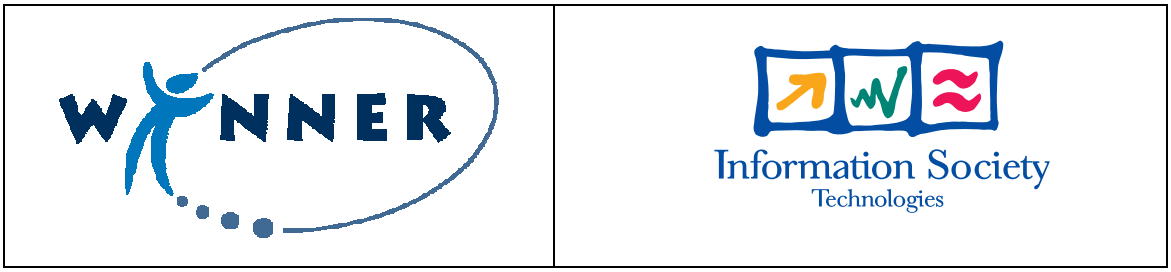


IST-2003-507581 WINNER

D2.4 ver 1.0

Assessment of adaptive transmission technologies

Contractual Date of Delivery to the CEC:	28.02.2005
Actual Date of Delivery to the CEC:	28.02.2005
Author(s):	Sorour Falahati - editor (UU/CTH), Daniel Aronsson (UU/CTH), Elena Costa (SM), Thomas Eriksson (CTH), Tony Ottosson (CTH), Stephan Pfletschinger (CTTC), Adam Piatyszek (PUT), Mikael Sternad (UU/CTH), Tommy Svensson (CTH), Peter Trifonov (SM), Michal Wodczak (PUT)
Participant(s):	CTH, CTTC, PUT, SM, UU/CTH
Workpackage:	WP2 – Radio Interface
Estimated person months:	15 PM
Security:	**Restricted**
Nature:	**R**
Version:	1.0
Total number of pages:	106

Abstract: This document contains a first assessment of adaptive transmission schemes and how these fit to the different WINNER scenarios. It focuses on an adaptation due to the channel properties that may vary both in time and frequency or, with respect to the selected antennas, in an OFDM-based air interface. Besides the evaluation by analysis, link simulation and multilink simulation, first detailed designs are presented of adaptive transmission feedback systems for the WINNER FDD and TDD physical layer modes.

Keyword list: Adaptive transmission, link adaptation, OFDM, channel estimation, channel prediction, channel state information, feedback delay, multiuser waterfilling, multicarrier CDMA, adaptive antenna selection.

Executive Summary

This document contains a first assessment of adaptive transmission schemes and how these fit to different Winner scenarios. It focuses on adaptation to channel properties that may vary in time, in frequency or with respect to the selected antennas, in an OFDM-based air interface. Besides evaluation by analysis, link simulation and multilink simulation, first detailed designs are presented of adaptive transmission feedback systems for the WINNER FDD and TDD physical layer modes.

Link adaptation is based on the use of channel state information at the transmitter, to adjust transmission parameters such as the modulation scheme, the code rate, the transmit power or the transmit antennas.

The current state of the WINNER system concept work, regarding the mode-dependent MAC layer and the assumed radio channels, is first outlined as a background for the subsequent investigations. Then, *link adaptation* strategies that are designed under various constraints and under different assumptions on the availability of channel state information at the transmitter are discussed. Strategies for transmit antenna selection are also outlined. The influence on the performance of code block lengths and channel errors at the receiver are studied. One particular problem is how to adapt to fast time-variations of the resources, when decisions are taken based on outdated information. A systematic solution is proposed, whereby adaptation constraints such as bit error rates can be fulfilled also when the adaptation is based on erroneous channel state information.

The document then considers *multi-link adaptation*, which presents two problems that may have to be solved jointly: The available resources are to be allocated among a set of data flows to/from terminals. Within the allocated resources, the transmission parameters are furthermore adjusted to optimise the transmission. Thus, we have a joint problem of resource scheduling and link adaptation. In the multilink case, the use of channel state information at the transmitter makes it possible to obtain a multiuser scheduling gain: The cell throughput will increase with the number of active users. This is in contrast to e.g. non-adaptive TDMA transmission where the throughput stays constant, or non-adaptive CDMA transmission where the cell throughput decreases with the number of active users in both downlinks and uplinks, due to imperfect orthogonality.

An adaptive transmission and multiuser scheduling design is outlined. It is based on TDMA/OFDMA, i.e. allocating individual time-frequency chunks in an OFDM system to different flows. An aim of this study is to take realistic design constraints into account at far as this is possible at the present state of the WINNER system evolution. Detailed schemes are therefore designed both for the wide-area FDD mode and for the short-range cellular mode based on TDD. The FDD downlink is evaluated by simulations, in which the multiuser scheduling gain, the effect of channel prediction errors, the impact of the size of allocated time-frequency chunks and the effect of using TDMA allocation instead of TDMA/OFDMA are evaluated. The link adaptation and scheduling is based on the use of channel state prediction to overcome the feedback delay. The performance of channel predictors is investigated systematically for the assumed transmission schemes. This leads to guidelines for in what transmission environments adaptive transmission can be expected to work well. The limit between using adaptive and non-adaptive resource scheduling for a particular flow is a function of the SINR and the vehicular velocity of the terminal. With an advanced channel predictor, it is possible to utilize the short-term fading at vehicular velocities at the highest considered carrier frequency within the WINNER project, at 5 GHz. Furthermore, a novel, simple and very efficient method for compressing the channel state feedback information is proposed and evaluated. This method solves the problem of reducing the feedback data rate to acceptable levels when adapting broadband TDMA/OFDMA multiuser transmission to the short-term fading of the channels.

A theoretical study, aimed at estimating the attainable cell performance when using multi-user waterfilling to adapt the transmission over uplinks, is presented. The method is then used to evaluate fundamental limits for the aggregate uplink capacity in WINNER cellular short-range and wide area modes, when using maximal throughput scheduling.

An adaptive multicarrier CDMA scheme that uses spreading over both time and frequency is also studied. An algorithm for adaptive allocation of users to subcarriers is proposed and evaluated. The algorithm enables optimisation of subcarrier sharing between users. In a simulation study, the optimisation criterion is the minimisation of transmit power under throughput constraints for each user. The results are evaluated as a function of the allowed computational complexity of the algorithm. The sensitivity of the performance to imperfect channel state information and channel variations is investigated.

Authors

Partner	Name	Phone / Fax / e-mail
UU/CTH	Daniel Aronsson	Phone: +46 18 471 3071 Fax: +46 18 555 096 e-mail: daniel.aronsson@signal.uu.se
SM	Elena Costa	Phone: + 49 89 636 448 12 Fax: + 49 89 636 455 91 e-mail: elena.costa@siemens.com
CTH	Thomas Eriksson	Phone: +48 31 1745 Fax: +46 31 772 1782 e-mail: tomas.eriksson@s2.chalmers.se
UU/CTH	Sorour Falahati	Phone: +46 18 471 3077 Fax: +46 18 555 096 e-mail: sorour.falahati@signal.uu.se
CTH	Tony Ottosson	Phone: +46 31 772 5189 Fax: +46 31 772 1782 e-mail: tony.ottosson@s2.chalmers.se
CTTC	Stephan Pfletschinger	Phone: +34 93 205 8561 Fax: +34 93 205 8399 e-mail: stephan.pfletschinger@cttc.es
PUT	Adam Piatyszek	Phone: +48 61 665 3936 Fax: +48 61 665 2572 e-mail: adam.piatyszek@et.put.poznan.p

UU/CTH	**Mikael Sternad**	
		Phone: +46 704 250 354
		Fax: +46 18 555096
		e-mail: mikael.sternad@signal.uu.se

CTH	**Tommy Svensson**	
		Phone: +48 31 772 1823
		Fax: +46 31 772 1782
		e-mail: tommy.svensson@s2.chalmers.se

SM	**Peter Trifonov**	
		Phone: +49 89 636 448 65
		Fax: +49 89 636 455 91
		e-mail: petert@dcn.infos.ru

PUT	**Michal Wodczak**	
		Phone: +48 61 665 3913
		Fax: +48 61 665 2572
		e-mail: mwodczak@et.put.poznan.pl

Table of Contents

List of Acronyms and Abbreviations

AP	Access Point
ARQ	Automatic Repeat reQuest
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CDMA	Code Division Multiple Access
CDCH	Common Data Channel
CM	Code and Modulation (scheme)
CNR	Channel Gain to Noise Ratio
COFDM	Coded Orthogonal Frequency Division Multiplexing
CSI	Channel State Information
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
FDD	Frequency Division Duplex
FEC	Forward Error Correction
FFT	Fast Fourier Transform
GCG	General Constant Gain
GMC	Generalized Multi-Carrier
HH	Hughes-Hartogs (algorithm)
IFFT	Inverse Fast Fourier Transform
IR	Impulse Response
ITW	Iterative Water-Filling
JCE	Joint Channel Estimation
LOS	Line-Of-Sight
LPC	Linear Prediction Coding
LS	Least Squares
MA	Multiple Access
MAC	Medium Access Control
MC-CDMA	Multi-Carrier Code Division Multiple Access
MC-TF-CDMA	Multi-Carrier Time Frequency Code Division Multiple Access
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
MSE	Mean Square Error
NLOS	Non Line-Of-Sight
NMSE	Normalized Mean Square Error
OFDM	Orthogonal Frequency Division Multiplexing
PSD	Power Spectral Density
PDU	Protocol Data Unit

QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RI	Radio Interface
RLC	Radio Link Control
RS	Resource Scheduling
RSB	Resource Scheduling Buffer
SISO	Single Input Single Output
SINR	Signal to Interference Noise Ratio
SLC	Service Level Control
SLCB	Service Level Control Buffer
SNR	Signal to Noise Ratio
STBC	Space-Time Block Code
SU	Scheduling Unit
TC	Transform Coding
TDCH	Targeted Data Channel
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TDMA/OFDMA	Allocation of individual time-frequency chunks to flows. (See [WIND26])
UT	User Terminal
VQ	Vector Quantization

1. Introduction

This document contains a first assessment of adaptive transmission schemes and how these fit to different Winner scenarios. The responsibility for the work and results reside within Task 2 of WINNER WP 2, but the work has been performed jointly with Task 4. The work builds on the survey and assessment of link adaptation strategies in [WIND21] and on discussions on adaptive methods for multiple access and multilink adaptation in [WIND26].

Adaptation to channel states, user requirements and changing traffic patterns is a general aim for many aspects of the evolving Winner system concept. This document focuses on one particular aspect: *Adaptation to channel properties* that may vary in time, in frequency or with respect to the selected antennas, in an OFDM-based air interface.

Link adaptation is based on the use of channel state information at the transmitter, to adjust link transmission parameters such as the modulation scheme, the code rate, the transmit power or the transmit antennas. Chapter 2 of the present document outlines some link adaptation strategies that are designed under various constraints and under different assumptions on the availability of channel state information at the transmitter. One particular problem is how to adapt to fast time-variations of the resources, where decisions are taken based on outdated information.

Chapter 3 then considers *multi-link adaptation*, which presents two problems that may have to be solved jointly: The available resources are to be allocated among a set of data flows to/from terminals. Within the allocated resources, the transmission parameters are furthermore adjusted to optimise the transmission. Thus, we have a joint problem of resource scheduling and link adaptation. In the multilink case, the use of channel state information at the transmitter makes it possible to obtain a *multiuser scheduling gain*: the cell throughput will increase with the number of active users. This is in contrast to e.g. non-adaptive TDMA transmission where the throughput stays constant[1], or non-adaptive CDMA transmission where the cell throughput decreases with the number of active users in both downlinks and uplinks, due to imperfect orthogonality. Chapter 3 also contains a section that outlines information-theoretic results on multi-user waterfilling, and proposes them as tools for investigating theoretical bounds on the uplink sum-of-rates capacity at an access point.

Chapter 4 draws conclusions for the implications for the radio interface design, and the applicability of adaptive transmission in different WINNER scenarios, in particular wide area cellular and short-range cellular transmission, and Chapter 5 summarizes the main conclusions of the document. Chapter 6 constitutes an appendix that contains additional results and discussions.

Link adaptation within the WINNER system has to be seen in the wider context of the evolving air interface.

1. First, as noted above, the problem of link adaptation will in a multiuser context be intimately related to that of scheduling the available resources among competing users/data flows. It will also be complemented by a link retransmission scheme, with which it should be jointly designed.
2. Second, while adaptive transmission is advantageous in many cases, it cannot always be used. The terminals may not support it, or the quality of channel state information may be too poor in some situations. Therefore, we must have the alternative to choose between adaptive transmission and a more traditional “averaging” strategy that uses coding, interleaving and antenna diversity to average out the variability of the channel to obtain reliable communication.
3. Third, the link adaptation and resource scheduling will be just a part of a total resource allocation strategy that involves cooperation among multiple protocol layers, multiple antenna resources and multiple WINNER modes.

To provide this context, Section 1.1 describes the current understanding of the relevant aspects of the evolving WINNER RI concept and MAC layer design, as described in the recent internal WINNER reports and in [WIND28].

Section 1.2 then outlines in more detail how the algorithms and results of the present document relate to the WINNER RI concept.

[1] Assuming a constant interference level.

1.1 Link adaptation and scheduling within the WINNER air interface

We focus on the allocation of transmission resources in radio communication links between an access point (AP) and a set of applications located in a set of user terminals (UT**)**. Multi-hop transmission over relays is not discussed in this document. These packet transmissions are termed ***flows***. A set of cooperating resource allocation modules control the resources in a geographical area. The scheduling is partitioned into two levels, called ***service level control (SLC)*** and ***resource scheduling (RS).***

1.1.1 Architecture of the MAC layer

Le us in more detail describe how traffic may be organized over a ***scheduled targeted data channel (TDC)*** suggested for use for point-to-point communication in [WIND28]. In the Radio Link Control (RLC) layer, per flow buffering of incoming IP packets is performed, with flows ordered by service class and terminal/user. The buffers are here denoted ***service level control buffers (SLCB).*** Inter-AP handover, inter-mode scheduling/handover, admission control and active queue management are used to control the queue lengths. A service level controller takes inter-flow fairness, service level contract agreements and total delay constraints into account. The service level controller works on a time scale characterized by the packet arrival rates, on the order of 10-1000 ms. The service level controller may request services via one or several access points, that may utilize different WINNER modes. It requests resources by copying IP packets into another buffer. This buffer is here called a ***resource scheduling buffer (RSB)***. Several RSB, each connected to one or several access points, may be available. The control strategy of the service level controller takes into account how much data that is already queued for transmission in the RSB. It may also assign different priorities to different flows. When a transmission has been acknowledged as successful, the corresponding entity is removed from the corresponding service level control buffer.

