G. Stuber, J. Barry, S. McLaughlin, Y. Li, M. Ingram, and T. Pratt, "Broadband MIMO-OFDM wireless communications," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 271–294, Feb 2004. [cited at p. 37]
- [Sha48] C. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, 1948. [cited at p. 115]
- [SIJT09] C. So-In, R. Jain, and A.-K. Tamimi, "Scheduling in IEEE 802.16e mobile WiMAX networks: key issues and a survey," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 2, pp. 156–171, 2009. [cited at p. 63, 126]
- [SL04] H. Shin and J. H. Lee, "Performance analysis of space-time block codes over keyhole Nakagami-m fading channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 351–362, March 2004. [cited at p. 47]
- [SMF05] T. Sorensen, P. Mogensen, and F. Frederiksen, "Extension of the ITU channel models for wideband (OFDM) systems," in *Vehicular Technology Conference, 2005. VTC-2005-Fall. 2005 IEEE 62nd*, vol. 1, Sept., 2005, pp. 392–396. [cited at p. 57, 65]
- [STMA07] A. Sayenko, V. Tykhomyrov, H. Martikainen, and O. Alanen, "Performance analysis of the IEEE 802.16 ARQ mechanism," in *MSWiM '07: Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems*. New York, NY, USA: ACM, 2007, pp. 314–322. [cited at p. 123]
- [SU09] A. Soysal and S. Ulukus, "Optimality of Beamforming in Fading MIMO Multiple Access Channels," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1171–1183, Apr. 2009. [cited at p. 6, 82]
- [TAP⁺08] S. Thompson, A. Ahmed, J. Proakis, J. Zeidler, and M. Geile, "Constant Envelope OFDM," *IEEE Trans. Commun.*, vol. 56, no. 8, pp. 1300–1312, August 2008. [cited at p. 159]

- [TD08] C. Tian and S. Diggavi, "Side-information scalable source coding," *Information Theory, IEEE Transactions on*, vol. 54, no. 12, pp. 5591–5608, Dec. 2008. [cited at p. 82]
- [TDN08] M. Tran, A. Doufexi, and A. Nix, "Mobile WiMAX MIMO performance analysis: Downlink and uplink," in *Proc. PIMRC*, Sept. 2008, pp. 1–5. [cited at p. 63]
- [Tho08] R. Thobaben, "Joint Network/Channel Coding for Multi-User Hybrid-ARQ," in *7th International ITG Conference on Source and Channel Coding*, 2008. [cited at p. 87]
- [TJC99] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999. [cited at p. 14, 15]
- [TJC00] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Correction to Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 314–314, Jan. 2000. [cited at p. 14]
- [TS02] L. Tassiulas and S. Sarkar, "Maxmin fair scheduling in wireless networks," in *Proc. IEEE Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies INFOCOM 2002*, vol. 2, 23–27 June 2002, pp. 763–772. [cited at p. 127]
- [TS05] —, "Maxmin fair scheduling in wireless ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 163–173, Jan. 2005. [cited at p. 127]
- [TSC98] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *Information Theory, IEEE Transactions on*, vol. 44, no. 2, pp. 744–765, Mar 1998. [cited at p. 11, 14, 15, 16]
- [TSZS09] L. Tan, G. Su, G. Zhu, and P. Shang, "Adaptive Proportional Fair Scheduling Based on Opportunistic Beamforming for MIMO Systems," in *Proc. 5th International Conference on Wireless Communications, Networking and Mobile Computing WiCom '09*, 24–26 Sept. 2009, pp. 1–4. [cited at p. 126]
- [TVZ04] D. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing trade-off in multiple-access channels," *Information Theory, IEEE Transactions on*, vol. 50, no. 9, pp. 1859–1874, Sept. 2004. [cited at p. 12, 26]
- [VLB04] A. Vielmon, Y. Li, and J. Barry, "Performance of Alamouti transmit diversity over time-varying Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1369–1373, Sept. 2004. [cited at p. 15, 44, 45]
- [vNP00] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Norwood, MA, USA: Artech House, Inc., 2000. [cited at p. 37, 38, 39, 53, 64]
- [VTL02] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *Information Theory, IEEE Transactions on*, vol. 48, no. 6, pp. 1277–1294, Jun 2002. [cited at p. 6, 82, 84, 88]
- [WFGV98] 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 *Signals, Systems, and Electronics, 1998. ISSSE 98. 1998 URSI International Symposium on*, 29 Sept.-2 Oct. 1998, pp. 295–300. [cited at p. 18]
- [WK09] M. Wigger and G. Kramer, "Three-user MIMO MACs with cooperation," in *Networking and Information Theory, 2009. ITW 2009. IEEE Information Theory Workshop on*, June 2009, pp. 221–225. [cited at p. 82]

- [WKP01] K. Witrisal, Y.-H. Kim, and R. Prasad, "A new method to measure parameters of frequency-selective radio channels using power measurements," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1788–1800, Oct. 2001. [cited at p. 41]
- [WL09] K. K. Wee and S. W. Lee, "Priority based bandwidth allocation scheme for WiMAX systems," in *Proc. 2nd IEEE International Conference on Broadband Network & Multimedia Technology IC-BNMT '09*, 18–20 Oct. 2009, pp. 15–18. [cited at p. 127]
- [WLG06] X. Wang, Q. Liu, and G. Giannakis, "Analyzing and Optimizing Adaptive Modulation-Coding Jointly with ARQ for QoS-Guaranteed Traffic," in *Communications, 2006. ICC '06. IEEE International Conference on*, vol. 3, June 2006, pp. 1008–1013. [cited at p. 53]
- [WLLC08] J.-H. Wen, S.-H. Lee, G.-R. Lee, and J.-T. Chang, "Timing and delay spread estimation scheme in OFDM systems," *IEEE Trans. Consum. Electron.*, vol. 54, no. 2, pp. 316–320, May 2008. [cited at p. 41]
- [WML02] K.-K. Wong, R. Murch, and K. Letaief, "Performance enhancement of multiuser MIMO wireless communication systems," *Communications, IEEE Transactions on*, vol. 50, no. 12, pp. 1960–1970, Dec 2002. [cited at p. 6, 82]
- [XYW09] N. Xu, S. Yu, and X. Wang, "A Combined Fair Scheduling Algorithm for Combined-Input-Crosspoint-Queued Switch," in *Proc. 5th International Conference on Wireless Communications, Networking and Mobile Computing WiCom '09*, 24–26 Sept. 2009, pp. 1–4. [cited at p. 126]
- [XZ09] C. Xiao and Y. R. Zheng, "Transmit Precoding for MIMO Systems with Partial CSI and Discrete-Constellation Inputs," in *Proc. IEEE International Conference on Communications ICC '09*, Jun. 14–18, 2009, pp. 1–5. [cited at p. 6, 82]
- [XZG04] P. Xia, S. Zhou, and G. Giannakis, "Adaptive MIMO-OFDM based on partial channel state information," *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 52, no. 1, pp. 202–213, Jan 2004. [cited at p. 27]
- [YA08] T. Yucek and H. Arslan, "Time Dispersion and Delay Spread Estimation for Adaptive OFDM Systems," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1715–1722, May 2008. [cited at p. 41]
- [Yag04] H. Yaghoobi, "Scalable OFDMA physical layer in IEEE 802.16 WirelessMAN," *Intel Technology Journal*, vol. 8, no. 3, pp. 201–212, 2004. [cited at p. 68, 70, 73]
- [YHK08] Z. Yan, L. Huang, and C.-C. Kuo, "Seamless high-velocity handover support in mobile WiMAX networks," in *IEEE Singapore International Conference on Communication Systems*, Nov. 2008, pp. 1680–1684. [cited at p. 63]
- [ZBDS06] L. Zhang, A. Burr, M. Darnell, and L. Song, "Performance evaluation of spatial multiplexing systems over finite scattering channel," in *Proc. ITW '06 Chengdu Information Theory Workshop IEEE*, 22–26 Oct. 2006, pp. 658–662. [cited at p. 18, 23]
- [ZLH04] H. Zheng, A. Lozano, and M. Haleem, "Multiple ARQ processes for MIMO systems," *EURASIP J. Appl. Signal Process.*, vol. 2004, pp. 772–782, 2004. [cited at p. 82]
- [ZT00] L. Zheng and D. N. C. Tse, "Sphere packing in the Grassmann manifold: a geometric approach to the noncoherent multi-antenna channel," in *Proc. IEEE International Symposium on Information Theory*, 25–30 June 2000, p. 364. [cited at p. 5]

- [ZT03] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *Information Theory, IEEE Transactions on*, vol. 49, no. 5, pp. 1073–1096, May 2003. [cited at p. 7]
- [ZVRLG03] X. Zhuang, F. Vook, S. Rouquette-Leveil, and K. Gosse, "Transmit diversity and spatial multiplexing in four-transmit-antenna OFDM," in *Proc. IEEE ICC*, vol. 4, 11-15 May 2003, pp. 2316–2320. [cited at p. 17]
- [ZXC07] W. Zhang, X.-G. Xia, and P. C. Ching, "High-Rate Full-Diversity Space-Time-Frequency Codes for Broadband MIMO Block-Fading Channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 25–34, Jan. 2007. [cited at p. 37]

Appendices

Appendix A

Closed-form SEP Calculation

In Chapter 2, the closed-form SEP for Alamouti STBC over time varying channels has been derived for M-order QAM modulations. First, the SNR distribution at the output of the Alamouti receiver has been derived for non-perfectly static channels. Then, the average SEP has been calculated integrating the instantaneous SNR over the fading distribution. The averaging was conducted using the MGF of the SNR distribution. In this appendix, the previous calculation technique is thoroughly analysed for a M-QAM modulation with a coherent receiver over a Rayleigh fading channel.

