D3.5 Location Based Cross-Layer Optimisation for PHY/MAC: Final

Nielsen, Jimmy Jessen; Monteiro, Valdemar; Rodriguez, Jonathan; Nascimento, Alberto; Dammann, Armin; Slock, Dirk; Zhang, Sheng; Yi, Na; Ma, Yi

Publication date:
2010

Document Version
Early version, also known as pre-print

ICT- 217033 WHERE

D3.5 Version 1.0

Location Based Cross-Layer Optimisation for PHY/MAC: Final

Contractual Date of Delivery to the CEC:	*M28*
Actual Date of Delivery to the CEC:	*22.05.2010*
Editor:	Valdemar Monteiro
Author(s):	Jonathan Rodriguez, Valdemar Monteiro, Alberto Nascimento, Jimmy Jessen Nielsen, Na Yi, Yi Ma, Armin Dammann, Dirk Slock, Sheng Zhang
Participant(s):	*AAU, DLR, IT, UNIS, EUR*
Work package:	*WP3*
Security:	**Public**
Nature:	**RE**
Version:	1.0
Total number of pages:	82

Abstract: This deliverable describes the work done by the WHERE partners in Work Package 3 Task 2. The topic of this deliverable is on using positioning information for allowing cross-layer optimization on the phyisical (PHY) and medium access control (MAC) layers. Initially, we propose a relaying scheme, which uses location information to allow simulationeous downlink transmissions from an access point to mobile destination nodes, by jointly optimizing the relay selection and transmit power in terms of maximizing the expected MAC layer throughput. Secondly, a link level interface model consisting of frame error rate results for different code rates and modulation alphabets at different receiver locations is provided for OFDM in cases with cellular diversity. These results are intended for use in upper layer simulations. After this, a detailed Dynamic Resource Allocation architecture for mobile WiMAX, which uses location information to achieve SDMA and improve overall throughput by use of beam forming is presented. Underlay cognitive radio (UCR) systems are investigated next, by assuming the availability of full or partial multiuser channel quality information (CQI). A new UCR physical-layer model is established through exploiting the location-aided approach for the scenario with partial multiuser CQI. Finally, SDMA beamforming design based on location information is considered. Sum rate degradations due to non line of sight components and position errors are analyzed. A simulation comparison with an LTE based approach is provided, showing significant potential for the position based approach.

Keyword list: Positioning, Location, PHY/MAC Cross-Layer Optimisation, Radio Resource Management, Link Level Interface, Dynamic Resource Allocation, Scheduling, Underlay Cognitive Radio, SDMA, Relay

Disclaimer:

Executive Summary

This document is the final deliverable for the Work Package 3 Task 3, which concerns location based cross-layer optimisation for PHY/MAC. The contributions in this document are organized in chapters and related to different technologies, including joint relay-selection and power adaptation, location dependent link level model to be used as input for higher layers simulations, a proposed Dynamic Resource Allocation (DRA) model based on Space Division Multiple Access (SDMA) for the mobile WiMAX, and location-based underlay cognitive radio.

In chapter 2 we investigate the achievable MAC throughput gain of letting two relay to destination transmissions occur simultaneously, when assuming that location information is available to the access point (AP) through a periodic collection mechanism. Further, we investigate if additional power adaptation can be used to enhance performance even further. In chapter 3 a link level interface based on the cellular Alamouti technique is investigated for convolutionally coded OFDM and frame error rates for different code rates and modulation alphabets at different receiver locations are provided. The frame error rates are intended to be used as a location dependent link level model input for higher layer simulations. Chapter 4 describes all the steps followed in the implementation of SDMA, in an 802.16e DRA module. Location information is used to estimate the direction of arrival of each mobile related to the base-station and it is assumed that the beamforming algorithm is optimal in steering one spatial beam into the direction of the desired mobile. In chapter 5 we study several fundamental physical-layer issues about location aided Gaussian Underlay Cognitive Radios (UCR). The contributions include full and partially knowledge of multiuser Channel Quality Information (CQI) and its relationship with primary and secondary users; the modeling of the UCR in the presence of full knowledge of multiuser CQI; and finally we propose new spectrum-access approaches for various underlay cognitive radio modes by assuming the availability of CQI about the secondary-primary user link. Chapter 6 considers again location based SDMA and provides sum rate degradation calculations due to the channel deviating from a pure line-of-sight case (Ricean channel) and due to position information errors. Comparitive simulations are provided for location based SDMA and an LTE style approach based on channel quantization and feedback. Already in a simple 2-user scenario as considered here, a factor two gain in sum rate can be obtained in the location based approaches over an LTE style approach. The chapter 7 concludes this document highlighting the main work carried out and the achievements.

Main conclusions are that a position-based relaying scheme where simultaneous relay to destination transmissions are used, can be beneficial in some cases, despite the cross-interference occurring when transmitting simultaneously. The proposed algorithm, named SimTX, jointly determines the best combination of relays and transmit power for the two simultaneous relay to destination transmissions. Results show that this approach can improve the throughput by 19% in the considered scenario when comparing to the cases where normal two-hop relaying would be used. In the link layer model technique investigated results provided frame error rates for a convolutionally coded OFDM system applying different modulation constellations and code rates were presented. The simulations have been carried out using real propagation measurements (from Work Package 4), rather than stochastic channel models. The results can serve as a basis for the generation of lookup tables for system level simulations, which then provide system performance measures for a real indoor scenario. Concerning DRA, an implementation of a SDMA-based DRA architecture for the mobile WiMAX was proposed and a set of system level simulations were performed. The results compare the gains achieved over simple scenarios where a non SDMA-based DRA architecture is implemented. For the proposed SDMA scheme the achieved gain on the throughput can reach up to 50 Mbps, a gain of roughly 38 Mbps over the simple non SDMA-based DRA scheme. In the cognitive radio fundamental relationship between the secondary user's achievable rate and capacity penalty to the primary user has been studied in four carefully classified UCR modes for full provided multiuser CQI. For partial multiuser CQI a new physical-layer model is established through exploitation of the location-aided approach and new spectrum access and power allocation strategies are investigated. Numerical examples are provided to show the performance of the UCR with full multiuser CQI and the proposed approach with partial multiuser CQI.

Authors

Partner	Name	Phone / Fax / e-mail
AAU		
	Jimmy Jessen Nielsen	Phone: +45 9940 9867 Fax: N/A e-mail: jjn@es.aau.dk
DLR		
	Armin Dammann	Phone: +49 8153 28 2871 Fax: +49 8153 28 1871 e-mail: Armin.Dammann@DLR.de
EUR		
	Dirk Slock	Phone: +33 4 9300 8106 Fax: +33 4 9300 8200 e-mail: slock@eurecom.fr
	Sheng Zhang	Phone: +33 4 9300 8249 Fax: +33 4 9300 8200 e-mail: sheng.zhang@eurecom.fr
IT		
	Jonathan Rodriguez	Phone: +351 234 377900 Fax: +351 234 377901 e-mail: jonathan@av.it.pt
	Valdemar Monteiro	Phone: +351 234 377900 Fax: +351 234 377901 e-mail: vmonteiro@av.it.pt
	Alberto Nascimento	Phone: +351 234 377900 Fax: +351 234 377901 e-mail: anascimento@av.it.pt
UNIS		
	Na Yi	Phone : +44 1483 68 4703 Fax : +44 1483 68 6011 e-mail : n.yi@surrey.ac.uk
	Yi Ma	Phone : +44 1483 68 3609 Fax : +44 1483 68 6011 e-mail : y.ma@surrey.ac.uk

Table of Contents

Executive Summary 2

Authors 3

Table of Contents 4

List of Acronyms and Abbreviations 6

1 Introduction 8

2 Joint relay selection and power adaptation for enabling simultaneous relay-transmissions 10

2.1 Introduction 10
2.2 Scenario 10
2.3 Methodology 12
2.3.1 Form pairs 12
2.3.2 Choose potential relays 12
2.3.3 Find max-throughput configuration 12
2.4 Throughput Model 13
2.4.1 BER calculation 13
2.4.2 FER calculation 13
2.4.3 Throughput calculation 15
2.5 Results and discussion 16
2.5.1 Throughput model results 16
2.5.2 Simultaneous transmissions results 18
2.6 Conclusions and Future Work 21

3 Link Level Interface Modelling Results for OFDM with Cellular Diversity 23

3.1 Cellular Diversity 23
3.2 Evaluation 23
3.2.1 Simulation Setup 23
3.2.2 Results 26
3.3 Summary 28

4 SDMA Dynamic Resource Allocation for the Mobile WiMAX (802.16e) 29

4.1 Introducation 29
4.2 Utility-based packet scheduler 30
4.3 Spatial Beamforming 31
4.4 SDMA Scheme for IEEE802.16e Mobile WiMAX 32
4.4.1 User Spatial Separability 33
4.4.2 SINR Estimation after Performing Beamforming 34
4.5 Proposed SDMA-Enabled DRA Architecture 35
4.5.1 Computation of Users Correlation 36
4.5.2 Computation of SINR for Resources in SDMA and in non-SDMA Zones 37
4.6 Joint Scheduling and Resource Allocation Algorithm Using SDMA Multiple Access 38
4.7 Results 40
4.7.1 Performance for Full Queue Traffic Users 41
4.7.2 Performance for VoIP and WWW Traffic Users 42
4.7.3 Performance for Voice Over IP Traffic Model 45
4.7.4 World Wide Web Traffic Model 47
4.8 Related Work 50
4.9 Conclusion 50

5 Location Aided Underlay Cognitive Radio 52

5.1 Motivation and Objective 52
5.2 System Model 54
5.2.1 Two-User UCR with Full Multiuser CQI 54
5.2.2 Two-User UCR with Partial Multiuser CQI 55
5.3 Main Results 56
5.3.1 The Individual Decoding (ID) Mode 56
5.3.2 The Secondary-User Side Multiuser Decoding (SSMD) Mode 57
5.3.3 The Primary-User Side Multiuser Decoding (PSMD) Mode 57
5.3.4 The Two Side Multiuser Decoding (TSMD) Mode 58
5.4 Numerical Resuts 59
5.5 Summary 62

6 Location based SDMA 63

6.1 Location aided Resource Allocation 63
6.2 SDMA considerations 63
6.3 SDMA via user feedback: ZF BF sum rate 65
6.3.1 Background and Motivation 65
6.3.2 System Model 66
6.3.3 Transmission Scheme with Oblivious Users 67
6.3.4 Uplink Training Phase 67
6.3.5 BS Transmission Strategy: ZF Precoding with Semi-Orthogonal User Selection (SUS) . 68
6.3.6 Downlink Training Phase 68
6.3.7 Coherent Data Phase 68
6.3.8 Sum Rate Lower Bound 69
6.4 Location based SDMA: ZF BF sum rate 70
6.4.1 Effect of LOS deviation on ZF BF sum rate 71
6.4.2 Effect of position error on ZF BF sum rate 71
6.5 Comparative Simulations of location based SDMA vs LTE quantization-feedback based SDMA 72
6.5.1 The Location Based LOS-Channel Model 72
6.5.2 The Location Based Linear Precoder 72
6.5.3 Performance Evaluation 73
6.5.4 Summary 76

7 Conclusions 77

References 79

List of Acronyms and Abbreviations

Term	Description
AAS	Adaptive Antenna System
ACK	Acknowledgement
AoD	Angle of Departure
AP	Access Point
BC	Broadcast Channel
BER	Bit Error Rate
BF	BeamForming
BPSK	Binary Phase Shift Keying
CAT	cellular Alamouti technique
CDMA	Code Division Multiple Access
CIR	channel impulse response
CQI	Channel Quality Information
CSI	Channel State Information
CSIT	CSI at the Transmitter
CTS	Clear To Send
CW	Contention Window
DoA	Direction of Arrival
DCF	Distributed Coordination Function
DL	DownLink
DPC	Dirty Paper Coding
DRA	Dynamic Resource Alloacation
EDCA	Enhanced Distributed Channel Access
FDD	Frequency-Division Duplex
FDMA	Frequency Division Multiple Access
FER	Frame Error Rate
GIC	Gaussian Interference Channel
GMAC	Gaussian Multiple-Access Channel
HOL	Head of Line
ID	Individual Decoding
LB	Lower Bound
LL	Log-Likelihood
LOS	Line-of-Sight
LTE	Long Term Evolution
MAC	Medium Access Control
MF	Matched Filter
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
ML	Maximum-Likelihood
MSDU	MAC Service Data Unit
MTMR	Multiple Transmitter Multiple Receiver
MU	Multi-User
NADA	Narrow Angle of Departue Aperture
NRT	Non-Real-Time
OFDM	Orthogonal Frequency Multiple Access
PHY	Physical
PSMD	Primary-User Side Multiuser Decoding
PUSC	Partial Usage sub-channelization Scheme
RT	Real-Time
RTS	Ready To Send
SDMA	Spatial Division Multiple Access
SINR	Signal to Interference and Noise Ratio
SNR	Signal to Noise Ratio
SSMD	Secondary-user Side Multiuser Decoding
STBC	Space-Time Block Code
SU	Single User

SUS	Semi-Orthogonal User Selection
TDD	Time Division Duplex
TSMD	Two Side Multiuser Decoding
UCR	Underlay Cognitive Radio
UL	UpLink
VoIP	Voice over IP
ZF	Zero-Forcing

1 Introduction

In the last decade mobile communications have witnessed an explosive increase in the amount of users and in the amount of data transferred over the air interface. High bandwidth demanding multimedia applications, originally intended for the fixed Internet, are currently being envisioned for the mobile wireless environment due to technology advances, which result in a more efficient utilization of the scarce bandwidth and the increase in the available capacity for data.

The WHERE project aims to perform the integration of communications and navigation, in the ubiquity paradigm of the wireless communications systems, and the WP3 mainly investigates advanced protocol/algorithms for OFDM based radio access networks exploiting positioning information. The task T3.2 focus on cross-layer optimisation strategies for radio resource management protocols that exploit positioning data, based on the underlying PHY layer. This deliverable corresponds to advanced stage of the task and positioning information is exploited for: enhancing relay systems, improving throughput through SDMA, and study PHY-layer issues about location aided cognitive radios.

Relaying system have been proposed and studied in order to improve the coverage, as meaning of better cell/AP coverage and throughput. Two-hop relaying as considered in [ZC07] and [LTN+07] can be used as a means to achieve better performance towards such destination nodes located close to the boundary of the AP's reach. Thereby the range of the AP can be extended, though, at the cost of a lower throughput. This idea of having simultaneous, spatially well-separated transmissions has been used for instance in [LLH06] where the authors improve the throughput of downstream multihop routed traffic near internet gateways by constructing a coordinate-based grid, which is used to determine if transmissions are well enough separated spatially to allow temporal concurrency.

The benefit incurred with the use of Adaptive Antenna Systems (AAS) is its ability to reduce interference by steering the beam to a specific user. Adaptive antenna elements can be used to distort the radiation pattern produced by an antenna array at the base station. The idea is to focus (beamform) the transmitted signal energy in the direction of the intended mobile receiver and, at the same time, steer neutralize signal energy in directions of co-channel mobile. This results in an enhancement of the perceived Signal to Interference plus Noise Ratio (SINR) at the desired mobile’s receiver. Comparable approaches are currently being standardized by 3GPP for UMTS or by the IEEE for 802.11n. These advanced antenna techniques have a significant impact on the capacity and service quality provided by wireless links and the efficient use of the available spectrum [CCP08].

Cognitive radio brought new perspective in the spectrum allocation. A good tutorial about cognitive radio techniques can be found in [ZS07] that is focused on signal-processing perspective, and in [JSM09] that is focused on information-theoretic perspective.

This deliverable includes contribution of the partners that works specifically in the T3.2 task and the contributions from each partner is organized in chapters.

In the chapter 2 we investigate the achievable MAC throughput gain of letting two relay to destination transmissions occur simultaneously, when assuming that location information is available to the AP through a periodic collection mechanism. Additionally we investigate whether additional power adaptation can be used to enhance performance even further. Results show that relaying schemes where simultaneous relay to destination transmissions are used are beneficial despite the cross-interference occurring when transmitting simultaneously in positioning–based relaying. The proposed algorithm, named SimTX, jointly determines the best combination of relays and transmit power for the two simultaneous relay to destination transmissions. Numerical results show that this approach can improve the throughput by 19% in the considered scenario when compared to the cases where normal two-hop relaying would be used.

Chapter 3 investigates a link level interface based on the cellular Alamouti technique for convolutionally coded OFDM and frame error rates for different code rates and modulation alphabets at different receiver locations are provided. The frame error rates are intended to be used as a location dependent link level model input for higher layer simulations. Provided results include frame error rates for a convolutionally coded OFDM system applying different modulation constellations and code rates, with simulations been carried out using real propagation measurements rather than stochastic channel models. These results can serve as basis for the generation of lookup tables for system level simulations, which then provide system performance measures for a real indoor scenario.

Chapter 4 describes all the steps followed in the implementation of SDMA in 802.16e based DRA module. Location information is used to estimate the direction of arrival of each mobile related to the

base-station and is assumed the beamforming algorithm is optimal in steering one spatial beam into the direction of the desired mobile. A full implementation of a SDMA-based DRA architecture for the mobile WiMAX was proposed and a set of system level simulations was performed to and results compare the gains achieved over simple scenarios where a non SDMA-based DRA architecture is implemented. Achieved throughput can reach up to 50 Mbps, a gain of roughly 38 Mbps over the simple non SDMA-based DRA scheme.

In the chapter 5 several fundamental physical-layer issues about location aided Gaussian Underlay Cognitive Radios are studied. The manifold contributions include full and partially knowledge of multiuser Channel Quality Information (CQI) and its relationship with primary and secondary users, the modelling of the UCR in the presence of full knowledge of multiuser CQI and propose new spectrum-access approaches for various underlay cognitive radio modes by assuming the availability of CQI about the secondary-primary user link. Fundamental relationship between the secondary user's achievable rate and capacity penalty to the primary user have been studied in four carefully classified UCR modes for full and partially provided multiuser CQI. For partial multiuser CQI a new physical-layer model through exploitation of the location-aided approach is established and new spectrum access and power allocation strategies investigated. Numerical examples are provided to show the performance of the UCR with full multiuser CQI and the proposed approach with partial multiuser CQI.

Chapter 6 considers again location based SDMA. The importance of precise channel state information at the transmitter (CSIT) is discussed and the standard mechanism of obtaining such CSIT is described and analyzed in detail for the TDD case (involving uplink training and channel reciprocity exploitation – the overall conclusions are similar in the FDD case). The operation of SDMA based on LOS channel models and position information is discussed and sum rate degradation calculations are provided for the case of the channel deviating from a pure line-of-sight case (Ricean channel) and for position information errors. Comparitive simulations are provided for location based SDMA and an LTE style approach based on channel quantization and feedback. Already in a simple 2-user scenario as considered here, a factor two gain in sum rate can be obtained in the location based approaches over an LTE style approach. The discrepancy can be expected to increase for the case of more users.

Chapter 7 concludes this document highlighting the main work performed and the achievements.

2 Joint relay selection and power adaptation for enabling simultaneous relay-transmissions

2.1 Introduction

In mobile wireless networks where the access points serve as internet gateways, the typical bottleneck when users are downloading and streaming content from the internet, is the AP [HT08][LTN+07]. In particular, performance degradations will occur for all users in a wireless network if one or more user nodes are forced to use a low bit-rate mode and experience retransmissions due to bad link quality as argued in [LTN+07]. This type of bad quality links typically exist when a user node is placed far away from the AP. In such cases, two-hop relaying as considered in [ZC07] and [LTN+07] can be used as a means to achieve better performance towards such destination nodes located close to the boundary of the AP's reach. Thereby the range of the AP can be extended, though, at the cost of a lower throughput.

In such cases, where the relay to destination transmission are at a distance from the AP, there may be cases where two or more relay to destination transmissions could be ongoing simultaneously if they are spatially well-separated in the sense that interference is kept at an acceptable level.

This idea of having simultaneous, spatially well-separated transmissions has been used for instance in [LLH06] where the authors improve the throughput of downstream multihop routed traffic near internet gateways by constructing a coordinate-based grid, which is used to determine if transmissions are well enough separated spatially to allow temporal concurrency. Their approach requires a stable network and uses RSS measurements to construct a virtual coordinate system of interference regions. This means that nodes cannot be mobile when used with this protocol.

Similarly, in [KLH06] the authors show that a higher capacity and performance improvements of up to 22 % can be achieved in a scenario with peer to peer transmission between neighbour nodes, by local tuning of transmit power, bit rate and carrier sense threshold based on the perceived interference level.

In [HT08] the CCMAC protocol is proposed, which is able to enhance network throughput for uplink transmissions to an AP. The scheme allows use of direct transmissions, two-hop transmissions, with and without coded cooperation. The scheme uses a Partially Observable Markov Decision Process model to schedule data transmissions, select transmission mode of nodes as well as select a suitable bit rate. An SNR-based cooperation table is used to keep track of which bit-rate can be used towards neighbours as in the CoopMAC protocol [LTN+07]. As the protocol relies on collecting information from previous transmissions as CoopMAC, we expect that it is similarly sensitive to mobility. A possible solution would be to exchange position information between nodes as discussed in [NMS10b] as it uses less overhead than link measurements. Further, position information would make it easier to predict interference.

This way of exploiting positioning information is explored in [NAA05]. Here an extension to the 802.11 DCF is proposed where location information is included in the 4-way RTS-CTS-DATA-ACK handshake. This makes wireless nodes able to better assess when they are able to transmit by use of a propagation model. Hereby, they reduce the amount of unnecessary blockings due to the exposed node problem in large scenarios with many node to node transmissions, which increases throughput by up to 22 %.

In this work we investigate the achievable performance gain of letting two relay to destination transmissions occur simultaneously, when assuming that location information is made available to the AP through a periodic collection mechanism as presented in [NMS10b]. Additionally we investigate if additional power adaptation can be used to enhance performance even further. Our work addresses the theme of this deliverable in the sense that we use location information to jointly optimize the relay selection (MAC) and transmit power (PHY) in terms of maximizing the expected MAC layer throughput.

2.2 Scenario

The considered case where two relay to destination transmissions are occuring simultaneously is sketched in Figure 2.1. The first two transmissions (AP to relays) are always done at highest transmit power level. Only for the simultaneous transmissions we assume that the relay nodes can choose from a fixed and discrete set of power levels (see Table 2.1) , as this is a means for limiting the cross-interference of the simultaneous transmissions. We assume that the MAC protocol is a modification of the 802.11 DCF, where the AP informs the relays R1 and R2, that they should initiate the simultaneous transmissions to D1 and D2, immediately after overhearing the ACK returned from R2 to the AP, that is without first

contending for access. This kind of privileged access can be gained by using a shorter inter-frame spacing, as in the EDCA mode of IEEE 802.11.

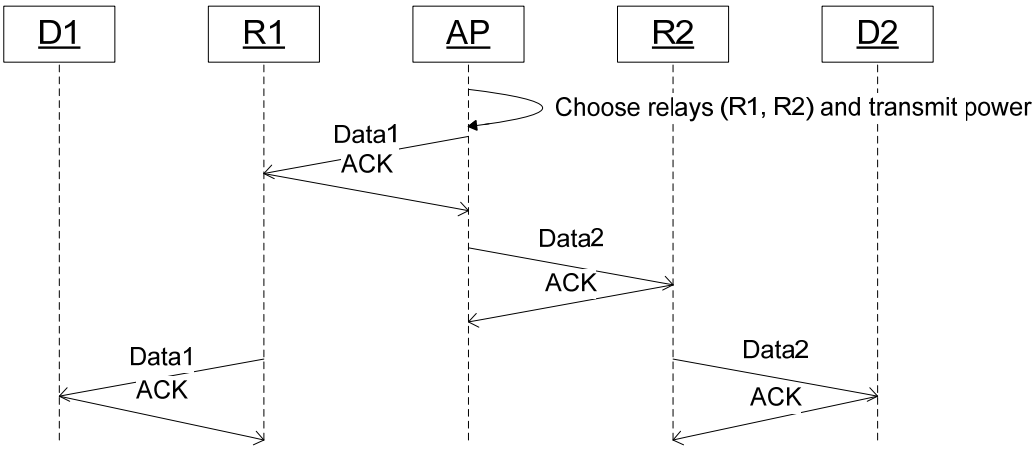

Figure 2.1: D1 and D2 are the destination nodes, R1 and R2 are the relays and AP is the source of the transmissions.

For simplicity we assume that all transmission are of the same payload length and that all entities use the same fixed modulation and coding scheme as specified in Table 2.1.

Parameter	**Value**
P_{tx} levels available	{0, 5, 10, 20, ... , 90, 100} mW
Path loss exponent (n)	2.9
$N0_{\text{floor}}$	$-86dBm$
Ricean K-factor (K)	15
Bit rate	6 Mbit/s
Modulation scheme	BPSK
Max no. of MAC retransmission (R)	8
B_{MSDU}	1024 bytes
Scenario size	100m x 100m
AP position	(0,0)
Number of nodes	30
Simulation runs	250

Table 2.1: Simulation parameters

An example of the possible node layout is shown in Figure 2.2. Here, the two destination nodes are shown in solid, and the corresponding candidate relays are shown with plusses and crosses.

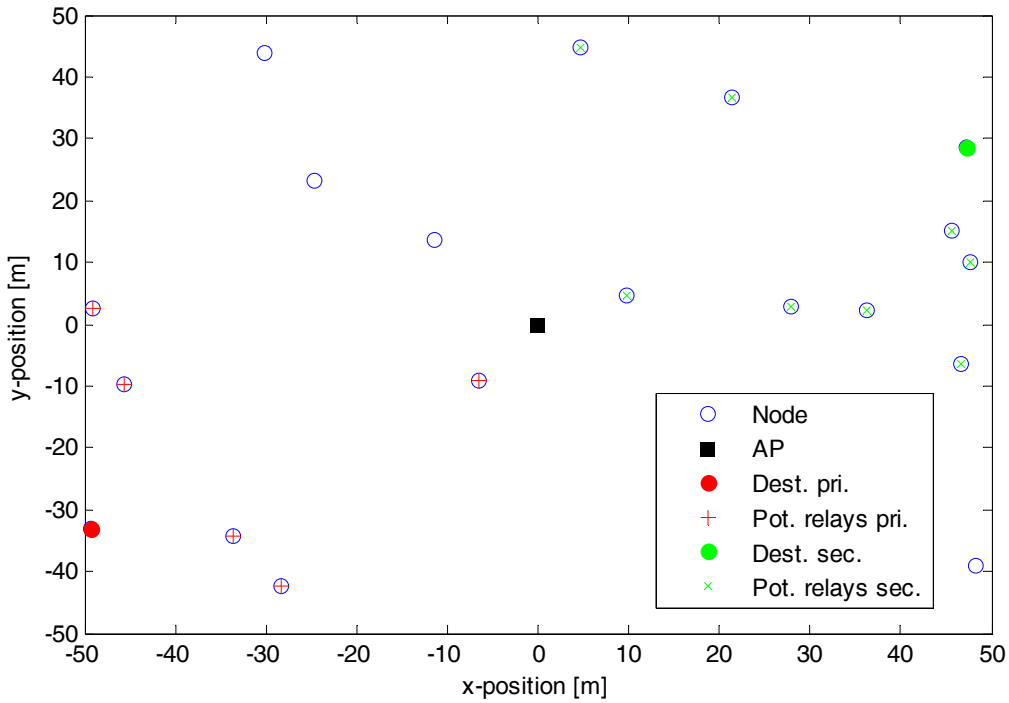

Figure 2.2: Example of the node layout, with destination nodes and potential relays.

2.3 Methodology

The proposed scheme for joint relay selection and power adaptation consists of three steps, namely: form pairs of destination nodes; pick out potential relays for these destinations; and find max-throughput configuration of relay nodes and transmit powers. These steps are explained in details in the following.

2.3.1 Form pairs

We first pick a random node in the scenario as the primary destination node. We are not very interested in the cases where direct transmissions are always clearly beneficial, so we pick the primary destination randomly from nodes that are at least 25m from the AP. Since we want to exploit the benefit of simultaneous transmissions in the cases where it is possible, we also consider a secondary destination node. We imagine that the AP has a queue of packets to deliver, and that the AP knows the positions of the corresponding destination nodes. Assuming that all nodes in the scenario are equally active and thus have at least one downlink transmission scheduled; we assume that we can pick out and upgrade a scheduled transmission to be the secondary transmission.

As we are interested in limiting the amount of interference, we should select the secondary destionation as a node that is spatially well separated from the primary destination. Initially, we simply consider the node that is closest to the coordinate obtained from mirroring the primary node's coordinates in the x and y axes: $(-x_{\text{pri}}, -y_{\text{pri}})$.

2.3.2 Choose potential relays

A group of potential relay nodes are selected for both the primary and the secondary destinations. The criteria for choosing a relay node is that both the distance to the AP and the distance to the destination are less than the distance between the AP and the destination node. That is:

$$d_{\text{AP-d}} > d_{\text{AP-r}} \quad \wedge \quad d_{\text{AP-d}} > d_{\text{r-d}} \tag{2.1}$$

2.3.3 Find max-throughput configuration

In order to find the combination of choices of relay nodes for the two destinations and transmit power levels used by the two relays, we iterate over all combinations of primary relay, secondary relay, transmit power of primary relay, and transmit power of secondary relay; calculate the throughput of each combination using the model described in the following section 2.4; and select the combination that results in the highest throughput.

2.4 Throughput Model

The purpose of the model is to be able to calculate the estimated throughput for different node positions and transmission power settings. The estimated throughput is then used to guide the selection of the most suitable relay nodes and transmit power levels.

First, we derive a model of the BER, given node positions and transmit power levels. From this BER we estimate frame error rate (FER) and finally the the throughput delivered by the MAC layer.

2.4.1 BER calculation

The received power is calculated based on the path-loss model from [DRX98]:

$$P_{\mathrm{rx}}(d)\,[dBm] = P_{\mathrm{tx}} + PL(d_0) + 10n\log_{10}\left(\frac{d}{d_0}\right) \tag{2.2}$$

Where P_{tx} is the transmit power, $P_{\mathrm{rx}}(d)$ is the received power at the receiver, d is the distance between transmitter and receiver, $PL(d_0)$ is the path loss at a reference distance $d_0 = 1m$, and n is the path loss exponent.

In cases where only 1 entity is transmitting at a time, we consider the SNR, which is calculated as:

$$\gamma_{SNR}\,[dB] = 10\log_{10}\left(\frac{P_{\mathrm{rx}}(d)\,[mW]}{N_{\mathrm{floor}}\,[mW]}\right) \tag{2.3}$$

where $P_{\mathrm{rx}}(d)$ is the received power calculated above and N_{floor} is the assumed noise floor, both converted to mW. For the simultaneous transmissions, we take interference into account and consider γ as the SINR instead of the SNR:

$$\gamma_{SINR}\,[dB] = 10\log_{10}\left(\frac{P_{\mathrm{rx}}(d_{\mathrm{m}})\,[mW]}{P_{\mathrm{i}}(d_{\mathrm{i}})\,[mW] + N_{\mathrm{floor}}\,[mW]}\right) \tag{2.4}$$

where $P_{\mathrm{rx}}(d_{\mathrm{m}})$ is the power received at the receiver from the main transmitter at distance d_{m} from the receiver, whereas $P_{\mathrm{i}}(d_{\mathrm{i}})$ is the power received from the interfering transmitter at distance d_{i} from the receiver. The SNR/SINR values are then converted into average BER given the BPSK modulation scheme and the Ricean fading model, using the matlab function berfading.

2.4.2 FER calculation

In the following header and preamble bits are included in N_{data} bits. The sizes of the data frame and ack frame are given in Table 2.2. Since we use the basic rate (6Mbit/s) for data transmissions, we can assume that the same modulation/coding scheme is used and thus the same BER applies.

N_{ACK}	N_{data}
112	$(36 + \mathrm{MSDU}) \cdot 8$

Table 2.2: Frame sizes in bits.

T_{SIFS}	T_{DIFS}	T_{ACK}	T_{data}
9	34	44	$20 + 4\lceil (16 + 6 + 8 \cdot (34 + MSDU))/24 \rceil$

Table 2.3: Frame durations in µs for 802.11 MAC DCF acknowledged mode

We assume that the main cause of bit errors when allowing simultaneous transmissions is interference, and not collisions as considered in [B00] where a single collision domain is considered and simultaneous transmissions are always seen as collisions. Therefore, we have 3 possible outcomes of a transmission attempt: 1) successful reception of DATA and ACK (s), 2) failed during reception of DATA (fd), and 3) failed during reception of ACK (fa).

Assuming a constant BER which we denote P_b, the outcomes have the following probabilities:

$$P_s = (1-P_b)^{N_{data}+N_{ACK}} \tag{2.5}$$

$$P_{fd} = 1-(1-P_b)^{N_{data}} \tag{2.6}$$

$$P_{fa} = (1-P_b)^{N_{data}} \cdot (1-(1-P_b))^{N_{ACK}} \tag{2.7}$$

The failed outcomes differ in time as the ACK is not transmitted if the receiver cannot decode the data. Therefore we calculate the average time of the failed outcomes by normalizing the probabilities of the two failed outcomes and weight the time of the outcomes accordingly. The normalized probabilities are:

$$P_{fdn} = \frac{P_{fd}}{P_{fd}+P_{fa}} \tag{2.8}$$

$$P_{fan} = \frac{P_{fa}}{P_{fd}+P_{fa}} \tag{2.9}$$

From these we obtain the following average transmission times:

$$T_s(r) = T_{data} + T_{SIFS} + T_{ACK} + T_{DIFS} + \overline{T_{BO}}(r) \tag{2.10}$$

$$\overline{T_f}(r) = (T_{data} + T_{DIFS} + \overline{T_{BO}}(r)) \cdot P_{fdn} +$$
$$+ (T_{data} + T_{SIFS} + T_{ACK} + T_{DIFS} + \overline{T_{BO}}(r)) \cdot P_{fan} \tag{2.11}$$

Here, $\overline{T_{BO}}(r)$ is the average back-off time, which depends on the number of the current retry attempt r. Therefore also the transmission time variables $T_s(r)$ and $\overline{T_f}(r)$ depend on *r*. According to [IEEE07], the contention window CW in IEEE 802.11 is a uniform random value between $CW_{min} = 15$ and $CW_{max} = 1023$. For each consecutive retransmission attempt the CW is set according to:

$$CW(r) = min(1023, 2^{4+r} - 1) \tag{2.12}$$

We assume the average waiting time due to back-off is:

$$\overline{T_{BO}}(r) = T_{slot} \cdot \frac{CW(r)}{2} \tag{2.13}$$

where $T_{slot} = 10\mu s$ is the slot time used in IEEE 802.11a.