Per-flow queuing is performed in the resource scheduling buffers. There is one resource scheduler for each RSB and it determines which queues are to be drained. It works on the time-scale of the ***medium access control (MAC) slot***, which may consist of one or a few chunk durations, see Section 1.1.2 below. The queues are drained with a granularity defined by the ***scheduling unit (SU)***, which may differ from the IP packet size. For each SU, an ***L1 packet*** is formed by coding and bit interleaving the SU. L1 packets are then mapped onto the physical channel resource units. A link ARQ functionality performs retransmission of failed SUs. Thus, the SU is the entity (packet) used for link retransmission.

Within the resource scheduler, two basically different allocation strategies can be used: *Adaptive* with respect to the fine-grained channel state information and *non-adaptive*, that uses averaging with respect to channel variations. These two types of transmission differ fundamentally, and they require resources with different properties. The adaptive transmission, which is the focus of the present document, uses channels state information to utilize the channel and interference variations for different terminals. By selecting the best resources for each flow, the throughput is increased and multiuser scheduling gains may be realized. Non-adaptive averaging strategies, on the other hand, are designed to combat and *reduce the effect of the variability* of the SINR, by e.g. interleaving, space-time coding and diversity combining. Non-adaptive transmission is required when fast channel state feedback is unreliable due to e.g. a high terminal velocity or a low SINR, or when the terminal does not support the adaptive transmission. Non-adaptive transmission is also mostly required for point-to-multipoint communication (common data channel (CDC) flows in [WIND28]). The resource scheduler will thus contain two separate algorithms for these two purposes. At each instant, a flow is controlled by one of these algorithms, but the other algorithm may take over control according to need or changed circumstances.

See Figure 1.1 for a summary of the scheduled downlink transmission. (Note that this figure mixes functionalities that in [WIND32] are separated into a control plane and a user plane.)

The scheduling architecture may look exactly the same for uplink traffic if grants are issued per flow, with the main difference that the header compression, ciphering, active queue management, buffers and ARQ are located in the UT whereas the schedulers and flow classifier are located somewhere in the network. In case resources are granted per UT instead, the UT also need to implement a modified service level controller to allocate the given resources to the individual flows.

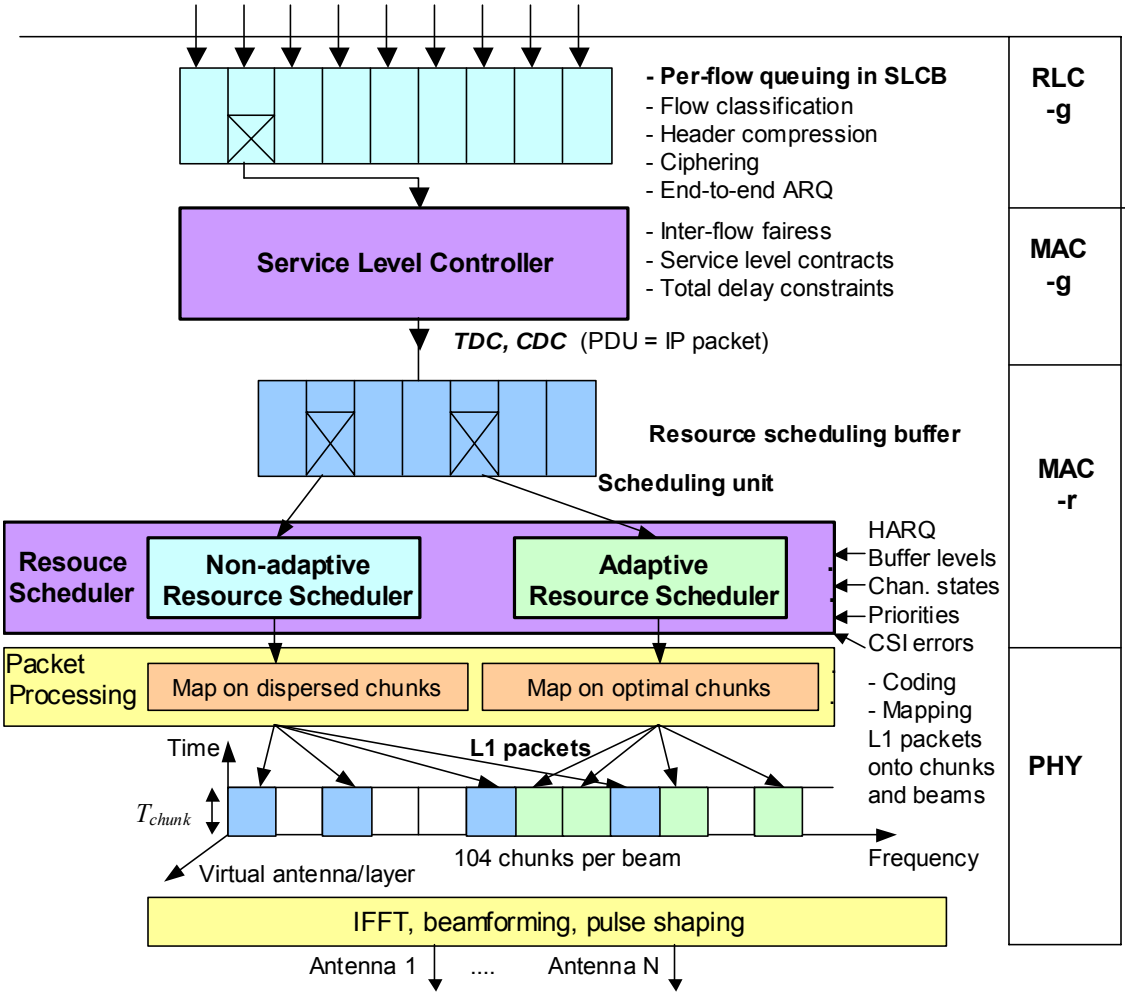


Figure 1.1: Downlink scheduled data channels. Summary of buffers, main control blocks, scheduling algorithms and sublayers. Here, MAC-g represents the generic (mode-independent) medium access control sublayer, while MAC-r is the mode-specific MAC layer

We now introduce the time-frequency resources that are at the disposal of the resource scheduler. Four Winner system modes are selected for closer study by WINNER Task 2.7. They utilize two different physical layer modes:

- In the ***wide-area cellular mode***, it is assumed that FDD transmission is performed over paired 20.0 MHz bands, using 1024 subcarriers, of which 832 (16.25 MHz) are utilized. With 19531 Hz subcarrier spacing and 5.0 μs guard interval per OFDM symbol, we have a total duration of the symbol and guard interval of 56.2 μs. In downlinks, OFDM transmission is assumed. In uplinks, different variants of ***generalized multicarrier*** (***GMC)*** transmission, which include OFDM and single carrier transmission as special cases [WIND21], are under investigation.
- The ***feeder mode, short-range cellular and peer-to-peer modes*** all assume transmission over a single (unpaired) band of up to 100 MHz, using TDD and OFDM with 2048 subcarriers, of which 1664 (81.25 MHz) are utilized. The subcarrier spacing is thus 48828 Hz, the assumed guard interval and symbol rolloff is 0.80 μs and the symbol + guard interval duration is 21.28 μs.

For simplicity, these two physical layer modes are denoted ***wide-area*** and ***short-range,*** respectively, in the following. The basic time-frequency resource unit is denoted ***chunk***. It consists of a rectangular time-frequency area that comprises a number of subsequent OFDM symbols and a number of adjacent subcarriers. The time-frequency interval of one OFDM symbol time and one subcarrier will be denoted one ***(constellation) symbol***. It can be loaded with a number of bits, depending on the modulation constellation used. A chunk contains payload symbols and pilot symbols. It may also contain control symbols in order to minimize feedback delays. This is called in-band control signalling. The number of

offered payload bits per chunk will depend on the utilized coding-modulation formats, on the sizes of the chunks, specified below, and on the number of control symbols.

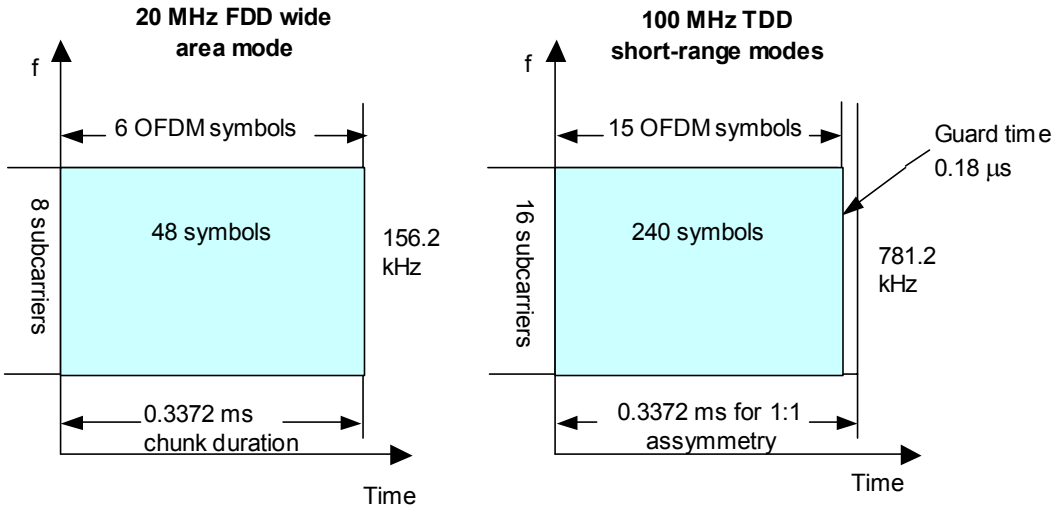


Figure 1.2: Summary of the chunk sizes in the wide area mode and the three short-range modes (short-range cellular, feeder to relays and peer-to-peer)

Table 1.1: Basic parameters used for simulation of wide-area and short-range physical layer modes

Parameter	Wide-area FDD	Short-range TDD (Asymmetry 1:1)	Units/notes
Centre frequency	5.0 +/- 0.384	5.0	GHz
Number of subcarriers in OFDM	1024	2048	Equals length of FFT
FFT BW	20.0	100.0	MHz
Signal BW	16.25, paired	81.25	MHz
Number of subcarriers in use	832	1664	[-416:416] and [-832:832] Subcarrier 0 not used
Subcarrier spacing	19531	48828	Hz
OFDM symbol length (Excluding cyclic prefix)	51.20	20.48	µs
Cyclic prefix length and rolloff	5.00	0.80	µs
Physical chunk size	156.24 x 337.2	781.25 x (319.2+18)	kHz x µs (TDD guard interval 18 µs)
Chunk size in symbols	8 x 6 = 48	16 x 15 = 240	(Frequency x time)

The chunk width is set to *8 subcarriers, or 156.2 kHz* in the wide-area mode and to *16 subcarriers, or 781.2 kHz* in short-range modes. Thus, the usable 832 subcarriers in the wide-area mode as well as the usable 1664 subcarriers in the short-range modes are both partitioned into 104 chunks. The chunk sizes have been selected as reasonable values, but have not been optimised. The suitability for adaptive transmission is a prime parameter in this design, so effects of changing the chunk size should be studied. In particular, variations of chunk sizes in the FDD-based wide-area mode are evaluated in Subsection 3.1.5.4 below.

Assuming TDD asymmetry factor 1 (equal time duration of uplink and downlink chunks in the short-range modes), the slot size T_{chunk} has in by Task 2.7 preliminarily been set to approximately 0.33 ms, to accommodate the basic delay requirements. It is here proposed to be specified more precisely to 0.3372

ms. Thus, in short-range TDD modes, the Uplink +Downlink +duplex guard times should equal 2 x 0.3372 ms = 0.6744 ms. The same slot size of 0.3372 ms is used in the wide-area mode, to simplify inter-mode cooperation. Thus, in the wide-area mode downlink that uses OFDM, 0.3372 ms corresponds to 6 symbol times of 56.2 μs. These chunks thus contain 8 x 6 = 48 constellation symbols. The same holds for wide-area uplinks that use OFDM transmission. For single-carrier uplinks that use the whole 16.25 MHz, a slot length of 0.3372 ms would contain 5480 constellation symbols, some of which would be used as pilots (training sequences). For the short-range modes, the total chunk length 0.3372 ms for TDD asymmetry 1:1 is used for 15 symbols of 21.28 μs, plus a duplex guard time of 18.0 μs. The chunks of short-range modes thus have size 16 x 15 = 240 constellation symbols at asymmetry 1:1. All these parameters are to be regarded as preliminary choices used for simulation. They are summarized in Figure 1.2 and Table 1.1.

1.1.2 Adaptive resource scheduling

Resource scheduling is performed for each MAC slot, which consists of one or several chunk times of 0.3372 ms. The choice of MAC slot size will be affected by several aspects: limits on the complexity of the scheduler, total delay constraints in the WINNER requirements and the predictability of channels for vehicular users.

The general outline of an adaptive resource scheduler below is an example. It is based exclusive allocation of chunks to flows, and on the simplifying assumption that link adaptation and multi-user resource scheduling can be solved in two subsequent steps. (This separation is consistent with an optimal resource allocation for some allocation criteria and constraints, but not for others such as power allocation [WIND21]. Some algorithms in later chapter will consider the two problems jointly.)

Transmission of flows controlled by the adaptive resource scheduler could then proceed as follows. First, separate link adaptations are calculated for *all allowed combinations of users and chunks*. Then, the chunks are allocated to flows:

1. For each terminal, the code, modulation and multiantenna transmit scheme is determined for each chunk, under the assumption that this chunk will be used exclusively by the terminal, and that a pre-defined transmit power will be used. This single-user rate adaptation is done based on SINR predictions and their uncertainty. They are determined for *each chunk* and for *each spatial channel* within the sub-bandwidth allocated to the UT.[2] In adaptive transmission, *the same code and modulation rate* is to be used for all payload symbols within a time-frequency chunk of a spatial channel.
2. Chunks are then allocated to flows that are assigned to adaptive transmission. The proposed code and modulation rates, as well as priorities, buffer levels and constraints on the chunk allocation policy due to the MA scheme are taken into account. The buffers are then drained.
3. The bits to be transmitted are partitioned into one or several sets, the scheduling units (SU).
4. CRC code bits and short sequence numbers are added to the payload bits of each SU. The bits are then FEC coded and bit interleaved, to form one L1 packet for each SU.
5. The L1 packets are mapped onto the allocated chunks. The code and modulation schemes used are those that were determined in Step 1. The so formed chunks are then mapped/beamformed on antenna streams and transmitted.
6. After reception, the chunks are demodulated and decoded and the L1 packets are extracted from them. Each received L1 packet is decoded, and the received SU's are recombined into a packet on the SLC level (an IP packet). When this is complete, the IP packet is forwarded to the MAC-g sublayer.
7. In the case of a failed transmission of an L1 packet, a retransmission is requested. The scheduling units represent the packets to be retransmitted, so they have to be saved at the transmitter, until there is a successful transmission or a timeout.