The technique is based on the alternative expression of the Q-function found by Craig in 1991, [Cra91], and on the alternative expression of the squared Q-function as derived in [SA05].

A.1 SEP Calculation using Moment Generating Function

Closed-form average SEP evaluation is usually difficult; the average SEP is obtained integrating the symbol error rate over the instantaneous SNR distribution. The conditional SEP is, in general, an expression containing non-linear functions of the SNR. The non-linear integrand depends on the modulation and coding scheme adopted, but even in the case of a simple BPSK, the error rate depends on the Q-function, which is an indefinite integral of an exponential function with no closed-form integral*. The MGF of the channel fading is the Laplace transform of the fading distribution. The SEP integral can be evaluated more easily plugging the MGF of the SNR distribution for the following reasons:

- the integrand function is in the form of a rational function (instead of non-linear Q function)
- the integration interval is finite avoiding the calculation of the indefinite integral over all the instantaneous SNR distribution.

A.1.1 SEP for M-QAM Modulations in AWGN Channels

As well-known, the SEP for QAM modulation of order M is obtained as the quadrature combination of two Amplitude Modulations (AM) with half

*The Q-function is defined as: $Q(x) = 1/\sqrt{2\pi} \int_x^\infty \exp(-y^2/2) dy = 1/2 \operatorname{erfc}(x/\sqrt{2})$.

symbol energy and order \sqrt{M} :

$$SEP(E_S)|_{M-QAM, E_S} = 1 - \left[1 - SEP(E_S)|_{\sqrt{M}-AM, E_S/2}\right]^2 \quad (A.1)$$

Recalling the SEP for AM:

$$SEP(E_S)|_{M-AM, E_S} = 2 \left(\frac{M-1}{M}\right) Q \left(\sqrt{\frac{6E_S}{N_0(M^2-1)}}\right) \quad (A.2)$$

the M-QAM SEP for the AWGN channel can be expressed as:

$$\begin{aligned} SEP(E)|_{M-QAM, E_S} &= 4 \left(\frac{\sqrt{M}-1}{\sqrt{M}}\right) Q \left(\sqrt{\frac{3E_S}{N_0(M-1)}}\right) \\ &- 4 \left(\frac{\sqrt{M}-1}{\sqrt{M}}\right)^2 Q^2 \left(\sqrt{\frac{3E_S}{N_0(M-1)}}\right) \end{aligned} \quad (A.3)$$

For the calculation of the average SEP, the Q-function and Q-squared function are involved as integrands of an indefinite integral. An alternative Q-function expression is found by Craig in 1991, [Cra91] via a variable transformation; it can be obtained extending the classical Q-function over two dimensions in polar coordinates and then integrating over half plane. The alternative form is written as:

$$\begin{aligned} Q_{AL}(x) &= \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{y^2}{2}\right) dy \\ &= \frac{1}{\pi} \int_0^{\pi/2} \exp\left(-\frac{x^2}{2\sin^2\theta}\right) d\theta \end{aligned} \quad (A.4)$$

For the Q-squared function, based on the previous derivation and following Appendix 4A in [SA05], a similar alternative form can be obtained as the following:

$$\begin{aligned} Q_{AL}^2(x) &= \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{y^2}{2}\right) dy \cdot \int_x^\infty \exp\left(-\frac{y^2}{2}\right) dy \\ &= \frac{1}{\pi} \int_0^{\pi/4} \exp\left(-\frac{x^2}{2\sin^2\theta}\right) d\theta \end{aligned} \quad (A.5)$$

The Q-squared function integration can be obtained simply changing the integration interval in the expression of (A.4). Despite its apparent simplicity, however, this change of interval leads to closed-forms that can be far more complicated compared to $Q(x)$. We will see in the next section, that a simple closed-form is available when the channel fading has Rayleigh distribution.

Using the alternative functions reported in (A.4) and in (A.5), the SEP for AWGN channel can be readily calculated as:

$$\begin{aligned} SEP(E) &= \frac{4}{\pi} \left(\frac{\sqrt{M}-1}{\sqrt{M}}\right) \int_0^{\pi/2} \exp\left(-\frac{E_s}{N_0} \frac{3}{2(M-1)\sin^2\theta}\right) d\theta \\ &- \frac{4}{\pi} \left(\frac{\sqrt{M}-1}{\sqrt{M}}\right)^2 \int_0^{\pi/4} \exp\left(-\frac{E_s}{N_0} \frac{3}{2(M-1)\sin^2\theta}\right) d\theta \end{aligned}$$

A.1.2 SEP with Generic Channel Fading

In the case of fading channel, the instantaneous SNR distribution needs to be integrated in order to obtain the average SEP. For the narrowband channel analysed in Section 2.6, the received signal amplitude is multiplied by the channel fading amplitude which is a random variable β with a mean-square value of Ω . The Probability Density Function (PDF) of the fading amplitude ($p_\beta(\beta)$) is characteristic of the radio propagation. After being multiplied by the channel fading, the signal is disturbed with the additive noise at the receiver. The noise at the receiver is usually considered independent from the fading statistic and defined by a one-sided power spectral density of $N_0(W/Hz)$. Following the assumption of statistic independence among fading and added noise at the receiver, an instantaneous SNR can be defined as:

$$\gamma = \beta^2 \frac{E_b}{N_0} \quad (\text{A.6})$$

where E_b is the energy per bit. The PDF of the SNR can be derived with the variable transformation defined in (A.6) as:

$$p_\gamma(\gamma) = \frac{p_\beta(\sqrt{\Omega\gamma/\bar{\gamma}})}{2\sqrt{\gamma\bar{\gamma}/\Omega}} \quad (\text{A.7})$$

where the average SNR is: $\bar{\gamma} = \Omega E_b/N_0$.

The conditional SEP for a certain instantaneous SNR, is obtained with the AWGN SEP where E_s/N_0 is replaced by $\gamma \log_2(M)$; $\log_2(M)$ is the bits per symbol which depends on the QAM modulation order. The average SEP is obtained from the conditional SEP integrated over all the real axis and the following has to be solved:

$$I = \int_0^\infty Q(a\sqrt{\gamma})p_\gamma(\gamma)d\gamma \quad (\text{A.8})$$

where a is a constant which depends on the modulation adopted.

Let us first define the integral expressions for Q and Q-squared functions over the fading distribution.

$$\begin{aligned} I_Q(a, \gamma) &= \int_0^\infty Q(a\sqrt{\gamma})p_\gamma(\gamma)d\gamma \\ &= \int_0^\infty \frac{1}{\pi} \int_0^{\pi/2} \exp\left(-\frac{a^2\gamma}{2\sin^2\theta}\right) d\theta p_\gamma(\gamma)d\gamma \\ &= \frac{1}{\pi} \int_0^{\pi/2} \left[\int_0^\infty \exp\left(-\frac{a^2\gamma}{2\sin^2\theta}\right) p_\gamma(\gamma)d\gamma \right] d\theta \end{aligned} \quad (\text{A.9})$$

The MGF[†] of the SNR distribution is defined as:

$$M_\gamma(s) = \int_0^\infty p_\gamma(\gamma)\exp(s\gamma)d\gamma \quad (\text{A.10})$$

and is the Laplace transform of the SNR probability density. The MGF permits to easily evaluate the SEP for several channel statistics and modulation and coding schemes avoiding the complex indefinite integral with non-linear integrand.

[†]The name Moment Generating Function comes from the fact that the the n^{th} moment of the random variable $X(t)$ can be obtained simply deriving n times the MGF and evaluating for $t = 0$: $E\{X^n\} = \frac{d^n}{dt^n} M_X(t)|_{t=0}$.