As we have now determined the time and probability of a single successful or failed transmission, we now derive the expected throughput delivered by the MAC layer service, when taking MAC layer retransmissions into account. In IEEE 802.11 the default maximum number of retransmission attempts, here denoted R is 8. After R attempts the frame transmission fails and an error will be returned from the MAC layer. In this work we only consider the MAC throughput, which may be different from the throughput achieved by overlying transport protocols and applications, due to for example time-out mechanisms as used in TCP to judge when a segment has been lost and needs to be retransmitted [KM05]. The time that a transmission attempt takes is represented by the random variable T_{tx}. The value of this variable is presented below. Here we need to distinguish whether the last retransmission attempt is successful. Let *n* be the number of retries that are needed. Then the value of the random variable is:

$$T_{\text{tx}}(n) = \begin{cases} \left[\sum_{r=0}^{n} T_{\text{f}}(r)\right] + T_{\text{s}}(n+1) & for \begin{cases} 0 \le n \le R-1 \\ n = R,\ success \end{cases} \\ \sum_{r=0}^{n+1} T_{\text{f}}(r) & for \quad n = R,\ failure \end{cases} \tag{2.14}$$

where the probabilities of each of these outcomes are:

$$P_{\text{tx}}(n) = \begin{cases} (1-P_{\text{s}})^{n} \cdot P_{\text{s}} & for \begin{cases} 0 \le n \le R-1 \\ n = R,\ success \end{cases} \\ (1-P_{\text{s}})^{n+1} & for \quad n = R,\ failure \end{cases} \tag{2.15}$$

From this, we can compute the expected time spent on a transmission attempt as:

$$E[T_{\text{tx}}] = \left[\sum_{n=0}^{R-1} T_{tx}(n) P_{tx}(n)\right] + T_{tx}^{s}(R) P_{tx}^{s}(R) + T_{tx}^{f}(R) P_{tx}^{f}(R) \tag{2.16}$$

Where $T_{tx}^{s}(R)$, $P_{tx}^{s}(R)$, $T_{tx}^{f}(R)$, and $P_{tx}^{f}(R)$ are the expected spent transmission time per attempt and frame delivery probabilities for the successful and failed cases where $n=R$. Also we can calculate the probability of a frame being delivered successfully by the MAC layer as:

$$P_{\text{suc}} = 1 - (1-P_{s})^{R+1} \tag{2.17}$$

2.4.3 Throughput calculation

For comparing the considered algorithms: direct transmission, relayed transmission, and simultaneous transmission, we need to calculate the average throughput obtained for the primary (pri) and secondary (sec) transmissions. Keeping in mind that throughput is $\frac{\text{Transmitted data}}{\text{Transmission time}}$, we can calculate the throughput in *Mbit/s*, given the payload size B_{MSDU}=1024 bytes. For the direct algorithm, the throughput is:

$$S_{\text{dir}} = \frac{(P_{\text{suc}}^{\text{pri}} + P_{\text{suc}}^{\text{sec}}) \cdot B_{\text{MSDU}} \cdot 8 \cdot 10^{-6}}{E[T_{\text{tx}}^{\text{pri}}] + E[T_{\text{tx}}^{\text{sec}}]} \tag{2.18}$$

In the following, we use the numeric indices 1 and 2 to indicate the AP to relay and relay to destination transmissions, respectively. The throughput for the relaying algorithm in Mbit/s is calculated as:

$$S_{\text{rel}} = \frac{(P_{\text{suc}}^{\text{pri},1} P_{\text{suc}}^{\text{pri},2} + P_{\text{suc}}^{\text{sec},1} P_{\text{suc}}^{\text{sec},2}) \cdot B_{\text{MSDU}} \cdot 8 \cdot 10^{-6}}{E[T_{\text{tx}}^{\text{pri},1}] + E[T_{\text{tx}}^{\text{pri},2}] + E[T_{\text{tx}}^{\text{sec},1}] + E[T_{\text{tx}}^{\text{sec},2}]} \tag{2.19}$$

Finally, we consider the simultaneous transmission algorithm. The throughput model assumes that the first transmission and all retransmissions experience the same SINR. However, in the case of simultaneous transmissions, when one transmission is successful and the other is failed, then only the failed transmission will lead to a retransmission. Thus, the actual SINR for this retransmission will be higher than the previous transmission, which means that the actual P_{tx} for the retransmission will be higher too. The result obtained using the model is therefore expected to be slightly pessimistic. We calculate the throughput as:

$$S_{\text{sim}} = \frac{(P_{\text{suc}}^{\text{pri},1} P_{\text{suc}}^{\text{pri},2} + P_{\text{suc}}^{\text{sec},1} P_{\text{suc}}^{\text{sec},2}) \cdot B_{\text{MSDU}} \cdot 8 \cdot 10^{-6}}{E[T_{\text{tx}}^{\text{pri},1}] + E[T_{\text{tx}}^{\text{sec},1}] + E[\max(T_{\text{tx}}^{\text{pri},2}, T_{\text{tx}}^{\text{sec},2})]} \tag{2.20}$$

$E[\max(T_{\text{tx}}^{\text{pri},2}, T_{\text{tx}}^{\text{sec},2})]$ is calculated by using that for 2 independent random variables X and Y the probability distribution of their maximum is:

$$P(\max(X,Y) \le c) = P(X \le c \text{ and } Y \le c)$$

$$= P(X \le c)P(Y \le c) = F_Y(c)F_X(c) \tag{2.21}$$

This is true when the two random variables can be assumed to be independent. However, as described above, the case when one transmission is successful and the other is failed will only lead to one retransmission, which is then more likely to succeed than the previous transmission. Therefore we will get slightly pessimistic results when assuming that $T_{\text{tx}}^{\text{pri},2}$ and $T_{\text{tx}}^{\text{sec},2}$ are independent. However, we will make this assumption in the following for simplicity. Hereby we get:

$$F_{T_{\text{tx}}^{(2)}}(t) = P(\max(T_{\text{tx}}^{\text{pri},2}, T_{\text{tx}}^{\text{sec},2}) \le t) = F_{T_{\text{tx}}^{\text{pri},2}}(t)F_{T_{\text{tx}}^{\text{sec},2}}(t) \tag{2.22}$$

where $F_{T_{\text{tx}}^{(2)}}(t)$ represents the product of the probability distributions of the time spent per transmission attempt on the simultaneous relay transmissions. Following, we are able to compute the expectation of the maximum, by considering the corresponding density function as:

$$f_{T_{\text{tx}}^{(2)}}(t) = \frac{dF_{T_{\text{tx}}^{(2)}}(t)}{dt} = \frac{dF_{T_{\text{tx}}^{\text{pri},2}}(t)F_{T_{\text{tx}}^{\text{sec},2}}(t)}{dt} \tag{2.23}$$

and now we arrive at the expected value as:

$$E[\max(T_{\text{tx}}^{\text{pri},2}, T_{\text{tx}}^{\text{sec},2})] = \int_{t=0}^{\infty} f_{T_{\text{tx}}^{(2)}}(t) \tag{2.24}$$

2.5 Results and discussion

In this section we will first discuss performance of the throughput model on a single link, and second we will consider the results of the proposed algorithm for direct, relayed and simultaneous transmissions.

2.5.1 Throughput model results

The results of the throughput model described in the previous section is shown together with the result of a custom MAC layer simulation for verification. The simulation is based on similar assumptions as the model, namely: no contending nodes, backoff and retransmission upon a frame error according to IEEE 802.11 DCF, and up to 8 retransmissions. These simulation results are based on 1,000,000 frame transmission attempts for each considered value of Rmax and BER.

In Figure 2.3 we show how the time spent per transmission attempt increases with the BER. This is expected, as bit-errors will trigger retransmissions in the case of frame errors. A different maximum retransmission limit (here Rmax) has been used in each of the curves in the figure. Notice that when the number of retransmissions reaches Rmax, another error will mean that the frame is dropped and not successfully delivered.

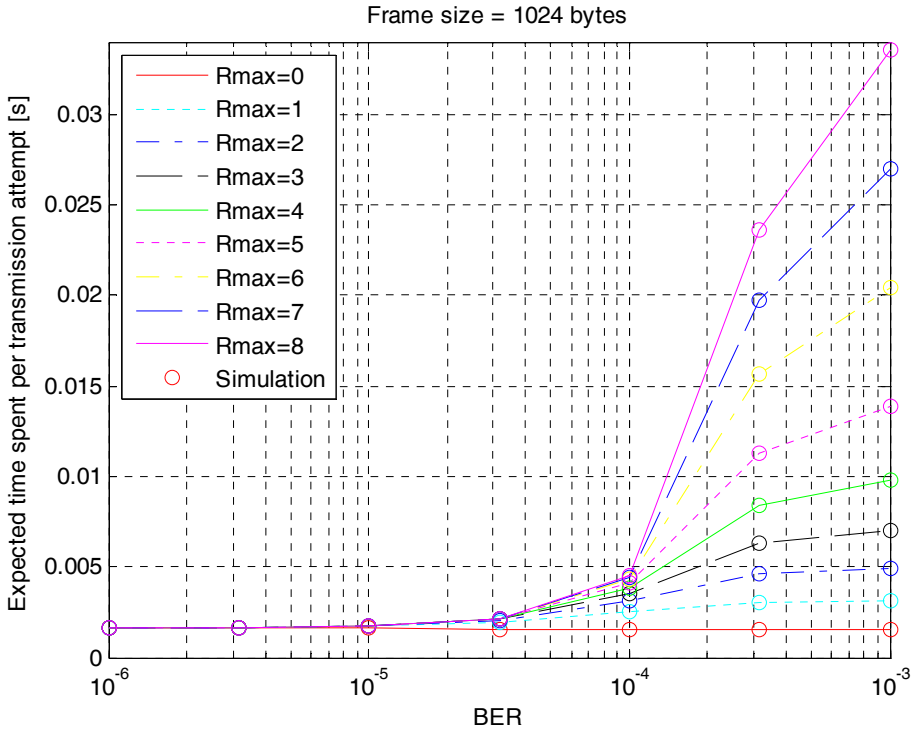

Figure 2.3: Average time spent per transmission attempt in seconds for different values of the maximum retransmission limit, given a frame size of 1024 bytes.

This is shown explicitly in Figure 2.4, where the frame delivery failure rate (FER) is plotted. Also here we show different curves for each value of Rmax. As expected, the frame error rate is less for high values of Rmax, and increases with an increasing BER. In the lower part of the figure the simulation results are not completely aligned with the model results. Considering that the results are based on 10^6 frame transmission attempts, it is clear that increasing the number of attempts would resolve this issue.

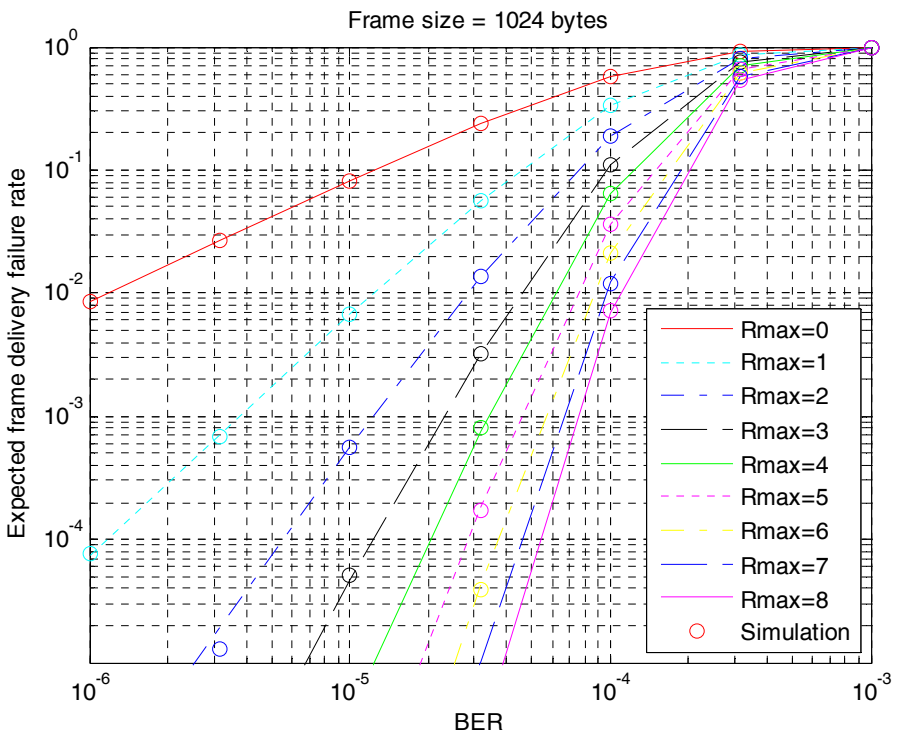

Figure 2.4: Average frame error rate for different values of the maximum retransmission limit, given a frame size of 1024 bytes.

Finally we combine these two metrics in Figure 2.5 and plot the expected throughput achievable on a single link. In this plot the throughput is slightly different for the different Rmax values and largest for Rmax=0. However, as this plot does not reflect the throughput achieved by an application, but merely the MAC throughput, one should not select Rmax=0 because it achieves a higher throughput. If no retransmissions are allowed on the MAC layer, the upper layers will have to make retransmissions or other recovery actions, which is likely to be less efficient.

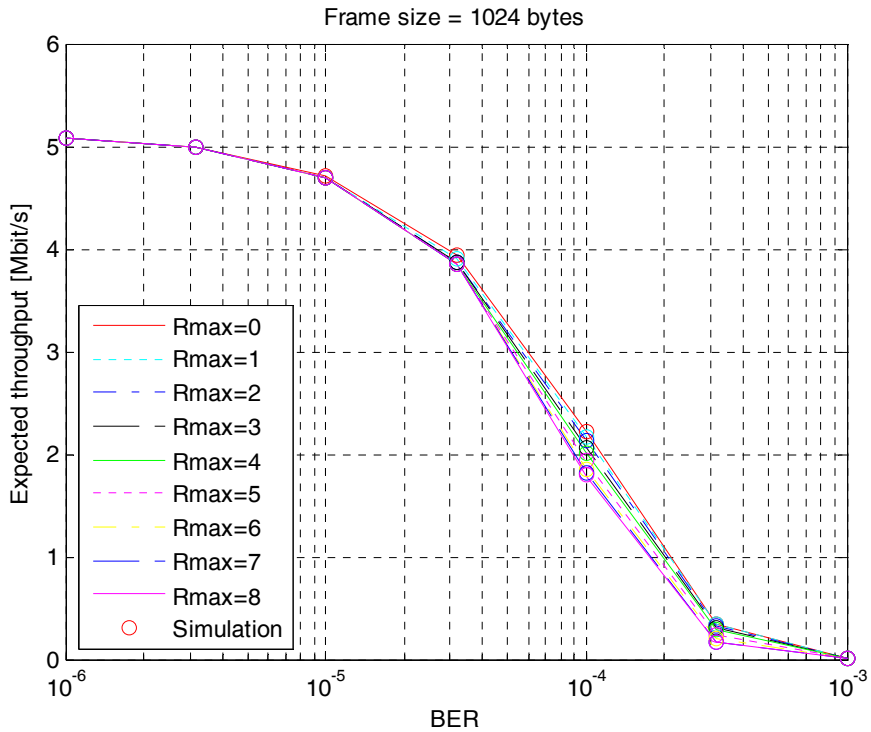

Figure 2.5: Expected average throughput for different values of the maximum retransmission limit, given a frame size of 1024 bytes.

2.5.2 Simultaneous transmissions results

Secondly, we compare the performance in the considered scenario for the following relaying schemes:

- **Direct**: Uses direct transmissions with maximum transmission power (100mW)
- **Relaying**: Uses two-hop relaying with maximum transmission power (100mW)
- **SimTX**: Relaying with simultaneous transmissions and power adaptation on relay to destination transmissions. The AP to relay transmission are done with maximum transmission power (100mW).

The results in Figure 2.6, Figure 2.7, and Figure 2.8 show example results of the first 75 runs of the 250 runs, whereas the results in Figure 2.9 and Figure 2.11 are based on all 250 runs.

Considering initially the achieved throughput for each run plotted in Figure 2.6, we can observe that the different algorithms have different maximum throughput levels. The throughput obtained with the Direct algorithm is fluctuating a lot, since some destination nodes are so far away that they cannot be reached in a one-hop transmission, however there seems to be a maximum level just below 5 Mbit/s. For the Relaying algorithm, the throughput is quite steady slightly below 2.5 Mbit/s. Having in mind that two-hop relaying is basically two direct transmissions, it makes sense that the throughput is only half of the Direct algorithm. Finally, the SimTX algorithm has a higher level than Relaying, around 3.2 Mbit/s, thanks to the ability of transmitting simultaneously. However, the throughput is a bit more fluctuating than for Relaying.

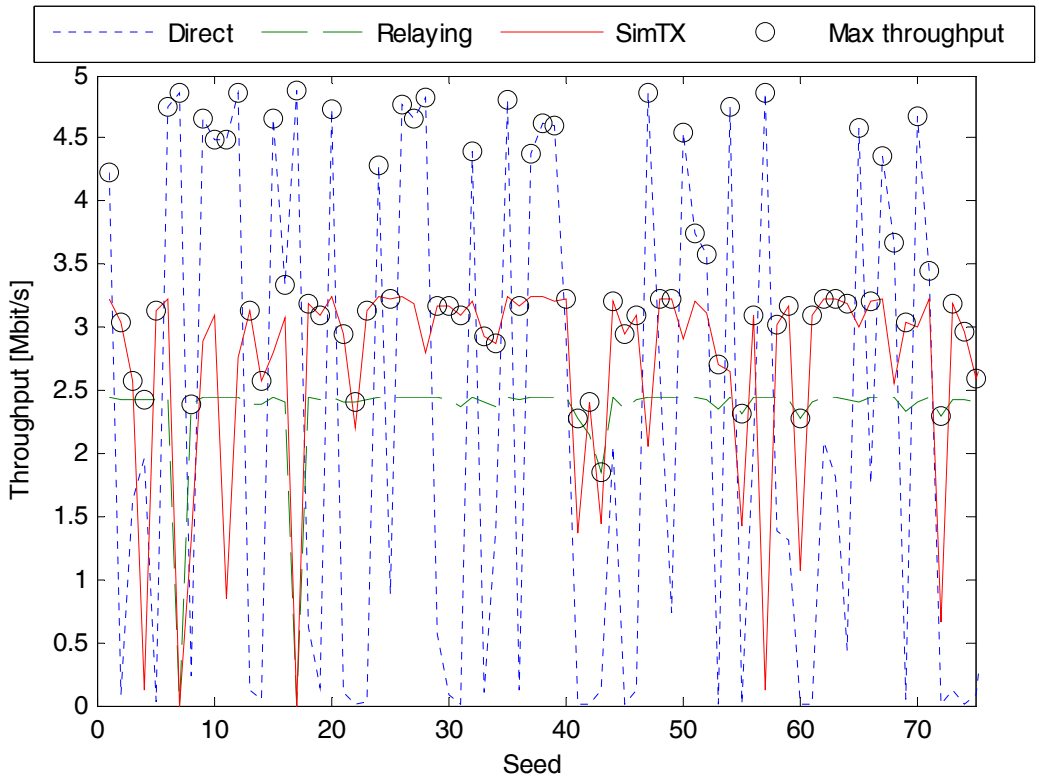

Figure 2.6: Throughput for the considered algorithms.

Figure 2.7 shows which algorithm achieves the highest throughput for each run. In this example we see that the SimTX algorithm is very often preferred to the Relaying algorithm. As we can see from Figure 2.6, the "gain" in obtained throughput for the SimTX algorithm compared to the Relaying algorithm is just around 0.5 Mbit/s in these cases.

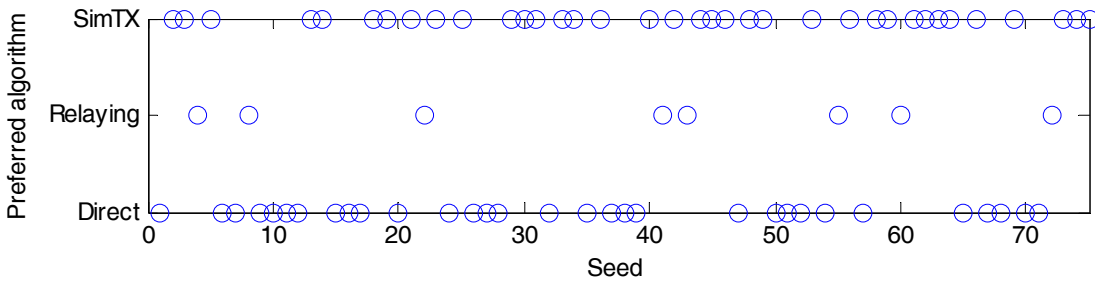

Figure 2.7: Preferred algorithm. Preference is based on highest throughput.

Focusing on the SimTX algorithm and its ability to adapt to interference using power adaptation, we show in Figure 2.8 the power levels selected by the SimTX algorithm, in the cases where it is the preferred algorithm. That is, the algorithm achieved the best throughput with this particular combination of transmit power settings, given the positions of the primary and secondary relay and destination nodes. Here we can notice that the SimTX alorithm does indeed take advantage of the possibility to adapt the transmit power. Further, we can see that at least on of the transmitters is using the maximum transmission power of 100 mW. In some cases, even both transmitters are using the maximum power. This tells us that the SimTX algorithm works as intended, as it is actually balancing the interference created by the two transmitters.

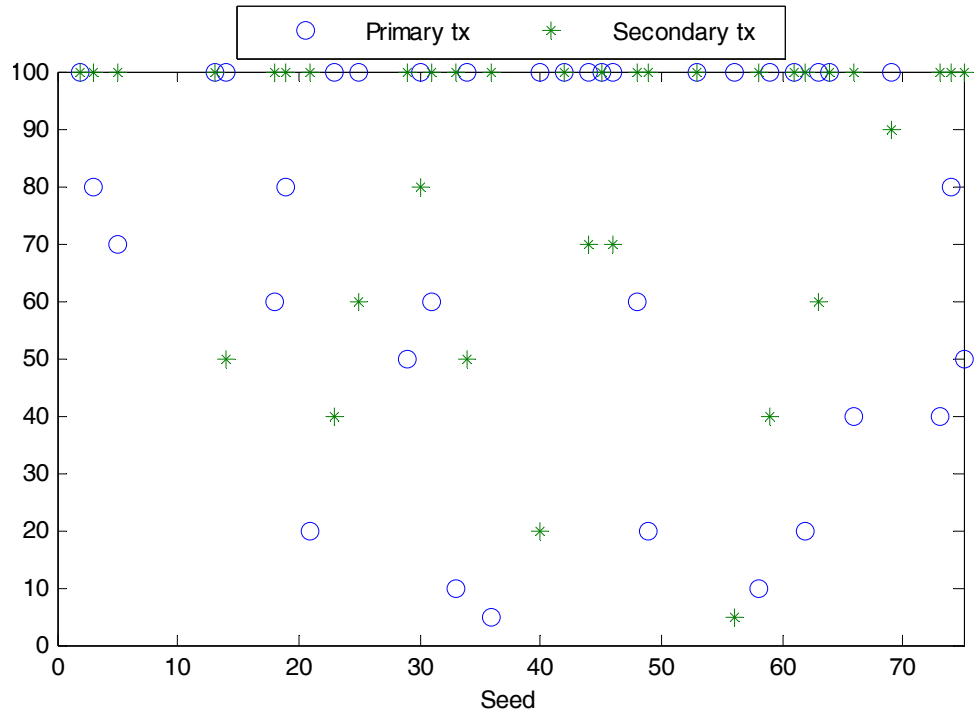

Figure 2.8: Transmission power selection by SimTX algorithm in cases where SimTX is the preferred algorithm.

Considering now the spatial layout of the considered scenario, Figure 2.9 shows us the positions at which each algorithm is preferred. Starting from the positions nearest to the AP, we see that the Direct algorithm is preferred at the positions that are within approximately 35 meters of the AP, since it is able to achieve a high throughput. Beyond this distance, both the Relaying and SimTX algorithms are represented, and they seem to be able to extend the range of the AP to reach any position in the 100x100m^2 map.

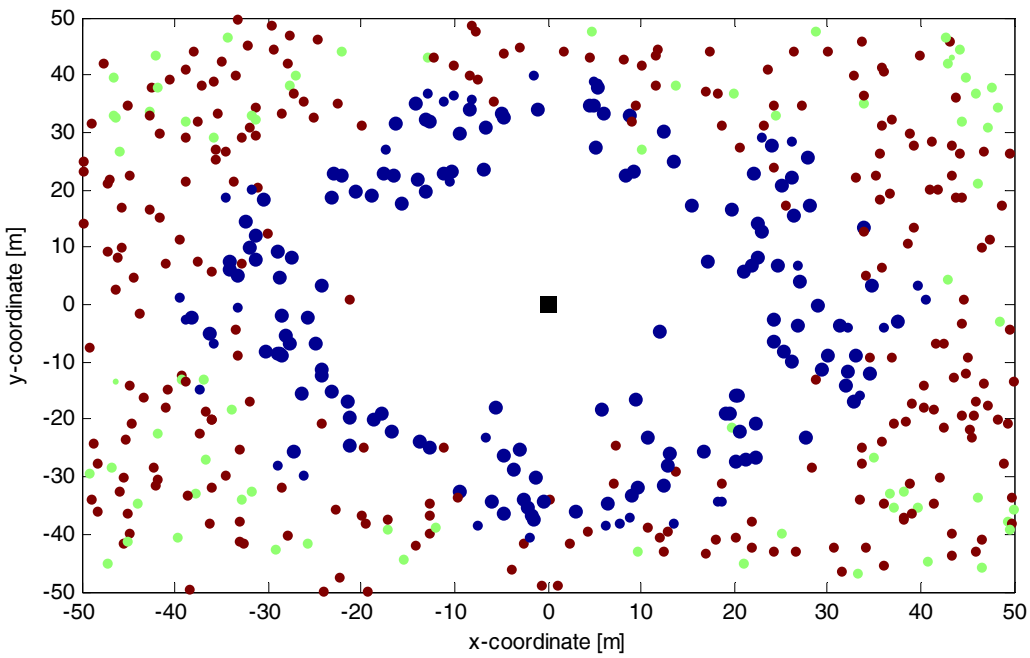

Figure 2.9: Points show destination node positions, where the color indicates the preferred algorithm (blue: direct, green: relaying, red: SimTX), and the size of the point is proportional to the throughput.

A different view of this picture is given in Figure 2.10, where we see the density of destination nodes for each of the 3 algorithms. Here we also clearly see the how the Direct algorithm is preferred close to the

AP in (0,0), but in this figure it is easy to see that SimTX is preferred in a ring at a larger distance than the Direct whereas the Relaying algorithm is mainly preferred in the corners.

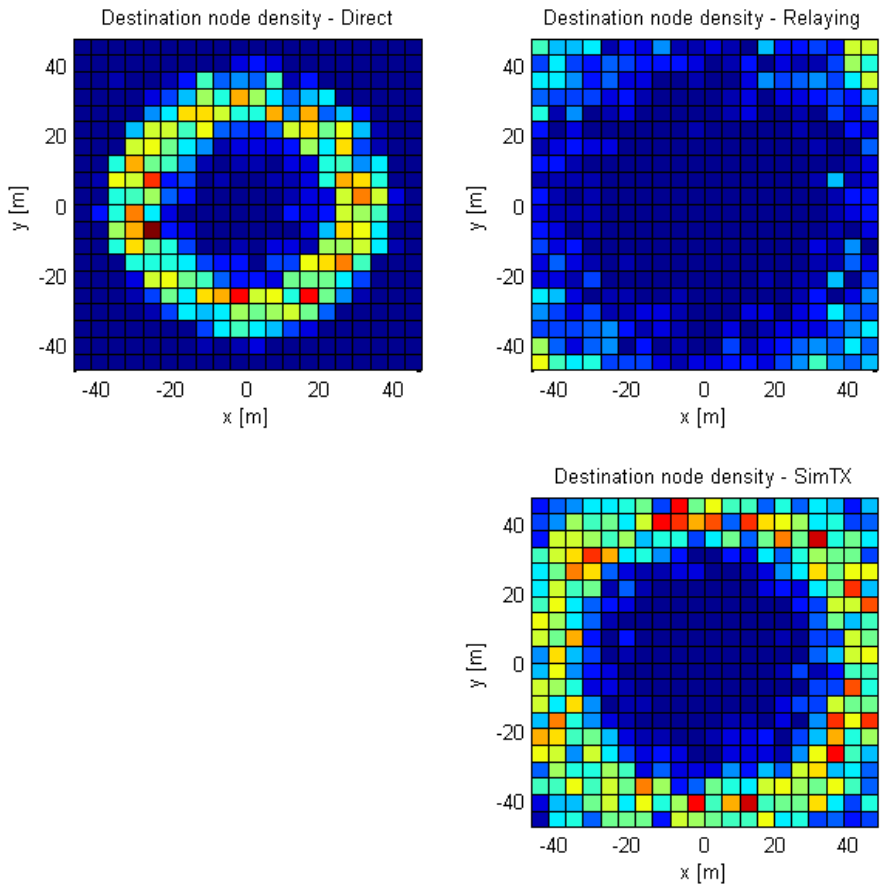

Figure 2.10: Density of destination node positions for each algorithm, based on 2500 runs.

Figure 2.11 shows a summary of the achieved throughput, when considering the cases where either Relaying or SimTX are the preferred algorithms, which allows us to compare these two algorithms. We see that in the considered scenario, the SimTX algorithm achieves an 11% (from 2.34 Mbit/s to 2.60 Mbit/s) better throughput than the Relaying algorithm. However, if we consider the improvement when going from Relaying to the best of Relaying and SimTX in each case, we go from 2.34 Mbit/s to 2.79 Mbit/s and get an improvement of 19%.

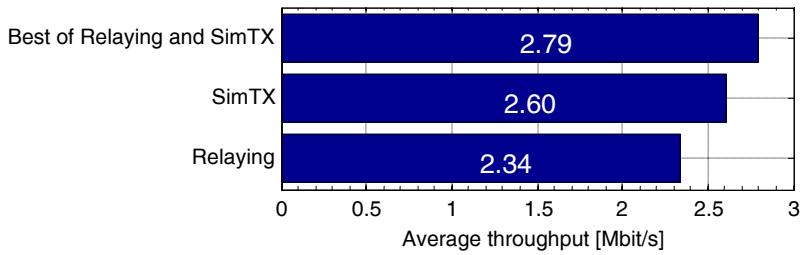

Figure 2.11: Comparison of the average throughput of Relaying and SimTX algorithms, when considering only the cases where direct transmissions are not preferred.

2.6 Conclusions and Future Work

In this work we have first developed a model to calculate the expected MAC layer throughput when taking into account the BER and maximum limit of retransmissions. By simulation we have verified the results of the model. This model has been used to show that a relaying scheme where simultaneous relay to destination transmissions are used is beneficial despite the cross-interference occuring when transmitting simultaneously.

Specifically, the proposed SimTX algorithm jointly determines the best combination of relays and transmit power for the two simultaneous relay to destination transmissions. Our results show that this approach can improve the throughput by 19% in the considered scenario when comparing to the cases where normal two-hop relaying would be used.

An obvious future work item would be to extend the joint relay selection and power adaptation, to include rate adaptation so that other, faster bit-rates than 6 Mbit/s are possible.

In this work, we have assumed that the proposed SimTX scheme has perfect knowledge of the positions of network entities. Considering the work on analysing the impact of mobility, inaccurate positioning information and inaccurate path-loss model parameters on two-hop relaying schemes in [NMS10a] and [NMS10b], also described in the WHERE D3.6, it would be interesting to make a similar analysis of the proposed SimTX scheme.

3 Link Level Interface Modelling Results for OFDM with Cellular Diversity

In this section, cellular diversity – in particular the cellular Alamouti technique – is investigated for convolutionally coded OFDM. The investigations are based on an indoor scenario [D1.1, Scenario M2]. For this scenario, channel measurements within WHERE have been performed [D4.1, Scenario M2]. These measurements have been used for channel propagation in simulations. The aim of this section is to provide frame error rates for different code rates and modulation alphabets at different receiver locations in that particular scenario. These frame error rates can be used as a location dependent link level model input for higher layer simulations.

3.1 Cellular Diversity

In the attempt of B3G standards significant improvements in the signal-to-interference-and-noise ratio (SINR) at the mobile terminal are required. The concept of coordinated multi-point transmission/reception (CoMP) is envisioned which implies a coordination of the transmission from multiple transmission points in the downlink [PDF08]. Figure 3.1 shows that principle, where a mobile station is served by 3 neighbouring base stations in addition to the original serving base station of the cell in which the mobile resides. In this section the cellular Alamouti technique is introduced to take advantage of the constellation of neighbouring base stations (BSs) serving the same area, especially cell border regions.

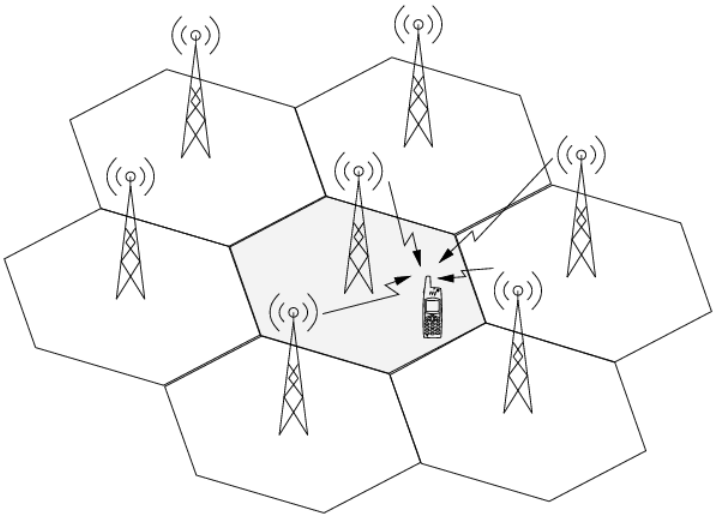

Figure 3.1: Mobile user in a cellular communication system served by several neighbouring base stations.

The concept of transmit diversity using the space-time block codes (STBCs) from orthogonal designs, namely the Alamouti code [Ala98] can be also applied for a cellular environment. Since the space dimension in STBCs is given by multiple transmit antennas at the transmitter, the space dimension can be also shifted to simultaneously transmitting BSs in a cellular communications system. The use of the simple Alamouti code as a special case of STBCs leads to the cellular Alamouti technique (CAT). Therefore, CAT explicitely implements macro diversity [Pla08]. This means performance improvements due to transmitting specific user related signals from different BSs. The requirements on synchronization among BSs are those of a single frequency network, i.e., it must be guaranteed that no or at least no significant level of intersymbol interference occurs.