The FEC coding mentioned in step 4 can be performed in two ways. The payload symbols to be allocated to each chunks can be regarded as separate blocks to be coded. Separate code rates, and modulation formats, can then be used within each chunk. The CRC code performs error detection over the whole

[2] Sub-bandwidths may be used to restrict the required channel prediction computational loads and the feedback rates.

(possibly multi-chunk) SU block. Alternatively, one may perform coding of the whole SU block, and then map it on several chunks, which each may use individually adjusted modulation levels for their payload symbols. The former method will be the primary method used here, but the latter is also of interest. Advantages and disadvantages are briefly discussed in Section 3.1.6.

The optimal multiple access scheme, or combination of schemes, is a topic under active investigation, and no firm decision has yet been reached. Three schemes, TDMA/OFDMA, OFDM-based TDMA and multicarrier CDMA with time-frequency spreading, will be utilized and partly compared in the present document. Seen from the perspective of the resource scheduler, these three schemes can be viewed as different sets of constraints on the allocation:

- TDMA/OFDMA is the most flexible scheme, in which individual chunks are allocated to flows.
- TDMA corresponds to the scheduling constraint that all chunks within the frequency band (or perhaps a sub-band) must be allocated to the same flow (or perhaps only to the aggregated flows for clients residing within the same user terminal).
- Multicarrier (MC)-CDMA using time-frequency spreading can be represented by the simultaneous allocation of sets of chunks to each user, where the number and placement of these chunks depends on the time-frequency spreading scheme used. L1 packets from several flows may then be multiplexed by spreading within each of these sets.

Within T2.4, a summary has been made of the drawbacks and advantages that can be seen with different multiple access schemes prior to large-scale evaluation of their performance by multilink simulation and system simulation. Combinations of schemes that preserve the advantages and avoid the weaknesses have been identified. This work has led to the following conclusions.

Multicarrier transmission:

One set of chunks is allocated to the adaptively scheduled flows, and another set to the non-adaptively scheduled flows. Different MA schemes are preferred within these sets:

- TDMA/OFDMA, i.e. allocation of individual chunks to flows, is used for the *adaptively allocated flows*. With the chunk sizes discussed in Section 1.1.1, the frequency-domain channels will be almost constant within chunks. The number of payload symbols per chunks is appropriate with respect to the expected content of RSB queues, whose inflows are IP packets. The possibilitys for attaining multiuser scheduling gains, with a fine-grained allocation over time, frequency and spatial channels, is maximized by using TDMA/OFDMA.
- For *non-adaptively allocated flows*, the assignment of sets of widely dispersed chunks to flows tends to result in too large resource blocks. TDMA/OFDMA is therefore not appropriate here. Instead, a more fine-grained allocation is used *within* the chunks assigned for the non-adaptive flows. Three alternatives are under present investigation for this purpose:
    1. TDMA. Individual OFDM symbols within the chunks are allocated to flows.
    2. OFDMA. Individual subcarriers within the chunks are allocated to flow, in a way that provides high diversity. (If IFDMA/FDOSS is to be used, cf. [WIND26], then regularly spaced subcarriers must be allocated.)
    3. MC-CDMA with time-frequency spreading, over the whole set of chunks allocated for non-adaptive transmission. The allocation of spreading codes to flows gives a fine-grained rate matching.

Single carrier uplinks: Here, the considered alternatives are TDMA and CDMA.

In the present document, we make the following simplifying default assumptions that tie down some parameters that are still under active investigation in the work on the WINNER system concept.

- The scheduling algorithm used for comparison purposes in multi-link simulation will mostly be the *Proportional Fair* algorithm that allocates the chunk(s) to the terminal with the best signal to interference and noise ratio (SINR), relative to its own average. In experiments where all terminals have the same average SINR, it reduces to the *maximum throughput scheduling*, which gives the chunks to users with highest (predicted) SINR. Thus, buffer levels and different priorities between flows will not be taken into account. An assumption of full buffers is used, so the behaviour of the service level control is not modelled in detail.

- The whole bandwidth (104 chunks in both wide-area and short-range modes) is used exclusively for adaptive transmission and its related control signalling. Thus, we have no mix of adaptively and non-adaptively allocated flows and contention-based channels in our simulation examples.
- Scheduling units of fixed size in bits are used. Packet error rates refer to packets of that size.
- As coding scheme, we utilize convolutional coding. The coding is performed individually within each chunk (or set of allocated chunks in MC-CDMA). No outer L1 packet FEC code is used.

1.2 Outline of contributions, from the perspective of the Winner air interface

Chapter 2 described and evaluates methods for adapting links in a single-user context. Section 2.1 describes an algorithm for rate- and power allocation over OFDM subcarriers, the Hughes-Hartogs algorithm. It then describes an adaptive coding and modulation scheme that combines convolutional coding with M-QAM modulation. The adaptive scheme can be used either for minimizing the transmit power or maximizing the throughput. It is here designed to maximize throughput under bit error constraints. The theoretically attainable gain in throughput when adjusting to short-term Rayleigh fading is illustrated by a comparison with fixed modulation schemes. The impact on performance of the codeword length, the number of rates, imperfect channel state information at the receiver and outdates channel state information at the transmitter are investigated.

Section 2.2 then introduced channel *prediction,* to mitigate the problem that feedback delays and computation delays will result in adaptation decisions that are based on outdated information. However, channel predictions will not be perfect. The section then discusses how to modify the adaptive M-QAM scheme of the type discussed in Section 2.1 to enable design constraints to be fulfilled in the presence of channel prediction errors. In particular, the adaptation rate limits are modified so that bit error constraints are fulfilled when the selection of the code and modulation rate is based on SNR predictions with known inaccuracy. Adaptive M-QAM rate limits optimised with these tools are then used in the multi-link adaptation system presented in Section 3.1.

Section 2.3 discusses methods for antenna selection, i.e. selecting an optimal subset of the available transmit antennas. It is discussed how this method could be applied in a cooperative diversity context, where the antennas are distributed in space.

Section 3.1 outlines and evaluates an adaptive transmission and multiuser scheduling proposal that is based on TDMA/OFDMA, i.e. allocating individual chunks to different flows. An aim of this study is to take realistic design constraints into account at far as this is possible at the present state of the WINNER system evolution. Detailed scheme are therefore designed both for the wide-area FDD mode and for the short-range cellular mode based on TDD. The FDD downlink is evaluated by simulations, in which the multiuser scheduling gain, the effect of channel prediction errors, the effect of modifying the chunk size and the use of TDMA instead of TDMA/OFDMA are evaluated. The link adaptation and scheduling is based on the use of channel state prediction to overcome the feedback delay. The performance of channel predictors is investigated systematically for the assumed transmission schemes. This leads to guidelines for in what transmission environments adaptive transmission can be expected to work well. The limit between using adaptive and non-adaptive resource scheduling for a flow will turn out to be a function of the SINR and the vehicular velocity of the terminals. It will be investigated to what extent adaptive transmission is possible at vehicular velocities. Furthermore, a novel, simple and very efficient method for compressing the channel state feedback information is proposed and evaluated.

Section 3.2 is a theoretical study of subcarrier bit- and power allocation in a multiuser setting, in the OFDM uplink. The tools can be utilized for estimating the attainable cell performance when using multi-user waterfilling to adapt the transmission over uplinks. This method is then in Chapter 4 used to evaluate fundamental limits for the aggregate uplink capacity in WINNER cellular short-range and wide area modes, when using maximum throughput scheduling without imposing fairness constraints.

Section 3.3 studies an adaptive multicarrier CDMA scheme that uses spreading over both time and frequency. This technique was proposed in [WIND26]. Here, an algorithm for adaptive allocation of users to subcarriers is proposed and evaluated. The algorithm enables optimisation of subcarrier sharing between users. In a simulation study, the optimisation criterion is the minimisation of transmit power, under individual throughput constraints for each UT. The results are evaluated as a function of the allowed computational complexity of the algorithm. The sensitivity of the performance to imperfect channel state information and channel variations is investigated.

2. Link Adaptation Algorithms

2.1 Optimum multi-coded OFDM transmission

Figure 2.1 shows a simplified scheme of the considered adaptive multi-coded *orthogonal frequency division multiplexing* (OFDM) system.

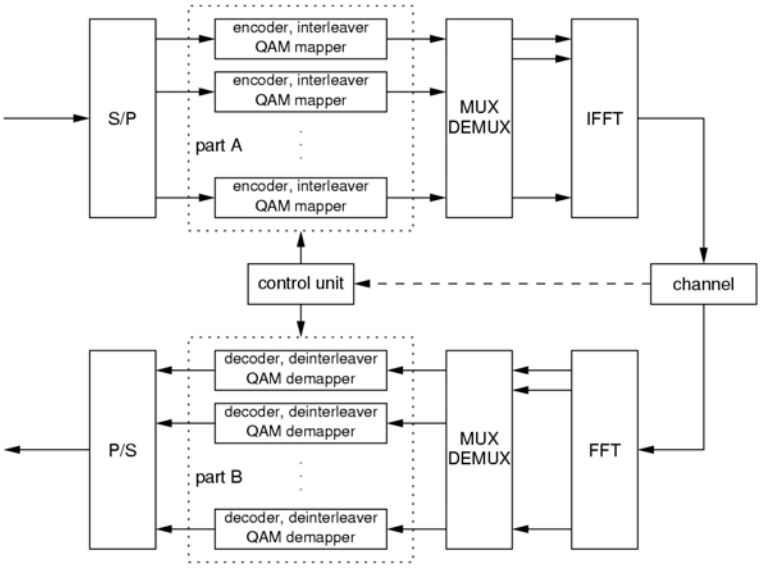


Figure 2.1: Simplified adaptive multi-coded OFDM system scheme

The control unit block is the brain of the system. It is responsible for the adaptation of various system parameters according to a set of input parameters. These input parameters are: the required average bit error rate, power consumption limit of the transmitter and, of course, temporary channel conditions. Throughput maximisation is one of the considered criteria of the adaptation. The throughput is here represented by the average number of information bits transmitted per constellation symbol (BPSK or M-QAM), so the results are independent of the system bandwidth. As an output, the control unit produces a set of parameters, i.e. the code rate, constellation size and power assigned to each subcarrier used for transmission [PWW05].

The proposed system can be realized in the form of a bank of parallel encoders and modulators in the transmitter and demodulators and decoders in the receiver, each of them used for different channel conditions. However, it is not possible to implement and use an infinite number of different coding/modulation schemes for each SNR value. In practice, the number of implemented coding/modulation schemes can be smaller than the number of subcarriers in the OFDM system. From the complexity point of view one should implement as few encoders/decoders as possible. It seems that using only one mother code with different puncturing patterns and a limited set of modulations gives a reasonable number of possible coding/modulation schemes.

The possible realization of the encoding and decoding sections of the transmitter and receiver, with a single mother code, is shown in Figure 2.2 [DPW04].

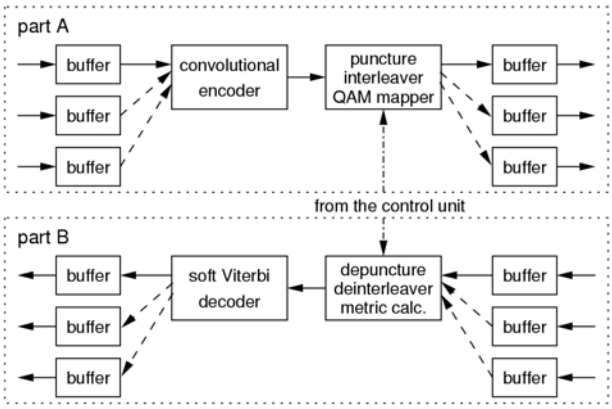


Figure 2.2: Encoding and decoding parts of the proposed adaptive system with only one mother code

One of the distinguishing features of OFDM with N subcarriers is that it decomposes a frequency-selective channel into N parallel flat subchannels:

$$y_n(k) = H_n \cdot x_n(k) + w_n(k) \quad (n = 1,\ldots,N) \tag{2.1}$$

where $x_n(k)$ is the signal on subchannel n before the IFFT in the transmitter in Figure 2.1, H_n is the frequency-domain channel coefficient and $w_n(k)$ is AWGN. The signal and noise powers are defined by:

$$\begin{aligned} p_n &= \mathrm{E}[|\, x_n(k)\,|^2] \\ \sigma_n^2 &= \mathrm{E}[|\, w_n(k)\,|^2] \end{aligned} \tag{2.2}$$

In general, the noise power can be different on each subcarrier, although in the following we will mostly assume that the noise is white over the whole transmission bandwidth, *i.e.* $\sigma_n^2 = N_0$ for all $n = 1,\ldots,N$. For the following algorithms, it is sufficient to describe the channel by its *channel gain to noise ratio* (CNR):

$$T_n = \frac{|\, H_n\,|^2}{\Gamma \cdot \sigma_n^2}, \tag{2.3}$$

where the SNR gap Γ is initially set to one. The capacity of the channel described by Eq. (2.3) is given by:

$$C = \sum_{n=1}^{N} \log_2\left(1 + p_n T_n\right) \tag{2.4}$$

This capacity is achieved by allocating the power on the subchannels according to the well-known water-filling theorem [Gal68], [CT91]:

$$p_n = \left[\, p_0 - \frac{1}{T_n} \right]^+, \quad p_0 \text{ such that } \sum_{n=1}^{N} p_n = p_{\max} \tag{2.5}$$

2.1.1 Code, modulation and power loading algorithm description

The water-filling solution from Eq. (2.5) gives the optimum power allocation in terms of channel capacity, but it does not give details about the modulation and coding formats. In order to make use of this result for (coded) QAM modulation, we can introduce the SNR gap Γ, which provides the link between the information-theoretic channel capacity and the achievable bit rate in QAM modulation. For uncoded QAM with symbol error probability $P_{\mathrm{e,S}}$, the SNR gap is given by:

$$\Gamma = \frac{1}{3}\left(Q^{-1}\left(\frac{P_{e,S}}{4} \right) \right)^2 \tag{2.6}$$

Let us denote by b_n the number of bits per QAM symbol and by B the set of possible values for b_n, i.e. $b_n \in \mathrm{B}$. In other words, if the possible modulation schemes include BPSK, 4-QAM, 16-QAM and 64-QAM, this set is given as $\mathsf{B} = \{1, 2, 4, 6\}$.