Noting that:

$$M_\gamma(s)|_{s=-\frac{a^2}{2\sin^2\theta}} = \int_0^\infty \exp\left(-\frac{a^2\gamma}{2\sin^2\theta}\right) p_\gamma(\gamma) d\gamma \quad (\text{A.11})$$

is the MGF of the instantaneous SNR distribution evaluated in $s = -\frac{a^2}{2\sin^2\theta}$, the averaged Q-function can be written as:

$$I_Q(a, \gamma) = \frac{1}{\pi} \int_0^{\pi/2} M_\gamma\left(-\frac{a^2}{2\sin^2\theta}\right) d\theta \quad (\text{A.12})$$

Using (A.5), the Q-squared function integrated over the SNR distribution $p_\gamma(\gamma)$ is:

$$\begin{aligned} I_{Q^2}(a, \gamma) &= \int_0^\infty Q^2(a\sqrt{\gamma}) p_\gamma(\gamma) d\gamma \\ &= \frac{1}{\pi} \int_0^{\pi/4} M_\gamma\left(-\frac{a^2}{2\sin^2\theta}\right) d\theta \end{aligned} \quad (\text{A.13})$$

Recalling $SEP(E_S)|_{M-QAM, E_S}$ from (A.3), substituting E_S/N_0 with $\gamma \log_2(M)$ and integrating over the fading distribution, yields to:

$$SEP(\bar{\gamma}) = 4 \cdot \left(\frac{\sqrt{M}-1}{\sqrt{M}}\right) I_Q(a, \gamma) - 4 \cdot \left(\frac{\sqrt{M}-1}{\sqrt{M}}\right) I_{Q^2}(a, \gamma) \quad (\text{A.14})$$

where for M-QAM modulations, $a^2 = 3 \frac{\log_2(M)}{M-1}$.

A.1.3 Rayleigh Fading Channel Statistic

In this section, the previous expressions are calculated for the particular case of Rayleigh fading. The Rayleigh fading is used to model non LoS channel with multipath fading. The fading statistic is expressed as:

$$p_\beta(\beta) = \frac{2\beta}{\Omega} e^{-\frac{\beta^2}{\Omega}} \quad (\text{A.15})$$

and is defined for $\beta \geq 0$. The instantaneous SNR probability density is thus:

$$p_\gamma(\gamma) = \frac{1}{\bar{\gamma}} e^{-\frac{\gamma}{\bar{\gamma}}} \quad (\text{A.16})$$

again defined for $\gamma \geq 0$. The MGF for the Rayleigh statistic is:

$$M_\gamma^{Ray}(s) = \frac{1}{1 - s\bar{\gamma}} \quad (\text{A.17})$$

Other common fading models (such as Nakagami-m, Rice, Hoyt) and relative MGFs can be found in [SA05].

In this case, the followings can be derived:

$$\begin{aligned} I_Q^{Ray}(a, \bar{\gamma}) &= \frac{1}{\pi} \int_0^{\pi/2} M_\gamma^{Ray}(s) d\theta \\ &= \frac{1}{\pi} \int_0^{\pi/2} \left(1 + \frac{a^2\bar{\gamma}}{2\sin^2\theta}\right)^{-1} d\theta \\ &= \frac{1}{2} \left(1 - \sqrt{\frac{a^2\bar{\gamma}/2}{1 + a^2\bar{\gamma}/2}}\right) \end{aligned} \quad (\text{A.18})$$

For the Q-squared function with Rayleigh fading, the following derivation holds:

$$\begin{aligned}
 I_{Q^2}^{Ray}(a, \bar{\gamma}) &= \frac{1}{\pi} \int_0^{\pi/4} M_{\gamma}^{Ray}(s) d\theta \\
 &= \frac{1}{\pi} \int_0^{\pi/4} \left(1 + \frac{a^2 \bar{\gamma}}{2 \sin^2 \theta} \right)^{-1} d\theta \\
 &= \frac{1}{4} \left[1 - \sqrt{\frac{a^2 \bar{\gamma}/2}{1 + a^2 \bar{\gamma}/2}} \left(\frac{4}{\pi} \tan^{-1} \sqrt{\frac{1 + a^2 \bar{\gamma}/2}{a^2 \bar{\gamma}/2}} \right) \right]
 \end{aligned} \tag{A.19}$$

Finally, considering that $a^2 = 3 \frac{\log_2(M)}{M-1}$ and $\bar{\gamma}_s = \bar{\gamma} \cdot \log_2(M)$, the SEP for Rayleigh fading with coherent QAM modulation, reported also in Fig. A.1, is obtained as:

$$\begin{aligned}
 SEP(\bar{\gamma}) &= 2 \left(\frac{\sqrt{M} - 1}{\sqrt{M}} \right) \left(1 - \sqrt{\frac{3\bar{\gamma} \log_2(M)/2}{M - 1 + 3\bar{\gamma} \log_2(M)/2}} \right) \\
 &\quad - \left(\frac{\sqrt{M} - 1}{\sqrt{M}} \right)^2 \left[1 - \sqrt{\frac{3\bar{\gamma} \log_2(M)/2}{M - 1 + 3\bar{\gamma} \log_2(M)/2}} \right. \\
 &\quad \left. \left(\frac{4}{\pi} \tan^{-1} \sqrt{\frac{M - 1 + 3\bar{\gamma} \log_2(M)/2}{3\bar{\gamma} \log_2(M)/2}} \right) \right]
 \end{aligned} \tag{A.20}$$

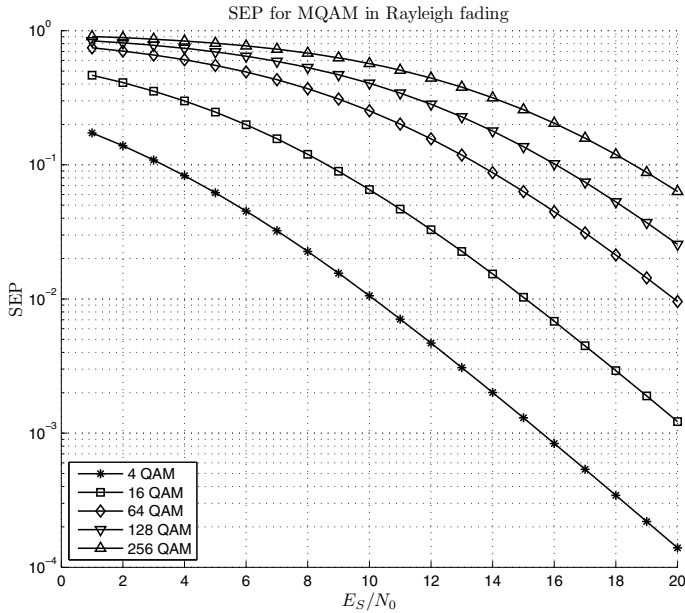


Figure A.1: SEP for various order of Quadrature Amplitude Modulations

Appendix B

IMT-A Requirements

The definitions and requirements for the 4G networks are defined globally by the International Communication Union (ITU). ITU-R provides a set of requirements collected in a series of recommendations that define the International Mobile Telecommunications Advanced (IMT-A). Though 4G is a broader term and could include standards outside IMT-A, 4G is generally associated with ITU IMT-A specification. ITU also provides the guidelines for the evaluation of the systems proposed as candidates to the IMT-A.

Currently, the major candidates for IMT-A are IEEE802.16 developed inside the WiMAX Forum and the LTE-A developed inside the 3rd Generation Partnership Project (3GPP) consortium.

ITU has published the specifications for IMT-A candidates in November 2008, [ITU08]. The overall IMT-A requirements, as to the date of publication, are briefly discussed in this section. It is important to note that these requirements are being enhanced in line with the technology developments. ITU compiles the following performance levels based on the ongoing activities and including the latest technical developments.

It is worth mentioning that the system specified in the ITU-R M.1457 recommendation (3G -IMT 2000 Release 1) was already including the 3GPP LTE Release 8 and the IEEE 802.16e-2005 ("Mobile WiMAX"). In that recommendation, many of the main 4G requirements were already specified: support of fixed and mobile terminals supported, all-IP network, flat network architecture, OFDMA, MIMO, Hybrid Automatic Repeat Request (HARQ) at the PHY layer, multi-megabit speeds delivered to users, just to cite few.

This means that both Mobile WiMAX and LTE are officially designated 3G technologies evolving to 4G inside their respective standards bodies. The race to 4G has been accelerated by a number of factors; the early deployment of Mobile WiMAX is one of the most important.