3.2 Evaluation

3.2.1 Simulation Setup

Figure 3.2 shows the transmitter block diagram of the simulation system. We consider 2 base stations which are the TX antennas for the CAT and serve one receiver. The transmitter is convolutionally coded OFDM system with an Alamouti space-frequency coding entity for the implementation of the CAT. Information bits are encoded using a punctured terminated convolutional code in order to achieve code rates of 1/2, 2/3 and 3/4. The encoded bit stream is interleaved and modulated onto a 4-, 16- or 64-QAM

symbol constellation. Such a frame forms one subchannel, which consists of a subset of the available subcarriers of one OFDM symbol. A multiplexer merges multiple frames/subchannels. The resulting N symbols are interleaved and build the N used subcarriers of one OFDM symbol. Both the bit- and the symbol interleaver are random interleavers. The resulting symbols are space-frequency encoded using the Alamouti scheme. Finally the two symbol streams out of the Alamouti encoder are OFDM modulated. For OFDM modulation we apply an IFFT with subsequent guard interval insertion. Table 3.1 summarizes the main parameters of the considered system.

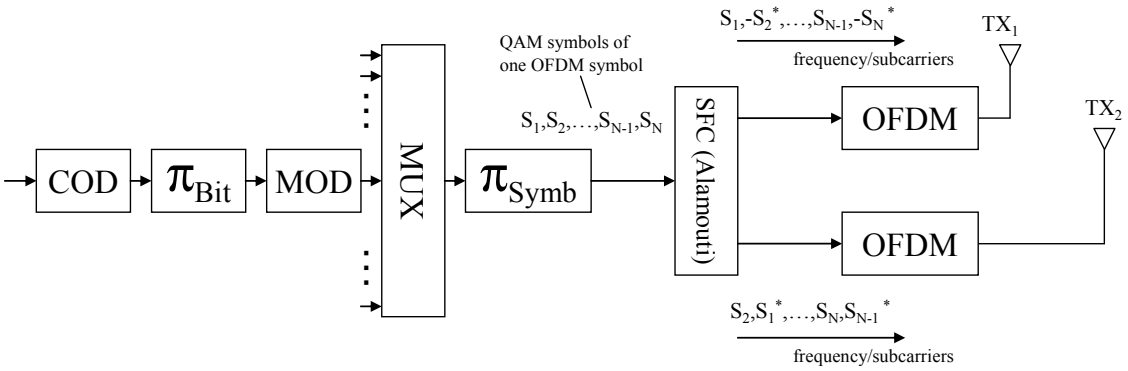

Figure 3.2: Block diagram of simulation.

At the receiver side we revert the TX sided signal processing. First the received signal is OFDM demodulated (guard interval removal, FFT). A subsequent Alamouti space-frequency decoder and the symbol deinterleaver provide the input for the QAM demodulator, which generates log-likelihood (LL) values for the codebits using the Max-Log-MAP algorithm. These LL values are deinterleaved an fed into a Viterbi decoder. The decoded information bits are finally used to evaluate the frame error rate.

Subcarrier spacing	Δf_{SC} = 10.94 kHz
FFT length	1024
Subcarriers	720
Subcarriers per subchannel	N_S = 24
Subchannels (per OFDM symbol)	30
Guard interval length	32 samples
Channel code, Code rate	$(171; 133)_{oct}$ conv. code, memory M = 6 Coderate R = 1/2, 2/3, 3/4
Modulation	4-QAM (N_{BpS} = 2 Bits per Symbol), 16-QAM (N_{BpS} = 4 Bits per Symbol), 64-QAM (N_{BpS} = 6 Bits per Symbol)
Demodulation	Max-Log-MAP
Decoding	Soft decision maximum likelihood (Viterbi)
Channel estimation	perfect

Table 3.1: System parameters.

Table 3.2 shows the number of information bits K and the number of codebits N per frame. Note, as the used convolutional code has a memory of 6, we need 6 bits for termination of the convolutional code trellis. Thus, the values for K and N are

$$\begin{aligned} N &= N_S\, N_{BpS} \\ K &= N\,R - M \end{aligned} \tag{3.1}$$

Since coding is in frequency direction, the duration of a frame equals the duration of an OFDM symbol, which is

$$T_{\text{Frame}} = T_{\text{OFDM}} = \frac{N_{\text{FFT}} + N_{\text{G}}}{N_{\text{FFT}}\,\Delta f_{\text{SC}}} = \frac{1024+32}{1024\times 10940\,\text{Hz}} = 94.26\,\mu\text{s}. \tag{3.2}$$

The subchannel bit rate therefore is

$$R_{\mathrm{SC}} = \frac{K}{T_{\mathrm{OFDM}}}, \tag{3.3}$$

which varies from 191 kBps (4-QAM, R=1/2) up to 1.1MBps. The different values for N and K are summarized in Table 3.2.

	R = 1/2	**R = 2/3**	**R = 3/4**
4-QAM	N = 48 K = 18	N = 48 K = 26	N = 48 K = 30
16-QAM	N = 96 K = 42	N = 96 K = 58	N = 96 K = 66
64-QAM	N = 144 K = 66	N = 144 K = 90	N = 144 K = 102

Table 3.2: Number of information bits K and code bits N per frame. Grey backgrounded combinations are used for simulations.

For modelling the channel propagation we use channel measurement data coming from an indoor channel measurement campaign within WHERE [D4.1, Scenario M2]. For one TX position TX_1 there are 19 RX locations. For each of these locations a channel impulse response (CIR) has been measured. Each CIR is characterized by a set of $\alpha_{k,i}$ and $\tau_{k,i}$, $k = 1,\ldots, N_i$, which are the complex amplitudes and path delays for each of the observed N_i paths at receiver location RX_i. Figure 3.3 shows the considered scenario. The 19 RX locations are on a straight line with equal distances of 0.5 m. Measurements have been performed for one TX (TX_1) location. For a 2-TX scenario, we assume a symmetric building geometry, which allows to mirror RX and TX locations for virtually obtaining a second TX position (TX_2). Because of symmetry, we assume that the CIR from TX_1 to RX_i. are equal to that from TX_2 to RX_{20-i}. The CIRs are static and valid for the specific RX location. We allow a small variation in the RX position around the location of CIR measurement in order to achieve a locally averaged frame error rate measurement. For that reason, we consider the absolute values $|\alpha_{k,i}|$ of the measured CIRs as static in this small region. However, since phases are changing rapidly, we model them as uniformly distributed random process in the interval $[0,2\pi)$.

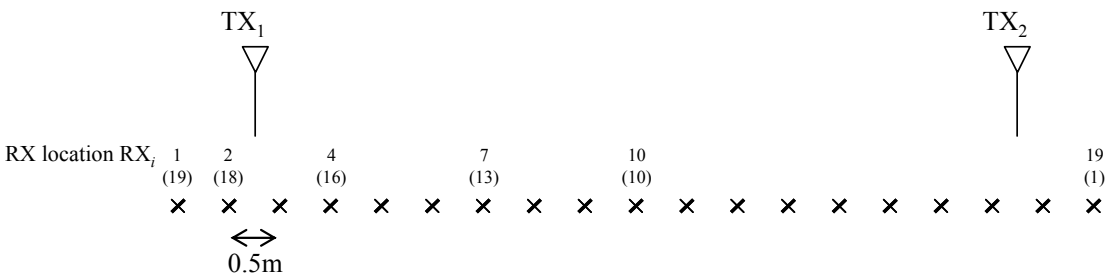

Figure 3.3: Scenario and RX locations.

From the CIR measurement data $\alpha_{k,i}$, we can obtain the power ratio of the signals at location RX_i coming from TX_1 and TX_2 as

$$\mathrm{SIR}_i = 10\log\frac{\sum_{k=1}^{N_i}\left|\alpha_{k,i}\right|^2}{\sum_{k=1}^{N_{20-i}}\left|\alpha_{k,20-i}\right|^2} \quad [\mathrm{dB}]. \tag{3.4}$$

Table 3.3 shows the SIR values for RX locations 1…10. Note, due to symmetry $\mathrm{SIR}_{20-i} = -\mathrm{SIR}_i$.

RX_i	1	2	3	4	5	6	7	8	9	10
SIR_i [dB]	12.3	17.1	21.4	15.5	6.2	3.9	-3.1	-2.7	10.3	0

Table 3.3: SIR values for different RX locations. Grey backgrounded locations are used for simulations

3.2.2 Results

Figure 3.4-Figure 3.7 show the frame error rates (FER) for RX locations 1, 4, 7, 10 for different combinations of coderate and modulation alphabets. The graphs are drawn versus the subcarrier SNR which is defined as

$$\mathrm{SNR} = 10\log\frac{\mathrm{E}\left\{|S|^2\right\}}{\sigma_{\mathrm{N}}^2} \quad [\mathrm{dB}], \tag{3.5}$$

where $\mathrm{E}\{|S|^2\}$ is the average received signal power at each used subcarrier and σ_{N}^2 is the variance of the complex valued additive white Gaussian noise observed at each subcarrier.

Dependent on the location of the receiver a different steepness of the FER graphs can be observed. This steepness is mainly determined by the amount of diversity which the propagation channel offers. A measure of diversity is the amount of observable signal propagation paths which are relevant in terms of their power. In general, CAT as the applied macro diversity technique is capable to increase and exploit such diversity, since the signal is received from two different TX sites, which in principle increases multipath and, therefore, diversity. The parameter SIR (see Table 3.3) describes the signal power ratio between the signals received from the two BSs. For this reason it is obvious that a high absolute value of SIR decreases diversity. The absolute value of SIR is principially low at the cell borders. When moving towards a BS, the SIR values increase.

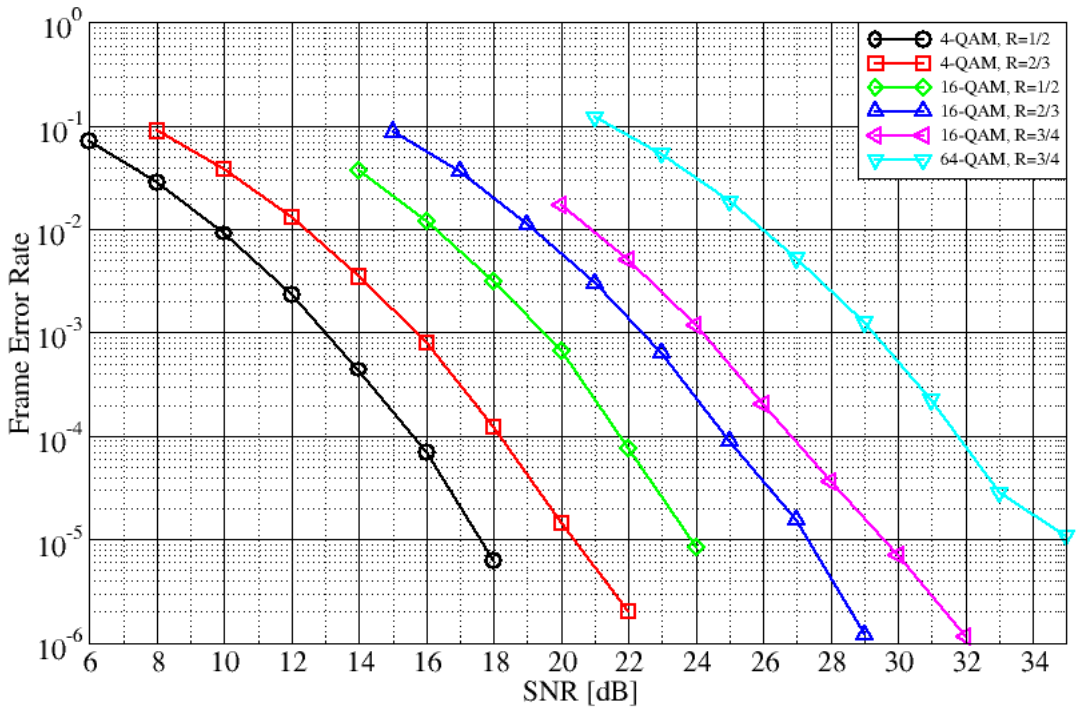

Figure 3.4: FER versus SNR for different coderates and modulation alphabets at RX position 1.

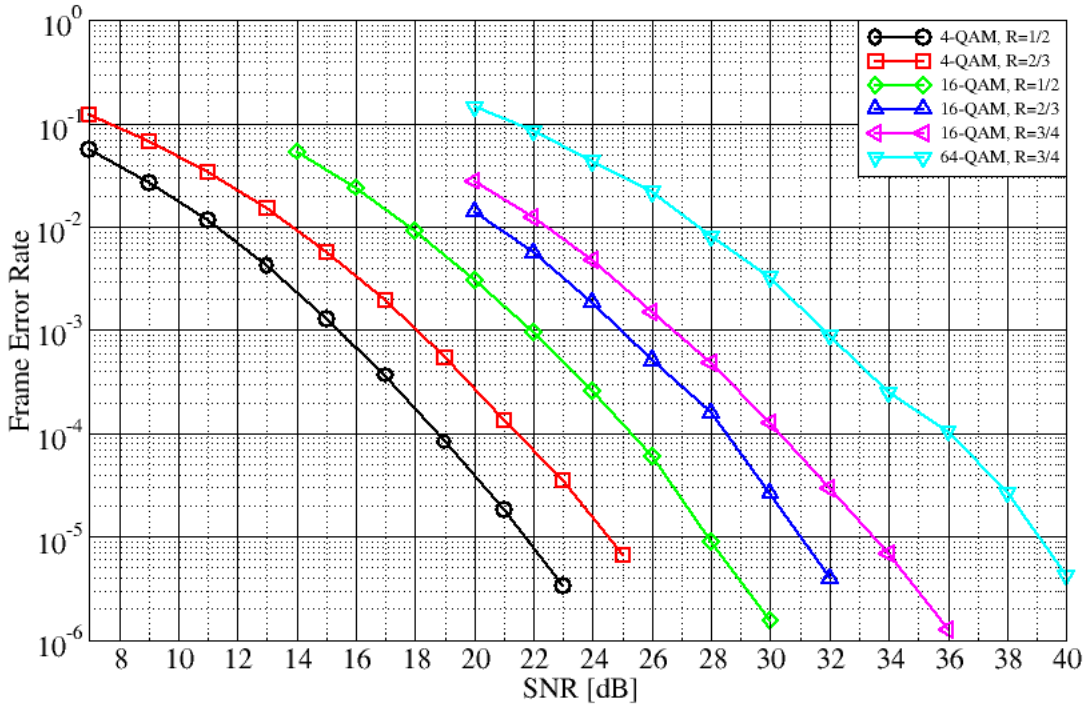

Figure 3.5: FER versus SNR for different coderates and modulation alphabets at RX position 4.

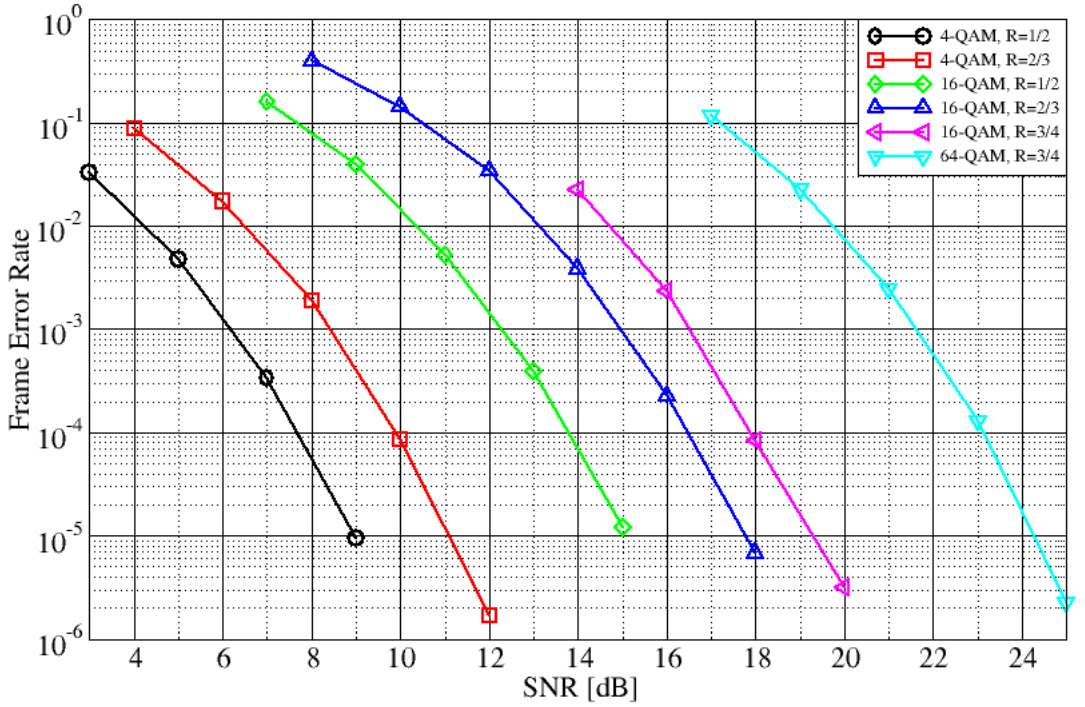

Figure 3.6: FER versus SNR for different coderates and modulation alphabets at RX position 7.

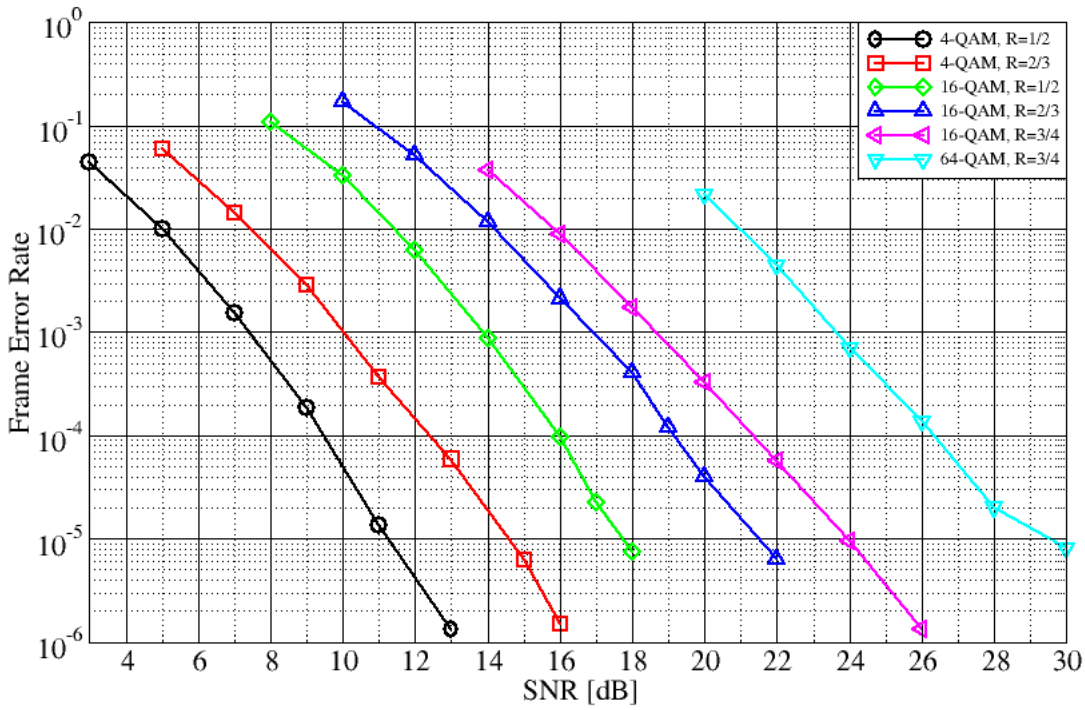

Figure 3.7: FER versus SNR for different coderates and modulation alphabets at RX position 10.

3.3 Summary

In this section we have provided frame error rates for a convolutionally coded OFDM system applying different modulation constellations and code rates. Simulations have been carried out using real propagation measurements rather than stochastic channel models. The results can serve as basis for the generation of lookup tables for system level simulations, which then provide system performance measures for a real indoor scenario.

4 SDMA Dynamic Resource Allocation for the Mobile WiMAX (802.16e)

4.1 Introducation

This chapter describes all the steps followed in the implementation of Space Division Multiple Access (SDMA), in the basic Dynamic Resource Allocation (DRA) module presented in D3.3 In Mobile WiMAX air interface SDMA is a technique resulting from the transmission along the same set of sub-channels and/or OFDM symbols, but using beam patterns separated in the space domain. SDMA provides another degree of freedom in the map of resources of the OFDMA-based air interface, compared to conventional multiple access schemes such as Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA) or Code Division Multiple Access (CDMA). The underlying idea behind SDMA is to divide the space into a number of orthogonal spatial beams which can be separated in the receiver. SDMA brings out another dimension in resource allocation, from physical (PHY) to medium access control (MAC) layer, in the transmission of information over the air interface.

The new SDMA-based DRA architecture is used in conjunction with a utility-based packet scheduler. Besides time and frequency domains, users can now be assigned spatial beams for information transfer, provided the degradation in signal quality for users already assigned slots in the map of resources, from new users which are assigned the same slots, but with different spatial beams. The performance is not affected to the point of compromising the user satisfaction for the quality of service accomplished by the network with the spatial separaion. Therefore, in the new DRA architecture, users in the scheduling priority list can now be assigned resources in space, besides time and frequency domains. Location information is used to estimate the direction of arrival of each mobile related to the base-station and is assumed the beamforming algorithm is optimal in steering one spatial beam into the direction of the desired mobile.

The standards of the IEEE 802.16 family adopt Adaptive Antenna System (AAS) as an option to enhance cell capacity and coverage. The benefit incurred with the use of AAS is its ability to reduce interference by steering the beam to a specific user. Adaptive antenna elements can be used to distort the radiation pattern produced by an antenna array at the base station. The idea is to focus (beamform) the transmitted signal energy in the direction of the intended mobile receiver and, at the same time, steer neutralize signal energy in directions of co-channel mobile. This results in an enhancement of the perceived Signal to Interference plus Noise Ratio (SINR) at the desired mobile's receiver. Comparable approaches are currently being standardized by 3GPP for UMTS or by the IEEE for 802.11n. These advanced antenna techniques have a significant impact on the capacity and service quality provided by wireless links and the efficient use of the available spectrum [CCP08].

The beamforming processing is accomplished by applying complex weights to the signals transmitted at each antenna elements at the base station antenna array (although it theoretically could be done also at the mobile station with the cost of added complexity). The weights are computed by using information regarding the Directions of Arrivals (DoA) of desired and interfering signals at the base station. The amount of main lobes to steer must be lower or equal to the amount of degrees of freedom, corresponding to the amount of antenna elements in the array. Due to the linear nature of the antenna array it is also possible to generate multiple patterns of signals to different mobiles, as a meaning of transmission simultaneously to different mobiles. The resulting pattern will be the linear combination of the different patterns.

In order to assign different spatial beams to mobiles they must be separated in space. This separation depends on the main lobe aperture, which is a function of the amount of antenna elements in the array, and on the spatial signatures from each mobile at the base station. The spatial signature captures the spatial characteristics (direction-of-arrivals or departures, number of multipath components and attenuation) associated with the mobile terminal. Of course, in reality, the beams are not completely orthogonal and the amount of intra-beam interference increases with the amount of users being assigned spatial beams for the same radio resource (time and/or frequency for example).

This chapter is organized as follows. Section 4.2 describes the utility function used to select users and packets for transmission. Section 4.3 introduces smart antenna techniques and their use in the SDMA multiple access scheme. Here, these techniques are based on linear processing of the antenna array at the base station, in order to result in non-overlapping antenna beams at each mobile attempting to transmit at the same frame interval. These non-overlapping beams avoid intra-beam interference over the same set of radio resources in the frame and increase the cell's capacity. Section 4.4 details the frame structure for the

implementation of AAS in Mobile WiMAX. It introduces different proposals available in the research literature for estimation of user's separability across different spatial beams and describes the steps followed in the SINR estimation, after user's assignment to spatial beams, over the same set of symbols and subchannels in the TDD frame. Section 4.5 is the core of this chapter as it explains the principle behind the SDMA-based DRA proposed for Mobile WiMAX MAC. The DRA computes the groups of users to transmit in the same radio resource (same set of slots in time and frequency domains), according to the estimated correlation among the set of channel matrices, over all data sub-carriers composing each resource. Section 4.6 describes the proposed SDMA-based DRA algorithm which comprises user's prioritization and resource allocation, according to SDMA multiple access. The algorithm used in the computation of the groups of users and assignment of spatial beams, over each radio resource, is presented as well as the scenario used in the system level simulations. Section 4.7 presents results from the performance evaluation of the proposed joint utility-based packet scheduler and SDMA multiple access scheme for Mobile WiMAX. Simulations were conducted separately for three types of traffic models: Full Queue, Voice over IP (VoIP) and 3GPP's World Wide Web (WWW). The scenario used for system level simulations was the one of a 4x2 MIMO channel antenna system configuration with Alamouti Space Time Block Coding (STBC). The gains achieved with this new DRA architecture are compared against standard solutions commonly referred to in the literature. Section 4.8 is about the related work available in research literature regarding the implementation of SDMA-based packet schedulers. Section 8 concludes the chapter.

4.2 Utility-based packet scheduler

The packet scheduler must efficiently allocate available radio resources in response to burst data traffic, time-varying channel conditions and service's QoS requirements. The scheduler functioning is based on the availability and access to contention-free radio resources from the Medium Access Control (MAC) layer. It is located in the base station in centralized systems and enables rapid response to traffic requirements and channel conditions. The scheduler processes data packets associated to service flows with well defined QoS parameters in the MAC layer, in order to correctly determine the packet transmission ordering over the air interface.

The scheduler behavior is strongly influenced by the type of traffic model being serviced:

- Real-Time (RT) services pose constraints on the maximum allowable delay for a data packet to be serviced and transmitted through the air interface. If a RT data packet violates the maximum delay bound it is dropped from the queue, because its transmission would result in a squander of the set of resources allocated for its transmission, as it is outdated for the receiver. Data packets of RT service flows are normally generated with a constant bit rate and are of fixed or variable size. The typical approach is to increase the priority of the service flow as its Head of Line (HOL) packet approaches the service deadline.
- Non-Real-Time (NRT) services are more relaxed regarding the satisfaction of a strict maximum delay bound for each packet sent through the air interface. Packets from applications of this type are not generated with a constant bit rate, that is, they are generated in bursts. The best approach to service a NRT user is to keep packets in the user's queue for it to fill up before start transmitting them over the air-interface. Packets can remain for some time in the queue waiting for the best possible user's channel state for transmission. These applications are particularly targeted for opportunistic schedulers, of which the Maximum C/I is a typical example, and benefit from the so called "multi-user diversity gain", achieved when the scheduler selects for transmission the user with the best channel in each frame period.

Utility functions were proposed for Voice over IP (VoIP) and WWW traffic respectivelly.

For VoIP traffic model a hybrid utility function is considered. The utility is kept constant until packet delay becomes higher than some priority timer. Whenever the delay becomes higher than this threshold the utility decreases with a significant rate. As VoIP packets are associated to traffic flows with some constant packet generation pattern, their transmission may be postponed until packet delay approaches the maximum delay bound allowed for the service. In this time interval the algorithm essentially behaves like the traditional maximum C/I algorithm, achieving efficiency in resource utilization and throughput maximization. The used utility function for VoIP traffic users is defined according to equation below.

$$U(\tau_j^l) = \begin{cases} 1 & \text{if } t_1 < T_{req} - T_{pri} \\ \left(T_{req} - t_l\right)/T_{pri} & \text{if } t_l \geq T_{req} - T_{pri} \\ 0 & \text{if } t_l > T_{req} \end{cases} \tag{4.1}$$

Where:

t_l is the waiting interval from the instant of arrival for the lth packet at the jth user's buffer.

T_{req} denotes the maximum allowable delay for the service. Whenever t_l exceeds T_{req} the packet is dropped due to time-out violation.

T_{pri} is the priority timer. It denotes the time interval between the instant when the priority

For WWW traffic model, a slowly decaying function of packet delay, with a smaller initial utility gain A is proposed. The idea behind this setting is to give more priority to voice packets in detriment of web ones, namely when voice packets approach the deadline.
The proposed utility function for WWW packets is defined according to equation below.

$$U(\tau_j^l) = \begin{cases} a - b(t_l / T_{req})^{1.3} & \text{if } t_1 \le T_{req} \\ 0 & \text{if } t_l > T_{req} \end{cases} \tag{4.2}$$

Where:
a is the initial gain of the utility function.

b is the step from this value at the maximum delay T_{req}.
Figure 4.1 plots both utility functions and Table 4.1 lists the values for the parameters considered in the definition of both utility functions.

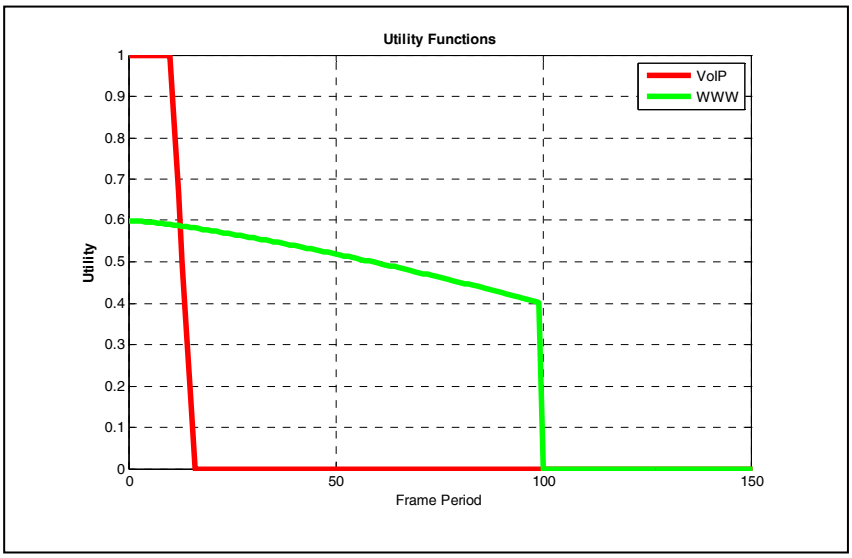

Figure 4.1: Proposed utility functions for VoIP and WWW users

	VoIP	WWW
Maximum Utility, A	1	0.6
Maximum Delay Bound T_{req}	80 ms	1 s
Priority Timer T_{pri}	50 ms	-
Utility step from maximum value	0	0.2

Table 4.1: Configuration of the proposed utility functions

4.3 Spatial Beamforming

Figure 4.2 depicts a typical antenna array system in which the spatial diversity is exploited for multiple access communications. In the absence of noise the response of an M-element antenna array to a signal s(t) is written as in equation below.

$$y(t) = (y_1(t), y_2(t), \ldots, y_m(t),)^T = as(t) \tag{4.3}$$

The vector $\mathrm{a} = (a(1), a(2),..., a(m),)^T$ is the array response that captures the spatial characteristics associated with the terminal and is designated commonly as "spatial signature". Assuming K mobiles communicate at the same time with the same base station the total signal output at the antenna array is given by equation below

$$y(t) = \sum_{k=1}^{K} \mathrm{a}_k s_k(t) + \mathrm{n(t)} = \mathrm{A}s(t) + \mathrm{n(t)} \tag{4.4}$$

Where the vector $\mathrm{A} = (\mathrm{a}_1,..., \mathrm{a}_k)$ is the array manifold whose columns represent the spatial signatures (one for each mobile terminal). Each entry in the vector n(t) represents the Gaussian noise in the respective antenna and each entry in the vector $s(t) = (s_1(t),..., s_K(t),)^T$ represents the signal transmitted by each one of the K terminals.

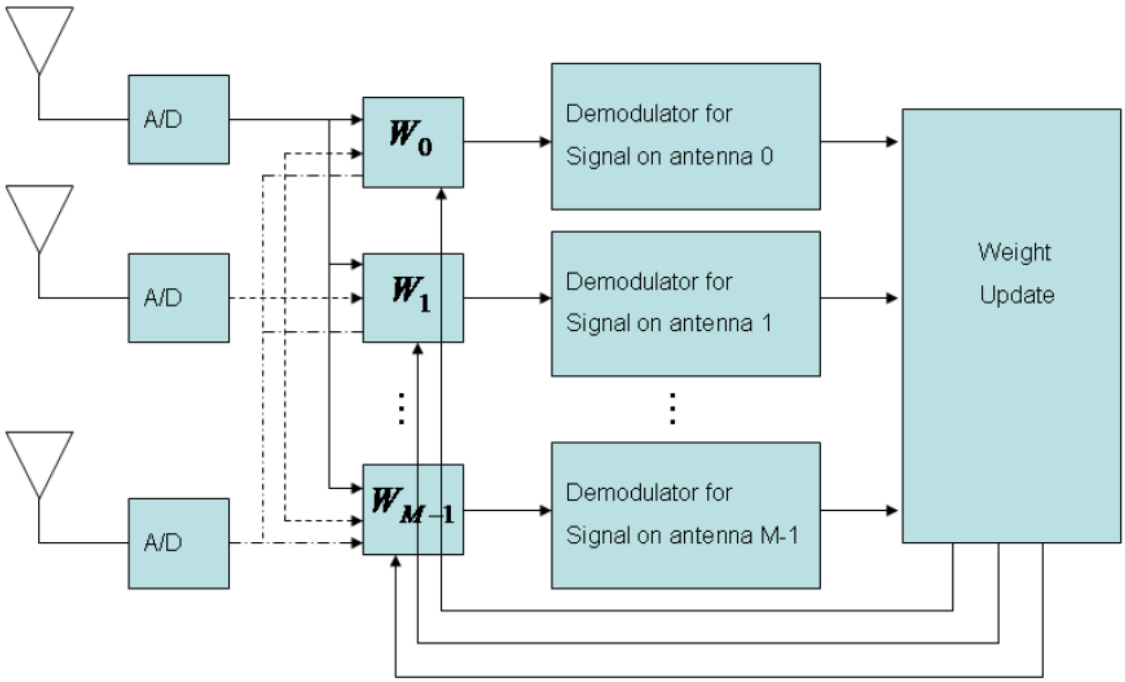

Figure 4.2: Spatial beamforming module for a linear antenna array

Although the spatial signatures from the K terminals are superimposed at the base station's antenna array, if they are spatially separable (orthogonal) they can be separated through spatial filtering processing. To retrieve a estimate of the individual signal $s(t)_i$ from mobile i, antenna outputs are weighted and summed with a set of complex weight coefficients (computed for each mobile), according to equation below.

$$\hat{S}_i(t) = \sum_{m=1}^{M} w_i^*(m) y_m(t) \tag{4.5}$$

The vector of weights $\mathrm{w}_i = (w_1(1),..., w_i(M))^T$ is designed to constructively combine the signal of interest from mobile i (steer the beam in this direction) and destructively combine the interference-plus-noise from other mobiles. If multiple beams are applied there must be one weight vector per beam.

4.4 SDMA Scheme for IEEE802.16e Mobile WiMAX

This section describes, in great level of details, the model describing the integration of SDMA with the OFDMA multiple access schemes used in Mobile WiMAX.