The two fundamental optimisation criteria are:

1) Maximum bit rate:

$$\begin{array}{ll} \max & \sum_n b_n \\ \text{subject to} & \sum_n p_n \le p_{\max},\ P_{\mathrm{e,b}} \le P_{\mathrm{e,max}} \end{array} \tag{2.7}$$

2) Minimum transmit power:

$$\begin{array}{ll} \min & \sum_n p_n \\ \text{subject to} & \sum_n b_n \ge b_{\min},\ P_{\mathrm{e,b}} \le P_{\mathrm{e,max}} \end{array} \tag{2.8}$$

An algorithm, which can be applied for both criteria, is the *Hughes-Hartogs* (HH) algorithm. We denote by c the index to the set B, such that B(c) is the c-th element of B and $c_{\max} = |\mathrm{B}|$ is the index for the highest number of bits per QAM symbol. The HH algorithm is located in the control unit and calculates the code rate and power allocation for the OFDM transmitter in Figure 2.1. It has been developed originally for uncoded QAM, but it can be generalized to coded QAM. One of the most important steps in this algorithm is the calculation of the "incremental power", which is the power required to transmit additional bits at the assumed BER. For uncoded QAM, this number of bits per QAM symbol is always integer and the incremental power can be calculated analytically.

The steps of the HH algorithm might be written as follows:

$p_{\mathrm{tot}} = 0, b_n = c_n = 0 \quad (\forall n = 1, \ldots, N)$
$\Delta p_n = f(0, T_n) \quad \forall n$
while $\sum_n b_n \le b_{\min} \quad \wedge \quad p_{\mathrm{tot}} \le p_{\max}$
  $i = \arg\min_n(\Delta p_n), \quad p_{\mathrm{tot}} = p_{\mathrm{tot}} + \Delta p_i$
  if $p_{\mathrm{tot}} < p_{\max}$
    $c_i = c_i + 1, \quad b_i = \mathrm{B}(c_i)$
    if $c_i < c_{\max}$
      $\Delta p_i = f(c_i, T_i)$
    else
      $\Delta p_i = \infty$
    end if
  end if
end

The function $f(c,T_n)$ calculates the incremental power that is required to increase the rate from B(c) to B(c+1) for a subchannel with CNR T_n. For uncoded QAM, this is equivalent to

$$f(c,T) = \frac{1}{T}\left(2^{\mathrm{B}(c+1)} - 2^{\mathrm{B}(c)}\right) \tag{2.9}$$

In case of coded transmission, the number of information bits per QAM symbol (which is equivalent to the overall code rate per subchannel) can be fractional and is given as the product of the coded bits per symbol and the code rate of the outer code. Hence, for coded QAM, the values for $f(c,T_{\mathrm{n}})$ can be stored in a look-up table, which is derived by simulations with the considered *code & modulation* (CM) schemes as depicted in Figure 6.1.

The above algorithm works in the same way for *coded OFDM* (COFDM). The main difference is that the set B contains non-integer *code rates* (the number of information bits per QAM symbol) instead of just the number of bits per symbol (see Section 6.1.1 in the Appendix).

The HH algorithm is known to provide the best solution; however, it is not the most efficient from an implementation point of view. This issue has been addressed by various authors who have presented suboptimum, computationally efficient solutions, see e.g. [Cam99], [CCB95], [Czy96], [FH96], [KRJ00],

[LCL+99], [Son02], [SS00], [YC01]. For the purposes of this deliverable, which focuses on multi-user adaptation schemes, suboptimum single-link adaptation will not be considered further. This is done in order to ensure better comparability in the multi-user case and to prevent degradation effects that stem from simplifications in the single-link adaptation stage.

2.1.2 Space-time block coding and MIMO extension of the presented adaptive approach

For the purposes of improving the overall performance of the presented adaptive scheme, multiple antennas at both the transmitter and receiver sides of a wireless link are considered. Such approach allows to exploit an additional gain from transmit and receive diversity by means of *multiple input multiple output* (MIMO) radio channel orthogonalisation. Moreover, in order to obtain the diversity gain, the space-time block coding (STBC) is proposed in concatenation with an outer convolutional code [PWW05].

The modified scheme of such system is presented in Figure 2.3.

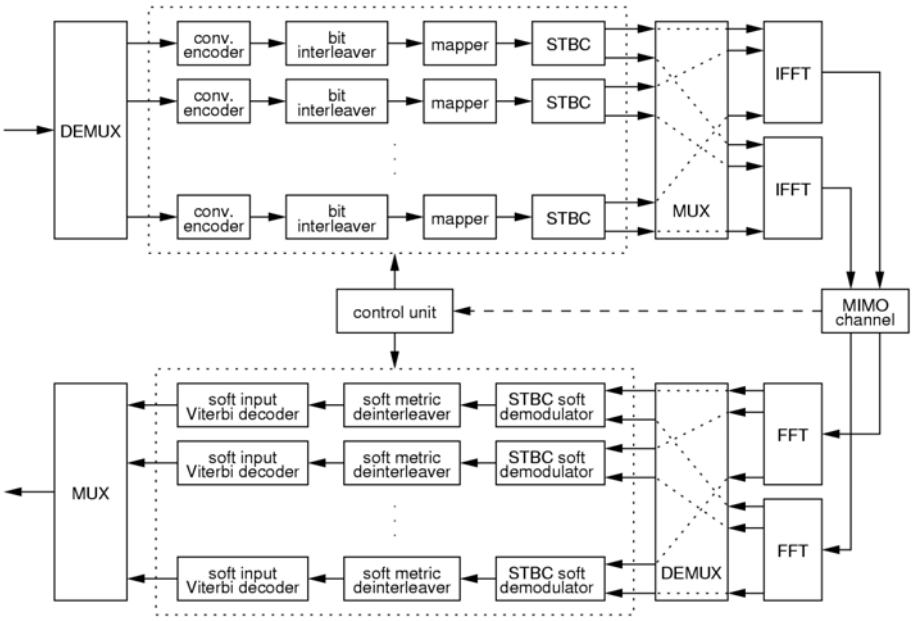


Figure 2.3: Adaptive multi-coded OFDM system with STBC for MIMO transmission.

The main advantage of the space-time block coding is that it utilises the spatial domain at the transmitter and, as a result, it suffices to increase the number of transmit antennas only in order to improve the reliability of the decisions at the receiver side. However, in order to obtain better results, multiple antenna utilisation at the receiver side might be taken into account as well.

In the proposed system a standard $\mathbf{G}_2$ space-time block code is used (see Eq. (2.20)), which was initially introduced by Alamouti [Ala98]. This code employs two transmit antennas and due to the fact that the resulting space-time coded sequence occupies two time slots only, the code rate is equal to one.

In fact, in case of the presented multi-coded OFDM system, the space-time block encoder should be perceived as a bank of $\mathbf{G}_2$ encoders (see Figure 2.3). The two output signals from each such $\mathbf{G}_2$ encoder are transmitted on a particular subcarrier by distinct antennas. Let us assume that the system is equipped with two transmit and N receive antennas. Then, the signal vector $\mathbf{r}_j$, received by antenna j in two consecutive time slots, can be written as:

$$\mathbf{r}_j = \mathbf{H}_j \mathbf{c} + \mathbf{v} \tag{2.10}$$

where: $\mathbf{r}_j = [r_{1,j} \quad r_{2,j}^*]^T$ are the signals received in the first and second time slots respectively, $\mathbf{c}_j = [c_1 \quad c_2]^T$ contains the encoded sequence and $\mathbf{v}_j = [v_{1,j} \quad v_{2,j}^*]^T$ is the noise vector, where $v_{i,j}$ represents the additive white Gaussian noise with zero mean and power spectral density $N_0/2$ W/Hz. The channel matrix is then defined as [ACN02]:

$$\mathbf{H}_j = \begin{bmatrix} h_{1,j} & h_{2,j} \\ -h_{2,j}^* & h_{1,j}^* \end{bmatrix} \tag{2.11}$$

where $h_{i,j}$ is the channel coefficient between the transmit antenna i and receive antenna j. In this situation the ML decision rules for decoding c_1 and c_2 are [PWW05]:

$$dec(c_1) = \min_{c_1} \left\| \sum_{j=1}^{N} (h_{1,j}^* r_{1,j} + h_{2,j} r_{2,j}^*) - c_1 \sum_{i=1}^{2} \sum_{j=1}^{N} \left| h_{i,j} \right|^2 \right\| \tag{2.12}$$

$$dec(c_2) = \min_{c_2} \left\| \sum_{j=1}^{N} (h_{2,j}^* r_{1,j} - h_{1,j} r_{2,j}^*) - c_2 \sum_{i=1}^{2} \sum_{j=1}^{N} \left| h_{i,j} \right|^2 \right\| \tag{2.13}$$

2.1.3 Performance evaluation

The simulation results presented in Sections 2.1 and 6.1 have been obtained with the use of the following channel models:

- WP5 Urban Macro channel model (wide area scenario) [BGS+04],
- IEEE 802.11n case 'C', non line-of-sight (NLOS) channel model (short range scenario) [IEEE04],

The first one has been used for *single input single output* (SISO) simulations only, whereas the latter one has been employed in both the SISO and MIMO evaluations. Both channel models are quasi-stationary. It means that the channel coefficients do not change during one OFDM symbol (or two for MIMO transmission). In case of MIMO simulations, the transmission between each pair of transmitter and receiver antennas is assumed as uncorrelated. Therefore, independent channel coefficients are generated for each link.

Channel coding is performed by one mother convolutional code $(177,133)_{oct}$, with the constraint length $L = 7$. The other code rates: 2/3, 3/4, 5/6 and 7/8, are achieved by puncturing. Each block of information bits is extended with six '0' tail bits. The decoder uses the standard soft input Viterbi algorithm. The following four modulation schemes have been used: BPSK, 4-QAM, 16-QAM and 64-QAM.

In each OFDM symbol maximum 832 (wide area scenario) or 1664 subcarriers (short range scenario) can be allocated for data transmission. The FFT sizes of the OFDM modulator are set to 1024 or 2048 respectively. The rest of the subcarriers are switched off and used as guard bands.

The results include the bit error rate and number of information bits per constellation symbol (overall code rate r_n) vs. average SNR measured in the channel.

In each simulation case, at least 10 million information bits have been transmitted for a particular SNR value. This provides sufficient precision in terms of bit error rates at the level of 10^{-5}.

2.1.3.1 Impact of the codeword length on general system performance

One of the drawbacks of the joint code and modulation adaptation is the fact that the resulting BER does not only depend on the average SNR value, but also on the length of encoded block. This is true, especially, when bit interleaving is applied for each block independently. In case of a very short encoded block, the actual BER is then worse than expected. This implies additional constraints to the loading algorithm, because it employs look-up tables with the SNR values taken from experimentally generated bit error rate curves (see an example of reference BER curves in Figure 6.1).

Figure 2.4 shows how big are the differences in coding gain depending on the length of a coded block.

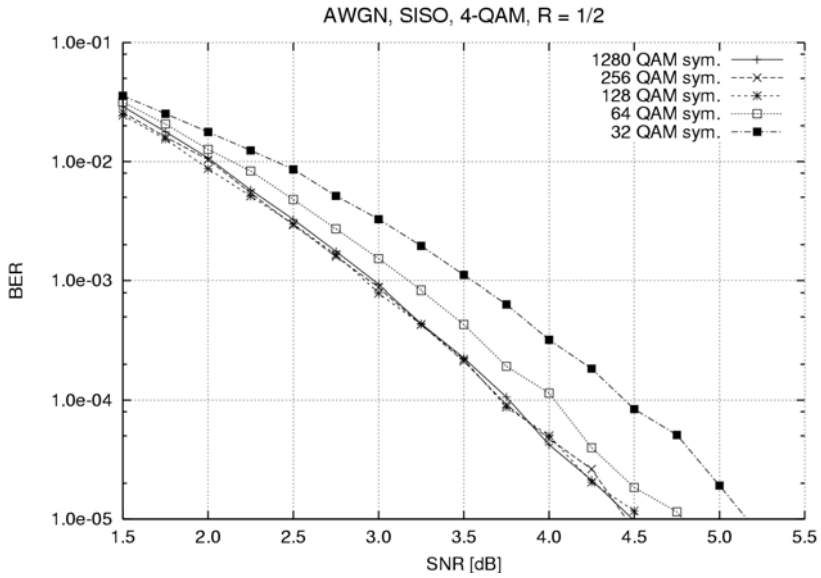


Figure 2.4: Impact of the codeword length on the bit error rate curves in case of SISO transmission in the AWGN channel (R = 1/2, 4-QAM)

One can observe that in order to achieve comparable bit error rate results for 4-QAM modulation and code rate $R = ½$, the resulting codeword size should have at least 64 constellation symbols, i.e. 128 encoded bits. This restriction is even stronger for higher constellation sizes. For instance, in case of 16-QAM modulation and code rate $R = 3/4$ the difference in coding gain depending on the codeword length is almost 1.7 dB at the BER level of 10^{-4} (see Figure 6.2 in the Appendix, Section 6.1). For such code and modulation one should consider the use of at least 128 or even 256 QAM symbols, i.e. 1024 coded bits, as the minimum size of encoded block. That is why it was assumed that the shortest codeword should be at least equal to the length of 64 constellation symbols, i.e. 64 bits for BPSK, 128 bits for 4-QAM, 256 bits for 16-QAM and 384 bits for 64-QAM respectively.

Concluding, it makes no sense using shorter codewords than 64 QAM symbols for two reasons:

1. The loss in coding gain will certainly exceed the improvement in the overall code rate from the adaptivity itself.
2. It makes no sense to transmit shorter packets than a few hundred information bits.

However, it is important to remember that the longer the minimum codeword length is, the less optimum adaptation one can achieve.

2.1.3.2 Selection of a number of different coding schemes used in adaptation unit

It is stated in [DPW02] that the more different code and modulation combinations are used, the more optimum adaptation such approach can provide. However, this thesis is only true if there is no power loading at the transmitter. The optimum algorithm described in Section 2.1.1 employs power loading and due to this fact it is possible to reduce the number of different coding and modulation schemes used by the control unit.

Figure 2.5 shows a comparison of the overall code rate results obtained for various numbers of code rates exploited by the adaptation algorithm.

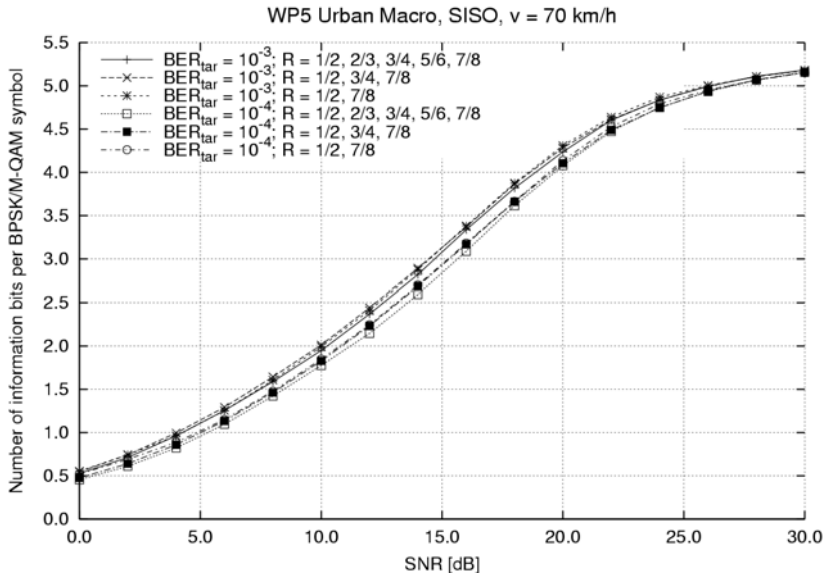


Figure 2.5: Impact of the number of coding schemes used in adaptation algorithm on the overall code rate for the WP5 Urban Macro SISO channel model (*v* = 70 km/h, perfect CSI, no feedback channel delay)

The results have been obtained in the WP5 Urban Macro channel model for the velocity equal to 70 km/h and for SISO case. The target bit error rate has been set to 10^{-3} and 10^{-4} accordingly. There are three types of curves depending on the number of code rates used:

- for all possible code rates (solid lines with '+' or '□'),
- for rates 1/2, 3/4 and 7/8 (dashed lines with 'x' or '■'),
- and finally, for R = 1/2 and R = 7/8 (dash-dot lines with '*' or 'o').