Since its kick-off in late 2005, WiMAX has been deployed by almost 400 operators spanning 133 countries. Although not an independent source, WiMAX forum predicts a worldwide WiMAX subscriber base of about 130 million by 2012 with a prediction of a strong base of about 27 million subscribers in India. Chip manufacturers like Intel have committed to the availability of embedded WiMAX chips for laptops, PCs and mobile handsets by the end of 2009.

The competitor LTE provides an evolution path for advancement from today's 3G Universal Mobile Telecommunications System (UMTS) and

3.5G High-Speed Downlink/Uplink Packet Access (HSXPA) cellular technologies to the future 4G networks. Given its voice-centric vision and its evolutionary path, LTE has been supported by several mobile operators worldwide. This endorsement makes LTE the most likely standard to be dominant worldwide in the short term for mobile voice with seamless roaming across different operators all over the globe. One key advantage of LTE is the possibility to leverage the 700MHz spectrum for efficient support of voice services with higher in-building penetration and across longer distances.

However, some consider the development road-map of LTE-A too slow (the name itself - Long Term Evolution - is not a guarantee of fastness) and the cellular industry can leverage on LTE as a way of slowing down WiMAX commercial spreading without being committed to providing a real user experience improvement. The advantage of the "evolutionary development" of LTE is also to be verified. In reality, LTE is more a "forklift upgrade" proposing a new Radio Access Technology (RAT), different network architecture and a totally different PHY layer configuration and BS capabilities. LTE modem pricing could be another issue considering an extrapolation of the already available 3G modems: 3G-HSPA modems are priced two to three times higher than a Mobile WiMAX modems delivering a peak rate of 14Mbit/sec against the 40Mbit/sec supported in WiMAX.

RAT is one of the key blocks for wireless networks and 4G specifications require high-performing PHY layers with an increases spectral efficiency, capacity and scalability. TDMA and/or Frequency Division Multiple Access (FDMA), used since 1G cellular networks, have their own advantages and drawbacks. In particular, TDMA is less efficient in serving the high data rate connection as it is requiring large guard periods to alleviate the multipath impact while FDMA has low spectral efficiency since more bandwidth is needed for guards to avoid ICI. In the 2G networks, one set of standards used the combination of FDMA and TDMA and another group of standards introduced a new access scheme based on Code Division Multiple Access (CDMA). The latter increased the system capacity as this access scheme is efficient in handling the multipath channel. However CDMA suffers from poor spectrum flexibility, poor spectrum exploitation (due to bandwidth spreading factors) and scalability. CDMA is the base of 3G extensions such as: Information System 2000 (IS-2000), UMTS, HSXPA, 1x Evolution-Data Only (1xEV-DO), Time Division CDMA (TD-CDMA) and Time Division Synchronous CDMA (TD-SCDMA).

A big shift in the RATs came with the application of OFDM. Thanks to simple implementation via DFT and IDFT, OFDM has been universally inserted in 4G candidates. OFDM permits very easy channel equalisation, efficient signal processing and very flexible MU resource allocation. A variant of pure OFDM radio access technology is the OFDMA where users can be allocated in parts of the time-frequency transmission frame (resource chunk). Variants of OFDM have been also proposed for reduced Peak to Average Power Ratio (PAPR) such as: Single Carrier Frequency Division Multiplexing (SC-FDMA) and Interleaved Frequency Division Multiple Access (IFDMA). On another side, Multi Carrier Code Division Multiple Access (MC-CDMA) has been developed and chosen for some standards as IEEE802.20 (MBWA).

WiMax uses OFDMA in the DL and in the UL. In LTE-A, OFDMA is used for the DL while SC-FDMA is used in UL. SC-FDMA in LTE-A UL has been considered one of the factors that could lead to a success of LTE over WiMAX due to a reduced PAPR and lower cost BSs. This argument seems not so strong as this issue has been addressed by the research community and solved. Several solutions have been presented, the

most recent in [TAP⁺08] where phase rotations create a constant envelope OFDM signal solving the PAPR problem for non SC-FDMA uplinks, as in WiMAX.

Based on these considerations, WiMAX has been chosen as the reference standard for the work presented in this thesis. However, it is important to note that the technical solutions proposed here are applicable in both candidates. This is allowed mainly because the RATs share several technical aspects: OFDMA (or its variants as SC-FDMA) are used, users are allocated in portions of the time-frequency grid in the same conceptual way (each user is assigned to a piece of bandwidth for a certain time inside a frame) and identical MIMO schemes are used. Further, both candidates have a connection-oriented MAC layer.

B.1 4G Technical Requirements and Challenges

4G aims at several performance enhancements in respect to the current 3G systems; various areas of research are vibrantly active and regard:

- Higher throughput per sector (both peak rates and average)
- Higher MT speeds supported
- Better coverage with reduced signal fading
- All-IP network
- Flat and simple network architecture
- Native VoIP support
- Lower latency (access < 10ms and handover < 30ms)
- Better cell edge throughput
- Enhanced broadcast and multicast services
- Enhanced location based services
- Deployment flexibility - Self Organising Networks (SON)

Due to spectrum segmentation, operators will have licensed spectrum in widely separated frequency bands with heterogeneous bandwidths. 4G broadband wireless systems should provide a single radio bearer channel with the flexibility to aggregate non-uniform and non-contiguous channels. Dynamic bandwidth allocation and scaling will be an important key for providing the capacity demanded by multimedia services in a dynamic mobile environment. The spectrum will be not anymore a fixed resource but dynamically adapted to traffic and user requirements. MT will be equipped with multiple transceivers operating over several radio channels co-existing with low self-interference (Multi-Radio Multi-Band Terminals). Such technology opens the problem of the added degree of freedom for resources allocation and traffic scheduling. A common Radio Resource Management (RRM) have been proposed in [MCM⁺09].

4G network will be deployed as a Single Frequency Reuse (SFR) network in order to maximise the spectral utilisation over large geographical areas. For this reason, interference will degrade the quality of the connections especially for the users at the cell edges with more challenges during UL connections. While in DL a maximum of 8-10 interferers are expected, in the UL this number could easily reach more than one hundred. Techniques

to manage and reduce interference will be needed jointly with resources and users scheduling.

3G networks have two parallel infrastructures for circuit switched and packet switched traffic. Instead, 4G will be based only on packet switching. This requires very low latency times in the control and in the data planes. The IP network in 4G will be supporting native IPv6.

The evolution of smart or intelligent antennas will deeply impact on the future of wireless technologies. Multiple antennas at both sides of the radio link can lead to improved capacity with SM transmission modes. In SM, independent streams can then be transmitted simultaneously from all the antennas increasing the data rate times a number equal to minimum of the number of receive and transmit antennas. The independence of the fading statistic for each antenna is the key for SM. Especially in small devices, low correlation among antenna elements is hard to achieve. Solutions based on antenna selection and Multiple Input Single Output (MISO) systems have been already proposed. Antenna selection provides a sensible cost reduction due to a single transceiver chain attached to a multi-antenna port and the MISO systems allow to have expensive antenna systems concentrated at the BSs, where the antenna elements can be separated several wavelengths achieving very low correlation. MIMO can be exploited not only for increased capacity but also for high reliability for lower error rates and longer-range communications. Reliability is increased with transmit or receive diversity, especially with Space Time Code (STC) which does not require the CSIT. When CSIT is needed at the transmitter, the transmitted signal can be pre-processed in order to have the best signal at the antenna ports at the MT, [GKH⁺07].

B.2 IMT-Advanced Performance Requirements

The performance requirements of IMT-A candidates are analysed in detail, referring to [ITU08] ITU recommendation. The key features to be fulfilled by the candidate system are:

- a high degree of commonality of functionality worldwide while retaining the flexibility to support a wide range of services and applications in a cost efficient manner
- maintained compatibility inside IMT systems and fixed networks
- capability of interworking with non IMT radio access systems
- worldwide roaming capability
- high quality mobile services
- user equipment suitable for worldwide use
- user-friendly applications, services and equipment
- enhanced peak data rates to support advanced services and applications.

Quantitative specifications are divided in minimum requirements and target requirements. The data rate for the mobile environment has been set to 100Mbit/s for high and 1Gbit/s for low mobility as targets for research. The followings are the minimum requirements for the candidates.

Test environment	Downlink (<i>bit/s/Hz/cell</i>)	Uplink (<i>bit/s/Hz/cell</i>)
Indoor	3	2.25
Microcellular	2.6	1.80
Base coverage urban	2.2	1.4
High speed	1.1	0.7

Table B.1: IMT-A Cell Spectral Efficiency

Cell Spectral Efficiency

The Cell spectral efficiency η is defined as the aggregate throughput of all users (the number of correctly received bits, i.e. the number of bits contained in the service data units (SDUs) delivered to Layer 3, over a certain period of time) divided by the channel bandwidth divided by the number of cells. The channel bandwidth for this purpose is defined as the effective bandwidth times the frequency reuse factor, where the effective bandwidth is the operating bandwidth normalised appropriately considering the UL/DL ratio.