A cellular network composed of three tiers of tri-sectored base stations is considered. Only the downlink connection is used for beamforming. Each base station is equipped with a linear uniform antenna array with N_T transmitting antennas and each mobile terminal is equipped with a linear uniform antenna array with N_R receiving antennas. With N_T antennas in the transmitting array there can be $M_N \leq N_T$ beams on the nth sub-carrier of the spectrum of each OFDM symbol. Based on the computed SINR on beam m and sub-

carrier n, the estimated Block Error Rate (BLER) is obtained from the look-up tables which implement the interface to the link layer. The supportable data rate, $r_{n,m}$ on beam m and sub-carrier n is derived from the estimated BLER.

The channel quality is periodically reported by the mobile in order to the transmitter adapt the transmission mode according the channel state. The amount of channel quality feedback reported by each mobile station may be extremely high if there is a need to compute the set of weights for each sub-carrier in the spectrum of each OFDMA symbol of the Time Division Dulex (TDD) frame. The periodicity of this report depends also on the type of environment in the mobile radio channel. For mobile terminals moving with high velocity, the coherence time results in the de-correlation among blocks of consecutive OFDM symbols, requesting feedback reports inside the same TDD OFDMA frame. However, for the mobile speeds and channels used in the set of simulations conducted under the scope of this work, the mobile channel is assumed as being constant along each frame period, and the report is performed periodically for each radio frame.

In the frequency domain the degree of correlation among sub-carriers depends on the type of sub-channelization used:

- For DL-PUSC sub-channelization, according to the way data sub-carriers are distributed in frequency it turns out to be not very practical for the implementation of beamforming because reports must be performed for all sub-carriers in each sub- channel and different vector of weights must be computed for each sub-carrier.
- For DL-AMC sub-channelization mode, sub-carriers are allocated continuously to form a bin and there is a significant correlation between adjacent sub-carriers inside the same bin, which means that the vector of weights may be estimated for the pilot sub-carrier and the same vector applied to all data sub-carriers inside the bin. This reduces the feedback rate.

4.4.1 User Spatial Separability

In Mobile WiMAX active users under the area of coverage of each base station are allocated slots, mapped into both time and frequency domains in the map of resources of the TDD OFDMA frame. SDMA expands the available capacity (in number of slots) by allowing multiple mobile stations to transmit in the same slot, through the assignment of different nonoverlapping spatial beams to the set of mobile stations. It adds another dimension to the map of resources by expanding each slot into another set of slots in the space domain.

In the Mobile WiMAX standard the base station can reserve a set of pilot sub-carriers and symbols in the TDD OFDMA radio frame for channel sounding of the uplink transmission from the desired mobiles. These pilot sub-carriers are used in the computation of the correlation matrices.

Assuming downlink connection, the size of the space domain is defined by the number of degrees of freedom associated to the number of antenna elements at the base station. If there are N_T antenna elements in the antenna array then up to N_T mobiles can transmit in the orthogonal beams associated to the same slot, in time and frequency. This translates into an N_T -fold increase of system throughout in an ideal scenario (up to N_T orthogonal beams can be created). These spatial beams are steered by the beamforming algorithm in the base station.

However, in most practical situations the number of active users in the area of coverage of a base station is greater than the maximum number of spatial beams which can be created for each slot. Also, in reality, the ability to capture multiple transmissions in different beams of the same slot depends critically on the spatial configuration of the co-slot mobiles, because not all users can be separated in space. Intuitively, by assigning “most-orthogonal” mobile terminals to the same slot, the average performance of the system can be improved.

The performance of the combined SDMA/OFDMA system is highly affected by the strategy followed in the separation of users into non-overlapping groups, in space. Therefore, different schemes were proposed in the literature.

In [PES07], a spatial grouping algorithm is proposed to group users according to their spatial separability. The result of the grouping is a set of groups of spatial users (separable in space), which overall result (all sets joitly) is a grid of users that can be separed by time or space. Since only users from the same set/group can be separated spatially, the users from distinct groups have to be allocated in the time domain.

For users inside the same group they can be serviced in the same resources in time and/or frequency at the same time. A different application of user’s spatial grouping is presented in [YK05] for an adaptive

resource allocation algorithm implemented on a MIMO/OFDMA multiple access system. The objective is the minimization of the overall transmission power, given user's QoS requirements in terms of bit error rate and data rate. The proposed scheme computes the correlation among users for each sub-carrier in order to control the sub-carrier's sharing according to the user's spatial separability. This makes it possible to decouple the multiuser joint resource-allocation probleminto simple single-user optimization ones.

In [SS04] a heuristic metric is proposed for the computation of the degree of separation for each pair of users. This metric is based on the normalized Frobenius norm of the product of their MIMO channel matrices and is used in the definition of the groups of users who can be separated in space. The metric that estimates the cost of putting two users together in the same group is given by equation below:«

$$\xi_{i,j} = \frac{\left\| \mathrm{H}_i \mathrm{H}_j^2 \right\|_F^2}{N_{Ri} N_{Rj}} \tag{4.6}$$

Where H_i and H_j are, respectively, the channel matrices for users i and j, and N_{Ri} , N_{Rj} are the number of antennas at the receiver of mobiles i and j, respectively. After computing all $\xi_{i,j}$ for all possible combinations, a Compatibility Optimization Algorithm (COA) is used to find the group that optimizes the space allocation, according to equation below:

$$\sum_{g=1}^{G} \left(\sum_{i,j \in G_8} \xi_{i,j} \right) \tag{4.7}$$

Where G is the number of groups in the system, corresponding to the number of time or frequency slots available for user multiplexing and G_g is one particular group in this set.

In [KTG08] the same principle is used in the computation of the degree of separation of two users transmitting in the same sub-carrier. Two users are called spatially separable on a sub-carrier if they share the same sub-carrier and the SINR requirements at corresponding receivers are satisfied. The normalized scalar product between the channels of each pair of users is computed according to equation below:

$$\eta_{i,j} = \frac{\left| \mathrm{H}_{i,n}^H \mathrm{H}_{j,n} \right|}{\left\| \mathrm{H}_{i,n} \right\| \left\| \mathrm{H}_{j,n} \right\|} \tag{4.8}$$

Where: $H_{i,n}$ and $H_{j,n}$ are, respectively, the channel matrices for users i and j on sub-channel n. The larger the metric is the more correlated their channels are, and more power is needed to meet the SINR requirements. For each user to be allocated, the one resulting in the minimum of the set containing the maximum of the scalar products, for all already allocated users is selected, as in equation below:

$$k^* = \arg \min_{i \in K\text{-}K_n} \max_{j \in K_n} \eta_{i,j} \tag{4.9}$$

Where K is the set of users in the cell and K_n is the set of users allocated to sub-carrier n.

In [SKY04] an SDMA based wireless packet cellular system is proposed. Users are assigned the same time-slot provided their angular separation in space is above a given threshold. The architecture encompasses a packet scheduler and a resource allocator. Users which cannot be spatially separated are assigned to an empty time-slot.

4.4.2 SINR Estimation after Performing Beamforming

After a given mobile station is assigned an empty beam in a given slot in the map of resources, a new SINR value will result for both the new mobile as well as for the mobiles already assigned empty beams in the same slot. The new value for the SINR ratio will be affected from both the intra-cell interference, resulting from all mobiles transmitting on different spatial beams associated to the same slot, as well as from the new value of the desired signal resulting from the beamformed pattern in the direction of the mobile.

As the new antenna pattern is steered into the direction of the mobile, the beamforming process results into an improvement of the SINR measured at the receiver. In order to estimate the correct SINR all relevant signals must be differentiated. Figure 4.3 illustrates the relevant signals during SDMA transmission. We are analysing the signals in one cell with SDMA transmission to the mobile stations A

and B. The dashed lines refers to pattern pointed to MS B, which has small interference in MS A and vice-versa. The arrows pointing to the MS A refers to the inter-cell interference in the MS A.

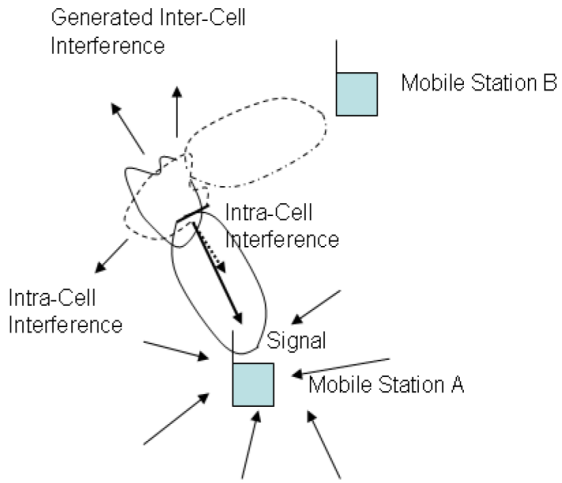

Figure 4.3: Estimation of the SINR ratio for antenna beamforming

An optimized antenna pattern is applied at the base station to produce a main lobe (spatial beam) in the direction of mobile station A in order to maximize the desired signal at the receiver. The other spatial beams are optimized for concurrent transmission, in order to minimize intra-beam interference in mobile station A. Neighboring cells produce inter-cell interference but the attenuated side lobes, resulting from the beamforming operation, result in a small contribution to the interference at mobile station A.

The SINR at mobile station is computed according to equation below:

$$SINR_{SDMA} = \frac{S}{\sum_{allBeams} I_{intra} + I_{Inter} + N_0} \tag{4.10}$$

Where S is the power of the desired signal, I_{Inter} is the contribution due to inter-cell interference, I_{Intra} is the contribution to the interference from each other spatial beam in the same cell and N_0 is the power of the thermal noise at the receiver.

4.5 Proposed SDMA-Enabled DRA Architecture

The principle behind the SDMA-based DRA architecture proposed for Mobile WiMAX MAC layer is now described. The SDMA architecture defines the groups of users to transmit in the same radio resource (group of slots in time and frequency domains), based on the correlation among the set of channel matrices over all data sub-carriers composing each resource.

As described in the D3.3, in the DL-PUSC sub-channelization mode the map of radio resources associated to the TDD OFDMA frame is made up of 15 resources. Each resource comprises 10 sub-channels and 6 OFDM symbols, which result in a total of 30 slots available per resource, since each one slot comports one sub-channels and two symbols in the DL-PUSC mode. The map of resources in the SDMA architecture is illustrated in Figure 4.4.

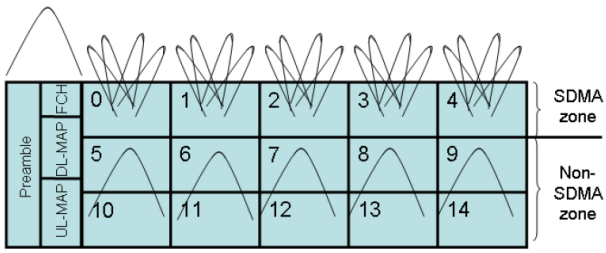

Figure 4.4: WiMAX TDD frame for SDMA implementation

For the implementation of SDMA, the original map of resources was divided in two zones:

- **SDMA zone**: each resource can be assigned to different mobiles by means of spatial beams. This zone comprises the resources in the first row of the map (indexes from 0 to 4), and, therefore, they are designated as SDMA-mode resources.
- **Non-SDMA zone**: there can be no spatial beam assignment and therefore SDMA is not applied to their resources. Resources in the other two rows of the map (indexes from 5 to 14) constitute the non-SDMA zone and, therefore, they are designated as non-SDMA mode resources.

The reason behind this subdivision of radio resources into two zones has to do, mainly, with complexity issues. If all resource units were allowed to transmit in SDMA mode the computational effort increase substantially, because correlations and verifications of SINR compatibility checks would have to be performed for all resources in the frame.

As can be seen from the figure, two types of radiation patterns were used in resource assignment for packet transmission and signalling broadcast:

- One radiation pattern with a larger beamwidth for broadcasting the preamble and control signals in the downlink sub-frame (FCH, DL-MAP, UL-MAP) and the control signals in the uplink sub-frame. This radiation pattern is also used for broadcasting resources in the non-SDMA zone in the downlink sub-frame (see Table 4.2).
- A second radiation pattern with a narrower beamwidth corresponding to the beamformed pattern used in the broadcast of the resources in the SDMA-zone. This beamformed pattern can be steered in the direction of the desired mobile (see Table 4.2).

4.5.1 Computation of Users Correlation

In the uplink sub-frame of the Mobile WiMAX system the base station defines a permutation zone for channel sounding [EMS02]. The correlation is computed from pilot symbols specially programmed for channel sounding in this zone, and for all data sub-carriers in each sub-channel composing each resource in SDMA mode in the map of resources.

Assuming user *i* is to be assigned a resource, the correlation of its channel matrix, H_n^i, with the channel matrix, H_n^j, of user *j* (which is already assigned to one spatial beam in the same resource), for the nth data sub-carrier, is simply the scalar product of both channel matrices as given by equation below.

$$\rho_{i,j}^{n} = \frac{\left\| \left(H_n^i\right)^H H_n^j \right\|_F}{\left\| H_n^i \right\|_F \left\| H_n^j \right\|_F} \tag{4.11}$$

Then, the algorithm performs the average over all data sub-carrier in the vector of correlations, as in equation below:

$$\overline{\rho}_{i,j}^{n} = \frac{\sum_{i=1}^{N} \rho_{i,j}^{l}}{N} \tag{4.12}$$

Where N is the number of data sub-carriers in the resource.

It is important to mention that a better approach would be to correlate the matrices resulting from stacking the channel matrices for all data sub-carriers as given in equation below.

$$\overline{\rho}_{i,j} = \frac{\left\| \mathrm{H}_i \mathrm{H}_j^H \right\|_F}{\left\| \mathrm{H}_i \right\|_F \left\| \mathrm{H}_i \right\|_F} \tag{4.13}$$

Where $\mathrm{H}_i = stack\left[\mathrm{H}_1^i \ldots \mathrm{H}_n^i \ldots \mathrm{H}_N^i\right]$ and $\mathrm{H}_j = stack\left[\mathrm{H}_1^j \ldots \mathrm{H}_n^j \ldots \mathrm{H}_N^j\right]$ are the matrices resulting from stacking the vectors of channel gains for each data sub-carrier. The problem with such an approach would be the increased degree of complexity in the computations. These would increase the computational effort. However, for the DL-PUSC subchannelization mode there is a kind of frequency diversity in the allocation of data sub-carriers due to their pseudo-random distribution over the spectrum. This seems to be a good reasoning for the proposed metric.

4.5.2 Computation of SINR for Resources in SDMA and in non-SDMA Zones

All active mobile stations in the cell read the preamble of the TDD OFDMA frame to estimate their channel quality (CQI). Also, besides being used for CQI estimation, the preamble is used in time and frequency synchronization between the mobile and its serving cell. It is broadcasted with the most robust MCS scheme for even mobiles in the edge of the cell to be able to decode its modulated pattern. For this reason, the preamble must be transmitted in the nonbeamforming configuration. For the computation of the SINR for all mobiles being assigned non-SDMA mode resources the antenna pattern that is considered in the serving and interfering base station was a beam with 70 degrees of half power bandwidth and 20 dB of maximum attenuation.

For all mobiles being assigned SDMA-mode resources, the SINR will depend on the type of antenna radiation pattern and on the beamforming algorithm. To simplify the simulations, it is assumed that the beamforming algorithm performs optimally in the determination of the line-ofsight direction between each mobile station and its serving cell, and on the computation of the set of weights to point a beam in this direction. The SINR must also be computed for each data sub-carrier in the radio resource. Ignoring time, mobile and used SDMA-mode resource indexes and assuming the desired mobile station is assigned a spatial beam with index l and that the serving base station has index i, the SINR for the nth data sub-carrier is given by equations (4.14- 4.17).

$$SINR_{SDMA}(n) = \frac{P^{i,l}_{Signal}(n)}{P^{i,l}_{Inter}(n) + P^{i,l}_{Intra}(n) + N_0 W_i F_{MS}} \tag{4.14}$$

$$P^{i,l}_{Inter}(n) = \sum_{j=1}^{N_{Inter}} \frac{G_{MS} \sum_{k=1}^{N_{Beams}} \frac{P_{BS_j}}{N_{beams}} \left| H^k_{BS_j \to MS(n)} \right|^2 G^k_{BS_j}}{PL_{BS_j \to MS} SH_{BS_j \to MS} L_{loss}} \tag{4.15}$$

$$P^{i,l}_{Intra}(n) = \frac{G_{MS} \sum_{k=1,k\neq l}^{N_{Beams}} \frac{P_{BS_i}}{N_{beams}} \left| H^k_{BS_i \to MS(n)} \right|^2 G^k_{BS_i}}{PL_{BS_i \to MS} SH_{BS_i \to MS} L_{loss}} \tag{4.16}$$

$$P^{i,l}_{Signal}(n) = \frac{G_{MS} \sum_{k=1,k\neq l}^{N_{Beams}} \frac{P_{BS_i}}{N_{beams}} \left| H^l_{BS_i \to MS(n)} \right|^2 G^l_{BS_i}}{PL_{BS_i \to MS} SH_{BS_i \to MS} L_{loss}} \tag{4.17}$$

Where:

- $P^{i,l}_{Signal}(n)$ is the desired signal from the serving cell i for the mobile transmitting in beam l.
- $P^{i,l}_{Inter}(n)$ is the contribution from the neighbouring cells to the inter-cell interference in the mobile station assigned the spatial beam l in cell i.
- $P^{i,l}_{Intra}(n)$, is the contribution from the other spatial beams in the same serving cell i in the same resource to which the mobile station is assigned to.
- $\left| H^k_{BS_i \to MS(n)} \right|^2$ is the channel gain from the desired mobile station to base station i in the kth spatial beam.
- G^l_{BSi} is the gain of the antenna array at base station *i* for the lth spatial beam.
- $PL_{BSj \to MS}$ and $SH_{BSj \to MS}$ are, respectively, the path loss and shadowing from base station *i* to the desired mobile station

- L_{loss} combine the losses due to cable attenuation, penetration loss, etc.
- $\frac{P_{BS_i}}{N_{beams}}$ is the power assigned to each spatial beam in each resource for base station i. It is important to remember that the power is divided uniformly over the set of resources.

The estimated SINR is computed from the vector of SINR values for each sub-carrier, using the

Effective Exponential SINR Mapping (EESM) compression method.

4.6 Joint Scheduling and Resource Allocation Algorithm Using SDMA Multiple Access

The proposed SDMA-based DRA algorithm comprises the following steps:

A. Prioritization (scheduling)

The first step is performed by the scheduler, which outputs a list with mobiles sorted by their decreasing order of priority. The utility-based scheduler is used as a reference.

B. Resource allocation

For each user returned from the priority list, the number of resources required is computed from the amount of information in its buffer and from the MCS scheme returned from the Link Adaptation module. If there are any SDMA-mode resources available, the resource allocator starts with the resource corresponding to the smallest accumulated spatial correlation to all mobiles in the SDMA-zone. In the sequence the algorithm verifies if the mobile can be assigned to the given resource.

Among the users that can be separated spatially, the selected (allocation of resources) is conducted in such a way that minize intra-beam interference. Assuming each user is assigned one spatial beam, new users are not allowed to transmit in empty spatial beams of the same resource, to which other users were already assigned into, if they might cause intra-beam interference high enough to degrade the CQI value corresponding to the selected MCS scheme for these users already assigned. In other words, for users which are too close to be spatially separable, the beam would be assigned only to the user with the highest priority. In order to use resources efficiently the mobile has to operate above the MCS SINR threshold to avoid excessive packet errors, which force the predicted BLER to be lower than the threshold. A new mobile may be allocated the same resource as long as the resulting compressed SINR ratio, after beamforming, is kept above the same threshold value associated to the selected MCS scheme from the CQI of the preamble.

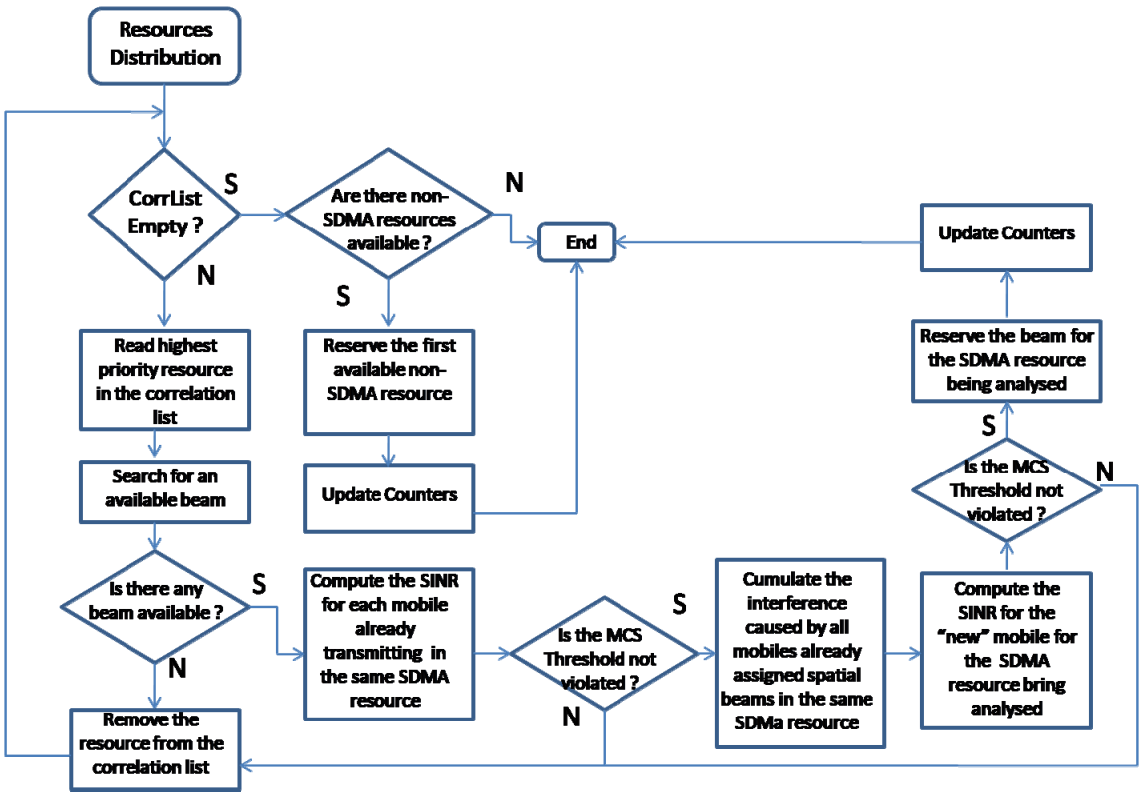

Figure 4.5: Fluxogram of the algorithm followed in the atribution of spatial beams in SDMA

The principle behind beam assignment to new users is the following:

1. The resource allocator determines the total amount of resources needed by the next mobile in the priority list.
2. For each resource in the SDMA-zone the resource allocator computes the spatial correlation between its channel matrix and the channel matrix corresponding to each one of the mobiles already allocated in the SDMA-mode resource. The resulting spatial correlations are sorted in increasing order of their value.
3. If there are resources available for allocation in the SDMA zone the resource allocator first tries to assign spatial beams to SDMA-mode resources. Only then it allocates resources in the non-SDMA zone.
4. The process starts with the resource resulting into smaller correlation among channel matrices. For this resource the algorithm computes the new SINR value (with the intrabeam interference) from the new mobile to each one of the already assigned mobiles. If the SINR level of each mobile is degraded to a level lower than the threshold, corresponding to the MCS scheme selected by the mobile already assigned, the process stops and the next resource in the correlation list is checked.
5. If the new mobile can be assigned an empty beam without affecting other already assigned users, the SINR of the new mobile is computed according to the amount of intrabeam interference from the already assigned mobiles into the new one. If the SINR level of the new mobile is degraded to a level lower than the threshold, corresponding to the MCS scheme selected, the process stops and the next resource in the correlation list is checked. Otherwise, an empty beam is assigned for this mobile in the SDMA-resource. In this way it is assured Quality of Connection for all mobiles transmitting spatially in the same resource.

Cell Layout	Hexagonal Grid, 19 cell sites, 3 sector BSs, 1 cell reuse in a cloverleaf layout with wraparound to simulate interference to edge cells
Number of Users	200 (Traffic Model Used); 72 (Full Queue Used)
Cell radius and Minimum MS to BS distance	900m; 35m minimum distance to the BS
BS Antenna Model for Horizontal Pattern	$A(\theta) = -\min\left[12\left(\frac{\theta}{\theta_{3dB}}\right)^2, A_m\right]$, Antenna Gain: 15dBi

BS Antenna Pattern in non-SDMA zone θ_{dB}, A_m	$\theta_{dB} = 70^{o}, A_m = 20dB$
BS Antenna Pattern in SDMA zone θ_{dB}, A_m	$\theta_{dB} = 8{,}75^{o}, A_m = 29dB$
BS Antenna bore-sight gain in non-AAS zone	3dBi
BS Antenna bore-sight gain in AAS zone	15dBi
Number of Beams per SDMA resource	4
MS Antenna Gain	Omni-directional with 0dBi
BS Maximum Transmission Power	43dBm
Propagation model	$L = 128.1 + 37.6\log_{10}(R)$, R in Km
Penetration Loss	10dB
Log-Normal and Shadowing Correlation	Standard Deviation = 8dB; 0.5 for sectors of different BSs and 1 for sectors of the same BS
Channel Model	ITU PedB and PedA with 3km/h; ITU VehA with 30km/h
Mobile Station Receiver	MRC
Traffic Models	VoIP and WWW
Duplex Mode	TDD
Operating Frequency	2.5GHz
Bandwidth and FFT size	10MHz; 1024 sub-carriers
Frame Duration	5ms
Number of OFDM Symbols in DL	35 (30 symbols available for data transmission)
Preamble, FCH, DL/UL MAP overhead	5 symbols
Sub-channelization	DL-PUSC
Number of Resources non-SDMA zone	10
Number of Resources SDMA zone	5
Burst Size	10 sub-channels x 6 symbols (15 resources)
Moving average filter length	1.5s
T_{CQI} (CQI feedback delay)	1 frame period in CQICH UL control channel
$T_{ACK,NACK}$ feedback	1 frame period in ACK/NACK control channel
Discard Timer and Priority lengths	15 and 3 frame periods respectively
N_{ret} Maximum Number of Retransmissions	4
Number of HARQ processes per MS	4
BLER Threshold for Link Adaptation	10%
Admission Threshold Parameter	-5dB
Scheduler	Full Queue – Max C/I; VoIP and WWW – Utility

Table 4.2: Simulation Setup Configuration

4.7 Results

System level simulations were conducted in the system level simulator for mobile WiMAX by incorporating the utility-based packet scheduler as detailed in 4.2.

Simulations were conducted separately for three types of traffic models: Full Queue, VoIP and 3GPP's WWW. The delay bound for VoIP and WWW are respectively 80 ms and 0.5 s. Full queue is used jointly with the simple and opportunistic Maximum C/I algorithm in order to estimate the gains achieved with the SDMA-based DRA, against the system throughput resulting from the non-SDMA-based DRA.

The scenario used for system level simulations was the one of a 4x2 MIMO channel antenna system configuration with Alamouti STBC. Simulations were performed both for the SDMA configuration, with a maximum of 4 beams per SDMA resource, and with the normal mode, without implementation of SDMA. Five resources were configured to be used in the SDMA zone and the remaining 10 resources were configured to be used in the non-SDMA zone.

Since in this work we intend to evaluate the achievable capacity of the proposed SDMA with positiong information, it was assumed that perfect knowledge of positioning of all mobiles is available, and that the direction of arrival from each mobile is perfectly estimated and that the beamforming algorithm is optimal in steering one spatial beam into the direction of the desired mobile. Power is uniformly distributed among the set of beams in each resource and neighbouring cells are assumed as transmitting with

maximum power (with full load) in non-SDMA configuration. They are considered only for the generation of inter-cell interference.

Table 4.2 lists the parameters used in the setup of the scenario used in the system level simulations.

We define some of the performance metrics used.

Service Throughput: is the average service throughput per cell is defined as the sum of the total amount of bits successfully received by all active users in the system, divided by the product of the number of cells simulated and the simulation duration.

Over-The-Air Throughput per cell: is the average OTA throughput per cell is defined as the sum of the total amount of bits being successfully and unsucessfully received by all active users in the system divided by the product of the number of cells simulated and the simulation duration.

3GPP Over-The-Air Throughput per cell: is defined as the sum of the total amount of bits being successfully received by all active users in the system divided by the product of the number of cells being simulated in the system and the total amount of time spent in the transmission of these packets.

4.7.1 Performance for Full Queue Traffic Users

In order to infer about the system capacity using the SDMA multiple access scheme, system level simulations were conducted for the Full Queue traffic model with the opportunistic Maximum C/I scheduler (schedule for transmission the mobiles with best channel conditions) [VJA05] for a total amount of 72 users. Figure 4.6 shows the average Over-The-Air (OTA) throughput (peak bit rate), the 3GPP Over-The-Air throughput, the average service throughput and the average offered load, per sector in Mbps.

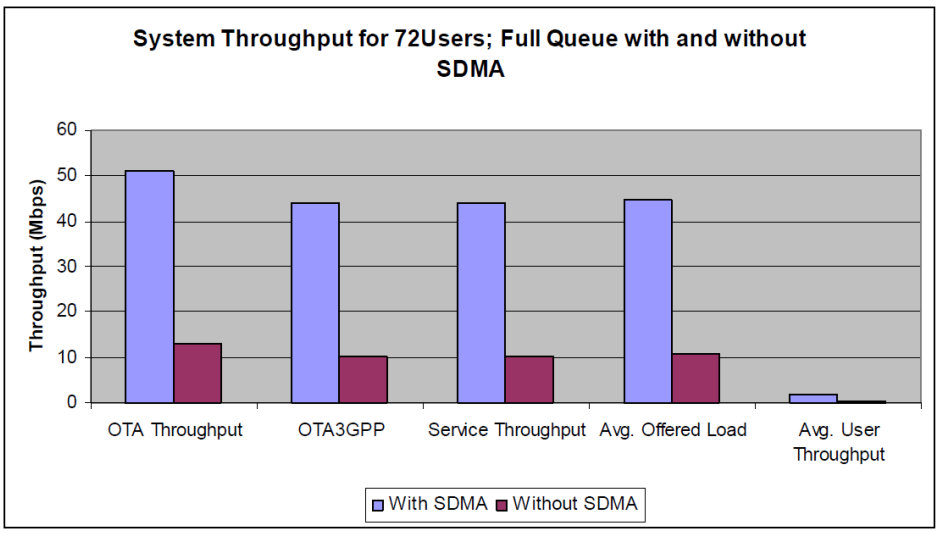

Figure 4.6: System throughput: SDMA and non SDMA-based DRA, for full queue and with 72 users load

It is noticed that if SDMA multiple access is implemented into the DRA the OTA throughput for the full queue traffic model can reach up to 50 Mbps, resulting in a throughput gain of roughly 38 Mbps over the simple non SDMA-based DRA scheme. It is worth mentioning that the OTA throughput measures the total amount of bits transmitted over the air interface while the service throughput measures the total amount of successfully transmitted bits over the air interface, both within the same simulation time. The discrepancy between service and OTA throughput, although slightly, is due to users being serviced at the cell edge which will normally experience poor signal quality arising from inter-cell interference from neighbouring cells and also due to intra-cell interference arising from beam assignment in the space domain. It is evident the gain achieved with an SDMA-based DRA architecture in terms of system

throughput.

In spite of the significant performance gains achieved these are obtained with the Maximum C/I scheduler, which does not provide QoS requests and for the theoretical full queue traffic model. Therefore, it is to be expected that different figures would result from other types of “more practical” schedulers and traffic models.

4.7.2 Performance for VoIP and WWW Traffic Users

The performance of the SDMA-based DRA was evaluated with the utility-based packet scheduler for both VoIP and WWW traffic models.

Figure 4.7 plots the user satisfaction ratio for VoIP traffic users using either SDMA or non SDMA based DRA versus the offered load. The user satisfaction ratio is measured from the ratio of packets dropped due to bad channel quality and time-out overflow.

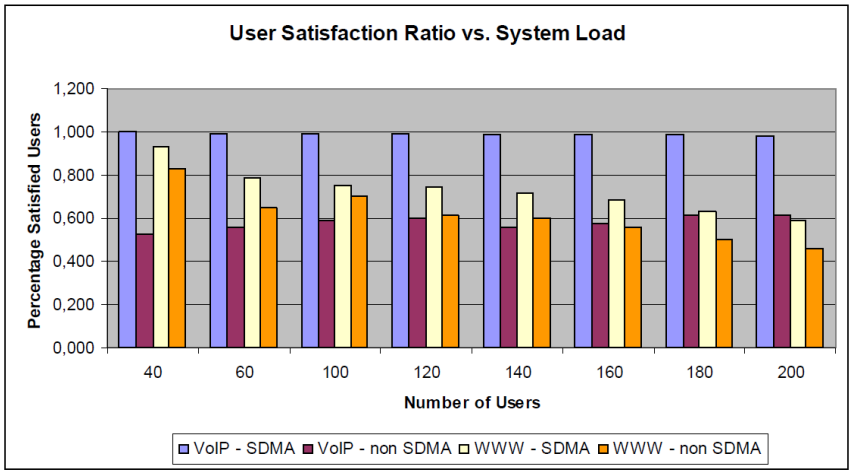

Figure 4.7: User satisfaction ratio: SDMA and non-SDMA based DRA for VoIP and WWW traffic models

Figure 4.7 plots the user satisfaction ratio for SDMA and non SDMA-based DRA for WWW traffic model versus the offered load. As expected, the SDMA-based DRA results into a higher percentage of satisfied users, compared to the non-SDMA based DRA. The reasons behind this performance gap between two modes are: (i) SDMA results in more resources available (in spatial domain); (ii) the algorithm followed in the allocation of spatial beams in the SDMAbased scheduler; (iii) WWW packets have to fragmented in order to fill radio resources, which translates into a small number of users transmitting over spatial beams in the SDMA zone of the map of resources per frame period and, as a consequence, this translates to an improvement inchannel quality because of the inherent smaller intra-cell interference.

As can be seen, for the VoIP users the user satisfaction ratio is roughly insensitive to the amount of users in the system as there are virtually no unsatisfied users if the SDMA-based DRA is implemented. This is because the system is relatively under-loaded for VoIP and, as the system load is not enough, additional radio resources cannot improve system performance due to small packet size and low utilization of radio resources. Nevertheless, one can observe a significant gain in terms of the satisfaction ratio for the non-SDMA based DRA. This is due to the smaller amount of resources available for allocation if the SDMA mode is not implemented.