The differences in the number of bits per BPSK/M-QAM symbol is negligible for all three cases. In fact, the results for all possible rates are slightly worst than for the reduced cases. This is due to the restriction on the minimum block length, described in the previous Section. The less coding schemes are chosen by the adaptive algorithm, the longer coded words are. Therefore, it is more probable that the minimum block length constraint is fulfilled. Otherwise, the adaptation algorithm has to regroup subcarriers allocated with a certain CM scheme. In practice it means that the overall BER is improved and as a result the overall rate is slightly degraded.

The corresponding bit error rate results are presented in Figure 2.6.

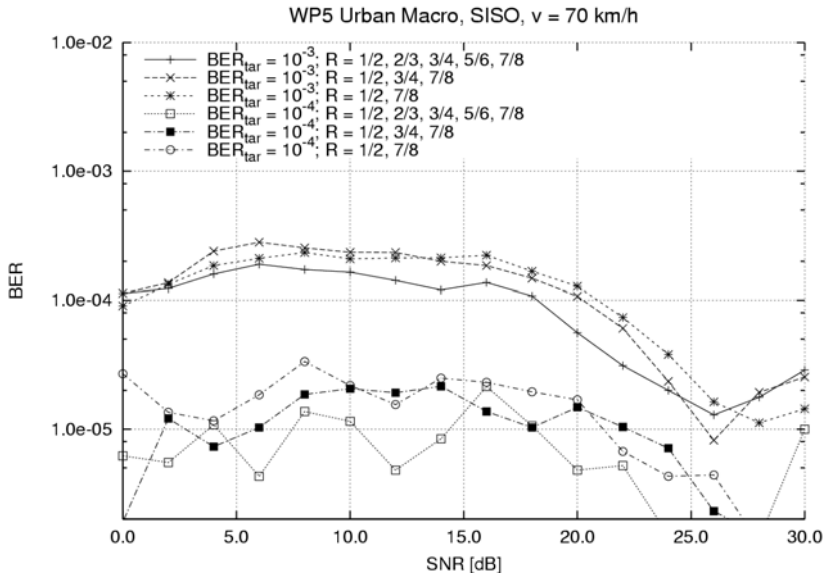


Figure 2.6: Impact of the number of coding schemes used in adaptation algorithm on the bit error rate for the WP5 Urban Macro SISO channel model (*v* = 70 km/h, perfect CSI, no feedback channel delay)

For both target BER values (10^{-3} and 10^{-4}) the output bit error rates are more or less comparable. It means that the power loading algorithm compensates for the lack of intermediate code and modulation schemes.

Analogous results have been obtained for IEEE 802.11n channel model (case 'C', NLOS) for the velocity equal to 10 km/h. The corresponding plots can be found in the Appendix in Figure 6.3 and Figure 6.4.

Summarising, one can assume that it is sufficient to employ only the maximum and minimum code rates combined with all considered constellation schemes. This considerably reduces the amount of information that has to be prepared and stored in the form of look-up tables in the control unit [PWW05].

2.1.3.3 Imperfect channel state information

An important issue in an adaptive transmission is the impact of imperfect *channel state information* (CSI) on the general system performance. It is assumed that the channel estimation error $\varepsilon_i = H_i - \hat{H}_i$ ($i = 1, ..., N$) is a Gaussian random variable with variance equal to the *mean square error* (MSE) of the estimator, and decision feedback errors are neglected. This assumption is valid for pilot aided channel estimation schemes [WIND21].

Thus, the effective model of the received OFDM signal, which include the MSE is as follows:

$$Y_i = H_i X_i + \eta_i \tag{2.14}$$

where η_i represents the noise term with variance $\sigma_\eta^2 = N_0 + \mathrm{MSE} \cdot E_s$ and zero mean, assuming a good estimator with a small bias.

In simulations, the MSE has been set to N_0/E_s, which is the worst case value for the *least squares* (LS) estimation, minus arbitrary chosen additional estimator gain $G_N = 6$ dB. The possible impact of an interpolation error on the MSE estimation curve has been neglected (see [WIND21]).

Figure 2.7 shows an impact of various estimation error models on the bit error rate for the proposed adaptive approach in the IEEE 802.11n ('C', NLOS) channel model.

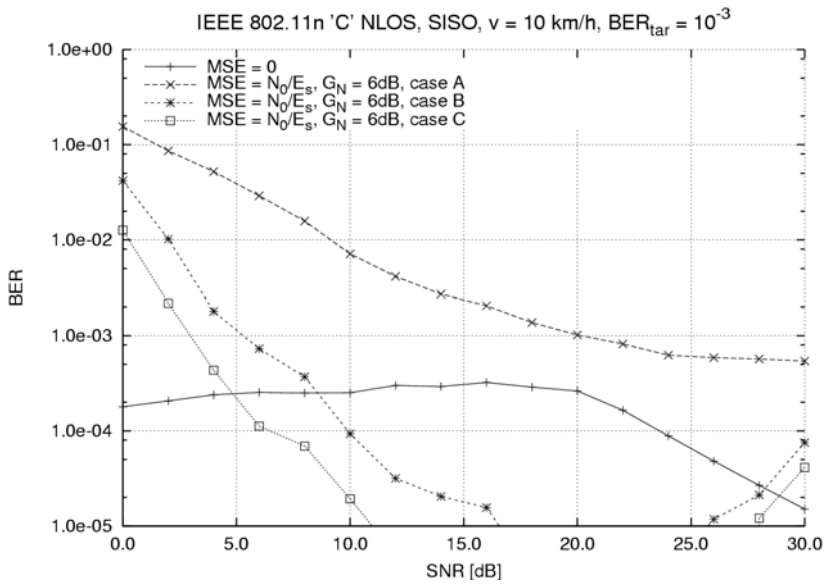


Figure 2.7: Impact of the estimation error on the bit error rate for the IEEE 802.11n 'C' NLOS SISO channel model (v = 10 km/h, $BER_{tar} = 10^{-3}$, no feedback channel delay)

The target BER is equal to 10^{-3} and the velocity is 10 km/h. Apart from the result for a perfect estimation (MSE = 0) there are plots for three different cases:

- case A includes the MSE = N_0/E_s and estimator gain G_N = 6 dB, but the BER curves that are used for adaptation (see Figure 6.1) has been generated in the AWGN channel only,
- case B includes the MSE = N_0/E_s and G_N = 6 dB, however this time the reference BER curves have been generated using this particular MSE value improved by the gain G_N,
- and finally case C uses the same values of MSE and G_N, but the reference curves from case A have been shifted by 3 dB, which represents the worst-case LS estimation with no gain (see [WIND21]).

One can notice in Figure 2.7 that the performance of case A is not acceptable, because it does not ensure the assumed target BER level for wide range of SNR values (0 - 20 dB). On the other hand, in cases B and C a huge gap between the assumed and actual BER in the middle of the presented SNR range can be observed. Moreover, the estimation error causes that the transmission at the assumed BER level is not possible for very low SNR values (< 5 dB). In cases B and C, the higher bit error level for high SNR values (25 - 30 dB) might be explained as a saturation effect. For such good channel condition, most of the subcarriers are loaded with maximum code rate and modulation available (7/8 rate and 64-QAM). So, the encoded block lengths fulfil the restriction described in Section 2.1.3.1, and the loading algorithm operates optimally. Whereas, for middle SNR values (10 –25 dB) the suboptimum rearrangement of CM schemes due to too short allocated block lengths causes such high improvement in terms of the BER.

The loss in the overall code rate due to the estimation error impact, for the same channel model and parameters settings, is presented in Figure 2.8.

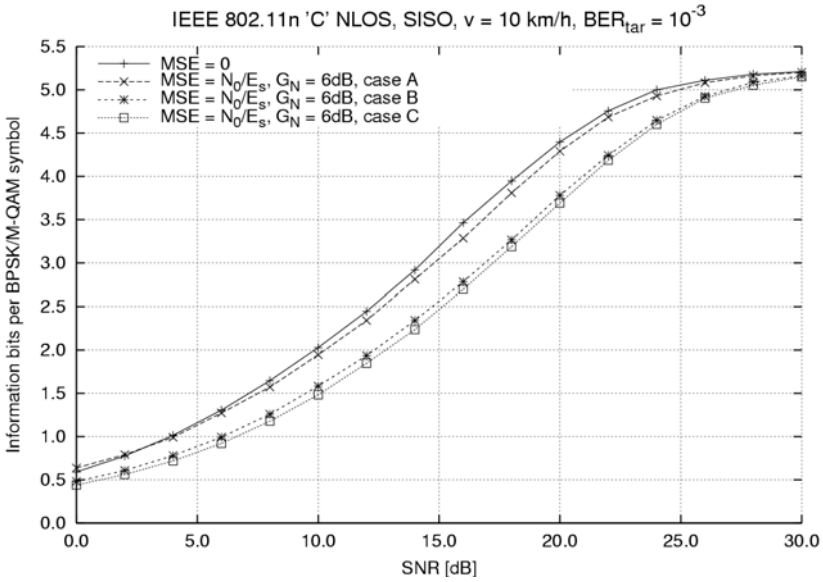


Figure 2.8: Impact of the estimation error on the overall code rate for the IEEE 802.11n 'C' NLOS SISO channel model (v = 10 km/h, $BER_{tar} = 10^{-3}$, no feedback channel delay)

The difference between a perfect CSI case and case B or C is about 0.2 to 0.8 information bits per BPSK/M-QAM symbol, depending on the average SNR value. The highest loss is correlated to the highest gap between the actual and target BER level, shown in Figure 2.7.

Simulation results of the estimation error impact on adaptive transmission for the MIMO case can be found in the Appendix, Section 6.1 (Figure 6.7 and Figure 6.8).

2.1.3.4 Impact of the feedback channel delay

The delay of a feedback channel might cause a serious problem from the adaptation point of view. When the channel is fast varying in the time domain, it is important that the CSI or the selected code, modulation and power values are sent back to the transmitter in the period shorter than the coherence time of the channel. That is why it is important to measure the maximum delay of the feedback channel for which the performance of the adaptation is still satisfactory.

Figure 2.9 and Figure 2.10 present the impact of the feedback delay on the bit error rate and overall code rate appropriately.

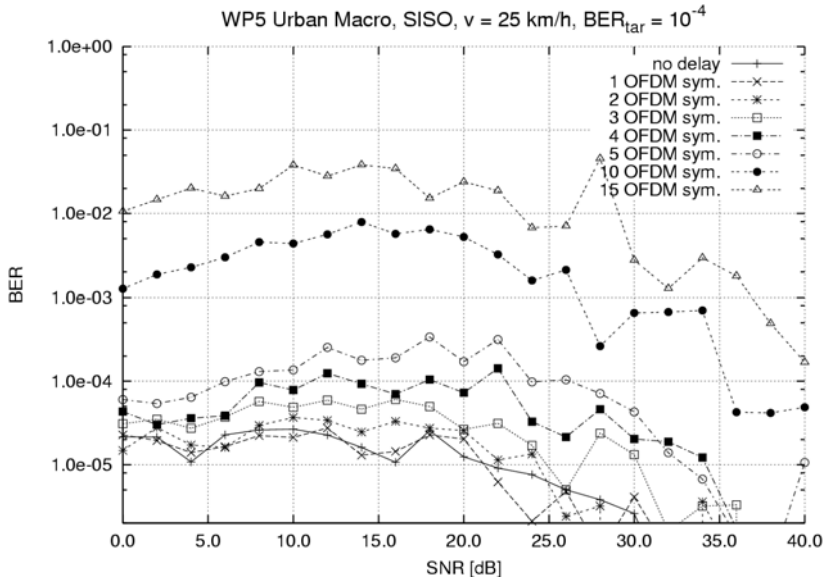


Figure 2.9: Bit error rate degradation for various delay of the feedback channel for the WP5 Urban Macro SISO channel model (v = 25 km/h, BER$_{tar}$ = 10^{-4}, perfect CSI)

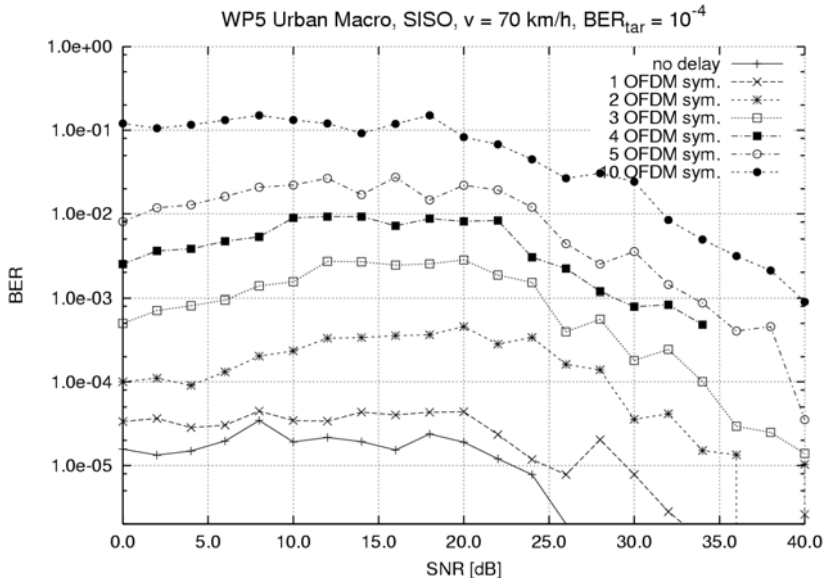


Figure 2.10: Bit error rate degradation for various delay of the feedback channel for the WP5 Urban Macro SISO channel model (v = 70 km/h, BER$_{tar}$ = 10^{-4}, perfect CSI at Rx)

The above results have been obtained in the WP5 Urban Macro channel model for the velocities equal to 25 km/h and 70 km/h, respectively, and the target BER level set to 10^{-4}. Moreover, a perfect channel state information at the receiver has been assumed, since the delay has been applied for the feedback signalling only.