Let X_i denote the number of correctly received bits by user i (DL or UL) in a system with N users and M cells, ω be the channel bandwidth and T the time over which the data bits are received. The system spectral efficiency, η , is then:

$$\eta = \frac{\sum_{i=1}^N X_i}{T \cdot \omega \cdot M} \quad (\text{B.1})$$

For each test environment, as defined in the ITU-R Report M.2135, the following minimum efficiency per cell to be fulfilled by the candidates is reported in Tab. B.1.

Peak Cell Efficiency

The peak spectral efficiency is the highest theoretical data rate, which is the received data bits assuming error-free conditions assignable to a single mobile station, normalised by bandwidth when all available radio resources for the corresponding link direction are utilised (i.e., excluding radio resources that are used for PHY layer synchronisation, reference signals or pilots, guard bands and guard times). The minimum requirements for peak spectral efficiencies are:

- DL peak spectral efficiency is 15 *bit/s/Hz*
- UL peak spectral efficiency is 6.75 *bit/s/Hz*.

These values were defined assuming an antenna configuration for DL of 4×4 and for UL of 2×4 .

Bandwidth

Scalable bandwidth is the ability of the candidate air interface to operate with different bandwidth allocations. This bandwidth may be supported by single or multiple RF carriers. The candidate shall support a scalable bandwidth up to and including 40 MHz even if proponents are encouraged to consider extensions to support operation in wider bandwidths up to 100 MHz.

Test environment	Downlink (<i>bit/s/Hz</i>)	Uplink (<i>bit/s/Hz</i>)
Indoor	0.1	0.07
Microcellular	0.075	0.05
Base coverage urban	0.06	0.03
High speed	0.04	0.015

Table B.2: Cell User Edge Spectral Efficiency

Edge Spectral Efficiency

The (normalised) user throughput is defined as the average user throughput, i.e. the number of correctly received bits by users (the number of bits contained in the SDU delivered to Layer 3) over a certain period of time divided by the channel bandwidth and is measured in *bit/s/Hz*. The channel bandwidth for this purpose is defined as the effective bandwidth times the frequency reuse factor, where the effective bandwidth is the operating bandwidth normalised appropriately considering the UL/DL ratio. The cell edge user spectral efficiency is defined as 5% point of the Cumulative Distribution Function (CDF) of the normalised user throughput.

With X_i denoting the number of correctly received bits of user i , T_i the active session time for user i and ω the channel bandwidth, the (normalised) user throughput of user i , c_i , is defined according to (B.2). Tab. B.2 lists the cell edge user spectral efficiency requirements for various test environments. The test environments are defined in the ITU-R Report M.2135.

$$c_i = \frac{X_i}{T_i * \omega} \quad (\text{B.2})$$

Latency

The latency requirements are defined for the control plane (C-plane) and the user data plane (U-plane). The C-plane latency is typically measured as the transition time from different connection modes, e.g., from idle to active state. A transition time (excluding wireline network signalling delay) of less than **100 ms** shall be achievable from idle state to an active state in such a way that the user plane is established. The U-plane latency (also known as transport delay) is defined as the one-way transit time between an SDU packet being available at the IP layer in the MT/BS and the availability of this packet, PDU, at IP layer in the BS/MT. U-plane packet delay includes delay introduced by associated protocols and control signalling assuming the user terminal is in the active state. IMT-A systems shall be able to achieve a U-plane latency of less than **10 ms** in unloaded conditions (i.e., a single user with a single data stream) for small IP packets (e.g., 0 byte payload + IP header) for both DL and UL.

Mobility

The following mobility classes are defined:

- Stationary: $v = 0\text{km/h}$
- Pedestrian: $0 \leq v \leq 10\text{km/h}$
- Vehicular: $10 \leq v \leq 120\text{km/h}$
- High speed vehicular: $120 \leq v \leq 350\text{km/h}$

Test Env.	Mobility Classes Supported
Indoor	Stationary, Pedestrian
Micro Cellular	Stationary, pedestrian, Vehicular (up to 30 km/h)
Base Urban	Stationary, Pedestrian, Vehicular
High speed	High speed vehicular, Vehicular

Table B.3: Mobility classes for test environments.

	SE (<i>bit/s/Hz</i>)	Speed (<i>km/h</i>)
Indoor	1.0	10
Microcellular	0.75	30
Base coverage urban	0.55	120
High speed	0.25	350

Table B.4: Traffic channels, link data rates

Handover type	Interruption time (ms)
Intra-frequency	27.5
Inter-frequency	
within a spectrum band	40
between spectrum bands	60

Table B.5: Handover Interruption Times.

For each test environment, a set of mobility classes shall be supported as reported in Tab. B.3. For each mobility class, a minimum spectral efficiency is required as reported in Tab. B.4. These values were defined assuming an antenna configuration of DL 4×2 , UL 2×4 .

Handovers

The Handover interruption times are defined in Tab. B.5 for intra- and inter- system handovers and for handovering within a spectrum band and among different spectrum bands.

VoIP capacity

The Voice over IP (VoIP) capacity for the IMT-A candidates is derived assuming a 12.2 kbit/s codec with a 50% activity factor such that the percentage of users in outage is less than 2% where a user is defined to have experienced a voice outage if less than 98% of the VoIP packets have been delivered successfully to the user within a one way radio access delay bound of 50 ms . The specified codec is used for measurements and does not need to be effective part of the candidate system. The VoIP capacity is the minimum of the calculated capacity for either link direction divided by the effective bandwidth in the respective link direction. These values were defined assuming an antenna configuration of 4×2 in the DL and 2×4 in the UL. The antenna configuration is not a minimum requirement and the conditions for evaluation are described in Report ITU-R M.2135. The

Test environment	Min VoIP capacity (Active users/sector/MHz)
Indoor	50
Microcellular	40
Base coverage urban	40
High speed	30

Table B.6: VoIP capacity

specifications indicate a minimum amount of VoIP connections to be supported, as in Tab. B.6.

B.3 WiMAX Mobile IMT-A Candidate

IEEE802.16m aims to provide an upgrade for previous IEEE802.16 standard to fulfill the requirements of IMT-Advanced next generation mobile networks for the licensed spectrum bands. IEEE802.16m provides continuing support for previous legacy 802.16d (WirelessMAN OFDMA) equipment. The working group is producing the followings core documents. The System Requirements Document (SRD), [IEE09d], contains a set of possible deployment scenarios and applications of the 802.16m standard. The target performance and system features are specified.

The Evaluation Methodology Document (EMD), [IEE09b], proposes a complete set of parameters, models, and methodologies for the link-level and system-level evaluations. A set of spatial channel model parameters are specified to characterise particular features of MIMO radio channels.

The System Description Document (SDD), [IEE09c] reports the architecture and design of the 802.16m air interface amendment.

The working group will produce a final IMT-A Proposal to be submitted at the ITU-R.

The requirements for IEEE802.16m are divided in three groups: functional requirements, baseline requirements and a set of target requirements fulfilling the IMT-A specifications.

The General requirements are intended to address and supplement the requirements specified by the ITU-R for IMT-Advanced systems and are the minimum performance requirements for the MTs and BSs.

The functional requirements contains system level functional requirements targeting higher peak rates, lower latency, lower system overhead as well as PHY/MAC features enabling improved service security, QoS and radio resource management (RRM) for a system comprised solely of IEEE802.16m MTs and BSs.

The Baseline requirements describe the performance requirements for IEEE802.16m systems specified in terms of absolute performance and relative performance with respect to that of the WirelessMAN-OFDMA Reference System.

B.3.1 Functional Requirements

IEEE802.16m systems will operate in RF frequencies less than 6 GHz and be deployable in licensed spectrum allocated to the mobile and fixed broadband services and shall be able to operate in frequencies identified for IMT-A. The IEEE802.16m operative frequency bands are reported in Tab. B.7.

Frequency Bands for IEEE802.16m
450-470 MHz
698-960 MHz
1710-2025 MHz
2110-2200 MHz
2300-2400 MHz
2500-2690 MHz
3400-3600 MHz

Table B.7: Frequency Bands for Mobile WiMAX

The frequency bands have been identified for IMT and/or IMT-2000 by World Administrative Radio Conference of 1992 (WARC92), World Radio Conference of 2000 (WRC2000) and World Radio Conference of 2007 (WRC07) and are reported in Tab. B.7 and are described in Recommendation ITU-R M.1036-3. New frequency bands have been identified in the WRC07 but further work is ongoing inside the framework of ITU-R. IEEE802.16m shall support scalable bandwidths from 5 to 40 MHz. This bandwidth may be supported by single or multiple RF carriers.