Figure 4.8 is the plot of the average packet drop rate per user, averaged over all active users in the

network, versus the offered load

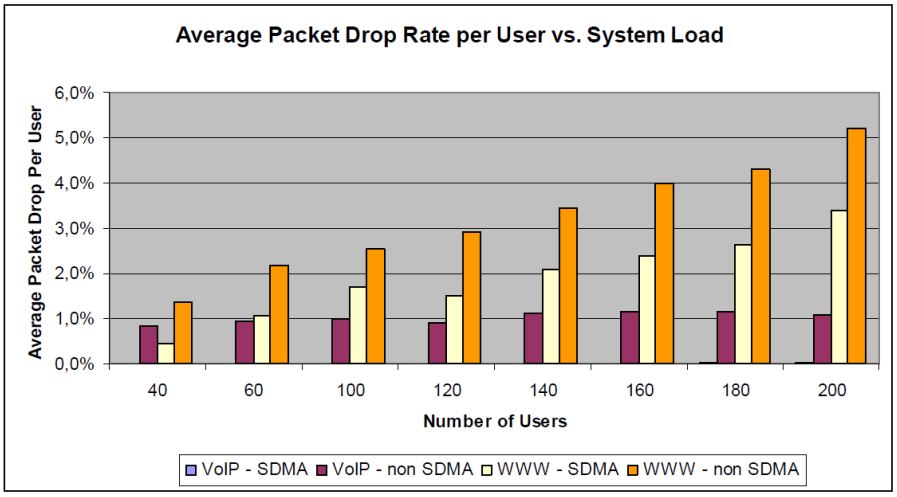

Figure 4.8: Average packet drop rate per user for SDMA and non-SDMA based DRA for VoIP and WWW vs number of users

As expected, the ratio of packets dropped due to time-out violation and/or bad channel quality is almost zero for the SDMA-based DRA. The difference in percentage to the non-SDMA mode is also small and remains constant with the increase in the system load, because, as mentioned, the system is under-loaded. The gap between both modes is due to the allocation strategy followed in the SDMA mode resource allocation.

Figure 4.9 is the plot of the average packet delay per user averaged over all active users in the network versus the offered load. It corroborates this same reasoning as it can be seen that the average packet delay remains roughly insensitive to the increase in the system load for both SDMA and non-SDMA modes.

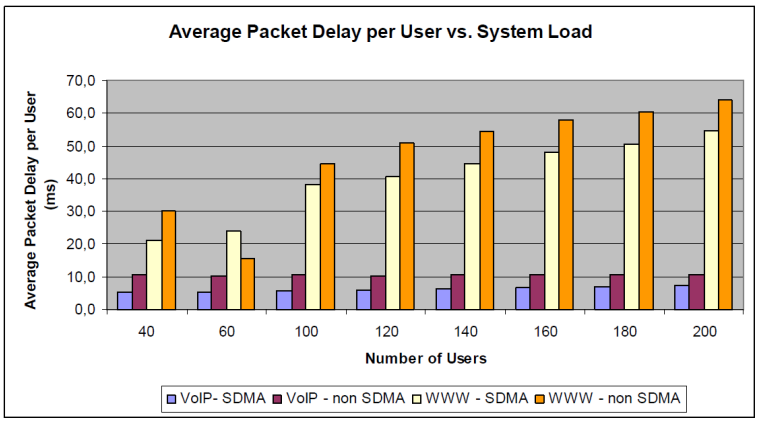

Figure 4.9: Average packet delay per user for SDMA and non-SDMA based DRA for VoIP and WWW vs number of users

Figure 4.10 is the plot of the average service throughput averaged over all active users in the network versus the offered load.

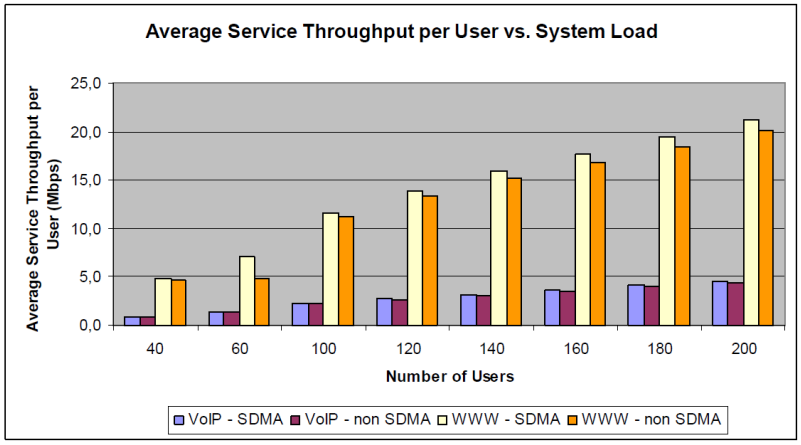

Figure 4.10: Average service throughput per user for SDMA and non-SDMA based DRA for VoIP and WWW vs number of users

As expected, the difference in the service throughput is almost zero for both modes and for all traffic loads. This has to do with: (i) under-loaded system; (ii) VoIP is a real time service with stringent delay constraints. Therefore packets cannot remain in buffer waiting for better channel conditions, as they are dropped whenever they achieve their maximum allowable delay. This results in a small gain for the multi-user diversity when using the utility-based scheduler.

Figure 4.11 shows the average OTA throughput per user averaged over all active users in the network versus the offered load. It can be seen that, for the same load, as the amount of resources increases the resulting multi-user diversity gain is smaller for the SDMA mode. Differently, for the non SDMA based DRA the multi-user gain is higher and this translates into a higher OTA throughput. As expected also, the OTA throughput increases with the amount of active users (system load) in the system

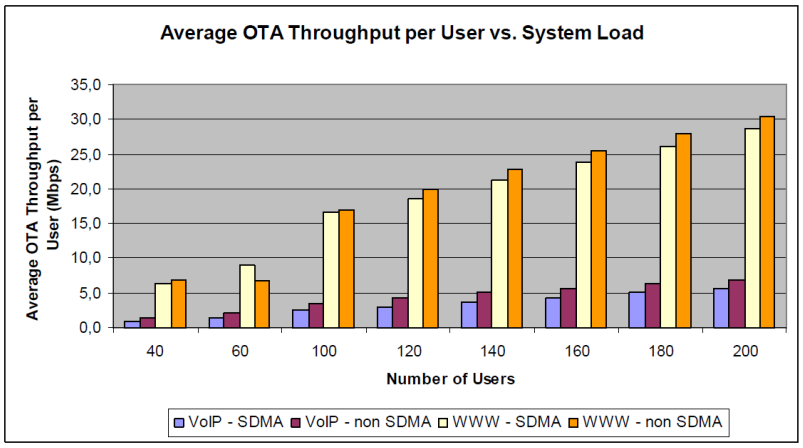

Figure 4.11: Average OTA throughput per user for SDMA and non-SDMA based DRA for VoIP and WWW vs number of users

Figure 4.12 is the plot of the average SINR per received packet averaged over all active users in the system versus the offered load. As expected, the average SINR is significantly better for the SDMA mode and it is roughly insensitive to the increase in the number of users in the system. This is due to the scheme followed by the utility based scheduler in the elaboration of the priority lists.

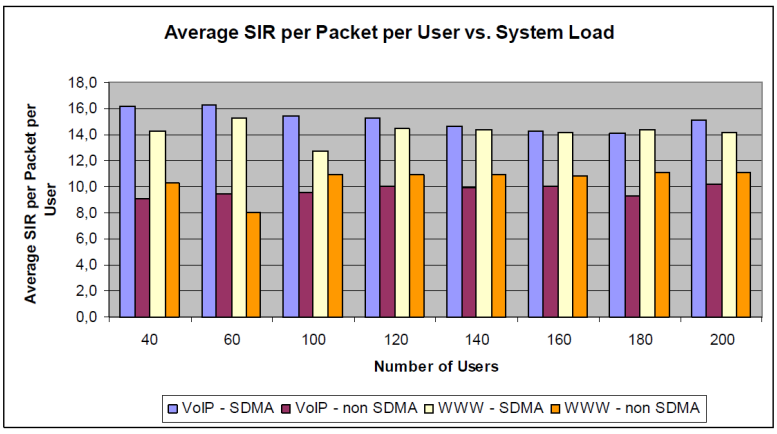

Figure 4.12: Average SINR per packet per user for SDMA and non-SDMA based DRA for VoIP and WWW vs number of users

4.7.3 Performance for Voice Over IP Traffic Model

Figure 4.13 is the plot of the CDF of the average packet drop rate per user for the highest load used in the simulations (200 users). It can be observed the significant difference in performance between both modes. This has to do with the strategy followed in the allocation of spatial beams in the SDMA-based DRA: users are assigned spatial beams according to two conditions: (1) the smallest value of the channel matrix correlation and also if (2) the SINR condition is not violated due beam-related intra-cell interference.

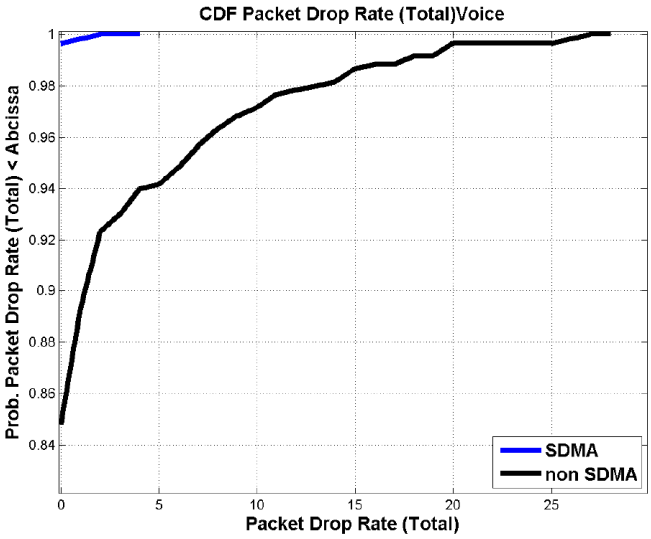

Figure 4.13: CDF of average packet drop rate for VoIP users

Figure 4.14 is the plot of the CDF of the average packet delay per user for the highest load used in the simulations (200 users). As expected, the SDMA-based DRA results in small packet delays per user. But the difference is not very significant due to: (i) VoIP packets are highly constrained in terms of delay, which means packets cannot wait in buffer for a better transmission opportunity and the utility scheduler prompts them for transmission as the delay approaches the delay bound; (ii) the system is under-loaded.

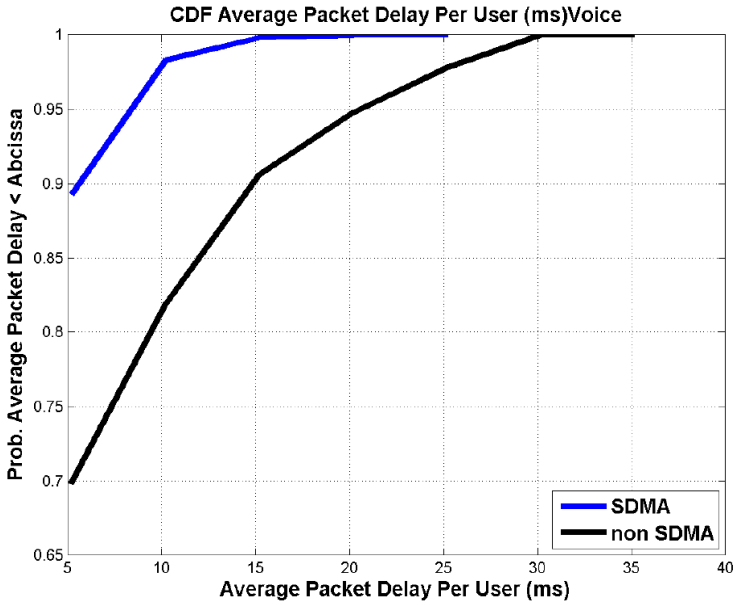

Figure 4.14: CDF of the average packet drop rate per VoIP user

Figure 4.15 is the plot of the CDF of the average service throughput per user for the highest load used in the simulations (200 users). As can be seen, some users in the edge of the cell have smaller service throughput with the non-based SDMA DRA.

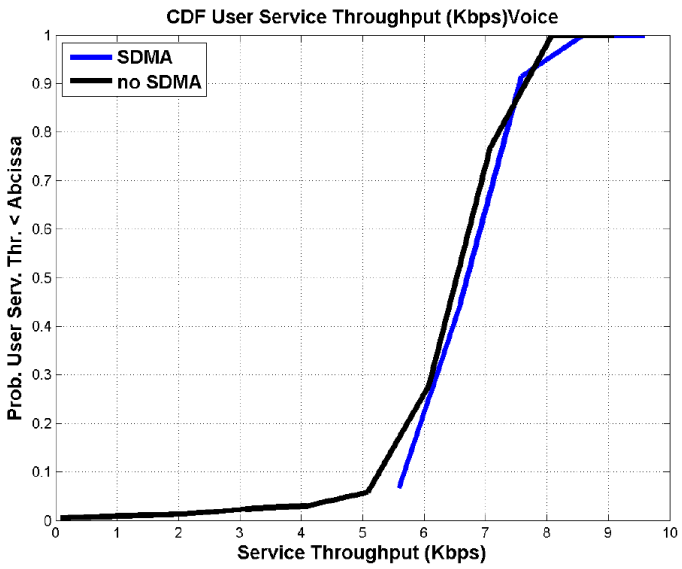

Figure 4.15: CDF of the average service throughput per VoIP user

To evaluate the impact on the cell edge, is presented in the Figure 4.16 the average service throughput in function of the geometric factor, which is the ratio of channel from the serving cell loss over the sum of channel of all interferring cells. Due to the mechanism followed in the allocation of empty beams in the SDMA-based DRA, users in the edge of the cell have more transmission opportunities and this can be confirmed from the plot in Figure 4.16 for the highest system load used in the simulations (200 users).

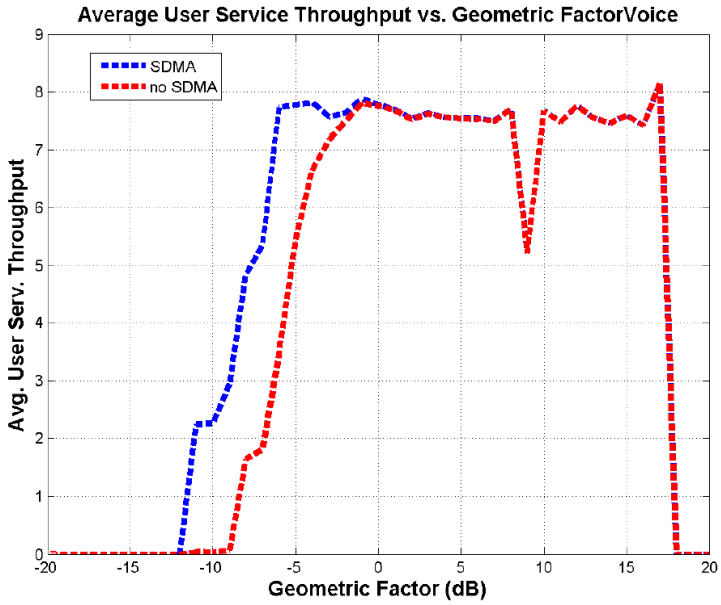

Figure 4.16: Average user service throughput versus geometric factor

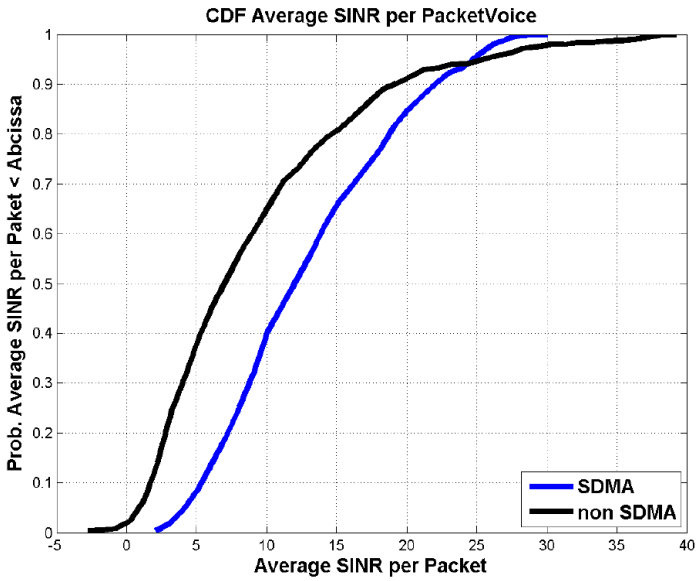

Figure 4.17: CDF of user average SINR per packet for VoIP users

The plot in Figure 4.17 is the CDF of the average SINR per received packet averaged over all active users in the system, for the highest load used in the simulations (200 users). The plot resulting from the non-SDMA mode has a heavy tail in the highest SINR zone. This tail translates to a higher maximum SINR for the non-SDMA mode. This is because a DRA with no SDMA mode is inherently more loaded than if SDMA is implemented, which results in a roughly higher gain of multi-user diversity which can pick users with better channel conditions for transmission. Also, the non SDMA mode uses a higher transmission power per resource, which is lower for the SDMA mode.

4.7.4 World Wide Web Traffic Model

Packets generated according to the WWW traffic model are of much larger size than VoIP ones. Therefore, for the same amount of users, the system is more loaded if WWW traffic model is used rather than VoIP. The higher load results in a significant higher multi-user diversity gain. Also WWW traffic model is much more tolerant in terms of packet delay than VoIP. This means that packets can remain for longer periods of time in buffer waiting for better channel conditions and performing more transmission attempts. Therefore, resources are used more efficiently with WWW traffic model. For the WWW traffic model the system can be assumed as having enough load. With the increase in the load there is a corresponding decrease in the amount of satisfied users because more packets are dropped due to bad channel quality and/ordelay overflow.

Figure 4.18 plots the CDF of the average packet drop rate, residual and time-out respectively, for both SDMA and non SDMA modes, with the WWW traffic model and the highest load assumed in the simulations (200 users). It is evident from these two plots that the main contribution to the difference in the average packet drop rate for the two schedulers is due to the amount of residual packets, i.e., packets which are dropped when the maximum number of transmission attempts is reached, which is significantly higher for the non SDMA-based scheduler. SDMA mode achieves much better quality transmissions according to the reasoning above.

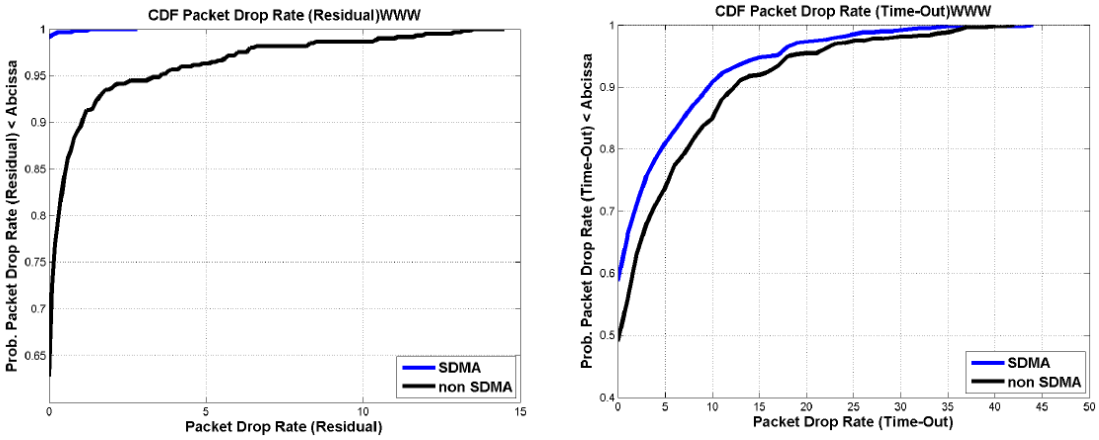

Figure 4.18: CDF of the average packet drop rate for (a) residual (b) time-out for WWW users

Figure 4.9 is the plot of the average packet delay per user, averaged over all active users in the network, versus the offered load, for the WWW traffic model. As expected, there is an increase in the average packet delay with the amount of users in the network: as the SDMA-based scheduler provides more resources for allocation, with the SDMA-based DRA the average packet delay is smaller for all values of the load. Figure 4.19 plots the CDF of the average delay per packet for all active users in the network and for the highest system load of 200 users used in the simulations.

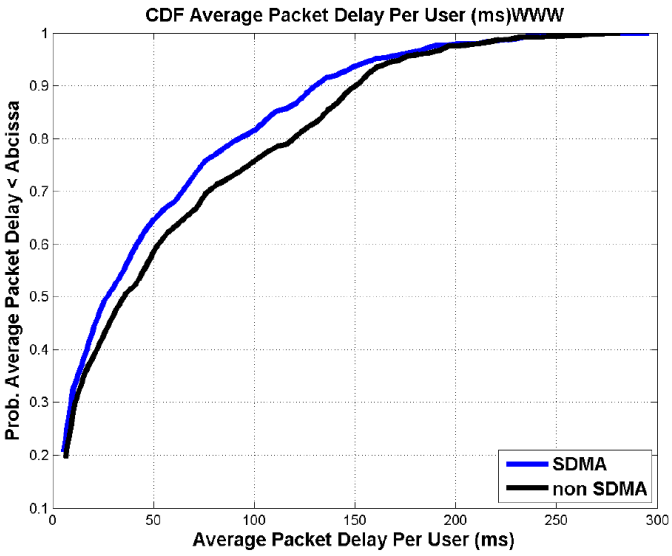

Figure 4.19: CDF of the average packet delay for WWW users

As can be seen from the plot in Figure 4.10, representing the average service throughput averaged over all active users in the system, the SDMA-based DRA results into higher service throughput for all system loads assumed in the simulations. The main contribution for this difference arises from the amount of transmission opportunities for users in the edge of the cell, which are allocated empty beams with enough quality for transmission (limiting the intra-beam interference) in the SDMA-based scheduler. This can be seen from the plots in Figure 4.19 and Figure 4.20 of the CDF of the user average service throughput and of the variation of the average service throughput with the geometric factor, respectively, for the maximum load assumed in the simulations. The difference in performance is not so significant because the size of the transport block with the most robust MCS scheme is very low and users in the edge of the cell (with low geometric factors) are assigned the most robust MCS scheme for transmission.

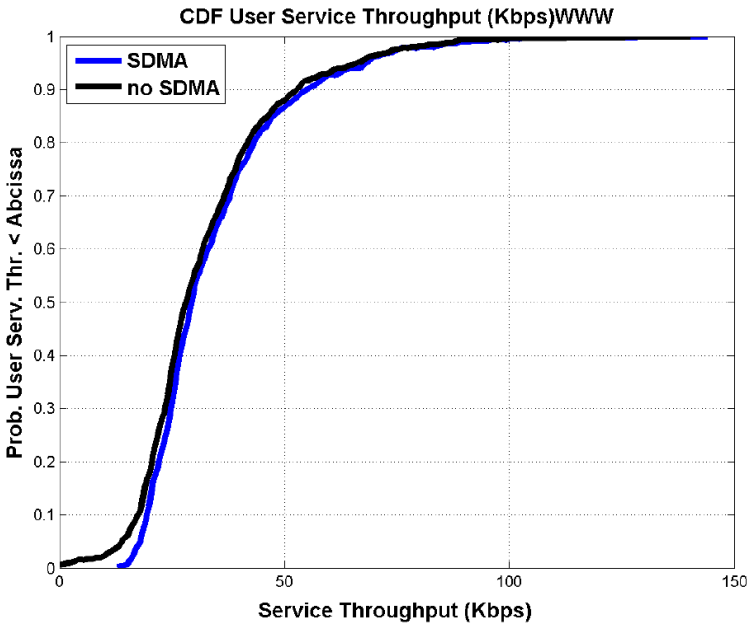

Figure 4.20: CDF of the average user service throughput for WWW users

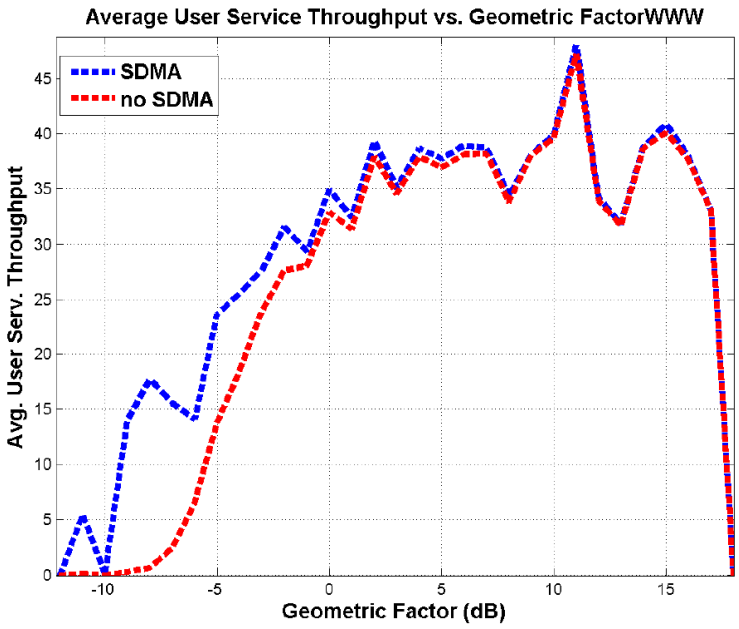

Figure 4.21: Average user service throughput versus the geometric factor for WWW users

As can be seen from Figure 4.11 the average OTA throughput increases with the increase in the amount of active users in the cell. This is the result of the multi-user diversity gain arising from the opportunistic channel access. This effect is evident with the WWW traffic model because WWW packets are more tolerant to delay than VoIP packets. Together with the shape of the utility function used in the utility scheduler, it is a consequence of the fact that under these conditions the scheduler behaves much like the opportunistic maximum C/I scheduler. Another reason for this behavior is that the offered load (in bits) is more significant for WWW packets, which are much larger in size than the packets from VoIP, and this results in a performance enhancement by creation of more radio resources in the SDMA zone.

As can be seen in Figure 4.12, due to the strategy followed in the allocation of empty beams in the SDMA-based DRA, the SINR of packets received correctly in the mobile is much better than without SDMA allocation. This is the same result which was verified with the VoIP traffic model and the same reasoning behind the increase in the service throughput with the amount of users in the system applies.

Figure 4.22 is the plot of the CDF of the average SINR per received packet for both SDMA modes and for the maximum load assumed in system level simulations.

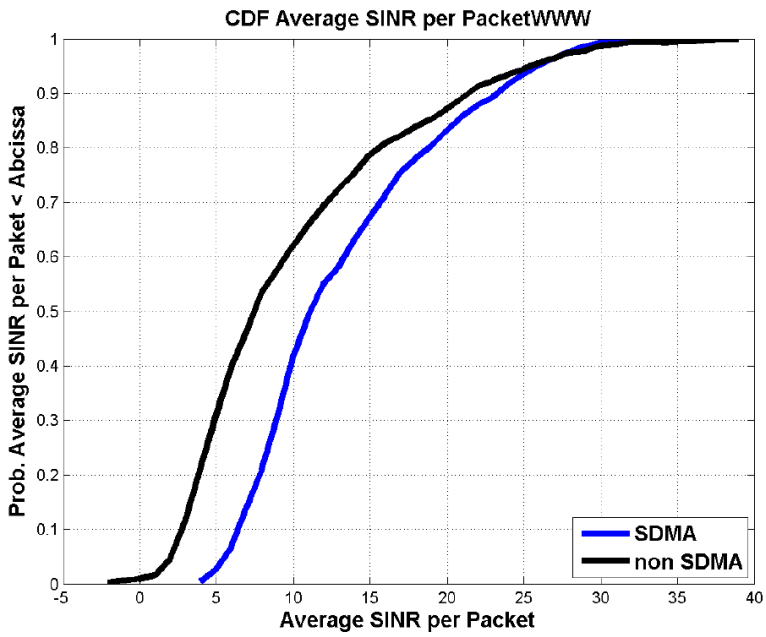

Figure 4.22: CDF of the average SINR per packet for WWW users

4.8 Related Work

There are few proposals in the research literature regarding the implementation of SDMA in WiMAX networks, in particular for the Mobile WiMAX standard.

In [HY03], a multimedia SDMA/TDMA scheduling algorithm is proposed. The algorithm consists of a packet scheduler and a packet allocator. The scheduler prioritizes packets from different users according to the requested QoS, defined in terms of the minimum SINR and time-out values. The allocation is performed for the time-slot which results in the minimum degradation of the SINR, achieved by already allocated users in each time-slot and according to the user's required SINR. The authors claim that their algorithm results in the satisfaction of OoS and in the maximization of the cell's throughput.

In [HWP06]-[HEW06] an exhaustive analysis, regarding the support of SDMA techniques in the MAC layer of WiMAX standard is presented together with validation results. To the best of the our knowledge, [CCP08] is the only work, up to now, in the literature regarding the proposal of a practical SDMA algorithm implemented in the Mobile WiMAX system. Should be noticed that in this work, System level simulations were conducted to reveal the gains resulting from the proposed dynamic resource allocation with SDMA.

4.9 Conclusion

This chapter details all the steps performed in the design and performance analysis of a new DRA architecture based on the addition of space domain as a new degree of freedom for resource allocation in Mobile WiMAX networks.

In the PHY layer radio resources are available in time, frequency and space domains by the joint implementation of OFDMA and SDMA over the same air interface. The MAC layer assigns spatial beams as long as new users do not significantly affect already assigned users in the same set of symbols and frequency channels. SDMA is made possible in Mobile WiMAX with the use of beamforming antennas which is standardized as AAS in the IEEE 802.16e standard for Mobile WiMAX networks. The original utility-based packet scheduler is used in conjunction with the proposed DRA.

A set of system level simulations was performed to infer on the gains achieved over simple scenarios where a non SDMA-based DRA architecture is implemented. The gain in capacity which results from the implementation of a SDMA-based DRA architecture over the non SDMA one is less pronounced whenever traffic models are used in detriment of the theoretical full queue traffic model. Traffic models with a pronounced activity factor and large packet sizes increase the SDMA cell throughput, since users with the highest channel propagation conditions would be continuously served with the highest MCS scheme, avoiding underutilization of resources. This is the reason for the observed differences in the results obtained for WWW and VoIP. Improved results can be attained by using greater density of users

to improve the multiuser diversity gain and a priority-based scheduler to control the trade-off between maximization of cell throughput and keeping the number of satisfied users at an acceptable level.

Full queue together with the Maximum C/I algorithm is used in the estimation of the gain in capacity achieved with the SDMA-based DRA. The OTA throughput can reach up to 50 Mbps, a gain of roughly 38 Mbps over the simple non SDMA-based DRA scheme.

The obtained gain was with the proposed SDMA, which included the selection of 5 slots over 15 for SDMA. Higher values can be obtained alocating more slots for the SDMA frame, in expense of complexity. The reliability in practical systems has to be tacken into account, due to fast scheduling requirements of high-speed packet based systems.

For VoIP users the user satisfaction ratio is roughly insensitive to the amount of users in the system. As a matter of fact, the small packet size and low utilization of radio resources from VoIP traffic demands a great amount of users to be simulated, in order to produce figures which could show the potential gain in system capacity from the implementation of an SDMA-based DRA for VoIP. For the same reasoning the average packet delay remains roughly insensitive to the increase in the system load for both SDMA and non-SDMA modes.

In order to have an idea of the potential gain of SDMA in such scenario, simulations were conducted for a load equal to 200 users. Average packet drop rate per user was produced from these simulations from which it was observed the significant difference in performance between both modes. This is because the implemented algorithm for SDMA attributes spatial beams to already assigned slots in the frame without compromising the channel quality sensed by those already assigned users, in such a way that the selected MCS scheme is not violated due to the assignment of beams in the same resource.

For higher loads SDMA-based DRA results in small packet delays per user. But the difference is not that significant due to the fact that VoIP packets are highly constrained in terms of delay, which means packets cannot wait in buffer for a better transmission opportunity and the utility scheduler prompts them for transmission as the delay approaches the delay bound. Therefore packets cannot remain in buffer waiting for better channel conditions, as they are dropped whenever they achieve their maximum allowable delay. This results in a small gain for the multi-user diversity when using the utility-based scheduler.

Due to the mechanism followed in the allocation of empty beams in the SDMA-based DRA, users in the edge of the cell have more transmission opportunities. The average SINR is significantly better for the SDMA mode and it is roughly insensitive to the increase in the number of users in the system. This is due to the scheme followed by the utility based scheduler in the elaboration of the priority lists.

Packets generated according to the WWW traffic model are of much larger size than VoIP ones. Therefore, for the same amount of users, the system is more loaded if WWW traffic model is used rather than VoIP. WWW traffic model is much more tolerant in packet delay than VoIP and packets can remain for longer periods of time in buffer waiting for better channel conditions and performing more transmission attempts. Higher load brings and less constringent packet delays brings in higher gains in system capacity due to multi-user diversity gain and, therefore, resources are used more efficiently. With the increase in the load there is a corresponding decrease in the amount of satisfied users because more packets are dropped due to bad channel quality and/or delay overflow. But the decrease in user satisfaction ratio is significantly smaller for the SDMA-based DRA.

It was observed that, for WWW services, which have large packet sizes and delay tolerances, the performance enhancement is contributed by creating more available radio resources in theSDMA zone. However, in the VoIP case, if the system load is not enough, additional radio resources cannot improve system performance due to small packet size and low utilization of radio resources. SDMA was proved to be an efficient means for increasing system capacity whilst still providing QoS requirements from user's applications. The advantages resulting from the implementation of such architecture are higher for higher system loads and/or for traffic models with higher inherent load in packet generation. Last but not least, it is worth mentioning that the architecture implemented with the DRA based on the SDMA access scheme is also innovative and no similar approach was conveyed in the research literature available up to now.

5 Location Aided Underlay Cognitive Radio

5.1 Motivation and Objective

Cognitive radio techniques convey a dynamic and flexible spectrum allocation policy that allows a secondary user to opportunistically access a primary user's spectrum through exploitation of multiuser side information such as user activity, channel quality information (CQI), message, codebook, location information... A good tutorial about cognitive radio techniques can be found in [ZS07] that is focused on signal-processing perspective, and in [JSM09] that is focused on information-theoretic perspective. One group of cognitive radios is known as the spectrum-sensing cognitive radio or called the interweave paradigm, which can opportunistically access temporary spectrum holes or white spaces existing in both licensed or unlicensed radio spectrum [MM99]. Fast and reliable spectrum sensing techniques are the key to the success of interweave cognitive radios. The other group of cognitive radios includes overlay and underlay paradigms, where the secondary user and the primary user form a cognitive interference channel in the primary spectrum (e.g. [MYK07], [DPT06],[WVA07]). Specifically, in the overlay paradigm, the secondary user employs advanced coding schemes such as the Gel'fand-Pinsker code presented in [GP80] and the dirty-paper corresponding code presented in [COS83] to pre-cancel interference to the primary user. In the underlay paradigm, the secondary user can enter the primary spectrum only when its activity will not cause considerable interference or capacity penalty to the primary user. Our considered cross layer framework is based on physical (PHY)-layer -centric cross layer, where optimization and decision is done at the PHY layer. Then the decision is provided to MAC layer for both layers to adaptively adjust to enhance spectrum efficiency. The goal of this section is to study several fundamental PHY-layer issues about location aided Gaussian underlay cognitive radios (UCRs).