According to [Rap96] the coherence time is defined as follows:

$$T_c = \sqrt{\frac{9}{16\pi f_{D_{\max}}}} \approx \frac{0.423}{f_{D_{\max}}} \tag{2.15}$$

where $f_{D_{\max}}$ is the maximum Doppler frequency. Assuming that the carrier frequency is equal to 5 GHz, the maximum Doppler frequency is about 116 Hz for $v = 25$ km/h and 324 Hz for $v = 70$ km/h. The coherence time is then equal to $T_c \approx 3.647$ ms (64 OFDM symbols) in the first case, and $T_c \approx 1306$ μs (23 OFDM symbols) in the latter one.

However, in practice, the maximum delay, for which the assumed BER is attained, is commensurate with the duration of 5 OFDM symbols (5*56,2 μs) for the lower mobility and only 2 OFDM symbols (2*56,2 μs) for v = 70 km/h. Longer delay results in the bit error rate much higher than expected 10^{-4} (see Figure 2.9 and Figure 2.10).

Corresponding plots with the overall rate results for the above presented velocities and the same channel model are included in the Appendix (see Figure 6.5 and Figure 6.6). No degradation of the overall rate occurs due to the delay impact.

The obtained bit error rate results seems to be quite pessimistic. However, one can reduce this crucial impact of the delay on the transmission in two ways:

- One might try to shift the BER reference curves by adding some SNR margin to them. This solution, in fact, has to be applied in order to mitigate the estimation error impact, as described in the previous Subsection.
- Another solution is the use of an additional channel prediction scheme. Such approach will be briefly described in Section 2.2.

2.1.3.5 Space-Time Block Coding and MIMO performance

In Figure 2.11, the performance of the extended adaptive OFDM system with STBC and MIMO (see Section 2.1.2) is compared with a few results obtained for a traditional coded OFDM transmission.

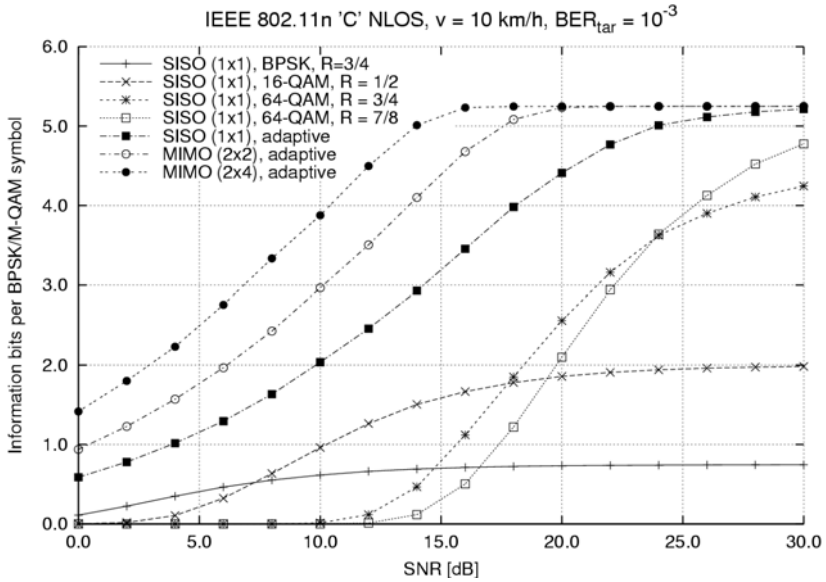


Figure 2.11: Overall code rate results of the adaptive approach as compared with a traditional OFDM transmission for the IEEE 802.11n 'C' NLOS SISO and MIMO channel models (v = 10 km/h, BER$_{tar}$ = 10^{-3}, perfect CSI, no feedback channel delay)

The 'traditional COFDM' means here that there is only one fixed code and modulation combination used for the transmission of a given frame. The simulations have been performed for both SISO and MIMO cases (1x1, 2x2 and 2x4 antennas) of the IEEE 802.11n channel model, with the velocity equal to 10 km/h. The target BER level has been set to 10^{-4}.

It is noticeable, how much flexibility and gain in the overall rate one can achieve using the adaptive code and modulation approach with power loading, as compared to the traditional OFDM systems (see Figure 2.11). For the latter ones, the transmission is possible either at high overall rate for good SNR conditions or for lower SNR values, but with much more reduced overall code rate. Moreover, it is possible to improve the robustness of the transmission by a few dB (approximately 3 dB gain for doubled number of antennas at the receiver), by using this extended approach with space-time coding and MIMO reception [PWW05].

In the Appendix, Section 6.1, one can find two more simulation results of this MIMO extension including the impact of an estimation error on transmission performance (see Figure 6.7 and Figure 6.8).

2.1.4 Conclusions

The presented adaptive multi-coded OFDM transmission approach has its roots it the well known classical Hughes-Hartogs bit and power loading algorithm [Bin90]. It was shown that it is possible to do the adaptation on both the code and modulation jointly with power loading in a real wireless environment. Such approach results in much higher flexibility of the multicarrier transmission, as compared to the conventional COFDM systems in which fixed CM schemes are applied.

Furthermore, it was shown that it is possible to extend such adaptive technique with space-time processing in the form of Alamouti space-time block codes at the transmitter and combining the signals from multiple antennas at the receiver. This results in a quite high improvement of a general robustness and overall rate of the transmission, as compared to the SISO case.

A few detailed aspects of the adaptive transmission have been evaluated in the previous Sections, i.e.:

- sensitivity to the estimation errors,
- impact of the delay of the feedback channel on the overall performance,
- problem of the number of code and modulation schemes needed, and finally,
- dependency of the reference BER curves on the length of encoded blocks.

Summarising, the proposed approach is quite sensitive to the estimation errors, which is really important for a wireless environment. Also, the feedback channel delay seriously restricts this adaptation method for the considered channel environments. However, it is possible to make the system robust to such errors by adding some SNR margin in the adaptation algorithm and use some advance channel prediction techniques (see Section 3.1).

Moreover, it was indicated that the performance of the loading algorithm is not degraded after reducing the number of code and modulation schemes used for the adaptation. One might employ only the maximum and minimum code rates combined with all considered constellation schemes and should not observe any deterioration in overall code rate and bit error rate, by no means.

Finally, it was concluded that the impact of variable lengths of encoded blocks might not be a crucial problem in case of the use of convolutional codes, because it can be diminished by constraining the minimum block size to 64 QAM symbols, for example.

The adaptive multi-coded OFDM system complexity is not increased substantially, as compared to traditional OFDM systems, because only one mother code with two puncturing schemes needs to be implemented. Therefore, the increase in the complexity might be basically caused by the implementation of the loading algorithm.

2.2 Adaptive coding and modulation based on channel prediction

In Section 3.1, a system will be outlined in which a number of active user terminals within a cell utilize adaptive TDMA/OFDMA transmission. Each involved data stream is in competition for all time-frequency chunks within the contention band, a subset of the total bandwidth. To allocate the time-frequency chunks to users, the signal to interference and noise ratio (SINR) within *all* chunks has to be monitored for *all* users. The resources are then allocated to terminals (and flows within terminals) according to a criterion that takes the estimated SINR into account.

Since channel estimation, signalling and resource scheduling invariably involves delays, the resource allocation will be based on outdated information. As became evident in the last Section, this is acceptable for terminals moving at pedestrian velocities, but it will have severely adverse effects at vehicular velocities. By predicting the channels instead of just extrapolating current estimates, the performance of the adaptation can be increased significantly. It will be demonstrated in Section 3.1 that for a TDMA/OFDMA system designed in accordance with the WINNER first RI concept, predictive adaptive transmission can be performed for terminals moving at vehicular velocities, if the velocity is below highway speed. This means that adaptive transmission would be feasible for a large majority of terminals in the wide-area narrowband FDD mode as well as in short-range wideband TDD modes.

However, prediction of the channel- and interference power will not be perfect. Prediction does not remove feasibility limits for adaptation; it just expands them. Predictive adaptive transmission must therefore in practice be combined with a link retransmission mechanism that acts as a safety net. With the required short delays of the Winner air interface (delay over air interface targeted at 1 ms, with total RI round trip time below 2 ms, see [WIND71]), link retransmission can be designed as an efficient safety net also for delay sensitive streams.

To work effectively, channel prediction, the rate adaptation and link retransmission should be designed jointly. The present Section outlines a strategy for such a joint design: The CM rate limits are adjusted based on the prediction uncertainty, so that an appropriate bit error rate and packet error rate is attained. We then avoid situations where the prediction errors result in packet error rates close to 100%, that would lead to a low throughput and a large total transmission delays also with a low link delay. Furthermore, unnecessarily conservative designs that under-utilize the link retransmission, can be avoided when these design tools are available.

2.2.1 Taking channel state information uncertainty into account in link adaptation

In the following, we focus on the error in the prediction of the received signal power, while the noise and interference is modelled by additive complex Gaussian variables with zero mean and a known or assumed variance.[3] In the considered TDMA/OFDMA system, the time-frequency chunks contain a number of adjacent subcarriers and a number of subsequent OFDM symbols, see Section 1.1. The chunks are to be allocated to different users, as described in more detail in Section 3.1. Within each chunk, the frequency domain channel is described as a scalar complex gain. The received power within each chunk is predicted based on earlier measurements of the fading channel within these chunks and adjacent chunks.

Assume that the complex channel gain $h(t)$ is predicted at time t-L by a linear MMSE estimator which is designed to minimize the prediction MSE

$$\sigma_{\varepsilon_h}^2 = E\,|\,h(t) - \hat{h}(t\,|\,t-L)\,|^2 \tag{2.16}$$

[3] The performance and BER figures are thus averages over disturbance and noise properties that are assumed unpredictable. This is a conservative assumption. In a Winner RI, the interference is likely to show significant correlation in time and frequency, and should therefore be partly predictable. However, the predictability, and the burstiness, of the interference will depend on the detailed RI design and on the deployment scenario, for which good assumptions do not exist at the moment. We will not pursue the topic of interference prediction here, but instead concentrate on the channel prediction. When better models for the interference properties become available, the residual signals in the Kalman channel estimators discussed in Section 3.1 could be used as estimates of the interference and noise. Time-series of such interference and noise estimates could then be combined with prior information about the interference properties, to construct predictors also for the interference.

where L is the prediction horizon. The received channel power $p(t)= |h(t)|^2$ is assumed to be predicted by an unbiased quadratic power predictor [ESA02], [Ekm02]. For a given prediction of the complex channel gain, this prediction is obtained by the squared magnitude of the predicted channel gain, adjusted by adding the prediction MSE:

$$\hat{p}(t \mid t-L) = \mid \hat{h}(t \mid t-L) \mid^2 + \sigma^2_{\varepsilon_h} . \tag{2.17}$$

If $\hat{h}(t \mid t-L)$ is the MMSE estimate of the channel gain, then this unbiased quadratic predictor attains the minimal power prediction MSE among all quadratic functions of the channel gain prediction, or of the utilized regression variables [ESA02], [Ekm02]. For this and other related predictors, the statistical properties such as conditional pdf's are functions parameterized by the normalized prediction MSE

$$\tilde{\sigma}^2 = \sigma^2_{\varepsilon_h} / \sigma^2_h \tag{2.18}$$

where $\sigma^2_h = E \mid h(t) \mid^2$. See e.g. Eq. (16) and (17) in [FSE+04], derived in Chapter 8 of [Ekm02]. The normalized prediction MSE $\tilde{\sigma}^2$ will depend on the structure and implementation of the predictor $\hat{h}(t \mid t-L)$, on the SINR of the signal used for prediction and also on the required prediction horizon L. Performance results for frequency-domain Kalman prediction for particular structures of pilot symbols within chunks will be presented in Subsection 3.1.3.

What is important for the moment is that if $\tilde{\sigma}^2$ is assumed known, any selected code and modulation adaptation scheme can be adjusted to *attain specific design targets that will be fulfilled at the specified level of prediction uncertainty*. The optimisation provides ***rate limits*** (boundary values) for the predicted SNR that specify where to switch a CM scheme. Two main design principles may be distinguished.

1. Generalizing the approach of Andrea Goldsmith and co-workers [GC97], [CG01] we may adjust the rate limits, and perhaps also the transmit power in-between rate limits, to maximize the spectral efficiency. This can be done under a maximum or average BER constraint, and under various types of transmit power constraints. With uncertain predictions, the CM rate and transmit power is determined by the predicted SINR, which depends on the predicted power $\hat{p}(t \mid t-L)$. The performance is then evaluated and optimised by averaging (marginalizing) over the unknown true power $p(t)$, using the known conditional pdf $f(p(t) \mid \hat{p}(t \mid t-L))$. This method has been developed for uncoded adaptive M-QAM in [FSS+03a], [FSE+04]. It has also been applied to adaptive trellis-coded M-QAM in [FSS+03b].
2. Alternatively, we may consider the number of bits in correctly received packets. This has been used as criterion e.g. in the work on EDGE rate adaptation and in [WOS+03] for adaptive TDMA/OFDMA downlinks. The optimal CM rate is then determined by weighting the number of bits per transmission unit (chunk) by the probability of failed transmission. Also in this case, when introducing prediction uncertainty the CM rate is assumed determined from the predicted power $\hat{p}(t \mid t-L)$. The performance is evaluated by averaging over $p(t)$, using the conditional pdf $f(p(t) \mid \hat{p}(t \mid t-L))$. The described approach was applied to adaptive uncoded M-QAM in [SF04]. The bit error rate is then not targeted or constrained. It becomes automatically adjusted to whatever level provides the highest goodput (throughput within correctly received chunks).

Both of the described methods have been tested on the presently studied Winner RI with adaptive TDMA/OFDMA. While method 2 can attain a somewhat higher spectral efficiency, it was found that method 1 (BER constraints) provides a more precise control over the joint design of link adaptation schemes and link retransmission strategies. Therefore, the optimisation of the spectral efficiency that targets a specific BER constraint will be used and evaluated in the following.

To be specific, the rate limits are adjusted to maximise the spectral efficiency under the following constraints:

- A ***maximal BER constraint*** is used. This constraint will be fulfilled exactly at the selected code and modulation rate limits, while the BER will be lower in-between the limits.[4] The BER constraint is to be fulfilled when the scheme is applied on the unknown true channel. In other

[4] For that reason, the attained average BER is below the specified maximal BER limit. One could alternatively utilize an average BER constraint to target average BER directly, see [FSE+04].

words, when the predicted SINR is equal to a value that specifies a rate limit, the BER averaged over the unknown true SINR should equal the target value.

- A ***constant transmit power constraint*** is used.[5] (In uplinks, the transmit power would be determined by a slow power control algorithm that remains to be developed.)
- ***Modulation and coding is adjusted, while the transmit power is held constant in-between rate limits***. Compared to also fine-tuning the transmit power, as was done in the previous section, this results in a small performance decrease. However, we obtain the very important property that *link adaptation and multiuser resource scheduling can be solved in two subsequent simple steps*. That would *not* be the case if the transmit power of each user were introduced as an additional variable to be optimised, under a total power constraint.