IEEE802.16m shall support both Time Division Duplex (TDD) and Frequency Division Duplex (FDD) operational modes. The FDD mode shall support both full-duplex and half-duplex MT operation. A BS supporting FDD mode shall be able to simultaneously support half duplex and full duplex terminals operating on the same RF carrier.

IEEE802.16m shall support both unpaired and paired frequency allocations, with fixed duplexing frequency separations when operating in FDD mode. In TDD mode, the DL/UL ratio should be adjustable with the extreme of supporting DL-only configurations on a given carrier. In FDD mode, the UL and DL channel bandwidths may be different and should be configurable.

The IEEE802.16m standard defines minimum antenna requirements for the BS and MT. For the BS, a minimum of two transmit and two receive antennas shall be supported. For the MT, a minimum of one transmit and two received antennas shall be supported. This minimum is consistent with a 2×2 DL configuration and a 1×2 uplink configuration. IEEE802.16m shall support MIMO, beamforming operation or other advanced antenna techniques. IEEE802.16m shall further support single-user and multi-user MIMO techniques.

IEEE802.16m peak spectral efficiency achievable between a BS and an MT under ideal conditions is the highest theoretical data rate (normalised by bandwidth), which is the received data bits assuming error-free conditions assignable to a single mobile station, when all available radio resources for the corresponding link direction are utilised (that is excluding radio resources that are used for physical layer synchronisation, reference signals or pilots, guard bands and guard times). IEEE802.16m peak SEs is specified in Tab. B.8 where the baseline and the target specifications are reported. Latency, handovers timing, VoIP capacity and the other specifications are set to the same level as in the ITU IMT-A base requirements.

Requirement Type	Link	MIMO	Peak rate (<i>bit/s/Hz</i>)
Baseline	Downlink	2x2	8.0
Baseline	Uplink	1x2	2.8
Target	Downlink	4x4	15.0
Target	Uplink	2x4	6.75

Table B.8: Mobile WiMAX Normalized Peak Data Rate

Appendix C

MIMO Channel Capacity SVD Decomposition

In this appendix, MIMO capacity is analysed from the perspective of the Single Value Decomposition (SVD) of the channel matrix in order to highlight the impact of MIMO channel rank on the achievable MIMO capacity.

In 1999, Telatar showed in [Emr99] that the MIMO channel can be converted in a set of parallel and independent SISO channel via an SVD of the MIMO matrix*.

The SVD theorem states that for any complex $N_R \times N_T$ matrix \mathbf{H} with rank M , there exists an SVD of the form:

$$\mathbf{H} = U\Lambda V^\dagger \quad (\text{C.1})$$

where $U_{(N_R \times N_R)}$ and $V_{(N_T \times N_T)}$ are unitary matrices. The $N_R \times N_T$ matrix $\Lambda = [\lambda_{ij}]$ has $\lambda_{ij} = 0$ for all $i \neq j$, and $\lambda_{11} \geq \lambda_{22} \geq \lambda_{33} \geq \lambda_{44} \geq \dots \geq \lambda_{MM} \neq 0$ and $\lambda_{M+1, M+1} = \dots = \lambda_{aa} = 0$, where $a = \min(N_T; N_R)$.

The singular values of \mathbf{H} , $\{\lambda_{ii}; i = 1, 2, \dots, M\}$, are the nonzero eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$, and hence are uniquely determined.

Let $\{\lambda_{mm}\}$ be the M nonzero singular values of \mathbf{H} . Then, $\{s_m = \lambda_{mm}^2\}$ are the M nonzero eigenvalues of either $\mathbf{H}\mathbf{H}^\dagger$ or $\mathbf{H}^\dagger\mathbf{H}$. Since the number of transmitting or receiving antennas could differ, depending on the case where is $N_R > N_T$ or $N_R < N_T$, eigenvalues can be expressed in two ways. Following Telatar, a random matrix \mathbf{W} can be defined as:

$$\mathbf{W} = \begin{cases} \mathbf{H}\mathbf{H}^* & \text{if } N_R < N_T \\ \mathbf{H}^*\mathbf{H} & \text{if } N_R \geq N_T \end{cases} \quad (\text{C.2})$$

\mathbf{W} is an $M \times M$ matrix, where $M = \min(N_R, N_T)$ and it is a *non-negative definite* matrix with real non-negative eigenvalues λ_i^2 .

Each one of the set of parallel and independent SISO channel in which the MIMO channel can be decomposed, has the gain of one of the eigenvalues of matrix \mathbf{W} . The optimum power allocation that maximises the channel capacity is the water-filling allocation: the power to be transmitted

*SVD has several applications, i.e. in the field of matrix norm, rank and condition number calculations and for linear system solving, noisy signal filtering and time series analysis.

over each subchannel is given by:

$$P_i = \left(\mu - \frac{1}{\lambda_i} \right)^+ \quad (\text{C.3})$$

with $1 \leq i \leq \min(N_T, N_R)$, μ is the waterfill level, P_i is the power in the i^{th} eigenmode and the $(x)^+$ operator is defined as $\max(0, x)$. The total capacity is thus the sum of the capacities over each eigenchannel:

$$C_T = \sum_{i=1}^{\min(N_T, N_R)} \log(\mu \lambda_i^2)^+ \quad (\text{C.4})$$

SVD is not straightforward to be implemented since it has an high computation complexity, with an order of O^3 and needs perfect CSIR and CSIT. Several advantages can be devised in SVD-based MIMO schemes: there is no cross-talk between spatial channels (eigenmodes) and conventional channel codes are readily used. SVD can be applied to channel matrixes of any size and any rank, thus SVD can be applied seamlessly to spatially correlated channels. Finally, SVD is optimal in the information theoretical sense, since unitary filters preserve covariance of the input data vector.

C.1 MIMO Channel Rank

The most important metric for a MIMO channel is the rank of \mathbf{W} . The rank of \mathbf{W} is directly related to the number of possible independent parallel subchannels that SVD decomposition can create, i.e. the spatial independent paths that are available in the MIMO channel. The worst case is when a unique path is present between the transmitter and the receiver: $\text{rank}(\mathbf{W}) = 1$. When the rank of \mathbf{W} is maximum, the MIMO channel is said to be orthogonal and is:

$$\text{rank}(W) = \min(N_R, N_T) \quad (\text{C.5})$$

The rank corresponds to the maximum diversity gain d as defined in Equation (2.3) at page 14; the orthogonal channel represents the capacity upper limit and the maximum diversity gain achievable.

High rank MIMO channel occurs when there is a rich scattering channel with little correlation between spatial channel paths. When the rank of matrix \mathbf{W} is high, the capacity can be approximated showing two terms related to transmit array gain and SNR improvement due to the receiver array. The capacity can be expressed as:

$$C_{\text{high}} \approx \underbrace{\min(N_T, N_R)}_{\text{array capacity advantage}} \cdot \underbrace{\log \left(1 + \frac{\frac{E}{2\sigma^2} N_R}{N_T} \right)}_{\text{receiver antenna SNR advantage}} \quad (\text{C.6})$$

Low rank MIMO channel occurs when there is high correlation between the spatial subchannels. In the case of a MIMO channel with only one *non-zero* eigenvalue, the system is equivalent to a single antenna system with the same total power divided among N_T elements. A completely correlated \mathbf{H} matrix ($\mathbf{H} \in M_{N_R \times N_T}$) can be represented as:

$$\mathbf{H} = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \quad (\text{C.7})$$

and correspond to the case of no diversity gain and the SDMA gain is null. The unique *non-zero* eigenvalue is given by:

$$\lambda^2 = N_R \sum_{n=1}^{N_T} E \{h_{nn}^* h_{nn}\} = N_R N_T \quad (\text{C.8})$$

where $h_{nn} = 0 \forall n$ except for $n = 1$, i.e. $h_{11} \neq 0$. The capacity of the low-rank MIMO channel can be written as

$$\begin{aligned} C_{low} &= \sum_{n=1}^{\min(N_R, N_T)} \log \left(1 + \frac{\frac{E}{2\sigma^2} \lambda_n^2}{N_T} \right) \\ &\approx \log \left(1 + \frac{\frac{E}{2\sigma^2} \lambda^2}{N_T} \right) \end{aligned} \quad (\text{C.9})$$

The channel behaves like a point to point channel with N_R times the received signal power achieved by Maximum Ratio Combining at the receiver.