Figure 5.1 illustrates an example of two-user UCR system accommodating one primary transmitter (Tx_1) and receiver (Rx_1) pair in systems #1 and one secondary transmitter (Tx_2) and receiver (Rx_2) pair in system #2. The block diagram of this UCR system is depicted in Figure 5.2. In the flat-Gaussian channel, this UCR system can be described as the following linear model

$$Y_1 = a_{11}X_1 + a_{21}X_2 + V_1, \tag{5.1}$$

$$Y_2 = a_{12}X_2 + a_{22}X_2 + V_2, \tag{5.2}$$

where X_i stands for the signal sent by the transmitter Tx_i with power P_i and rate R_i, Y_j for the signal received at the receiver Rx_j, a_{ij} for the channel coefficient of the Tx_i-Rx_j link, V for the Gaussian noise with zero mean and variance N_o. This linear model shows that the UCR system is a special case of interference channels presented in [CAR78] and [SAT81], but the interference term $a_{21}X_2$ in (5.1) must not cause considerable capacity penalty to the primary user. According to multiuser decoding strategies, we can classify the UCR system into the following four typical modes. Detailed introduction about these four modes can be found in the following.

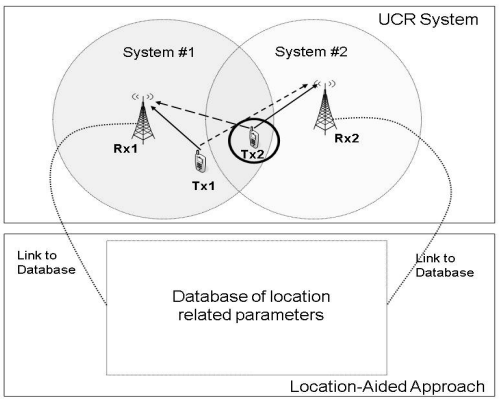

Figure 5.1 Illustation of an example about the two-user UCR system and a location-aided approach

Key physical-layer issues this work seeks to address are mainly in two folds:

- Issue 1:

  Provided full knowledge of CQI, what is the fundamental relationship between the secondary user's achievable rate, denoted by C_2, and capacity penalty to the primary user, denoted by ΔC_1? What are criteria for Tx_2 to perform efficient power allocation? Those questions require an answer for various UCR modes.

- Issue 2:

  In many practical environments, having full knowledge of CQI about all links of the UCR system is not a suitable assumption. What are more suitable assumptions in practice? What is the efficient UCR strategy under new assumptions? What is the secondary user's performance? Those questions require a satisfactory answer.

The primary objective of this work is to partially answer the above questions through an study from the information-theoretic viewpoint. In order to focus on the major technical issues, the following system configurations and assumptions are made throughout this section:

- A1): We consider a two-user UCR system accommodating one primary transmitter-receiver pair and one secondary transmitter-receiver pair.
- A2): Users in the system are synchronized in both the time and frequency.
- A3): Both receivers employ maximum-likelihood (ML) detector/decoder to offer the optimum decoding performance.

Contributions of this work are manifold:

- 1): The first work is to answer those questions listed in *Issue 1*. Provided full knowledge of multiuser CQI, the fundamental relationship between C_2 and ΔC_1 is investigated for four UCR modes. Criteria for efficient power allocation at Tx_2 are established. The produced results are the key to study new UCR strategies when the full knowledge of multiuser CQI is assumed not available.
- 2): As a starting point of the work towards *Issue 2*, we study modeling of UCR systems in the presence of full knowledge of multiuser CQI. After a careful justification, we establish an UCR system model, where the secondary user is assumed to have partial knowledge of CQI about the Tx_2 - Rx_1 link, and have full knowledge of CQI about the rest. Location-aided UCR is employed as an example to support our justification.
- 3): We propose new spectrum-access approaches for various UCR modes by assuming the availability of p.d.f. of CQI about the Tx_2 - Rx_1 link. Power allocation criteria are carefully investigated in terms of $C_2, \Delta C_1$, and capacity outage probability. Assuming the channel to be Rayleigh, numerical results are provided to show performance difference between the UCR with full knowledge of multiuser CQI and the proposed one with partial knowledge of multiuser CQI.

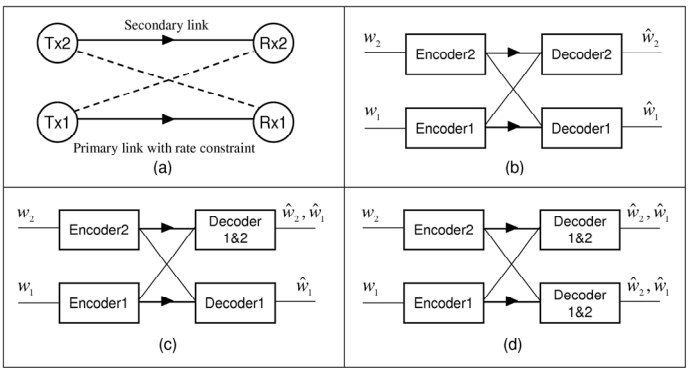

Figure 5.2 Block diagram of various UCR modes: (a) two-user UCR channel, (b) the individual mode, (c) the CSMD mode, (d) the TSMD mode.

5.2 System Model

This sub-section first presents capacity theorem about two-user Gaussian interference channel (GIC) and its relationship with the UCR system, and then presents modeling of the UCR with partial knowledge of CQI.

5.2.1 Two-User UCR with Full Multiuser CQI

The UCR system is a special case of interference channels. The information-theoretic research towards interference channels started from Carleial's work published in [CAR78]. Although lots of research efforts have been spent in the past 30 years, capacity region of interference channels has been found only for the case with strong interference [CG87]. To the best of our knowledge, the state-of-the-art capacity bound of two-user GIC has recently been reported in [KRA04] and [SKC09]. Next, we provide a brief review about capacity theorem of two-user GIC, which is the theoretical support to our further investigation towards the two-user UCR system.

1) Two-User GIC with Strong Interference: In the linear model (5.1)-(5.2), the case of strong interference denotes the scenario $|a_{12}| \geq |a_{11}|$ and $|a_{21}| \geq |a_{22}|$ [SAT81]. In this case, the two-user GIC is in fact a compound Gaussian multiple-access channel (GMAC), whose capacity region is known as the following union [CT06]

$$\bigcup \begin{pmatrix} R_1 < c[\gamma_{11}], R_2 < c[\gamma_{22}] \\ R_1 + R_2 < \min(c[\gamma_{21} + \gamma_{11}], c[\gamma_{12} + \gamma_{22}]) \end{pmatrix} \tag{5.3}$$

where $\gamma_{ij} = (P_i |a_{ij}|^2)/(N_o)$ denotes the instantaneous signal-to-noise ratio (SNR) and $c(x) = \log_2(1+x)$. Provided (5.3) and the assumption that users share their codebook, each receiver can recover the message sent by Tx_1 and Tx_2, respectively.

2) Two-User GIC with Weak or Mixed Interference: This scenario includes cases other than the case of strong interference. The closed-form of capacity region is unknown to this date. A look-up table (in one scenario, the sum-rate capacity is still unknown) about the channel capacity with respect to various channel conditions has been reported in [SKC09]. Alternatively, we can divide the two-user GIC into the following three groups with respect to the way of dealing with the interference. The following result is adequate for us to investigate the two-user Gaussian UCR system.

- *Group I*: Each receiver can reliably decode the message sent by Tx_1 and Tx_2, respectively. The achievable-rate region, denoted by $\Re^I$ is (5.3).
- *Group II*: Each receiver can only decode the message sent by the desired transmitter. The interference will be regarded as noise. The achievable-rate region, denoted by $\Re^{II}$ is

$$\Re^{II} = \bigcup \begin{pmatrix} R_1 < c\left[\dfrac{\gamma_{11}}{\gamma_{21}+1}\right] \\ R_2 < c\left[\dfrac{\gamma_{22}}{\gamma_{12}+1}\right] \end{pmatrix} \tag{5.4}$$

- *Group III*: One receiver can decode the message sent by both transmitters, and the other can only decode the message sent by the desired transmitter. In this group, the achievable-rate region, denoted by $\Re^{III}$ is

$$\Re^{III} = \bigcup \begin{pmatrix} R_j < c\left[\dfrac{\gamma_{jj}}{\gamma_{ij}+1}\right] \\ R < c[\gamma_{ii}+\gamma_{ji}] - R_j \\ R < c[\gamma_{ii}] \end{pmatrix}, i \neq j. \tag{5.5}$$

Provided $\gamma_{ij,i,j,=1,2}$, we can obtain the maximum sum-rate, $\max(R_1 + R_2)$, through a comparison between $\Re^{I}$, $\Re^{II}$, and $\Re^{III}$.

3) Two-User Gaussian UCR: The UCR system is a special case of interference channel, where the primary user wants to keep its interference-free capacity. But in some cases, the secondary user will cause capacity penalty ΔC_1 to the primary user. Hence, the primary user's capacity is expressible as

$$C_1 = c[\gamma_{11}] - \Delta C_1 \tag{5.6}$$

and the secondary user's achievable rate is

$$C_2 = \max(R_1 + R_2) - C_1. \tag{5.7}$$

Define

$$\Delta C_1 = \rho c[\gamma_{11}], \tag{5.8}$$

where ρ is a positive coefficient. Eqn. (5.6) is expressible as

$$C_1 = (1-\rho)c[\gamma_{11}]. \tag{5.9}$$

In order to keep the capacity penalty to be reasonably small, we should let $\rho << 1$.

5.2.2 Two-User UCR with Partial Multiuser CQI

In practice, the primary transmitter-receiver pair may operate in the frequency-division duplex (FDD) manner, where the transmitter Tx_1 periodically sends training sequences to support channel estimation and coherent detection/decode at the receiver Rx_1. Rx_1 informs Tx_1 regarding the CQI of Tx_1 - Rx_1 link through a feedback channel. On the secondary-user side, we assume that Rx_2 can communicate with Tx_2 through a low-rate feedback channel, which is employed for the purpose of signaling. This feedback channel is orthogonal to the primary user's frequency band. Based on the above system description, we provide the following justification of assumptions about the knowledge of CQI:

- The secondary receiver Rx_2 listens to the conversation between Tx_1 and Rx_1. Then, Rx_2 can estimate the CQI of Tx_1 - Rx_2 link.
- We assume that Rx_1 employs a simple common codebook such as repetition code to perform the feedback of CQI. Then, Rx_2 can obtain the CQI about Tx_1 - Rx_1 link through sensing of the primary user's feedback channel.
- At the beginning of cognitive communication, Rx_2 requests Tx_2 to send a training sequence over the primary spectrum. This offers the knowledge of CQI about the Tx_2 - Rx_2 link, but introduces a short burst of noise to the primary user. We argue that this burst of noise will not cause considerable performance loss to the primary user.
- Rx_1 may estimate the CQI of Tx_2 - Rx_1 link if appropriate, but does not show this information in its feedback channel due to an upper-layer protocol. In this case, Rx_2 cannot

know the CQI of Tx_2 - Rx_1 link. Then, our assumption is that the secondary user knows the probability density function (p.d.f.) of CQI about the Tx_2 - Rx_1 link. This assumption is suitable for a scenario such as where the secondary user has the location information about itself and the primary user. The secondary user can access a well-designed and maintained database, which records the p.d.f. of CQI between two locations. Figure 5.1 illustrates an example of location-aided UCR system, where Rx_1 and Rx_2 are fixed network nodes such as base-stations or access points, and Tx_1 and Tx_2 are mobile stations. The database has a look-up table about the p.d.f. of CQI between each fixed network node and a certain area such as the black circle with solid line. Provided the location of Tx_2, Rx_2 knows which circle Tx_2 is currently in, and thus can look up the database to find the p.d.f. of CQI about the Tx_2 - Rx_1 link. How to design and maintain the location-related database is becoming an important research topic. However, it is out of the scope of this section. Further information about location estimation and location-related database can be found in European ICT WHERE [ICT].

As a summary, when we investigate the UCR strategy with partialmultiuser CQI, the following assumptions are made in addition to A1-A3:

- A4): Rx_2 has full knowledge of CQI about the Tx_1 - Rx_1 link, the Tx_1 - Rx_2 link, and the Tx_2 - Rx_2 link, but only knows p.d.f. of CQI about the Tx_2 - Rx_1 link, denoted by $p\left(|a_{21}|^2\right)$, as well as the mean $\mathrm{E}\left(|a_{21}|^2\right)$.
- A5): Rx_2 determines the secondary user's power and transmission rate, and then informs Tx_2 through the feedback channel.

5.3 Main Results

5.3.1 The Individual Decoding (ID) Mode

Figure 5.2-(b) depicts the individual decoding (ID) mode where each receiver only wants to decode the message sent by its corresponding transmitter, and deals with the corresponding interference as noise. This mode is suitable for the following cognitive radio scenarios:

- Both the primary user and secondary user employ their private codebook;
- Even if both users employ a common codebook, each receiver cannot decode the other user's message due to various reasons such as channel conditions and upper layer protocols, etc.

In this situation, the UCR system can be regarded as a simple collection of individual links. This mode has recently received an intensive investigation in both the basic and system-level research, e.g. in [XMH07] and [HL06].

When full knowledge of CQI is avaliable, we have the following result,

Remark 1: A remarkable issue is that ΔC_1 is a monotonically decreasing function of γ_{11} due to the partial derivative $(\partial \Delta C_1)/(\partial \gamma_{11}) < 0$. This means that the primary user operating at a high-SNR scenario is less sensitive to the interference. Considering a high-SNR scenario that fulfills the conditions C1) $\gamma_{11} \geq 1$, and c2) $\gamma_{11} \geq \gamma_{21}$, then ΔC_1 approximates to

$$\Delta C_1 \approx c[\gamma_{21}]. \tag{5.10}$$

Plugging (5.8) into (5.10) leads to

$$P_2 \approx \frac{N_o}{|a_{21}|^2}\left((1+\gamma_{11})^{\rho} - 1\right). \tag{5.11}$$

This simplified result can be utilized to allocate the secondary user's power when the primary user operates in the high-SNR rage.

When partical multiuser CQI is avaliable, as shown in previous, provide the power P_2, the secondary user can employ (5.4) to configure its transmission rate. However, configure P_2 requires the knowledge about $|a_{21}|^2$, which supposes to be unknown in some situations. Next, we propose a new power-allocation criterion based on the assumption A4).

1) Step 1: Outage probability to the primary user is a function of the SNR mean of Tx_2 - Rx_1 link denoted by $\bar{\gamma}_{21} = \left(P_2 E(a_{21})^2\right)/N_o$. Motivated by this fact, the secondary user can first calculate the outage probability for a given $p(|a_{21}|^2)$, and then determine a threshold $\bar{\gamma}_t$ corresponding to O_t.

2) Step 2: The secondary user can access the primary spectrum for the condition $\bar{\gamma}_{21} \le \bar{\gamma}_t$. The maximum of P_2 is therefore given by

$$\max(P_2) = \frac{\bar{\gamma}_t N_o}{E(|a_{21}|^2)} \tag{5.12}$$

The secondary user's transmission rate can be calculated by applying (5.12) into (5.4)

5.3.2 The Secondary-User Side Multiuser Decoding (SSMD) Mode

Figure 5.2-(c) depicts the secondary-user side multiuser decoding (SSMD) mode where each receiver wants to decode the message sent by its corresponding transmitter. The secondary receiver Rx_2 will decode the primary user's message if appropriate. The primary receiver Rx_1 always deals with the interference term $(a_{21}x_2)$ as noise. This mode is suitable for the following cognitive radio scenario:

- The secondary user knows the primary user's codebook, and thus has a chance to decode the primary user's message. This is possible if the primary user is either using a common codebook or broadcasting its own codebook to support, for example, cooperative communications. On the other hand, the primary user may be not aware of the existence of secondary user, or does not know the secondary user's private codebook.

In this situation, the receiver Rx_2 can reliably decode the primary user's message only for the channel condition $|a_{12}| \ge |a_{11}|$, otherwise the SSMD mode reduces to the individual decoding mode presented as following. Therefore, the focus of SSMD mode is on the case $|a_{12}| \ge |a_{11}|$.

When full knowledge of CQI is avaliable, we have the following result,

Remark 2: For the high-SNR conditions C1), C2) and the case $|a_{12}| \ge |a_{11}|$, it can be obtained

$$C_2 \approx \min\left(\log_2\left(\frac{\gamma_{12}+\gamma_{22}}{\gamma_{11}}\right) + c[\gamma_{21}], c[\gamma 22]\right). \tag{5.13}$$

When partial multiuser CQI is avaliable, the UCR strategy for SSMD mode is the same as that for the ID mode, and the transmit-power P_2 is limited by (5.12). The secondary user's transmission rate is restricted by the result product by applying (5.12) into

$$C_2 = \min\left(c[\gamma_{12}+\gamma_{22}] - (1-\rho)c[\gamma_{11}], c[\gamma_{22}]\right). \tag{5.14}$$

5.3.3 The Primary-User Side Multiuser Decoding (PSMD) Mode

This mode is referred to as a scenario where the secondary user does not know the primary user's codebook, but share its own codebook through upper-layer protocols. In this case, the primary user has a chance to decode the secondary user's message, and thus has the potential to cancel the interference

caused by the secondary user. On the other hand, the secondary user has to deal with the interference term $(a_{12}X_1)$ as noise.

When full knowledge of CQI is available, we have the following result,

Theorem 1: Suppose $|a_{21}| \neq 0$ and $\rho = 0$, the secondary user's achievable rate is

$$R_2 \leq c\left[\frac{\gamma_{22}}{\gamma_{12}+1}\right] \tag{5.15}$$

for the channel condition

$$\frac{|a_{21}|^2}{|a_{22}|^2} > \frac{\gamma_{11}+1}{\gamma_{12}+1} := \lambda_1 \tag{5.16}$$

otherwise

$$R_2 \leq c[\gamma_{21}+\gamma_{11}] - c[\gamma_{11}]. \tag{5.17}$$

When partial multiuser CQI is avaliable, we have the following result. Define a threshold of probability denoted by ε, the secondary user will only access the primary spectrum for the following cases:

$$\text{Case 1: } \Pr\left(\frac{|a_{21}|^2}{|a_{22}|^2} > \lambda_1\right) > \varepsilon, \tag{5.18}$$

$$\text{Case 2: } \Pr\left(\frac{|a_{21}|^2}{|a_{22}|^2} < \lambda_2\right) > \varepsilon, \tag{5.19}$$

$$\text{Case 3: } \Pr\left(\frac{|a_{21}|^2}{|a_{22}|^2} \leq \lambda_1\right) > \varepsilon. \tag{5.20}$$

5.3.4 The Two Side Multiuser Decoding (TSMD) Mode

Figure 5.2-(d) depicts the TSMD mode where each user knows the other's codebook. Then, each user has the chance to decode the other user's message so as to cancel the interference.

When full knowledge of CQI is avaliable, we have the following results.

For the channel condition $|a_{12}| \geq |a_{11}|$ and $|a_{21}| \geq |a_{22}|$, the secondary user's capacity is

$$C_2 = \min(c[\gamma_{21}]+c[\gamma_{11}], c[\gamma_{12}+\gamma_{22}]) - c[\gamma_{11}]. \tag{5.21}$$

The transmit power P_2 is limited only by the local power constraint.

For the channel condition $|a_{12}| \geq |a_{11}|$ and $|a_{21}| < |a_{22}|$, if each user deals with the interference as message, the secondary user can access the primary spectrum without causing capacity penalty to the primary user, and the secondary user's transmission rate is (5.21). If the primary user deals with the interference as noise, then the secondary user's transmission rate is (5.14).

When partial multiuser CQI is avaliable, the TSMD mode reduces to the PSMD mode for the channel condition $|a_{12}| < |a_{11}|$. Therefore, the UCR strategy here is proposed only for the condition $|a_{12}| \geq |a_{11}|$. We have the following results.

- Case 1: Suppose $\Pr(|a_{21}|^2 \geq |a_{22}|^2) > \varepsilon$, the secondary user will access the primary spectrum at the transmission rate by replacing the term γ_{21} in (5.21) with $(P_2\varsigma)/N_o$ where $\varsigma > |a_{22}|^2$.

- Case 2: Suppose $\Pr\left(|a_{21}|^2 < |a_{22}|^2\right) > \varepsilon$, using (5.12) to determine $\max(P_2)$ with respect to a given capacity penalty ΔC_1, then calculating $C_2^{(14)}$ replacing P_2 in (5.14) with (5.12), and calculating C_2^{17} by replacing the term γ_{21} in (5.17) with $(P_2\varsigma)/(N_o)(\varsigma < |a_{22}|^2)$. Determine the secondary user's transmission rate via $R_2 \le \max(C_2^{(14)}, C_2^{(17)})$.

5.4 Numerical Resuts

In this subsection, the numerical results will be provided for the investigated UCR modes.

1) The ID mode: Based on the above analytical results, a visual example is used to exhibit the performance. In this example, the UCR system is configured as: $|a_{11}| = 1$, $|a_{22}| = 1$, $|a_{12}| = 1$. The primary user's power-to-noise ratio is $P_1 / N_o = 16$ dB. The secondary user's power-to-noise ratio is also limited by 16 dB. This ratio is one of typical configurations for high-data-rate systems. For the scenario with full multiuser CQI, we set $|a_{21}| = 0.1$. Figure 5.3 illustrates the secondary user's achievable rate (see (5.4)) against the capacity penalty ΔC_1 for cases with full or partial multiuser CQI. It is observed that the secondary user's achievable rate generally increases with the pay of capacity penalty to the primary user. Moreover, in the scenario with partial multiuser CQI, the secondary user shows increased achievable rate for the case of larger outage probability, e.g. $O_t = 0.1$ or smaller $E(|a_{21}|^2)$, e.g. $E(|a_{21}|^2 = 0.005)$.

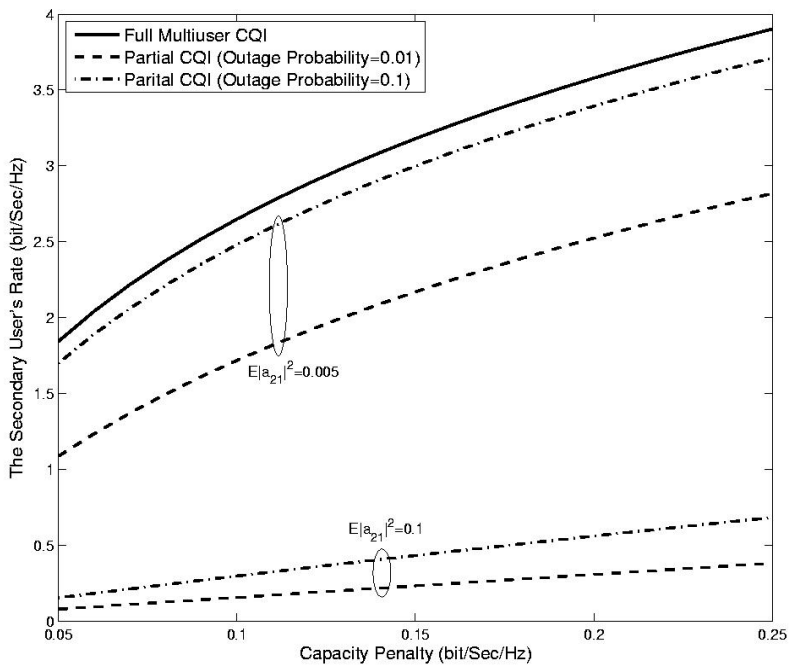

Figure 5.3: An example of capacity results for the ID mode. The secondary user's achievable rate vs. capacity penalty to the primary user.

2) the PSMD mode: Based on the above analytical results, a visual example is used to exhibit the performance. The system configuration is the same setup. For the scenario with full multiuser CQI, Figure 5.4 shows the secondary user's achievable rate as a function of the ratio $|a_{21}|^2 / |a_{22}|^2$. For the scenario with partial multiuser CQI, Figure 5.5 shows the secondary user's achievable rate as a function of the ratio $E(|a_{21}|^2)/|a_{22}|^2$, and outage probability $O_t = O_1 = 10\%$ and probability $\varepsilon = 90\%$. It is observed that Case 1 will happen

only for the condition $E(|a_{21}|^2)/|a_{22}|^2 > 300$, which often does not hold in practice. Case 3 requires the primary user to operate at a SNR larger than 15 dB. However, in this case, the secondary user cannot gain more than 1 bit/sec/Hz at $P_2/N_o = 16$ dB. Finally, Case 2 shows a comparable performance with the corresponding scenario ($|a_{21}|^2/|a_{22}|^2 < \lambda_2$) in Figure 5.4.

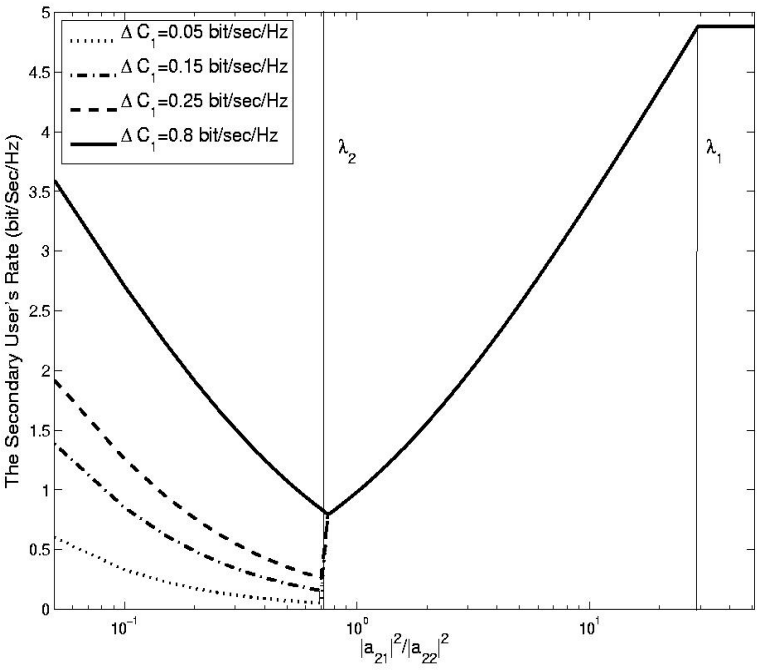

Figure 5.4: An example of capacity results for the PSMD mode with full multiuser CQI.

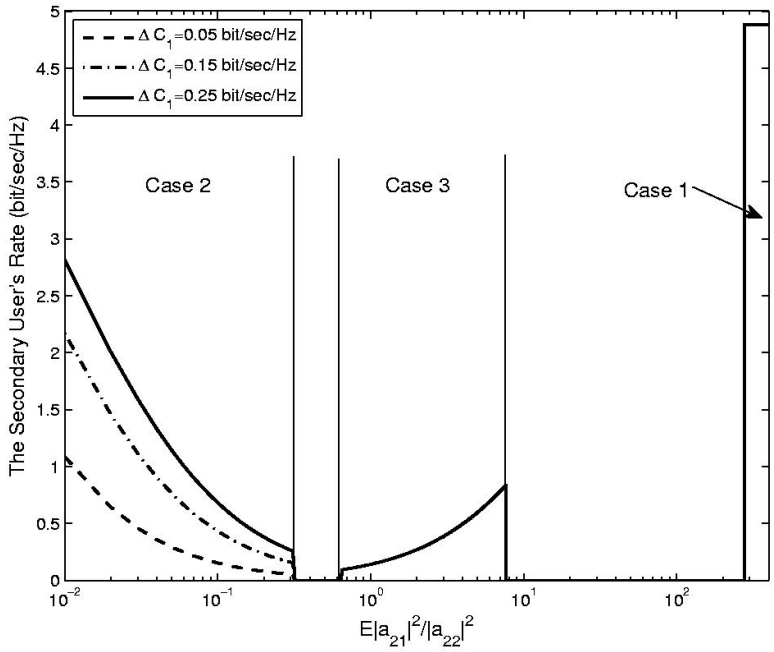

Figure 5.5: An example of capacity results for the PSMD mode with partial multiuser CQI.

3) the TSMD mode: Figure 5.6 and Figure 5.7 show a visual example for scenarios with full or partial multiuser CQI, respectively. The system configuration is almost the same, but we set $|a_{12}|^2 = 4$ to fulfil the condition $|a_{12}| > |a_{11}|$. For the scenario with partial multiuser CQI, we set $O_t = 10\%$ and $\varepsilon = 90\%$ as an example. It is observed that Case1-2 in Figure 5.7 offers comparable performance with the corresponding scenario in Figure 5.6.

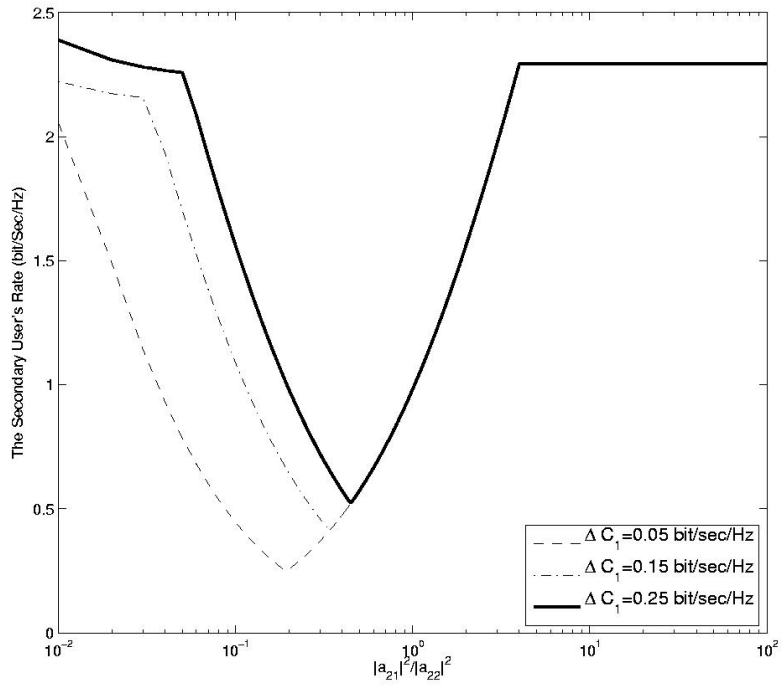

Figure 5.6: An example of capacity results for the TSMD mode with full multiuser CQI.

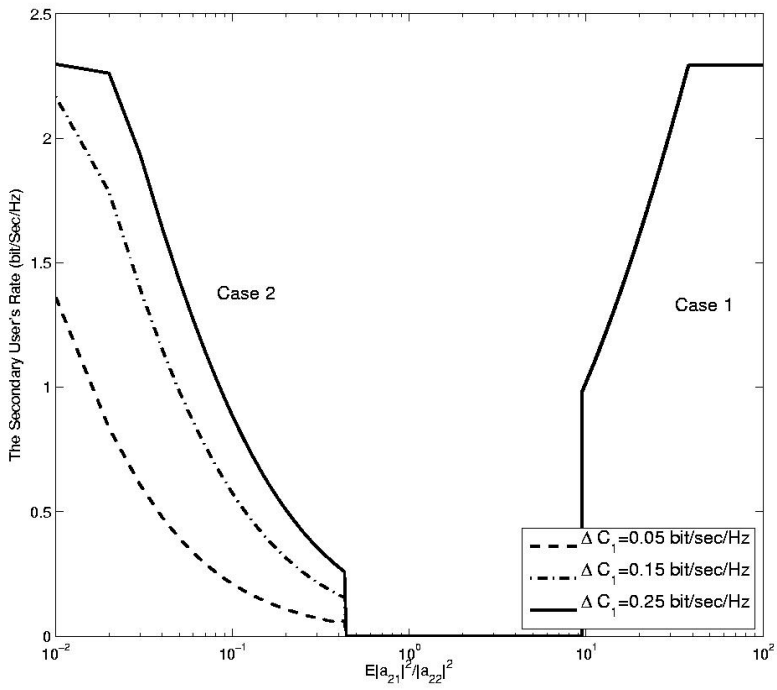

Figure 5.7: An example of capacity results for the TSMD mode with partial multiuser CQI.

5.5 Summary

In this section, we have investigated two-user Gaussian UCR systems by assuming the availability of full multiuser CQI or partial multiuser CQI. Provided full multiuser CQI, we have studied the fundamental relationship between the secondary user's achievable rate C_2 and capacity penalty to the primary user ΔC_1 in four carefully classified UCR modes. For the scenario with partial multiuser CQI, we first established a new physical-layer model through exploitation of the location-aided approach. Then, new spectrum access and power allocation strategies have been investigated in terms of C_2, ΔC_1, and capacity outage probability. Numerical examples are provided to show the performance of the UCR with full multiuser CQI and the proposed approach with partial multiuser CQI. It is observed that the secondary user's achievable rate generally increases with the pay of capacity penalty to the primary user. Moreover, in the scenario with partial multiuser CQI, the secondary user shows increased achievable rate for the case of larger outage propability.

6 Location based SDMA

6.1 Location aided Resource Allocation

Location based information can be exploited to aid Resource Allocation in a variety of ways. Some possibilities are:

• Single user aspects: the environment (profile) and mobility information can be exploited to aid (time,frequency) allocation in OFDMA.

• Multi-user MIMO: Use environment information to preselect users, to limit channel feedback to a reduced set of preselected users. The user preselection can e.g. involve: users with similar RSS, users with rank 1 MIMO channels (close to LOS), ...

• Multicell aspects (interference coordination) or for Cognitive Radio (interference from secondary systems to primary systems): the interference level can be predicted from position based information.

See also the other sections in this deliverable for the implementation of some of these and other ideas. A transversal aspect is also that location tracking can lead to location prediction. This leads in turn to slow fading predictibility (and not just fast fading prediction, which can be done simply from past channel response observations).

In this section, we shall focus on the Spatial Division Multiple Access (SDMA) problem, which in Information Theory is called the Broadcast Channel (BC). The SDMA terminology dates from the eraly nineties. These days it is referred to as the multi-user MISO (or MIMO) communications problem, and we shall particularly focus on the more difficult downlink.

6.2 SDMA considerations

Whereas single user (SU) MIMO communications represented a big breakthrough and are now integrated in a number of wireless communication standards, the next improvement is indeed multi-user MIMO (MU MIMO). This topic is nontrivial as e.g. illustrated by the fact that standardization bodies were not able to get an agreement so far on the topic to get it included in the 3G+ LTE standard. MU MIMO is a further evolution of SDMA, which was THE hot wireless topic in the early nineties. The MU MIMO area has now sufficiently evolved to allow us to understand the following key elements:

• SDMA is a suboptimal approach to MU MIMO, with transmitter precoding limited to linear beamforming, whereas optimal MU MIMO requires Dirty Paper Coding (DPC).

• Channel feedback has gained much more acceptance, leading to good Channel State Information at the Transmitter (CSIT), a crucial enabler for MU MIMO, whereas SDMA was either limited to TDD systems (channel CSIT through reciprocity) or Covariance CSIT. In the early nineties, the only feedback that existed was for slow power control.