2.2.2 Rate limits for adaptive M-QAM schemes adjusted to the prediction uncertainty

Optimisation for a target BER was performed for two link adaptation schemes, that will be used and evaluated in Section 3.1:

- ***Coded M-QAM***: A scheme using convolutional coding and M-QAM, with eight rates: BPSK rate ½, QPSK (4-QAM) rate ½, QPSK rate ¾, 16-QAM rate ½, 16-QAM rate 2/3, 16-QAM rate 5/6, 64-QAM rate 2/3 and 64-QAM rate 5/6. The mother convolutional code has rate ½, $(561,753)_{oct}$, with constraint length 9. Soft input Viterbi decoding is used.
- ***Uncoded M-QAM***: An uncoded scheme, using BPSK, QPSK, 16-QAM and 64-QAM.

The resulting rate limits are functions of the following three variables: target BER, prediction NMSE $\tilde{\sigma}^2$ and average SNR. The figures below show examples of results for coded M-QAM.

The prediction uncertainty turns out to have a rather small effect on the rate limits up to $\tilde{\sigma}^2 = 0.01$, while it becomes quite substantial above $\tilde{\sigma}^2 = 0.1$. The rate limits show an interesting behaviour. Limits that are active when predicting into a fading dip (when the predicted instantaneous SNR is lower than the average SNR) are raised markedly when the prediction uncertainty is increased. To preserve the ability to attain the BER constraint, the scheme becomes "afraid of fading dips", and uses more conservative code and modulation rates in such situations. The rate limits that would become active at fading peaks are affected much less. As explained in [FSE+04], the reason for this behaviour is that the *relative* prediction uncertainty is much higher during a fading dip than when the channel power is high. This has an interesting effect on the interplay with multiuser scheduling: If users are assigned by a proportional fair-like scheme, so they obtain access when having good channels relative to their own average, then they will fortunately also have relatively small prediction errors in such situations.

The relative performance deterioration due to prediction uncertainty can therefore be expected to decrease with an increasing number of adaptively scheduled streams within the cell.

Corresponding graphs for uncoded M-QAM are shown in Appendix 6.2, for an average of SNR 10 dB. The uncoded M-QAM provides lower capacity, as expected. The rate limits for comparable bit/symbol levels are located at higher instantaneous SNR. However, its sensitivity to the prediction uncertainty is smaller. The reason for this is that the BER-versus-SNR waterfall plots are steeper for coded modulation than for uncoded modulation. When averaging this function with respect to the unknown true SNR, a steeper plot will result in higher sensitivity to prediction errors.

[5]When optimising link adaptation schemes for single-user links, it has been more common to utilize an *average-power constraint*, which takes into account that no transmission occurs below an SINR cutoff rate, One then utilizes this to increase the power when transmission occurs [CG01]. However, this method cannot be used in our *multi-user* downlinks: When one user leaves the channel unoccupied, some other user typically will use it instead. In uplinks, an average power constraint could be used, but power control must there take other factors into account.

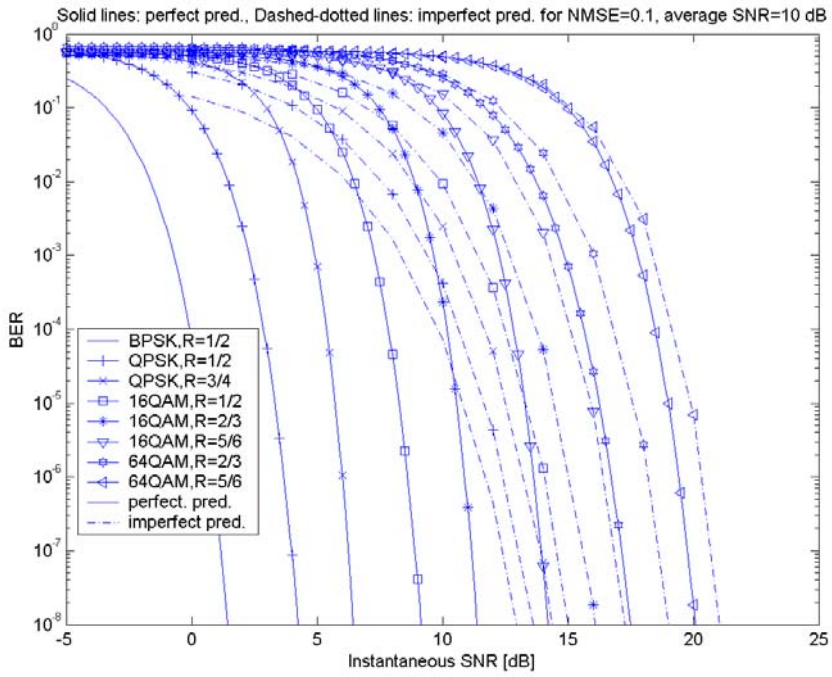


Figure 2.12: Bit error rates as functions of the instantaneous predicted SNR, for the eight selected combinations of M-QAM and convolutional coding. Solid lines: Results for perfect prediction. Dash-dotted lines: attained BER, averaged over the true SNR, when using a specific rate at constant transmit power. Average SNR 10 dB; prediction error NMSE $\tilde{\sigma}^2$ = 0.10

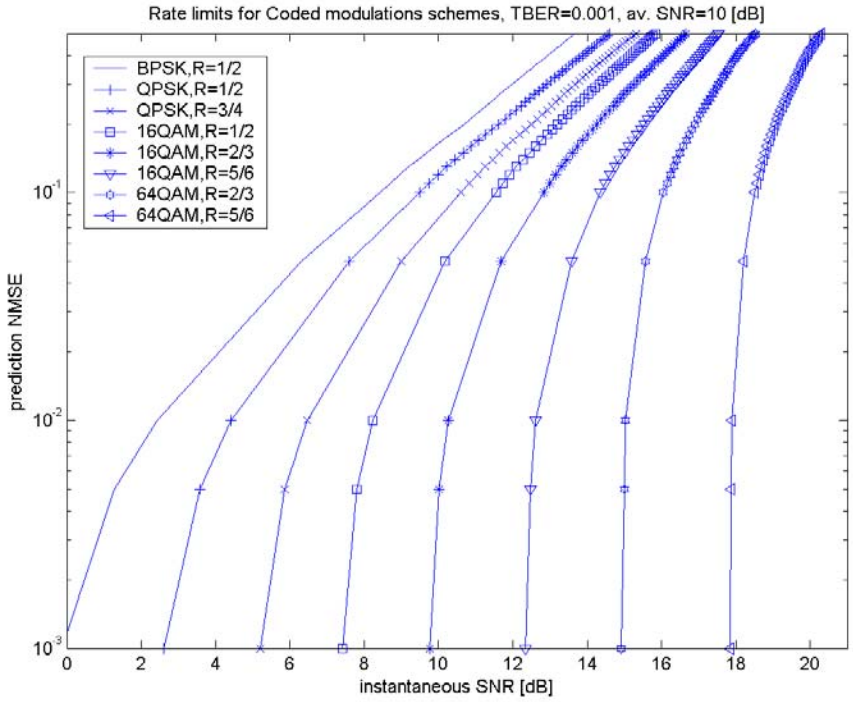


Figure 2.13: Rate limits for the predicted SNR per symbol for adapting coded M-QAM optimised for a target BER 0.001 and constant transmit power, shown as functions of the complex prediction error normalized MSE $\tilde{\sigma}^2$. Average SNR 10 dB

2.3 Adaptive approach to antenna selection and space-time block coding

Despite its name space-time coding may be generally perceived as a modulation rather than a coding technique. It was designed to provide an additional diversity gain [Ala98] in wireless systems exploiting multiple element antenna arrays for the purposes of enhancing the transmission reliability. Recently it has been noticed [Lan03] that space-time coding is also applicable to wireless ad-hoc networks, where mobile nodes may constitute virtual antenna arrays in order to improve the overall system performance. This issue is also known as the cooperative diversity. In the remainder of this chapter an adaptive approach to antenna selection and space-time block coding in context of the aforementioned virtual antenna arrays is investigated [Wod05].

2.3.1 Space-time block coding

A wireless system exploiting *N*-element transmit and *M*-element receive antenna arrays is presented in Figure 2.14.

Figure 2.14: A wireless MIMO system

The wireless MIMO channel is then defined by the channel matrix $\mathbf{H}_{NxM}$, containing the coefficients $h_{i,j}$, which correspond to the radio links between each transmit antenna i ($1 \le i \le N$) and each receive antenna j ($1 \le j \le M$):

$$\mathbf{H}_{NxM} = \begin{bmatrix} h_{1,1} & h_{1,2} & .. & h_{1,M} \\ h_{2,1} & h_{2,2} & \dots & h_{2,M} \\ \dots & \dots & \dots & \dots \\ h_{N,1} & h_{N,2} & \dots & h_{N,M} \end{bmatrix} \tag{2.19}$$

There are a few space-time processing techniques, which may be applicable to preprocessing of the transmitted signals so that they were more robust to the wireless radio channel impairments. Among others there is the aforementioned space-time block coding, which offers so called diversity gain but no coding gain. The simplest space-time block code $\mathbf{G}_2$ is defined with the use of the following matrix:

$$\mathbf{G}_2 = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix} \tag{2.20}$$

This code may be used in a system employing two transmit and any number of receive antennas. In the first timeslot the x_1 and x_2 symbols are sent by the first and second transmit antenna respectively and then, in the second timeslot, the $-x_2^*$ and x_1^* symbols are transmitted in the same manner. There are also other space-time block codes known in the literature [Lan03], which may be applicable when a larger transmit antenna array is used. Two of them, namely $\mathbf{G}_3$ and $\mathbf{G}_4$, which are investigated in this chapter, are presented below:

$$\mathbf{G}_3 = \begin{bmatrix} x_1 & x_2 & x_3 \\ -x_2 & x_1 & -x_4 \\ -x_3 & x_4 & x_1 \\ -x_4 & -x_3 & x_2 \\ x_1^* & x_2^* & x_3^* \\ -x_2^* & x_1^* & -x_4^* \\ -x_3^* & x_4^* & x_1^* \\ -x_4^* & -x_3^* & x_2^* \end{bmatrix} \quad \mathbf{G}_4 = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2 & x_1 & -x_4 & x_3 \\ -x_3 & x_4 & x_1 & -x_2 \\ -x_4 & -x_3 & x_2 & x_1 \\ x_1^* & x_2^* & x_3^* & x_4^* \\ -x_2^* & x_1^* & -x_4^* & x_3^* \\ -x_3^* & x_4^* & x_1^* & -x_2^* \\ -x_4^* & -x_3^* & x_2^* & x_1^* \end{bmatrix} \tag{2.21}$$

There is a trade-off between the robustness of each of these codes, which is strictly connected with the number of transmit antennas, and the code rate *R*, equal to 1 in case of the $\mathbf{G}_2$ code only. The $\mathbf{G}_3$ and $\mathbf{G}_4$ codes have lower code rate, equal to 0.5, but in turn they can offer lower error rates in comparison with the $\mathbf{G}_2$ one.

Each space-time block decoder may operate with a single receive antenna only. However, for the best performance, it is strongly recommended to use a larger antenna array at the receiver. The signal received by a receive antenna *j* at time *t*, may be written as:

$$r_t^j = \sum_{i=1}^{N} h_{i,j} s_t^i + \eta_t^j \tag{2.22}$$

where $h_{i,j}$ corresponds to the channel coefficient (2.19), s_t represents the symbol transmitted by antenna *i* and the noise samples η_t^j are modelled by the complex Gaussian process with zero mean and $N_0/2$ variance per dimension.

The main feature of space-time block codes is their orthogonality. This is the main condition, under which the decoding operation may be successfully performed, and is defined as:

$$\mathbf{G}_N \mathbf{G}_N^H = \left(\sum_{i=1}^{N} |x_i|^2 \right) \mathbf{I}_N \tag{2.23}$$

where *N* is equal to the number of transmit antennas, $\mathbf{G}_N^H$ indicates the Hermitian of the code matrix and $\mathbf{I}_N$ is an *NxN* identity matrix.

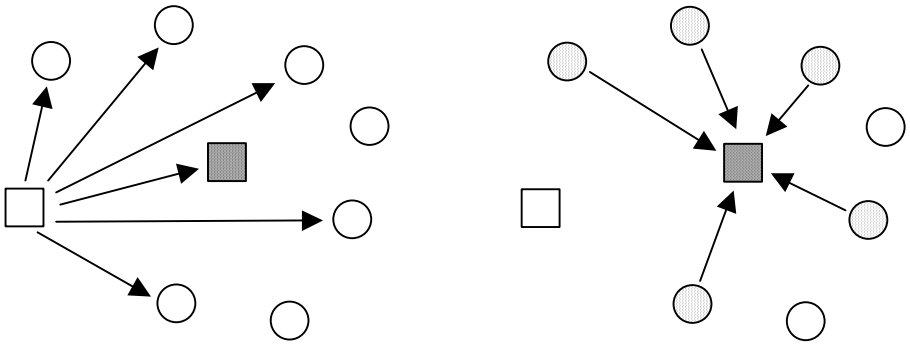
The general decision metric is then given by the following formula:

$$z = \sum_{t=1}^{l} \sum_{j=1}^{m} \left| r_t^j - \sum_{i=1}^{n} h_{i,j} s_t^i \right|^2 \tag{2.24}$$

For a given code, it suffices to find such a combination of potentially transmitted symbols, which minimizes this expression. However such approach is not optimum in the computational aspect and for the purposes of the simulation research the optimised metrics were used, which are given in [TJC99]. The diagram of the $\mathbf{G}_2$ coded system is depicted in the Appendix in Figure 6.11.

2.3.2 Cooperative diversity

The issue of cooperative diversity has attracted a lot of interest recently [Lan03]. The idea is based on the usage of the virtual antenna arrays, which are constituted in an ad-hoc manner by a set of mobile stations assisting the transmitting node in order to provide additional diversity and, as a consequence, to improve the reliability of decisions at the receiver side. To this end a special transmission mode is required [Lan03], which consists of two separate stages (Figure 2.15).



a) the first stage b) the second stage

Figure 2.15: Cooperation of mobile stations

In the first stage the transmitting node sends the signal, which is received by the receiving node as well as by its neighbours. The neighbours process the received signal cooperatively and resend it to the

destination node in the second stage. This is the moment when the aforementioned virtual antenna array is created. It may be built up by all the neighbours or by some of them . This way or another there is always a subset of nodes selected under some agreed constraints and the retransmitted signal may be cooperatively coded with the use of the suitable space-time block encoder. A generalized model for the cooperation of mobile stations can be found in [HZF04].