Appendix D

Publications

The list in Tab. D.1 reports the publications produced during the period of this PhD Thesis. Some of the works are not thoroughly presented in the previous chapters but are always referred in the text when relevant to the treated matters. Three symbols indicate the level of relevance for each chapter:

Relevance Level	
Core Material	*
Full Relevance	●
Relevance	○

Ref	Title	Chapter				
		1	2	3	4	5
[MMPR07]	D. Marabissi, F. Meucci, L. Pierucci, and L. S. Ronga., "Adaptive selection of MIMO schemes," in <i>IWCMC 2007 MIMO Systems Symposium</i> , Honolulu, Hawaii, USA, Aug. 2007	●	*		○	
[CVM ⁺ 07]	I. Cerutti, L. Valcarenghi, D. Marabissi, F. Meucci, L. Pierucci, L. S. Ronga, P. Castoldi, and E. Del Re, "Enhanced Operation Modes in IEEE 802.16 and Integration with Optical MANs," in <i>Proceedings of the 12th Wireless Communications CNIT Thyrrenian Symposium, 2007</i>	○	●			●
[MP08]	F. Meucci and L. Pierucci, "A real-time siren detector to improve safety of guide in traffic environment," in <i>European Conference on Signal Processing 2008</i> , Lausanne, Switzerland, Aug. 2008					
[CMCP08]	I. Cerutti, F. Meucci, P. Castoldi, and L. Pierucci, "An adaptive Cross-Layer strategy for QoS-guaranteed links in 4G networks," in <i>IEEE Globecom 2008 Wireless Networking Symposium</i> , New Orleans, LA, USA, Nov. 2008	●		*	*	●
[MMAPO8]	F. Meucci, A. Mihovska, B. Anggorojati, and N. Prasad., "Joint Resource Allocation and Admission Control Mechanism for an OFDMA-Based System," in <i>Proceedings of the 11th International Symposium on Wireless Personal Multimedia Communications (WPMC'08)</i> , Lapland, Finland, Nov. September 2008			●	○	●
[CVC ⁺ 09]	I. Cerutti, L. Valcarenghi, P. Castoldi, D. Marabissi, F. Meucci, L. Pierucci, E. Del Re, L. S. Ronga, R. Tka, and F. Kamoun, <i>Current Technology Developments of WiMax Systems</i> . Springer, 2009, Book Chap. Enhanced operation modes in IEEE 802.16 and Integration with optical MANs	○	*			*
[MVM ⁺ 09]	A. Mihovska, F. Velez, F. Meucci, O. Cabral, and N. Prasad, "Multi-Operator resource sharing scenario in the context of IMT-Advanced systems," in <i>Wireless Communication Society, Vehicular Technology, Information Theory and Aerospace & Electronics Systems Technology</i> , Aalborg, Denmark, May 2009					●
[MWP09a]	F. Meucci, S. A. Wardana, and N. R. Prasad, "Data Mules Networks Analysis with Semi-Markov-Process for Challenging Internets and Developing Regions," in <i>Proceedings of the 4th Africa ICT Conference</i> , Kampala, Uganda, May 2009					●
[MWP09b]	F. Meucci, S. A. Wardana, and N. R. Prasad, "Secure Physical Layer using Dynamic Permutations in Cognitive OFDMA Systems," in <i>Proceedings of the IEEE 69th Vehicular Technology Conference: VTC2009-Spring</i> , Barcellona, 2009	●				●
[MPDP09]	F. Meucci, L. Pierucci, E. Del Re, and R. Prasad, "Dynamic allocation of STBC for OFDM based on WSSUS channel model," in <i>Wireless Communication Society, Vehicular Technology, Information Theory and Aerospace & Electronics Systems Technology</i> , Aalborg, Denmark, May 2009	○	*			
[MCM ⁺ 09]	F. Meucci, O. Cabral, A. Mihovska, F. Velez, and N. Prasad, "Spectrum Aggregation with Multi-Band User Allocation over Two Frequency Bands," in <i>IEEE Mobile WiMAX Symposium 2009</i> , Napa Valley, California, USA, Jul. 2009					●
[PMZ ⁺ 09]	N. Pratas, F. Meucci, D. Zrno, N. Prasad, A. Rodrigues, and R. Prasad, "Cogito testbed: Wireless research evolved," in <i>Wireless Communication Society, Vehicular Technology, Information Theory and Aerospace & Electronics Systems Technology</i> , Aalborg, Denmark, May 2009					●
[MPCC10]	F. Meucci, L. Pierucci, I. Cerutti, and P. Castoldi, "Performance of dynamic service addition in mobile WiMAX networks," in <i>submitted to IEEE ICC 2010 - Wireless and Mobile Networking Symposium</i> , Cape Town, South Africa, 2010	○		*		●
[MDCP09]	F. Meucci, E. De Carvalho, and P. Popovski., "Multi-User MIMO ARQ strategies", in submission to <i>IEEE Trans. Wireless Commun.</i>	●		○	*	●

Table D.1: Publication list

List of Figures

1.1	MIMO Capacity with perfectly known CSIT and CSIR	3
2.1	STTD and LSTBC comparison	18
2.2	AAS Frame Structure in IEEE802.16	21
2.3	Alamouti transmission for SCa PHY layer profile	23
2.4	IEEE802.16 AMC System Diagram	25
2.5	Diversity – Spatial Multiplexing Tradeoff	26
2.6	Demmel Condition Number and SM-STBC Switching Condition	29
2.7	Functional Block of MT and BS	30
2.8	AMC Permutation Pilots Pposition	31
2.9	Alamouti and Hybrid STC Schemes Comparison	33
2.10	BER over SNR (dB) of a constrained Adaptive system with Multiplexing Gain of 2	33
2.11	Adaptive MIMO Link Adaptation for upper-bounded BER, Flat Fading Rayleigh Channel	34
2.12	Adaptive MIMO Link Adaptation for upper-bounded BER, ITU Veh channel A	35
2.13	Throughput for ideal channel estimation for a $BER_{max} = 0.01$	35
2.14	Influence of channel estimation over throughput for Alamouti and Spatial Multiplexing schemes	36
2.15	MIMO-OFDM system model with STBC encoder	39
2.16	OFDM Transfer Function Vehicular Channel	40
2.17	Expontential Decaying PDP - subcarriers correlation	42
2.18	Jakes model - subcarriers correlation	43
2.19	Two Rays model - subcarriers correlation	43
2.20	Uniform model - subcarriers correlation	44
2.21	SEP for Alamouti for M-QAM over narrow-band partially cor- related Rayleigh channel	46
2.22	Dynamic Allocation of STFBC	47
2.23	BER improvement for Vehicular Channel at 150 Km/h	48
3.1	ARQ Cross-layer optimisation algorithm	56
3.2	Stop-and-Wait ARQ protocol	56
3.3	ARQ: Percentage of bandwidth saving vs. Γ , for ACL and SL strategies, $\sigma = 0$ (no shadowing)	59
3.4	ARQ: Overall Number of OFDM Symbols used vs. time, for different values of σ	59
3.5	ARQ: ITU vehicular A channel: \bar{D} vs. time, for different QoS requirements of \bar{D}_{max} , $L_{max} = 10^{-4}$	60

3.6	ARQ: Flat fading channel: \bar{D} vs. time, for different QoS requirements of $D_{max}, L_{max} = 10^{-4}$	61
3.7	ARQ: ITU vehicular A channel: \bar{D} vs. time, for different values of σ	61
3.8	ARQ: ITU Vehicular A Vchannel: \bar{D} vs. time, for different values of λ	62
3.9	DSA: Time chart of DSA signalling	64
3.10	DSA: Block Fading: blocking probability vs. SER	69
3.11	DSA: Block Fading: signalling latency vs. SER	70
3.12	DSA: Signalling blocking vs. SNR for different values of v and n	71
3.13	DSA: Admission control blocking vs. SNR for different values of v and n	72
3.14	DSA: Latency vs. SNR for different values of v and n	72
3.15	DSA: Signalling blocking vs. SNR for $v = 5$ km/h, when UL and DL channels are either independents or fully reciprocal	73
3.16	DSA: Latency vs. SNR for $v = 5$ km/h, when UL and DL channels are either independents or fully reciprocal	74
3.17	DSA: Signalling blocking vs. SNR = 20 dB, for $v = 50$ km/h and different values of n and N_{fft}	74
3.18	DSA: Signalling blocking vs. SNR = 20 dB, for $v = 300$ km/h and different values of n and N_{fft}	75
3.19	DSA: Signalling blocking vs. SNR for $v = 5$ km/h and different values T_f	76
3.20	DSA: Latency vs. SNR for $v = 5$ km/h and different values T_f	76
3.21	DSA: Signalling blocking vs. SNR for different values of timeouts and $v = 5$ km/h	77
3.22	DSA: Latency vs. SNR for different values of timeouts and $v = 5$ km/h	77
3.23	DSA: Admission control blocking vs. λ for SNR = 20 dB and $v = 5$ and 300 km/h	78
4.1	MU MIMO-ARQ State Machine Diagram	91
4.2	MIMO-ARQ Receiver Diagram	94
4.3	Alamouti and VBLAST BERs over SNR	104
4.4	Average Rate for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems	106
4.5	Average Delay for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems	106
4.6	Delay Jitter for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems	107
4.7	Average Rate for S1, S2 and MU retransmission strategies	108
4.8	Average Delay for S1, S2 and MU retransmission strategies	108
4.9	Average Rate for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems, $\sigma_{SNR} = 2db$	109
4.10	Average Delay for MU MIMO-ARQ compared with non adaptive and BER-adaptive systems, $\sigma_{SNR} = 2db$	110
4.11	Delay Jitter for MU MIMO-ARQ compared with non adaptive and BER-adaptive system, $\sigma_{SNR} = 2db$	110
4.12	Average Rate for S1, S2 and MU retransmission strategies, $\sigma_{SNR} = 2db$	111
4.13	Average Rate for MIMO-ARQ in MU and SU modes for short channel coherence times	112
4.14	Average Delay for MIMO-ARQ in MU and SU modes for short channel coherence times	112
5.1	IEEE802.16 Protocol Stack	119