• Since SDMA, the concepts of multiuser diversity and user selection have emerged and their impact on the MU MIMO sum rate is now well understood. Furthermore, it is now known that user scheduling allows much simpler precoding schemes to be close to optimal.

• Whereas SU MIMO allows to multiply transmission rate by the spatial multiplexing factor, when mobile terminals have multiple antennas, MU MIMO allows to reach this same gain with single antenna terminals.

• Whereas in SU MIMO, various degrees of CSIT only lead to a variation in coding gain (the constant term in the sum rate), in MU MIMO however CSIT affects the spatial multiplexing factor (multiplying the log(SNR) term in the sum rate).

In the process attempting to integrate MU-MIMO into the LTE-A standard, a number of LTE-A contributors had recently become extremely sceptical about the usefulness of the available MU-MIMO proposals. The isue is that they currently do MU-MIMO in the same spirit as SU-MIMO, i.e. with feedback of CSI limited to just a few bits! However, MU-MIMO requires very good CSIT! Some possible solutions are:

• Increase CSI feedback enormously (possibly using analog transmission).

• Exploit channel reciprocity in TDD (electronics calibration issue though).

• Limit MU-MIMO to NADA (see below) users and extract essential CSIT from position information.

Narrow AoD Aperture (NADA) case

The idea here is to focus on the category of mobiles for which the angular spread seen from the BS is limited. This is a small generalization of the LOS case. In the NADA case, the MIMO channel $\mathbf{H}$ (assumed frequency-flat here or we assume a narrowband case (e.g. an OFDM subcarrier)) is of the form

$$\mathbf{H} = \sum_i \mathbf{h}_r(\theta_i)\mathbf{h}_t^T(\phi_i) = \mathbf{B}\,\mathbf{A}^T\ ,\ \ \mathbf{A} = \left[\mathbf{h}_t(\phi)\ \dot{\mathbf{h}}_t(\phi)\right] \tag{6.1}$$

where $\mathbf{h}_r(.)$ is the receive side antenna array response, $\mathbf{h}_t(.)$ is the transmit side antenna array response, θ_i is the Angle of Arrival (AoA) of path i and ϕ_i is the Angle of Departure (AoD) of path i. In the case of narrow AoD spread, we have

$$\phi_i = \phi + \Delta\phi_i \tag{6.2}$$

where ϕ is the nominal AoD and $\Delta\phi_i$ is small. Hence

$$\mathbf{h}_t(\phi_i) \approx \mathbf{h}_t(\phi) + \Delta\phi_i\ \dot{\mathbf{h}}_t(\phi)\,. \tag{6.3}$$

This leads to the second equality in (1). Hence $\mathbf{H}$ is of rank 2. The LOS case is a limiting cas in which the power of the $\dot{\mathbf{h}}_t(\phi)$ term becomes negligible and the channel rank becomes 1. The factor $\mathbf{A}$ in $\mathbf{H}$ depends straightforwardly on position (which translates into LOS AoD), only $\mathbf{B}$ remains random. Some possible (indirect) measures of Angular spread (aperture) could be

$$\frac{\lambda_2(\mathbf{H}^H\mathbf{H})}{\lambda_2(\mathbf{H}^H\mathbf{H})}\quad or\quad \frac{\dot{\mathbf{h}}_t^T(\phi)\mathbf{H}^H\mathbf{H}\dot{\mathbf{h}}_t^*(\phi)}{\mathbf{h}_t^T(\phi)\mathbf{H}^H\mathbf{H}\mathbf{h}_t^*(\phi)}\,. \tag{6.4}$$

where $\lambda_i(\mathbf{H}^H\mathbf{H})$ are the eigenvalues of $\mathbf{H}^H\mathbf{H}$ in decreasing order. Either measure becomes 0 in the case of LOS.

In what follows, we shall focus on the LOS limit for considerations of location based processing. We propose that location based MU MIMO transmission involves position based user selection (attenuation, nominal AoD, AoD spread) and associated beamforming (BF) and power control (PC).

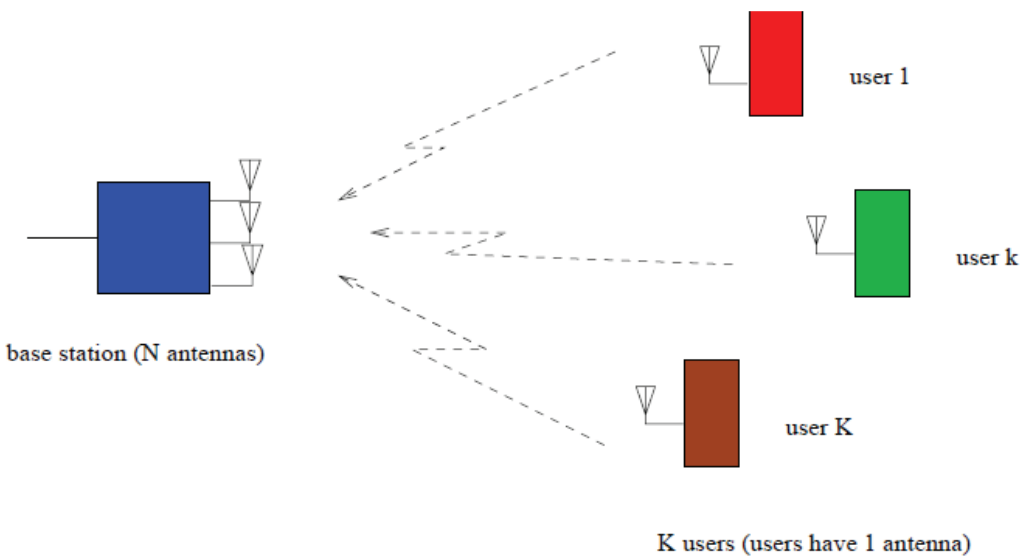

Figure 6.1 SDMA set-up (Broadast channel (BC), multi-user (MU) MISO downlink).

6.3 SDMA via user feedback: ZF BF sum rate

6.3.1 Background and Motivation

In a broadcast channel (BC) having a base station (BS) equipped with n_t transmit antennas and K ($K \geq n_t$) single antenna users, the dominant term of the sum capacity is $n_t \log(\mathrm{SNR})$ [CS03], [JG05], [Jin06], where SNR denotes the signal-to-noise ratio of the received signal and n_t is called the multiplexing gain. The BC enjoys another gain, coined as multi-user diversity [KH95], which is due to the possibility of user selection from a larger ($K > n_t$) pool of users. It has been shown in [SH05], [SH07] that the sum capacity of the Gaussian broadcast channel scales with the number of users as $n_t \log(\log(K))$, where K is the total number of users in the system whose channel information is available at the BS. Apart from the multiplexing gain and the multi-user diversity benefit, the BC enjoys two other advantages over the single user multiple-input multiple-output (SU MIMO) channel. It allows mobile users to have a single antenna each so user terminals can remain quite inexpensive and simple. The second advantage is that the BC channel matrix is often much better conditioned as compared to that of a SU MIMO link which may suffer from line-of-sight channel conditions and strong spatial correlation [GKH07]. These same advantages typically continue to hold compared to multi-user (MU) MIMO systems in which the total number of receive antennas equals or exceeds n_t.

These promising advantages of broadcast MIMO do not come for free as without channel state information at the transmitter (CSIT) and perfect channel state information at the receiver (CSIR), the dominant term of the sum capacity is only $\log(\mathrm{SNR})$ because of the optimality of transmitting to a single user in that case [CT06], [Cov72], [TV05]. Thus the CSIT of n_t users is indispensable to achieve the full multiplexing gain [Jin06], [CS03]. Furthermore, to capture the multi-user diversity benefit of $n_t \log(\log(K))$ in the sum rate, the BS should know the CSI of all these K users where normally K could be much larger than n_t.

With perfect CSIT, [YG06] shows that zero-forcing (ZF) precoding achieves the full multiplexing gain of n_t and the full multi-user diversity gain of $n_t \log(\log(K))$ of the broadcast channel if the number of users is asymptotically large, although the optimal transmission strategy for the Gaussian BC is dirty paper coding (DPC) [WSS06]. In [SH05], the authors introduced a very innovative scheme coined as Orthogonal Random Beam Forming (ORBF) where only a few bits of feedback are required from every user and the sum rate was shown to converge to the optimal DPC sum capacity [SH07] for a large number of users.

There is an enormous volume of research publications analyzing CSIT acquisition techniques and the associated feedback gains in different scenarios but the fundamental issue, which is usually ignored, is the feedback overhead cost of providing CSI to the BS which leads to reduced sum capacity. Both the gain and the acquisition overhead increase with the amount of feedback but there is an optimal operating point (optimal amount of feedback) that maximizes the net gain. To the best of our knowledge, there is no single contribution that has properly analyzed the **net gain of feedback** in a general broadcast setting (with $K > n_t$) which can be defined as the gain in downlink (DL) sum rate due to feedback, taking into account the uplink (UL) feedback load. For a given feedback load, the problem boils down to a secondary problem, namely partitioning of this feedback either to improve the CSIT quality (better interference cancellation for a given user selection) or to get the multi-user diversity gain (select from a larger pool). We take a step back and address the more fundamental problem of how much feedback should be there to achieve the feedback gain-cost trade-off. A very simple example showing the importance of this absolute gain would be the ORBF transmission scheme which requires as little feedback as $\log(n_t)$ bits plus a scalar from each user in the system, but considering the fact that ORBF requires the presence and the feedback from a large number of users, the absolute gain would become questionable.

The second fundamental aspect which often gets overlooked in the analysis of multi-user systems is the consideration of channel coherence time. Wireless channels have a finite coherence time and when multi-user transmission strategies with multiple rounds of training, feedback and data are devised, there is the possibility that the channel has significantly changed during the preliminary training and feedback intervals and that the channel information attained during these phases has become meaningless.

Contribution

In this work, we do not make any assumption of CSI. Hence initially, the BS and the users are ignorant of the channel realization but they can estimate/feedback the CSI as is done in practice. To analyze the cost incurred and the attainable benefit of feedback in a meaningful and tractable fashion, we simplify the problem by selecting a time-division duplex (TDD) broadcast channel with perfect reciprocity. TDD reciprocal channels simplify the CSIT acquisition through UL pilot transmission [Mar06], [MH06] in contrast to frequency-division duplex (FDD) systems in which the users first estimate the DL channel and then send its quantized version in an UL slot. We restrict the CSIT acquisition to be training based only, thanks to TDD reciprocity [MH06], [Mar06]. In the sequel, we use the terms training and feedback synonymously due to this restriction [MH06]. So we have a fixed resource (bandwidth and time slot) available, a BS having n_t transmit antennas and K ($K > n_t$) single antenna users. Now this fixed resource can be used for UL/DL data transmission or training/feedback. We assume that the users have no data to transmit in the UL direction. So the UL is solely reserved for channel training/feedback. But due to the TDD mode of operation, any UL transmission will come at the expense of having reduced DL transmission in the overall time slot, hence the cost of training/feedback gets properly accounted for. The users, who feed back, are chosen independently of their channel realizations (whence **oblivious users**). We derive a novel lower bound (and at the same time good approximation) for the sum rate obatined with ZF BF, capturing the gains and the costs of the CSIT acquisition, which shows explicitly the rate loss with respect to (w.r.t.) a perfect CSI system. The simplified expressions obtained can be used to maximize the DL sum rate (the performance metric considered here) achieving the cost-benefit trade-off of the feedback.

Notation: E denotes statistical expectation. Lowercase letters represent scalars, boldface lowercase letters represent vectors, and boldface uppercase letters denote matrices. $\mathbf{A}^H$, $\mathbf{A}^T$, $\mathbf{A}^*$ denote respectively the Hermitian transpose, transpose, and complex conjugate of matrix $\mathbf{A}$. The identity matrix of n_t dimensions is denoted by $\mathbf{I}_{n_t}$. The logarithm with base e is denoted by $\log(.)$. The cardinality of a set S is expressed as $|\mathsf{S}|$.

6.3.2 System Model

The frequency-flat system we consider consists of a BS having n_t transmit antennas and K ($K > n_t$) single-antenna user terminals. In the DL, the signal received by k-th user can be expressed as

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k, \quad k = 1,2,\ldots,K \tag{6.5}$$

where $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_K$ are the (complex conjugated) channel vectors of users 1 through K, with $\mathbf{h}_k \in \mathsf{C}^{n_t \times 1}$ ($\mathsf{C}^{n_t \times 1}$ denotes the n_t-dimensional complex space), $\mathbf{x} \in \mathsf{C}^{n_t \times 1}$ denotes the n_t-dimensional signal transmitted by the BS and $n_1, n_2, \ldots, n_K$ are independent complex Gaussian additive noise terms with zero mean and unit variances. We denote the concatenation of the channels by $\mathbf{H}_F^H = [\mathbf{h}_1 \mathbf{h}_2 \cdots \mathbf{h}_K]$, so $\mathbf{H}_F$ is the $K \times n_t$ forward channel matrix. The channel input from the BS must satisfy an (average) transmit power constraint of P, i.e. $\mathsf{E}[\|\mathbf{x}\|^2] \le P$. In this setting, the transmit power P is equal to the true signal-to-noise ratio at each user due to normalized noise variances.

The channel is assumed to follow a block fading model having a coherence length of T symbol intervals without channel variation, with independent fading from one block to the next [MH99]. The entries of the forward channel matrix $\mathbf{H}_F$ are independent and identically distributed (i.i.d.) complex Gaussian with zero mean and unit variance. Due to the no CSI assumption, initially all the users and the BS are oblivious of the channel realizations in each block.

For the power constraint at the user terminals, we mainly treat the case of peak power constraint. The peak power per user per channel use is bounded by P_{pk}. The noise in the UL at the BS is also assumed to be spatiotemporally white complex Gaussian with unit variance.

6.3.3 Transmission Scheme with Oblivious Users

In this scheme, the users who feed back are unaware of their channel state. So they might be selected in a round-robin or any other fashion, independently of their channel realizations. For the block fading channel with coherence length of T symbol intervals, we divide this interval in three phases (see Figure 6.2): 1) uplink training, 2) downlink training and 3) coherent data transmission with imperfect CSI. The first phase is the uplink training phase in which a certain number of users train the BS about their forward channels and the BS makes an estimate of the associated forward channel matrix. Based upon this channel information, the BS does the scheduling and chooses the transmit precoding which in general could be simple linear ZF, some non-linear strategy like vector perturbation or the optimal DPC. The second phase is the downlink training phase where the BS transmits pilots so that the scheduled users estimate their corresponding effective channels. When this second phase ends, both sides of the broadcast channel have necessary CSI, albeit imperfect. Hence in the third data phase, the BS transmits simultaneously to the selected users who can decode the data coherently.

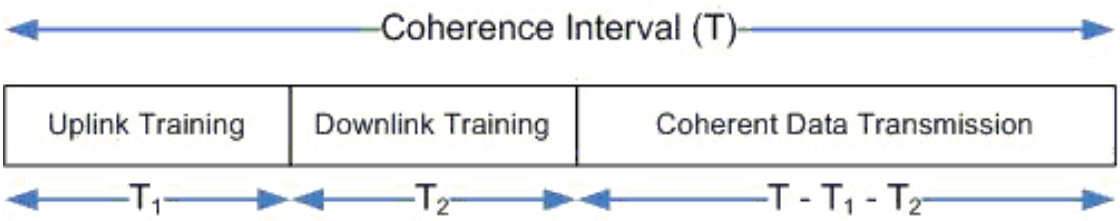

Figure 6.2 Transmission Phases for Oblivious Users.

Remark : When transmission is switched from UL to DL or vice versa, a guard interval must be inserted in practice. We do not take this guard interval into account as it does not affect our absolute feedback gain analysis.
Below we provide a detailed analysis of the three transmission phases and the necessary BS processing steps.

6.3.4 Uplink Training Phase

In a TDD system with perfect channel reciprocity, CSIT can be provided to the BS just by transmitting pilots from the users. The BS estimates the users' uplink channels and these are then also the forward channels due to the perfect reciprocity assumption. Suppose K users transmit pilots, then the length of this uplink training interval is $T_1 = \beta K$ where $\beta \geq 1$ *should assume a real value so that* T_1 *is an integer* ($\beta > 1$ can be used when we want a given number of users to transmit for more time and improve their channel estimates at the BS, $\beta < 1$ leads to significantly degraded channel estimates). Assuming orthogonal codes (which is optimal here) of length T_1, the users can transmit simultaneously with transmit energy per user equal to $P_{pk}T_1$. As each antenna at the BS receives the transmitted code from a particular user through the channel coefficient which links this antenna to that user, the energy received for each channel coefficient (of CSIT) would be $P_{pk}T_1$. Assuming the BS employs MMSE channel estimation, the resulting channel coefficient estimation errors are i.i.d. complex Gaussian with zero mean and mean-square error (MSE)

$$\sigma_h^2 = \frac{1}{P_{pk}T_1 + 1} = \frac{1}{P_{pk}\beta K + 1}. \tag{6.6}$$

For the k-th user, with channel $\mathbf{h}_k$, the channel estimate is denoted as $\hat{\mathbf{h}}_k$ and the corresponding estimation error is $\tilde{\mathbf{h}}_k$ according to $\mathbf{h}_k = \hat{\mathbf{h}}_k + \tilde{\mathbf{h}}_k$. The $\hat{\mathbf{h}}_k$ also have i.i.d. complex Gaussian entries with zero mean and variance $1 - \sigma_h^2$ (due to the orthogonality property of MMSE estimation). Note that due to training code orthogonality, the channel estimation quality improves with the number of users K.

The training length $T_1 \geq K$ is basically the price of obtaining CSIT at the BS through feedback which reduces the effective channel coherence time to $T - T_1$. Hence, the CSIT acquisition from a very large number of users may be very sub-optimal.

6.3.5 BS Transmission Strategy: ZF Precoding with Semi-Orthogonal User Selection (SUS)

It is known that DPC allows to achieve the full capacity region of the MIMO broadcast channel [WSS06] but this scheme is complex and its implementation is quite tedious. ZF linear precoding with user selection has been shown to behave quite optimally at high SNR achieving full multiplexing gain in the sum rate [Jin05]. Furthermore in [YG06], the authors showed that ZF preceded by semi-orthogonal user selection (SUS) achieves both the multiplexing gain and the multi-user diversity gain. SUS has been modified in [SJL08] to work with imperfect CSIT in a robust manner. Due to its simplicity, analytical tractability and attractive performance, we choose SUS and ZF precoding as the BS transmission strategy.

We adopt the SUS algorithm of [YG06] where user orthogonality is imposed at each selection stage. Suppose S denotes the set of selected users having cardinality $|\mathsf{S}| = n_t$ and $\hat{\mathbf{H}}(\mathsf{S})$ denotes the BS estimate of the channel matrix of the selected users. In ZF precoding, the unit-norm beamforming vector for the k-th selected user (denoted as $\overline{\mathbf{v}}_k$) is chosen to be orthogonal to the channel vectors of all other selected users, i.e., $\hat{\mathbf{h}}_j^H \overline{\mathbf{v}}_k = 0,\ j \in \mathsf{S} \setminus \{k\}$. If $\mathbf{W}(\mathsf{S})$ is the pseudo-inverse of $\hat{\mathbf{H}}(\mathsf{S})$, i.e.

$$\mathbf{W}(\mathsf{S}) = \hat{\mathbf{H}}(\mathsf{S})^H \left[\hat{\mathbf{H}}(\mathsf{S})\hat{\mathbf{H}}(\mathsf{S})^H\right]^{-1}, \tag{6.7}$$

then the precoding matrix $\overline{\mathbf{V}} = [\overline{\mathbf{v}}_1 \overline{\mathbf{v}}_2 \cdots \overline{\mathbf{v}}_{|\mathsf{S}|}]$ can be obtained from $\mathbf{W}(\mathsf{S})$ by normalizing all of its columns. For ZF with perfect CSIT, each user receives only the beam directed to it and no multi-user interference is experienced. For the imperfect CSIT case, there is some residual interference. If $\mathbf{u}$ represents the vector of information symbols (u_k intended for the k-th user), the transmitted signal $\mathbf{x}$ becomes $\mathbf{x} = \overline{\mathbf{V}}\mathbf{u}$ and the signal received by the k-th selected user (6.5) can be expressed as follows:

$$\begin{aligned} y_k &= \mathbf{h}_k^H \overline{\mathbf{V}}\mathbf{u} + n_k \\ &= \mathbf{h}_k^H \overline{\mathbf{v}}_k u_k + \sum_{j \neq k} \mathbf{h}_k^H \overline{\mathbf{v}}_j u_j + n_k. \end{aligned} \tag{6.8}$$

6.3.6 Downlink Training Phase

It was remarked in [SS08] that only one symbol interval is sufficient to let the $|\mathsf{S}|$ selected users learn their effective scalar channels $\mathbf{h}_k^H \overline{\mathbf{v}}_k$. In a very recent reference [JLM09], the authors show that this minimal training becomes optimal with joint pilot and data processing. As this DL training length has no relation with the number of users K present in the system or the number of BS antennas (n_t), we assume that the selected users are able to estimate their effective scalar channels perfectly even though we ignore the overhead of this phase. This simplifies the analysis without influencing the underlying cost-benefit trade-off of the feedback.

6.3.7 Coherent Data Phase

We adopt uniform power allocation. So the k-th user input signal u_k is i.i.d. Gaussian, $u_k : \mathsf{CN}(0, p)$, where p is the power allocated to k-th user data stream. The BS is bound to satisfy an average power constraint of P but it does not transmit during the entire coherence block due to the initial UL training phase of length T_1. Hence, for the rest of the coherence block, the BS is able to transmit an average per symbol power of $P\,T/(T - T_1)$ instead of P. So the power p allocated to each of the $|\mathsf{S}|$ selected users becomes

$$p = \frac{P}{|\mathsf{S}|}\frac{T}{T-T_1}. \tag{6.9}$$

6.3.8 Sum Rate Lower Bound

We are interested in getting an expression for the achievable sum rate of this broadcast channel which captures the gain and the cost associated with feedback. The received signal from (8) can be further written as

$$y_k = \hat{\mathbf{h}}_k^H \overline{\mathbf{v}}_k u_k + \tilde{\mathbf{h}}_k^H \overline{\mathbf{v}}_k u_k + \sum_{j \neq k} \tilde{\mathbf{h}}_k^H \overline{\mathbf{v}}_j u_j + n_k. \tag{6.10}$$

This is obtained by exploiting the fact that $\mathbf{h}_k^H \overline{\mathbf{v}}_j = \tilde{\mathbf{h}}_k^H \overline{\mathbf{v}}_j$ for $k \neq j$ due to ZF beamforming and by splitting the effective channel $\mathbf{h}_k^H \overline{\mathbf{v}}_k$ in two parts, where $\hat{\mathbf{h}}_k^H \overline{\mathbf{v}}_k$ is perfectly known at the BS. The above equation can be rewritten as

$$y_k = \hat{\mathbf{h}}_k^H \overline{\mathbf{v}}_k u_k + \sum_{j \in \mathsf{S}} \tilde{\mathbf{h}}_k^H \overline{\mathbf{v}}_j u_j + n_k. \tag{6.11}$$

From this equation, where we have relegated the signal part $\tilde{\mathbf{h}}_k^H \overline{\mathbf{v}}_k u_k$ into the interference and by treating all the interference terms as additional independent Gaussian noise as in [Med00] and [HH03], the SINR of the k-th user for particular channel estimates can be written as

$$\mathrm{SINR}_k = \frac{p\,|\hat{\mathbf{h}}_\mathbf{k}^\mathbf{H} \overline{\mathbf{v}}_\mathbf{k}|^2}{1 + p\sum_{j \in \mathsf{S}} \mathsf{E}_{\tilde{\mathbf{h}}_\mathbf{k}} |\tilde{\mathbf{h}}_\mathbf{k}^\mathbf{H} \overline{\mathbf{v}}_\mathbf{j}|^2}. \tag{6.12}$$

The variance of each interference coefficient $\tilde{\mathbf{h}}_k^H \overline{\mathbf{v}}_j$ can be computed based upon the fact that the BS performs MMSE estimation which makes the estimation error $\tilde{\mathbf{h}}_k$ (with variance σ_h^2 per channel entry) independent of any function of the channel estimates $\hat{\mathbf{h}}_k$ [Kay93], of which beamforming vectors are one particular instance:

$$\mathsf{E}_{\tilde{\mathbf{h}}_k} |\tilde{\mathbf{h}}_k^H \overline{\mathbf{v}}_j|^2 = \overline{\mathbf{v}}_j^H \mathsf{E}_{\tilde{\mathbf{h}}_k} \left(\tilde{\mathbf{h}}_k \tilde{\mathbf{h}}_k^H\right) \overline{\mathbf{v}}_j = \overline{\mathbf{v}}_j^H \left(\sigma_h^2 \mathbf{I}_{\mathbf{n_t}}\right) \overline{\mathbf{v}}_j = \sigma_h^2. \tag{6.13}$$

Furthermore, by introducing $\hat{\mathbf{h}}_k = \sqrt{1-\sigma_h^2}\,\mathbf{g}_k$ where $\mathbf{g}_k : \mathsf{CN}(\mathbf{0}, \mathbf{I}_{\mathbf{n_t}})$, the SINR becomes

$$\mathrm{SINR}_k = \frac{1-\sigma_h^2}{1 + p\,|\mathsf{S}|\,\sigma_h^2}\, p\,|\mathbf{g}_k^H \overline{\mathbf{v}}_k|^2. \tag{6.14}$$

For a system with perfect CSIT ($\sigma_h = 0$), $p\,|\mathbf{g}_k^H \overline{\mathbf{v}}_k|^2 = p\,|\mathbf{h}_k^H \overline{\mathbf{v}}_k|^2$ would be the SINR and hence its coefficient in the above expression represents the SINR loss factor w.r.t. a system with perfect CSIT. So during the data phase, the lower bound (LB) for the per symbol sum rate (assuming Gaussian inputs) can be written as

$$\mathrm{LB} = \sum_{k \in \mathsf{S}} \mathsf{E}_{\mathbf{g}_k} \log\left(1 + \frac{1-\sigma_h^2}{1 + p\,|\mathsf{S}|\,\sigma_h^2}\, p\,|\mathbf{g}_k^H \overline{\mathbf{v}}_k|^2\right). \tag{6.15}$$

Putting the value of p, we get

$$\mathrm{LB} = \sum_{k \in \mathsf{S}} \mathsf{E}_{\mathbf{g}_k} \log\left(1 + \frac{1-\sigma_h^2}{1 + P\frac{T}{T-T_1}\sigma_h^2}\frac{T}{T-T_1}\frac{P}{|\mathsf{S}|}|\mathbf{g}_k^H \overline{\mathbf{v}}_k|^2\right). \tag{6.16}$$

If the same system (K users and n_t BS antennas) had perfect CSI ($\sigma_h^2 = 0, T_1 = 0$), the sum rate obtained through SUS and ZF beamforming with equal power allocation would be [Jin06], [YG06]

$$R_{\mathrm{ZF}}(K,n_t,P)=\sum_{k\in\mathsf{S}}\mathsf{E}_{\mathbf{g}_k}\log\left(1+\frac{P}{|\mathsf{S}|}|\mathbf{g}_k^H\overline{\mathbf{v}}_k|^2\right). \tag{6.17}$$

So the lower bound of the sum rate from (6.16) can be written in terms of the sum rate of a perfect CSI system as

$$\mathrm{LB}=R_{\mathrm{ZF}}(K,n_t,\alpha P), \tag{6.18}$$

where α, the transmission power scaling due to imperfect CSIT, is given by

$$\alpha=\frac{(1-\sigma_h^2)\dfrac{T}{T-T_1}}{1+P\dfrac{T}{T-T_1}\sigma_h^2}. \tag{6.19}$$

An important subtlety however is that in (6.17), the expectation is over the $\mathbf{g}_k=\mathbf{h}_k$, the true channels, whereas in (6.16) the $\mathbf{g}_k$ are a per coefficient variance normalized version of the $\hat{\mathbf{h}}_k$, the channel estimates. With the assumptions taken here, of i.i.d. channel coefficients and noise elements, this does not make any difference here. However, this issue could potentially make a big difference in the case of spatially correlated channels and/or receiver noise at the BS. In that case also, (6.16) would allow for a more straightforward analysis of SUS, which is also based on the $\hat{\mathbf{h}}_k$!

By taking now into account the loss of coherence interval T due to feedback (training) interval of length $T_1=\beta K$, the per symbol sum rate lower bound for this oblivious scheme becomes

$$\mathrm{LB}=\frac{T-\beta K}{T}\sum_{k\in\mathsf{S}}\mathsf{E}_{\mathbf{g}_k}\log\left(1+\frac{(1-\sigma_h^2)\dfrac{T}{T-\beta K}}{1+P\sigma_h^2\dfrac{T}{T-\beta K}}\frac{P}{|\mathsf{S}|}|\mathbf{g}_k^H\overline{\mathbf{v}}_k|^2\right) \tag{6.20}$$

or hence

$$\mathrm{LB}=(1-\frac{\beta K}{T})R_{\mathrm{ZF}}(K,n_t,\alpha P). \tag{6.21}$$

The biggest virtue of this lower bound is that it gives the achievable sum rate of the scheme considered in terms of the sum rate $R_{\mathrm{ZF}}(K,n_t,P)$ of a perfect CSI system (employing SUS and ZF precoding) with loss appearing as an SNR reduction factor and as reduced multiplexing gain due to the feedback interval.

It now can easily be shown [SS10] that the optimal value for β is $\beta=1$. In [SS10] one can also find an evaluation of the tightness of the lower bound, as also results on the optimization of the sum rate over K. Larger K leads to larger multi-user diversity, but also larger rate loss due to training. Hence there exists an optimal finite K.

6.4 Location based SDMA: ZF BF sum rate

Although some extension to the more general NADA case could priobably be considered, we shall focus here on the LOS case. So a first restriction in the SDMA user selection process is that for MU-MISO purposes, users to be considered need to be in LOS mode. So in this case we get for the downlink channel to user k:

$$\mathbf{h}_k^H=\gamma_k\,e^{j\psi_k}\,\mathbf{h}_t^T(\phi_k) \tag{6.22}$$

where $\mathbf{h}_t(.)$ is the (unit norm) BS antenna array response, ϕ_k is the AoD for user k, which in the LOS case can be computed from the user's position, $\gamma_k>0$ is a complex attenuation factor, and ψ_k is a phase that is unimportant for transmission purposes. There are a variety of ways in which the information of γ_k can be obtained:

- User feedback of just the scalar γ_k.

• Infer γ_k from the uplink. Not only in TDD but even in FDD, in the case of a LOS channel, the channel gain should be reciprocal (because there is no frequency-dependent superposition of multipath contributions).

• Determine the attenuation from the position and simple (e.g. free space (LOS!)) propagation laws.

6.4.1 Effect of LOS deviation on ZF BF sum rate

In this case we can model the user's channels as

$$\mathbf{h}_k^H = \gamma_k \, e^{j\psi_k} \, \mathbf{h}_t^T(\phi_k) + \tilde{\mathbf{h}}_k^H \tag{6.23}$$

where $\tilde{\mathbf{h}}_k$ could in a first instance be modeled as random with i.i.d. zero mean components with variance σ_h^2. The ratio $\frac{\gamma_k^2}{n_t \sigma_h^2}$ could be considered as a Ricean factor. The reasoning leading to the sum rate LB (6.20) can be adapted to yield the following sum rate LB for location based ZF BF (for uniform transmit power)

$$\mathrm{LB}^{rice} = \sum_{k \in \mathsf{S}} \mathsf{E} \log\left(1 + \frac{1}{1+P\sigma_h^2} \frac{P}{|\mathsf{S}|} \gamma_k^2 |\mathbf{h}_t^T(\phi_k)\bar{\mathbf{v}}_k|^2\right) \tag{6.24}$$

where the expectation is over the distribution of the ϕ_k and the γ_k. Hence

$$\mathrm{LB}^{rice} = R_{\mathrm{ZF}}^{los}(K, n_t, \frac{1}{1+P\sigma_h^2}P) \tag{6.25}$$

where the perfect LOS case would be obtained by putting $\sigma_h^2 = 0$. In contrast to the training based approach, here the performance increases with the number K of users to choose from as then they can be better chosen to have close to orthogonal antenna array responses (note that K should grow with SNR if sum rate saturation at high SNR is to be avoided). Another contrast to the training based approach in which σ_h^2 is due to channel estimation error, in which case σ_h^2 in (6.6) decreases with the UL SNR, here σ_h^2, which is now due to LOS approximation error, is independent of SNR. The result of this is that at high SNR the sum rate will saturate and the spatial multiplexing factor will be lost. This only happens though at SNR above which the interference resulting from channel approximation error dominates the noise, i.e. when $P > \frac{1}{\sigma_h^2}$.

One remark is in order here about antenna spacing. For the purpose of DoA estimation, and considering a uniform linear array (ULA) of antennas, it is generally considered that an antenna spacing of $\lambda/2$ is good. However, for the purpose of SDMA, in which we would like the antenna array responses between different angles to be easily orthogonal, it is preferable that the antenna spacing is larger. Indeed, the larger the antenna spacing, the larger the number of angles within a sector for which the array response is orthogonal to the array response at a given angle in the same sector. This multiplicity of "orthogonal" angles on the other hand leads to ambiguities in the DoA estimation problem. In the case where the DoA is not estimated from received signal data but is computed on the basis of the position, these ambiguity problems are irrelevant and then antenna spacing should indeed be as large as possible (although not too large to invalidate the far field and narrowband assumptions).