5.2	MAC PDU content in 802.16e	120
5.3	Generic MAC Header Content	120
5.4	MAC Fragmentation	121
5.5	MAC Packing	121
5.6	IEEE802.16 ARQ TX state machine	122
5.7	ARQ Types in IEEE802.16	123
5.8	ARQ Packet Rearrangement	124
5.9	Scheduler Input Parameters	125
5.10	DL and UL Schedulers in IEEE802.16	126
5.11	Modified Protocol Stack with per-antenna ARQ	129
A.1	SEP for M-QAM – closed-form derivation	155

List of Tables

2.1	IEEE802.16d/e MIMO support	20
2.2	Upper-Bounded BER MIMO Switching, parameters	32
2.3	ITU Veh Type A Channel model PDP	32
2.4	Upper-Bounded BER MIMO Switching: transmission modes	32
2.5	PDP models and $\phi_h(\Delta f)$ correlation function	42
2.6	STFBC: system parameters	46
3.1	ARQ ACL strategy, system parameters	58
3.2	ARQ ACL strategy, mobility scenarios	59
4.1	MU MIMO-ARQ: case study	96
4.2	MU MIMO-ARQ: configurations	103
4.3	MU MIMO-ARQ: system paramters	104
5.1	IEEE802.16 mandatory QoS calsses and attributes	127
B.1	IMT-A Cell Spectral Efficiency	161
B.2	Cell User Edge Spectral Efficiency	162
B.3	Mobility classes for test environments.	163
B.4	Traffic channels, link data rates	163
B.5	Handover Interruption Times.	163
B.6	VoIP capacity	164
B.7	Frequency Bands for Mobile WiMAX	165
B.8	Mobile WiMAX Normalized Peak Data Rate	166
D.1	Publication list	172

Index

- $Q(x)$, 152
- $Q(x)^2$, 152
- 4G Requirements, xvii

- AAS, *see* WMAX, AAS20
- AMC, 24–25
- ARQ
 - Adaptive Cross-Layer ACL strategy, 53
 - Feedback types, 123
 - Fragmentation and packing, 120
 - Rearrangement, 123
 - State Machine, 122
- Arrival Process, 55

- Capacity
 - Dirty Paper Coding, 6
 - Ergodic, 3
 - MIMO
 - Perfect CSI, 2
 - With Fading, 4
 - Shannon Definition, 3
- Cellular Generations, xvii
- Channel Model
 - ARQ ACL strategy, 57
 - DSA
 - Block Fading, 68
 - Mobility, 71
 - Dynamic MIMO with Upper-Bounded BER, 31
 - MU MIMO-ARQ, 103
 - SU MIMO-ARQ, 86
- Channel Model, STFBC, 39
- Cross Layer Designs, 5
 - ARQ
 - Slow varying Channels, 53
 - ARQ optimisation, 52

- Definitions
 - Diversity Gain, 14
 - Frobenius Norm, 27
 - Optimal Diversity Gain, 26
 - Rate Gain, 18
- Demmel Condition Number, 7
 - Variation over time, 28
- Dirty Paper Coding, 6
- Diversity Order, *see* STBC, Diversity Gain Definition
- DL/UL Channel Reciprocity, 71
- Doppler Spectra, 41
- DSA Protocol, 63

- Activation without MT Knowledge, 67
- Admission Control Blocking, 67
- Model Validation, 68
- Performance, 71
- Signalling Analysis, 65
- Signalling Blocking, 67
- Timeouts, 65
- Timeouts Impact, 75
- Dynamic STBC Allocation, 37

- Fading
 - Av.SEP with Generic Fading, 153
 - Av.SEP with Rayleigh fading, 154
 - Nagakami Statistics, 47
- Frobenius Norm, *see* Definitions, Frobenius Norm

- LSTBC, 13
 - Non Orthogonality, 23
- LTV Systems, 39

- MGF, 45, 151
- MIMO
 - High Rank, 168
 - Impact of channel rank, 5
 - Low Rank, 168
 - ORBF, 6
 - Processing Gain, 2, 11
 - Protocol Gain, 2, 81
 - Waterfilling, 168
 - WiMAX Support, 19–23
- MIMO Switching
 - Based on ED, 26
 - Demmel Condition Number, 28
 - Demmel Number, 7
 - Dynamic DIV/SM, 25
 - Seamlessly, 85
 - SM-STBC, 29–36
 - STBC-SFBC, 37–48
 - Upper-Bounded BER, 7, 29
 - Performance, 45
- Mobile Internet, xvii
- Moore-Penrose Pseudoinverse, 16
- MU Diversity
 - User Selection, RAND/BEST, 8
- MU MIMO, 6
 - Advantages, 2
- MU MIMO-ARQ
 - MU Retransmission Strategies, 8, 89

- OFDMA Frame, 38
- Orthogonal Designs, *see* STBC, Orthogonal Designs
- Parameters
 - ARQ ACL strategy, 57
 - DSA Protocol, 70
 - MIMO Switching for Upper-Bounded BER, 30
 - MU MIMO-ARQ, 103
 - STFBC, 45
- Performance
 - ARQ ACL strategy, 57
 - DSA, 70
 - Dynamic MIMO, 31
- Power Control, 24
- Power Delay Profiles, 41
- Processing Gain, 2, *see* MIMO, Processing Gain
- Protocol Gain, 2, *see* MIMO, Protocol Gain
- Schedulers
 - Integration with MU MIMO-ARQ, 93, 128
 - MT scheduler, 126
 - Priority Based, 52
- Security with subcarrier permutation, 51
- SM, 12
 - first testbed, 1
 - Hybrid SM transmission, 19
 - Optimum Decoding Order, 18
 - OSIC, 18
 - Rate Gain Definition, 18
 - Scheme, 17–19
 - VBLAST Performance, 4
- Spaced-Frequency, Spaced-Time correlation function, 41
- Standards
 - Dispute between WiMAX and LTE-A, xviii
 - IEEE802.16m, 164
- State of the Art
 - MU MIMO, 82
- Statistical Multiplexing, 51
- STBC, 12, 14–15
 - Alamouti Scheme, 15
 - Channel Correlation Coefficient, 15
 - Diversity Gain Definition, 14
 - LSTBC Scheme, 16–17
 - Orthogonal Designs, 14–15
 - Complex Design, rate One, 15
 - Real Design, 15
 - Quasi-Orthogonal Schemes, 16
 - Rate Definition, 14
 - Time Varying
 - SNR Fading Statistic, 44
 - Time Varying Channels
 - Constellation Precoding, 37
- STFBC, 37–38
- SU MIMO-ARQ, 7
 - SNR based design, 86
 - Throughput based design, 86
- Subcarriers Grouping, 37
- SVD, 167
- System Model
 - ACL ARQ strategy, 53
 - DSA protocol, 64
 - Dynamic MIMO, 30
 - MU MIMO-ARQ, 88
 - STFBC over OFDMA, 38
- Time-Variant Channel Impulse Response, 40
- Turbo Coding, 1
- User Selection
 - Integration with MU MIMO-ARQ, 93, 128
- WiMAX
 - MAC Security Sublayer, 51
 - AAS, 20
 - AMC, 24
 - AMC Permutation, 31
 - ARQ, *see* AQ122
 - Dynamic Subcarrier Assignment, 24
 - PHY Configurations, 20
 - QoS Classes, 127
 - Schedulers, 125
 - STDD in MC, 22
 - STTD in SC, 22
 - UL Collaborative SM, 23
- WSSUS Channel, 38