6.4.2 Effect of position error on ZF BF sum rate

Assume a (2D) position error Δp_k for user k, with mean square value $\sigma_p^2 = \mathsf{E}\,||\Delta p_k||^2$ (assuming also the position error to be isotropic). The position error will lead to a AoD error

$$\Delta\phi_k = \frac{\Delta p_k}{d_k\sqrt{2}} \tag{6.26}$$

where d_k is the distance of user k from the BS, and $\sqrt{2}$ is not an exact representation but leads to the correct AoD error variance, accounting for the fact that AoD error only depends on the component of

Δp_k orthogonal to the LOS direction. The AoD error will lead to an error in the steering vector, which for small AoD error we can approximate by a first order Taylor series expansion (similar to the NADA case)

$$\mathbf{h}_t(\phi_k + \Delta\phi_k) \approx \mathbf{h}_t(\phi_k) + \Delta\phi_k \, \dot{\mathbf{h}}_t(\phi_k) \,. \tag{6.27}$$

Paralleling the reasoning in the previous cases, we can obtain a ZF BF sum rate LB

$$\mathrm{LB}^{loc} = \sum_{k\in\mathsf{S}} \mathsf{E} \log\left(1 + \frac{\frac{P}{|\mathsf{S}|}\gamma_k |\mathbf{h}_t^T(\phi_k)\bar{\mathbf{v}}_k|^2}{1 + \frac{P}{|\mathsf{S}|}\gamma_k \frac{\sigma_p^2}{2d_k^2}\sum_{j\in\mathsf{S}} |\dot{\mathbf{h}}_t^T(\phi_k)\bar{\mathbf{v}}_j|^2} \right). \tag{6.28}$$

The effect of the position error is hence to reduce the SNR for user k by a factor

$$1 + \frac{P}{|\mathsf{S}|}\gamma_k \frac{\sigma_p^2}{2d_k^2}\sum_{j\in\mathsf{S}} |\dot{\mathbf{h}}_t^T(\phi_k)\bar{\mathbf{v}}_j|^2 \quad . \tag{6.29}$$

6.5 Comparative Simulations of location based SDMA vs LTE quantization-feedback based SDMA

6.5.1 The Location Based LOS-Channel Model

Assume MIMO at first and both transmitters and receivers with ULAs. The location based LOS-channel model for one user is given by [LFSS06],

$$\mathbf{H} = \mathbf{H}_L + \mathbf{H}_N = \alpha_0 e^{j\varphi_0}\mathbf{h}_r(\theta_0)\mathbf{h}_t^T(\phi_0) + \sum_{i=1}^{B}\alpha_i e^{j\varphi_i}\mathbf{h}_r(\theta_i)\mathbf{h}_t^T(\phi_i), \tag{6.30}$$

where the receive antenna steering vectors $\mathbf{h}_r(\theta_i)$ and the transmit antenna steering vectors $\mathbf{h}_t(\phi_i)$ are

$$\mathbf{h}_r(\theta_i) = \frac{1}{\sqrt{n_r}}\begin{bmatrix} 1 & e^{-j2\pi\frac{d_r \sin\theta_i}{\lambda}} & \cdots & e^{-j2\pi(n_r-1)\frac{d_r \sin\theta_i}{\lambda}} \end{bmatrix}^T, \tag{6.31}$$

and

$$\mathbf{h}_t(\phi_i) = \frac{1}{\sqrt{n_t}}\begin{bmatrix} 1 & e^{-j2\pi\frac{d_t \sin\phi_i}{\lambda}} & \cdots & e^{-j2\pi(n_t-1)\frac{d_t \sin\phi_i}{\lambda}} \end{bmatrix}^T, \tag{6.32}$$

with the Angle of Arrival (AoAs) θ_i, the Angle of Departure (AoDs) ϕ_i, the path gains α_i, the random scattering phases φ_i, the Tx/Rx antenna separation d_t / d_r, the wavelength λ and the number of Tx/Rx antennas n_t / n_r. Note that the LOS gain is deterministic. The rank of the channel matrix is determined by the number of paths $B+1$, and also depends on the angular spread. The randomness of the channel is determined by the independent random scattering phases φ_i which are uniformly distributed on $[0,2\pi]$. The Ricean factor $\mathbf{k}$ is the power ratio between the LOS and the NLOS path gains, i.e., $\mathbf{k} = \frac{\alpha_0^2}{\sum_{i=1}^{B}\sigma_{\alpha_i}^2}$ where $\sigma_{\alpha_i}^2$ is the variance of α_i.

It is clear that a large Ricean factor results in a rank-1 channel, in which case only a single stream per user can be transmitted and a single receive antenna ($n_r = 1$) would be sufficient to detect the desired signal. Multiple-Input-Single-Output (MISO) systems are henceforth considered.

6.5.2 The Location Based Linear Precoder

Consider for a moment the 2 user case where the best pair of the users is selected by choosing (as much as possible) orthogonal transmit steering vectors, i.e.,

$$\mathbf{h}_t^H(\phi_p)\mathbf{h}_t(\phi_q) = \frac{1}{n_t}\sum_{i=0}^{n_t-1} e^{-j2\pi\frac{d_t}{\lambda}(\sin\phi_p-\sin\phi_q)i} = \frac{1}{n_t}\frac{1-e^{-j2\pi\frac{d_t}{\lambda}(\sin\phi_p-\sin\phi_q)n_t}}{1-e^{-j2\pi\frac{d_t}{\lambda}(\sin\phi_p-\sin\phi_q)}} = 0, \tag{6.33}$$

where ϕ_p and ϕ_q correspond to the p th and q th user. The necessary and sufficient conditions for array response orthogonality are given by

$$n_t\frac{d_t}{\lambda}(\sin\phi_p - \sin\phi_q) = k\,, \quad k \in \mathbb{Z}, \tag{6.34}$$

and

$$\frac{d_t}{\lambda}(\sin\phi_p - \sin\phi_q) \neq t\,, \quad t \in \mathbb{Z}. \tag{6.35}$$

Note again that a large antenna spacing d_t ($d_t > \lambda/2$) facilitates this orthogonality over a limited sector.

A first precoder that will be considered is the (normalized) MU channel matched filter (MF)

$$\mathbf{V}_{MF} = \mathbf{H}_F^H(diag[\mathbf{H}_F\mathbf{H}_F^H])^{-\frac{1}{2}}, \tag{6.36}$$

The location based Zero-Forcing (ZF) precoder is given by

$$\mathbf{V}_{ZF} = \mathbf{H}_F^H(\mathbf{H}_F\mathbf{H}_F^H)^{-1}(diag[(\mathbf{H}_F\mathbf{H}_F^H)^{-1}])^{-\frac{1}{2}}, \tag{6.37}$$

where the operation $diag$ only selects the diagonal entries. The normalization appearing here greatly reduces the power loss of ZF approaches.

The third system considered is a MU-MISO system inspired by SU-MIMO LTE operation. There are 2 operation modes of SU-MIMO spatial multiplexing: closed-loop and open-loop modes. In the closed-loop mode, a fixed codebook is stored at both the transmitter and the receiver. Without the assumption of channel reciprocity in the closed-loop spatial multiplexing mode, the receiver estimates the channel and feeds the optimal Precoding Matrix Indicator (PMI) to the base station. Additionally, the feedback also contains the Rank Indicator (RI), and the Channel Quality Information (CQI) where the CQI informs the modulation scheme and the power allocation, the RI indicates the number of spatial layers that can be supported by the current channel experienced at the UE. If the PMI is unreliable, the open-loop mode is switched on, where Cyclic Delay Diversity (CDD) is employed to further enhance the robustness of the precoder. The open-loop mode is not discussed here further.

6.5.3 Performance Evaluation

The system capacity of the location based MF, the location based ZF, and LTE codebook based precoders is evaluated in this subsection. For the location based approaches, position error is neglected; however, deviations from a pure LOS channel in the form of a Ricean channel are accounted for. The transmitter takes $n_t = 4$ antennas which is one of the cases considered in LTE and the antenna spacing is $\lambda/2$. For the case of $n_t = 4$, LTE uses a codebook containing 16 code vectors (possible channel quantized values). For the sake of simplicity, a uniform power allocation is used for the $|\mathbb{S}| = 2$ users. The frame length is 1000 and the simulations are averaged over 10000 channel realizations (for the Rayleigh part of the channel, the LOS part being kept constant).

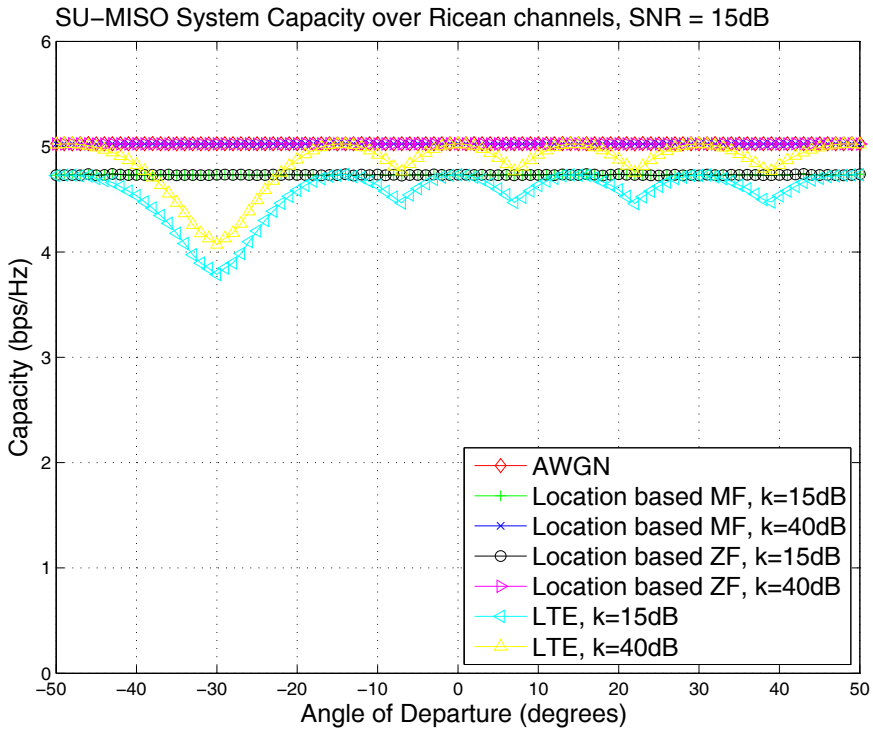

Figure 6.3 Capacity of a SU MISO system v.s. AoD.

The impact of the various AoDs and the Ricean factors in the SU-MISO (system) capacity is illustrated in Figure 6.3, where the SNR is 15dB. A lower Ricean factor leads to lower system capacity for all of the above algorithms. In the SU case, the location based MF and ZF coincide, and both of them are insensitive to the AoDs. For a high Ricean factor (40dB) and hence negligable NLOS paths, the system capacity achieved by the MF/ZF is identical to the AWGN channel capacity. However, the LTE system suffers a capacity loss that depends on the AoD, due to its limited codebook. The highest capacity degradation occurs at $\phi = -30$. Note that a limited range of AoDs ([-50° 50°]) are scanned because of the limited sector covered.

The capacity benefits enjoyed by the location based algorithms are more attractive in the 2-User MISO system, as shown in Figure 6.4. The AWGN capacity of the 2-User MISO system is less than twice that of the SU MISO system under the same transmit power constraint. We fix one user at AoD -30° and scan all the AoDs for the second user. At SNR=15dB, a Ricean factor of 15dB leads already to indistinguishable results from a pure LOS case (and hence from a Ricean factor of 40dB also). It is clear the angle pairs of the 2 users play a crucial role in the performance of all the discussed algorithms. The location based ZF is more stable than the other algorithms and leads to good performance from the moment that the two angles are sufficiently separated. The location based MF is worse than the location based ZF due to inter-user interference. However, due to steering vector orthogonality for the second user at 0° or at 30°, the capacity reduction disappears at those angles. The LTE system performs even worse than the location based MF because of the limited size of the codebook (even though it contains 16 code vectors).

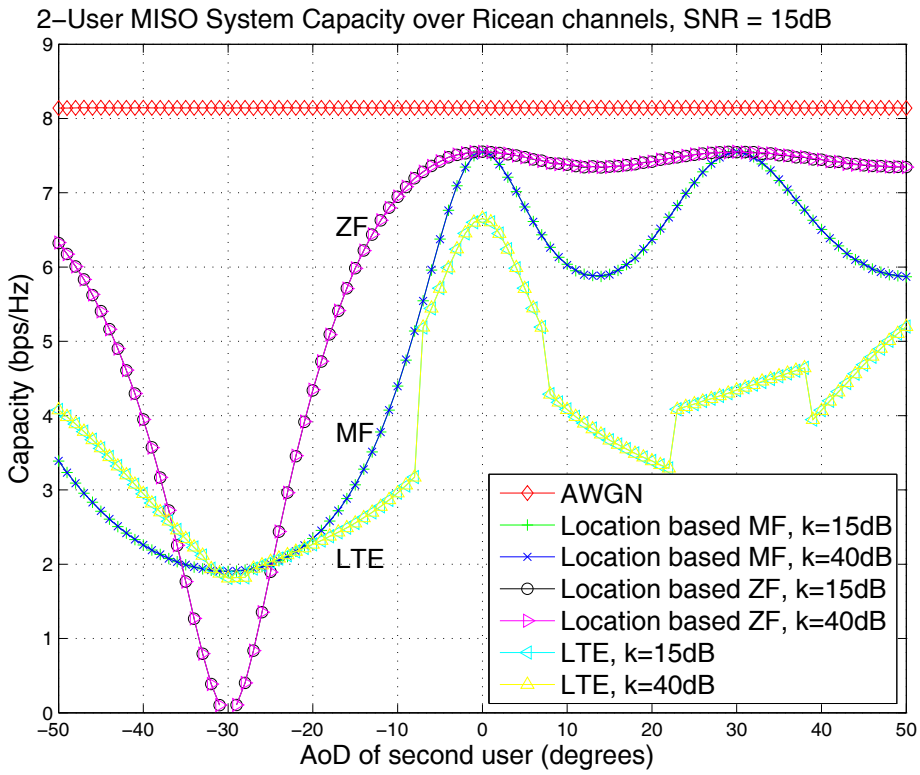

Figure 6.4 Capacity of a 2-User MISO system v.s. AoD of the second user ($AoD_1 = -30^o$).

Similar conclusions are drawn by varying the SNR, which is shown in Figure 6.5, where the achievable SU-MISO capacity of the algorithms is insenstive to the AoD except for the LTE system. In the LTE system, the best and worst precoder vectors correspond to the AoDs of 0° and -30° respectively (smallest (zero) and largest quantization error). The location based linear precoders indeed enhance the capacity siginificantly in the 2-User MISO system, as shown in Figure 6.6, where the AoD pairs of the 2 users are crucial to the system capacity. We select users with an optimal angle pair as close as possible to the orthogonality conditions (6.34) and (6.35). The location based MF and ZF achieve identical system capacity at various Ricean factors. Unfortunately, the LTE system yeilds up to two times smaller sum rate than the location based algorithms due to the limited codebook size, even with the 2 true channels being orthogonal to each other.

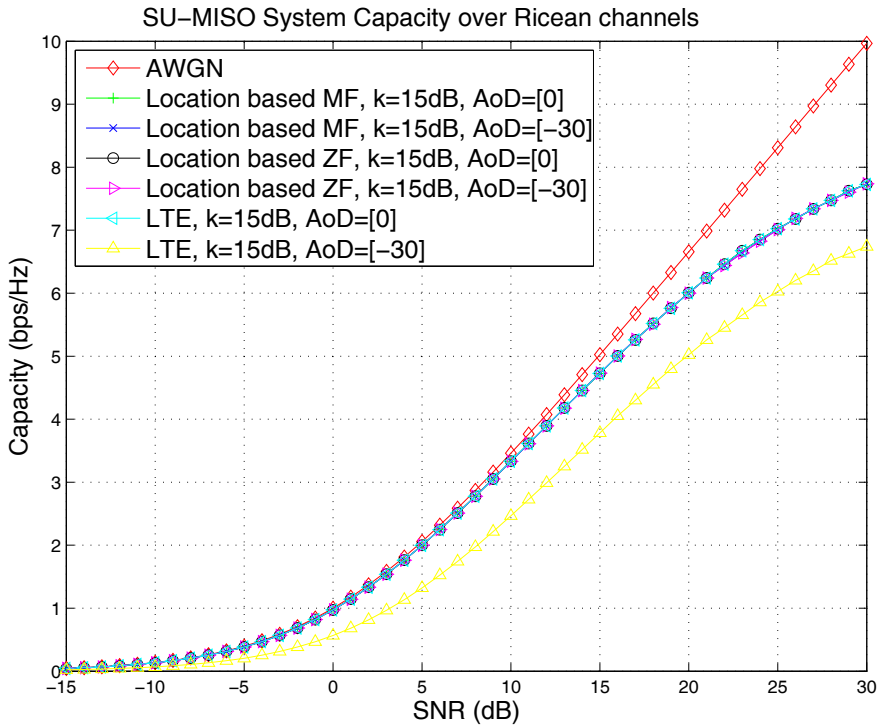

Figure 6.5 Capacity of a SU MISO system v.s. SNR.

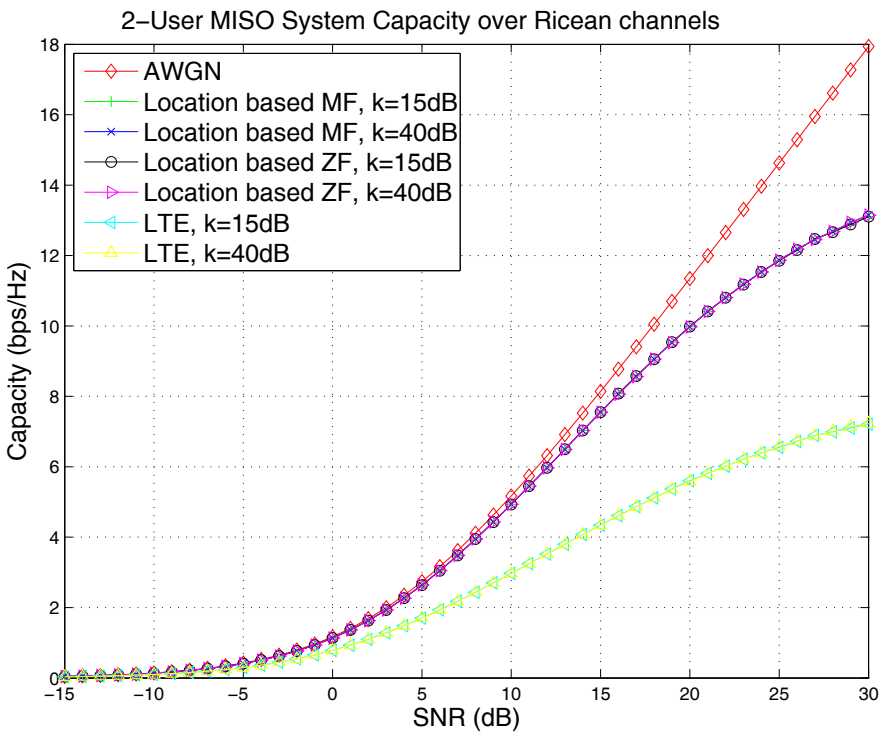

Figure 6.6 Capacity of a 2-User MISO system v.s. SNR.

6.5.4 Summary

In this section, the location information based LOS channel model and associated linear precoders are discussed. It can be concluded that the location based precoders are capable to achieve higher system capacity than a channel quantization and feedback based LTE system due to the limited codebook in the LTE system. The capacity enhancement by the location based linear precoders is significant already for a 2-User MISO system.

7 Conclusions

This documents presents the final version of location based cross-layer optimisation for PHY/MAC in OFDM based systems. The main contributions in this document include using positioning information to:

- Investigate the achievable performance gain of letting two relay to destination transmissions occur simultaneously, when assuming that location information is made available to the AP through a periodic collection mechanism (AAU);
- Provide frame error rates for different code rates and modulation alphabets at different receiver locations in a particular scenario, which can be used as a location dependent link level model input for higher layer simulations (DLR);
- Describe all the steps followed in the implementation of SDMA in 802.16e based DRA. It includes performance gain on the throughput and discuss the trends of such a SDMA implementation on the mobile WiMAX (IT);
- Study Gaussian Underlay Cognitive Radios with the propose of new spectrum-access approaches for various underlay cognitive radio modes, by assuming the availability of CQI about the secondary to primary user link (UniS);
- Investigate the gain in sum rate of position based SDMA (multi-user MIMO) over LTE style SDMA approaches plus provide an analytical description of the sum rate loss degradation due to non-LOS components and position information errors (EUR).

The technical fields analysed in this deliverable include cross from relay based systems to cognitive radio, passing to a DRA MAC design and link layer interface model. The main conclusions from each work focused in the deliverable are presented above and includes highlights for future work.

The results in chapter 2 have shown that positioning-based relaying schemes where simultaneous relay to destination transmissions are used can be beneficial despite the cross-interference occuring when transmitting simultaneously. Proposed algorithm, named SimTX, jointly determines the best combination of relays and transmit power for the two simultaneous relay to destination transmissions. Numerical results show that this approach can improve the throughput by 19% in the considered scenario when comparing to the cases where normal two-hop relaying would be used. Future work item would be to extend the joint relay selection and power adaptation, to include rate adaptation. Inaccuracy in the positioning information and inaccurate path-loss model parameters on two-hop relaying schemes is recommended to have more near-practical situations results.

A Link Level Model study in the chapter 3 includes frame error rates for a convolutionally coded OFDM system applying different modulation constellations and code rates, with simulations been carried out using real propagation measurements from work package 4, rather than stochastic channel models. These results can serve as basis for the generation of lookup tables for system level simulations, which then provide system performance measures for a real indoor scenario.

In the chapter 4 a whole a fully implementation of a SDMA-based DRA architecture for the mobile WiMAX was proposed and a set of system level simulations was performed. The reults compare the gains achieved over simple scenarios where a non SDMA-based DRA architecture is implemented. For the proposed SDMA scheme, the achieved gain on the throughput can reach up to 50 Mbps, a gain of roughly 38 Mbps over the simple non SDMA-based DRA scheme. Location information is used to estimate the direction of arrival of each mobile related to the base-station and is assumed the beamforming algorithm is optimal in steering one spatial beam into the direction of the desired mobile. Although the desired performance of such a system was provide, inaccuracy in the positioning information is recommended to have a near-practical situations results.

Concerning underlay cognitive radio in the chapter 5, fundamental relationship between the secondary user's achievable rate and capacity penalty to the primary user have been studied in four carefully classified UCR modes for full and partially provided multiuser CQI. For partial multiuser CQI a new physical-layer model is established through exploitation of the location-aided approach and new spectrum access and power allocation strategies are investigated. Numerical examples are provided to show the performance of the UCR with full multiuser CQI and the proposed approach with partial multiuser CQI.

Chapter 6 considers again location based SDMA and provides sum rate degradation calculations due to the channel deviating from a pure line-of-sight case (Ricean channel) and due to position information errors. Comparitive simulations are provided for location based SDMA and an LTE style approach based

on channel quantization and feedback. Already in a simple 2-user scenario as considered here, a factor two gain in sum rate can be obtained in the location based approaches over an LTE style approach.

References

[Ala98] S. Alamouti, "A simple transmit diversity technique for wireless communications," IEEE Journal on Selected Areas in Communications, 16(8):1451-1458, October 1998.

[B00] Bianchi, G. Performance analysis of the IEEE 802.11 distributed coordinationfunction. *Selected Areas in Communications, IEEE Journal on*, 18(3):535--547, 2000.

[CAR78] A. B. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 60-70, Jan. 1978.

[CCP08] Ching Yao Huang; Chieh-Yao Chang; Po-Han Chen; Ming-Hsien Wu; "Performance Evaluation of SDMA based Mobile WiMAX Systems" IEEE International Wireless Communications and Mobile Computing Conference, 2008, (IWCMC 2008), Aug. 2008, pp. 1042-1046.

[CG87] M. H. M. Costa and A. A. El. Gamal, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Trans. Inf. Theory*, vol. 33, no. 710-711, Sep. 1987.

[COS83] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439-441, May 1983.

[Cov72] T. Cover, "Broadcast channels", *IEEE Trans. on Information Theory*, vol. 18, pp. 2-14, 1972.

[CS03] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel", *IEEE Trans. on Information Theory*, vol. 49, pp. 1691-1706, 2003.

[CT06] T. M. Cover and J. A. Thomas, *Elements of Informaiton Theory*, Wiley, New York, Second Edition, 2006.

[D1.1] WHERE Deliverable D1.1 "Definition of the WHERE framework an scenarios". Online available at http://www.ict-where.eu

[D4.1] WHERE Deliverable D4.1 "Measurements of location-dependent channel features". Online available at http://www.ict-where.eu

[DPT06] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813-1827, May 2006.

[DRX98] Durgin, G., Rappaport, TS, and Xu, H. Measurements and models for radio path loss and penetration loss inand around homes and trees at 5.85 GHz. *IEEE Transactions on Communications*, 46(11):1484--1496, 1998.

[EMS02] Eklund, C.; Marks, R. B.; Stangwood K. L.; Wang, S.; "IEEE Standard 802.16: A Technical Overview of the Wireless MAN Air Interface for Broadband Wireless Access", IEEE Communications Magazine, June 2002, pp. 98-107.

[GKH07] D. Gesbert, M. Kountouris, J. R. W. Heath, C. B. Chae and T. Salzer, "From Single user to Multiuser Communications: Shifting the MIMO paradigm", *IEEE Sig. Proc. Magazine*, 2007.

[GP80] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9, pp. 19-31, Jan. 1980.

[HEW06] Hoymann C.; Meng H.; Ellenbeck J.; "Influence of SDMA-Specific MAC Scheduling on the Performance of IEEE 802.16 Networks", Proceedings of the 12th European Wireless Conference 2006, Athens, Greece, April 2006.

[HH03] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?", *IEEE Transactions on Information Theory*, vol. 49, pp. 2058-2080, 2003.

[HL06] A. T. Hoang and Y. C. Liang, "A two-phase channel and power allocation scheme for cognitive radio networks," in *Proc. IEEE PIMRC*, 2006, pp. 1-5.

[HT08] Z. Hu and C. Tham, "CCMAC: coordinated cooperative MAC for wireless LANs," in *Proceedings of the 11th international symposium on Modeling, analysis and simulation of wireless and mobile systems*. ACM New York, NY, USA, 2008, pp. 60–69.

[HWP06] Hoymann C.; "MAC Layer Concepts to Support Space Division Multiple Access in OFDM based IEEE 802.16", Wireless Personal Communications Magazine, pp. 363-385, Sep. 2006.

[HY03] Hujun Yin; "Performance of Space-Division Multiple-Access (SDMA) With Scheduling", IEEE Transactions on Wireless Communications, vol. 1, no. 4, Oct. 2003.

[ICT] EU-ICT FP7 WHERE Project, www.ict-where.eu.

[IEEE07] IEEE 802.11-2007, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, June 2007.2007.

[JG05] N. Jindal and A. Goldsmith, "Dirty paper coding versus TDMA for MIMO broadcast channels", *IEEE Trans. on Information Theory*, vol. 51, pp. 1783-1794, 2005.

[Jin05] N. Jindal, "A high SNR analysis of MIMO broadcast channels", *Proc. IEEE Int. Symp. Information Theory, Adelaide, Australia*, pp. 2310-2314, 2005.

[Jin06] N. Jindal, "MIMO Broadcast Channels with Finite Rate Feedback", *IEEE Trans. on Information Theory*, vol. 52, pp. 5045-5060, 2006.

[JLM09] N. Jindal, A. Lozano and T. Marzetta, "What is the Value of Joint Processing of Pilots and Data in Block-Fading Channels?", in *Proc. IEEE Int. Symp. Information Theory*, 2009.

[JSM09] S. A. Jafar, S. Srinivasa, I. Maric, and A. Goldsmith, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," *Proceeding of the IEEE*, vol. 97, no. 5, pp. 894-914, May 2009.

[Kay93] S. M. Kay, "*Fundamentals of Statistical Signal Processing - Estimation Theory*", Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.

[KH95] R. Knopp and P.A. Humblet, "Information capacity and power control in single-cell multiuser communications", in *Proc. International Conference on Communications, Seattle, USA*, pp. 331-335, 1995.

[KLH06] T. Kim, H. Lim, and J. Hou, "Improving spatial reuse through tuning transmit power, carrier sense threshold, and data rate in multihop wireless networks," in Proceedings of the 12th annual international conference on Mobile computing and networking. ACM, 2006, p. 377.

[KM05] Kesselman, A. and Mansour, Y. Optimizing TCP retransmission timeout. *ICN'05: Proceedings of The 4th International Conference on Networking*, pages 133--140, 2005. Citeseer.

[KRA04] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 581-586, Mar. 2004.

[KTG08] Kai Sun; Ying Wang; Tan Wang; Zixiong Chen; Guona Hu; "Joint Channel-Aware and Queue-Aware Scheduling Algorithm for Multi-User MIMO-OFDMA Systems with Downlink Beamforming", IEEE Vehicular Technology Conference, 2008, (VTC 2008), Fall, Sept. 2008, pp. 1-5.

[LFSS06] G. Lebrun, M. Faulkner, M. Shafi and P.J. Smith, "MIMO Ricean channel capacity: an asymptotic analysis", *IEEE Trans. on Wireless Communications*, vol. 5, pp. 1343-1350, 2006.

[LLH06] H. Lim, C. Lim, and J. Hou, "A coordinate-based approach for exploiting temporal-spatial diversity in wireless mesh networks," in *Proceedings of the 12th annual international conference on Mobile computing and networking*. ACM, 2006, p. 25.

[LTN+07] P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S. Panwar, "CoopMAC: A cooperative MAC for wireless LANs," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, p. 340, 2007.

[Mar06] T.L. Marzetta, "How much training is required for multiuser MIMO?", in *Proc. Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA*, pp. 359-363, 2006.

[Med00] M. Medard, "The Effect upon Channel Capacity in Wireless Communications of Perfect and Imperfect Knowledge of the Channel", *IEEE Transactions on Information Theory*, vol. 46, pp. 933-946, 2000.

[MH99] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communications link in Rayleigh flat fading", *IEEE Trans. on Information Theory*, vol. 45, pp. 139-157, 1999.

[MH06] T. Marzetta and B. Hochwald, "Fast Transfer of Channel State Information in Wireless Systems", *IEEE Trans. on Signal Processing*, vol. 54, pp. 1268-1278, 2006.

[MM99] J. Mitola III and G. Q. Maguire Jr., "Cognitive radio: making software radio more personal," *IEEE Personal Communs*., pp. 13-18, Aug. 1999.

[MYK07] I. Maric, R. D. Yates and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3526-3548, Oct. 2007.

[NAA05] T. Nadeem, L. Ji, A. Agrawala, and J. Agre, "Location enhancement to IEEE 802.11 DCF," in *Proceedings IEEE INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 1, 2005.

[NMS10a] Jimmy J. Nielsen, Tatiana K. Madsen , and Hans-Peter Schwefel. Mobility Impact on Centralized Selection of Mobile Relays. *Proceedings of the 6th IEEE Consumer Communications And Networking Conference CCNC 2010*, 2010. IEEE.

[NMS10b] J. J. Nielsen, T. K. Madsen, and H.-P. Schwefel, "Location-based mobile relay selection and impact of inaccurate path loss model parameters," in *Proceedings of the 2010 IEEE Wireless Communications and Networking Conference WCNC 2010*. IEEE, 2010.

[P95] Proakis, J.G. Digital Communications Third Edition. *McGrawHill Inc, New York, USA*, 1995.

[PDF08] S. Parkvall, E. Dahlmann, A. Furuskar, Y. Jading, M. Olsson, S. Wanstedt, K. Zangi, "LTE-Advanced – Evolving LTE towards IMT-Advanced," In *Proceedings 68th IEEE Vehicular Technology Conference (VTC 2008 - Fall), Calgary, Canada*, Sept. 2008.

[PES07] Pabst R.; Ellenbeck J.; Schinnenburg M.; Hoymann C.; "System Level Performance of Cellular WiMAX IEEE802.16 with SDMA enhanced Medium Access", IEEE Wireless Communications and Networking Conference, 2007, (WCNC 2007), March 2007, pp. 1820-1825, March 2007.

[Pla08] S. Plass. *Cellular MC-CDMA Downlink Systems - Coordination, Cancellation, and Use of Inter-Cell Interference*. PhD thesis, Ulm University, Germany, Aug. 2008. VDI Verlag Düsseldorf, Series 10, No. 788, ISBN 978-3-18-378810-1.

[SA05] M. K. Simon and M-S. Alouini, *Digital Communication over Fading Channel*, Wiley & Son, Canada, 2nd Edition, 2005.

[SAT81] H. Sato, "The capacity of Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. 27, no. 6, pp. 786-788, Nov. 1981.

[SH05] M. Sharif and B. Hassibi, "On the Capacity of MIMO Broadcast Channels with Partial Side Information", *IEEE Transactions on Information Theory*, vol. 51, pp. 506-522, 2005.

[SH07] M. Sharif and B. Hassibi, "A Comparison of Time-Sharing, DPC and Beamforming for MIMO Broadcast Channels With Many Users", *IEEE Transactions on Communications*, 5vol. 5, pp. 11-15, 2007.

[SJL08] O. Sjöbergh, E. Jorswieck and E. Larsson, "Greedy user selection for zero-forcing and MMSE multiuser beamforming with channel estimation errors", in *Proc. Int'l Conf. Acoustics, Speech and Sig. Proc. (ICASSP)*, Las Vegas, USA, 2008.

[SKC09] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689-699, Feb. 2009.

[SKY04] Shi S.; Amano Y.; Kawamoto K.; Yamagughi A.; Inoue T.; Takeuchi Y.; Kawazawa T.; "Forward link throughput evaluation of a SDMA based wireless packet cellular system", IEEE Vehicular Technology Conference, 2004, (VTC 2004) Fall, vol. 2, Sept. 2004, pp. 949-953.

[SS04] Spencer Q. H.; Swindlehurst, A.L.; "Channel allocation in multi-user MIMO wireless communications systems", IEEE International Conference on Communications, 2004, (ICC 2004), vol. 5, June 2004, pp. 3035-30-39.

[SS08] U. Salim and D. Slock, "Broadcast Channel: Degrees of Freedom with no CSIR", in *Proc. Allerton Conf. on Communication, Control, and Computing*, 2008.

[SS10] U. Salim and D.T.M. Slock, "How much Feedback is Required for TDD Multi-Antenna Broadcast Channels with User Selection?", *EURASIP Journal on Advances in Signal Processing*, 2010.

[TV05] D. Tse and P. Viswanath, "*Fundamentals of Wireless Communications*", Cambridge Univ. Press, 2005.

[VJA05] V. Monteiro; J. Rodriguez; A. Gameiro; " Dynamic scheduling for HSDPA systems ", WWRF meeting 15, Paris, France, 08-09 December 2005

[WSS06] H. Weingarten, Y. Steinberg and S. Shamai, "The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel", *IEEE Trans. on Information Theory*, vol. 52, pp. 3936-3964, 2006.

[WVA07] W. Wu, S. Vishwanath and A. Arapostathis, "Capacity of a class of cognitive radio channels: interference channels with degraded message set," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4391-4399, Nov. 2007.

[XMH07] Y. Xing, C. N. Mathur, M. A. Haleem, R. Chandramouli, and K. P. Subbalanshmi, "Dynamic spectrum access with QoS and interference temperature constraints," *IEEE Trans. Mobile Comput*., vol. 6, no. 4, pp. 423-433, 2007.

[YG06] T. Yoo and A. Goldsmith, "On the optimality of Multiantenna Broadcast Scheduling Using Zero-Forcing Beamforming", *IEEE Journal on Selected Areas in Communications*, vol. 24, 2006.

[YK05] Ying Jung Zhang; Khaled Ben Letaief; "An Efficient Resource-Allocation Scheme for Spatial Multiuser Access in MIMO/OFDMA Systems", IEEE Transactions in Communications, col. 53, no. 1, Jan 2005.

[ZC07] H. Zhu and G. Cao, "rDCF: A relay-enabled medium access control protocol for wireless ad hoc networks," in *Proceedings IEEE INFOCOM* 2005. 24th *Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 1, 2005.

[ZS07] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access, " *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 79-89, May 2007.