2.3.3 Adaptive approach to antenna selection

In this section the initial results are presented concerning the adaptive approach to antenna selection and space-time block coding. To this end a wireless system with 4-element fixed transmit antenna array and one receive antenna is investigated. The results were obtained with the use of the MIMO flat fading Rayleigh channel, where the fading coefficients for each wireless link between one of the transmit and the receive antenna are calculated according to the modified and optimised formulas given in [ZX03]. Due to the context of virtual antenna arrays constituted by fairly separated mobile nodes it is assumed there is no correlation among distinct radio links. The power transmitted by each transmit antenna is always normalized so that the overall power transmitted by all the used transmit antennas was always equal to 1. The received signal is perturbed by the additive white Gaussian noise with zero mean and $N_0/2$ variance per dimension, according to formula (2.22). Always 40 000 000 bits are transmitted and the QPSK modulation scheme is used. The channel state information is assumed known at the transmitter side and as a result no channel estimation is preformed. The reference results for the $\mathbf{G}_2$, $\mathbf{G}_3$,and $\mathbf{G}_4$ encoders are presented in the Appendix in Figure 6.12.

The adaptive approach is based on a space-time block encoder, which can operate in two separate modes: $\mathbf{G}_3$ and $\mathbf{G}_4$. As there are four transmit and one receive antenna, four channel coefficients, one for each pair of antennas, are taken into account. The $\mathbf{G}_3$ mode is selected if the minimum value out of the channel coefficients' moduli is lower than the given threshold β, where β is a real value, which is chosen in such a way that the minimum BER was provided. As a result one transmit antenna is not used and three other ones transmit signals in accordance with the $\mathbf{G}_3$ code matrix. Otherwise all transmit antennas are used and the encoder operates in the $\mathbf{G}_4$ mode. This way the worst radio link is discarded until its parameters meet the selection algorithm requirements again. The $\mathbf{G}_3$ and $\mathbf{G}_4$ encoders have been investigated for the purposes of this initial evaluation of the proposed adaptation strategy because they provide the same code rate and therefore can be easily compared. The example results for $\beta = 0.8$ in comparison with the reference results for the $\mathbf{G}_4$ encoder are presented in Figure 2.16. The gain achieved is about 1 dB.

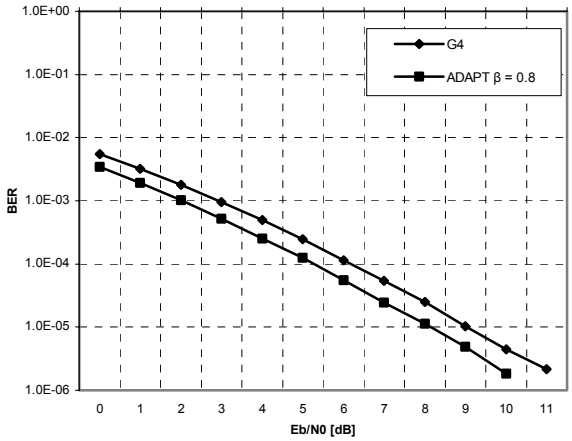


Figure 2.16: $\mathbf{G}_4$ encoder against the adaptive system (β = 0.8)

The detailed results regarding the bit error rate improvement at a specific E_b/N_0 for a given threshold β are presented in Figure 2.17.

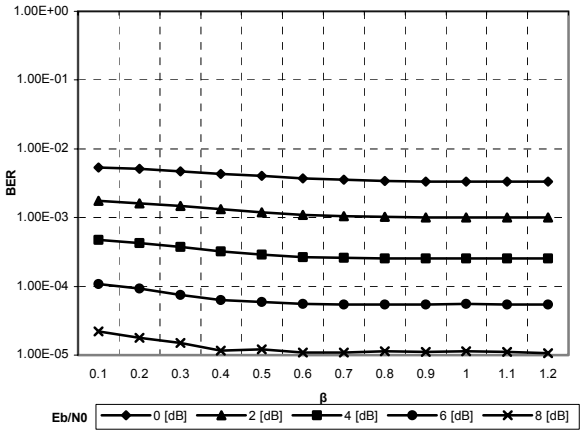


Figure 2.17: BER improvement at a specific E_b/N_0 for a given threshold β

This improvement may be observed regardless the E_b/N_0 value and the optimum results are achieved more or less between $\beta = 0.5$ and $\beta = 0.8$. For higher β values there is only a minor change. The $\mathbf{G}_3$ encoder usage percentage is presented in the Appendix in Figure 6.13. It is about 2 % for $\beta = 0.1$ and about 95 % for $\beta = 1.2$. It means that in the first case the adaptive system works almost always in the $\mathbf{G}_4$ mode, whereas in the latter one mostly three best radio links are selected and the signal is coded in accordance with the $\mathbf{G}_3$ code matrix.

2.3.4 Conclusions

The presented results were obtained for the fixed antenna array case and are the basis for further, extended research on advanced algorithms for optimum selection of the mobile relay nodes, assisting transmission in the second stage. Therefore, the presented solution and results may be applicable when cross issues between the *Radio Interface* (Work Package 2) and the *Radio Network* (Work Package 3) are taken into account. However, even in the case of fixed antenna matrices, the conclusion need not be that it is always the best solution to select only three antennas out of all four available ($\beta >> 1$). In the next phase of the research an adaptive system, additionally employing $\mathbf{G}_2$ ($R = 1.0$), $\mathbf{G}_3$ ($R = 0.75$) and $\mathbf{G}_4$ ($R = 0.75$) codes, along with different modulation schemes and outer coding will be investigated. It will be then possible to evaluate the optimum β value so that not only the lowest bit error rate was provided, but also the required code rate and hence the required throughput. This will be the basis for the final analysis of a mobile ad-hoc network, where virtual antenna matrices are constituted by relay stations and where the quality of service issues are taken into account. There may be many concurrent transmissions in such a network and the research will be focused on optimal selection of the nodes assisting each of them in such a way that the *quality of service* (QoS) requirements were met for each route. Obviously the interaction with the routing protocol will be necessary to make the physical layer aware of the required transmission parameters on the one hand and to indicate the routing layer about the available resources on the other hand.

3. Multi-link Adaptation

3.1 Predictive adaptive resource scheduling using TDMA/OFDMA

We will here describe the design of a channel state feedback system that enables fast link adaptation and multiuser scheduling using TDMA/OFDMA. The design and evaluation is intended to explore and push the limits of adaptive transmission in several respects:

- The use of TDMA/OFDMA provides the possibility to adapt to the frequency-domain variability of the channels. The frequency-selective channels are likely to differ even for closely co-located terminals. Therefore, a multiuser scheduling gain is obtained also for stationary terminals, by allocating frequency-domain subchannels optimally. However, this requires much more channel state information than the adaptive allocation of timeslots only (TDMA). *The feedback data rate* of TDMA/OFDMA might become overwhelmingly large. It must be taken into account, and means must be found for minimizing it. This issue is studied and solved in subsection 3.1.4.
- The feasibility of *fast link adaptation with respect to the short-term fading is explored for terminals at vehicular velocities*, using the best available predictors that are computationally feasible. The design and evaluation are performed at the highest considered WINNER carrier frequency of 5 GHz. This results in the most challenging test case: At the shortest carrier wavelength, the fading becomes the fastest for a given terminal velocity. The constraints on the feedback loop delays become tightest and the attainable prediction horizons become shortest.
- Adaptive TDMA/OFDMA is designed and evaluated not only for downlinks, but also for *multiuser uplinks*. This requires good frequency synchronization of all uplink transmissions, to minimize inter-carrier interference. In FDD wide-area uplinks, channel prediction of multiple-input single output (MISO) channels is furthermore required: Each potential user will have to send training data for all potential uplink chunks simultaneously. The predictor will then have to estimate and predict the uplink channels from all the active users simultaneously.

The research described in this section is based on, and extends, work performed within the SSF Wireless IP project [WIP]. An adaptive TDMA/OFDMA downlink for 1900 MHz UMTS bands was presented and analysed in [SOA+03] and [WOS+03]. These basic principles are here extended to the WINNER physical layer modes, and generalized to adaptive uplinks. The design of the adaptation schemes is explained for wide-area cellular FDD downlinks and uplinks in Subsection 3.1.1, as well as for short-range cellular TDD downlinks and uplinks in Subsection 3.1.2. The general WINNER MAC algorithms described in Section 1.1 is assumed. Thus, a non-adaptive fallback alternative is assumed to exist whenever the adaptive transmission turns out to be infeasible.

The proposed adaptation schemes are based on Kalman state-space filtering and predictors of the channel in the frequency domain. The performance of these predictors is evaluated in Subsection 3.1.3. We then combine the link adaptation strategy from Section 2.2, that takes the prediction uncertainty into account, with these channel predictors and the feedback control schemes, to obtain complete algorithms for the link adaptation and multiuser scheduling.

Subsection 3.1.4 then discusses compression of the channel state feedback. Appropriate algorithms for utilizing the correlation in time and frequency are proposed and evaluated, after discussing alternative schemes.

Predictive multiuser scheduling and link adaptation is evaluated for the FDD downlink in Subsection 3.1.5, with respect to the throughput, mutiuser scheduling gains, bit and packet error rates and effects of the scaling of chunk sizes. The section also contains an investigation relevant to the choice of MA scheme: How much of the multiuser scheduling gain obtainable by adaptation would be lost if TDMA were used instead of TDMA/OFDMA?

Subsection 3.1.6 summarizes the results and outlines remaining open issues. Section 4.1 will draw conclusions for the WINNER system design in different scenarios.

More details on the computational complexity of the algorithms and additional results can be found in Appendix 6.4.

3.1.1 Link adaptation and scheduling for wide-area FDD downlink and uplink

The wide-area FDD-based mode utilizes a narrower band of 20 MHz, as described in Subsection 1.1.1. It is assumed to use half-duplex FDD, so that terminals do not have to transmit and receive simultaneously. However, an access point can transmit and receive simultaneously. The terminals may use either full duplex (high-end terminals) or half duplex (cheaper and simpler terminals). The active half-duplex terminals are partitioned into two sets within a cell, where set 1 receives at time *j*, *j+2, ...*, and transmits at time *j+1, j+3,...,* while set 2 receives at time *j+1, j+3,...,* and transmits at time *j, j+2*. Full duplex terminals can be considered as equivalent to two terminals, one of each set. The uplink and downlink chunks are assumed to be synchronized. An OFDM-based air interface using TDMA/OFDMA is used in both uplink and downlink, with frequency-time chunk sizes of 8 x 6 = 48 constellation symbols in both cases. The design is intended also for multi-antenna transmission but will here be evaluated for SISO links.

A brief overview of the adaptation control loop is given here. An earlier but similar FDD design was presented in Section 6.1.5 of [WIND26]. Since the predictability of time-varying frequency-selective channels is limited, the adaptation feedback control loop has to be made as fast as practically possible. Therefore, a MAC frame here equals only one chunk duration. The feedback loop delay should ideally be only two chunk durations. Therefore, in-band control signalling is used: The control information is transmitted in chunks that also contain payload data, to minimize delays. The following constraints on the timing of the feedback loop have been assumed, in the FDD design as well as in the TDD design:

1. At least 0.06 ms (3 OFDM symbols in FDD) is reserved for the multiuser resource scheduling computation. This should be more than adequate for many high-performance algorithms.
2. At least 0.10 ms is reserved for the last steps of the Kalman channel prediction operation, that has to be performed after the latest measurement data has been received. This applies both when predictors are located in terminals and when predictors are situated at access points.
3. After the decision has been taken on what a specific OFDM symbol should contain, at least one OFDM symbol time interval is allowed for forming the symbol before it is transmitted.

Further discussions of timing issues and computational complexity issues can be found in Appendix 6.4.

The **downlink** works as follows. Regular pilot patterns are transmitted on the downlink. Based on channel measurements up to chunk time *i*, all active terminals predict the channel quality in *all chunks within a sub-band of interest* at the future chunk time *i+2*. These reports are source-coded and transmitted on uplink control symbols within the uplink chunks at time *i+1*. The appropriate code and modulation rate that could be used by each terminal in each chunk is then determined. As described in Section 2.2, this is done via the tables of optimised rate limits designed under a maximal bit error rate constraint, for a given average SINR and prediction accuracy (variance). For each antenna stream/beam, the adaptive resource scheduler then allocates each chunk at time *i+2* exclusively to one of the flows. The allocation is reported to the users by control symbols embedded within the so allocated downlink chunks at time *i+2*.

In each downlink chunk, we assume four pilot symbols and eight in-band control symbols. The number of payload symbols per chunk is thus 36, and the overhead factor due to the pilots and signalling is 36/48 = 0.75. When using adaptive modulation combined with convolutional coding, CM schemes between BPSK rate ½ and 64-QAM rate 5/6 are used, as described in Section 2.2. The number of payload bits per chunk may therefore vary between 18 and 180. Let the subcarriers within a chunk be enumerated from c = 1 to 8 with increasing frequency, and let the six OFDM symbols be enumerated from s = 1 to 6. On the downlink, the pilots and control symbols are then placed on subcarriers 3 and 7. The channel predictors utilize these subcarriers, and the channel estimation is improved by utilizing the control symbols in decision-directed mode. The locations of the downlink pilots and the control symbols are specified below:

- Symbols (c,s) = (3,1), (3,2), (7,1) and (7,2) are for in-band downlink control: They report which present chunks belong to which flows. Appropriately coded 4-QAM symbols are used.
- Symbols (c,s) = (3,3) and (7,3) are known pilots. At multi-antenna access points, different pilot symbols are transmitted from each antenna. The pilots are used for two purposes:
    - Channel estimation for coherent detection within chunks with payload of interest[6].

[6] Whenever possible, the pilots from neighbouring and preceding chunks are utilized in the channel estimation for coherent detection. Thereby, an interpolation of the time- and frequency varying channel at the payload symbol locations within the chunk can be performed, and extrapolation which increases the error is avoided..

- Prediction for all frequencies of interest for the future adaptive downlink transmission. For this purpose, subcarriers 3 and 7 in each chunk are used in a Kalman estimator of low complexity, with the state-space GCG algorithm derived in [SA03].

When OFDM symbol s = 3 within chunk *i* has arrived, the channel for chunk *i+2* is predicted. The required prediction horizon to the end of chunk *i+2* is thus 2.5 x 0.3372 ms = **0.843 ms**.

- Symbols (c,s) = (3,4), (3,5), (7,4) and (7,5) carry control information for the next uplink transmission. They report which out of the next uplink chunks have been appointed to which uplink flows. Appropriately coded 4-QAM symbols are used.
- Symbols (c,s) = (3,6) and (7,6) are known pilots, used for coherent detection and for updating the predictor states.

Higher order modulation than 4-QAM is not used for the control information. This information is broadcast and has to be detectable by all users. That property essentially defines the cell radius. The control symbols are all located on the same subcarriers as the pilots, to make use of them for the decision-directed channel estimation. Thereby, the channel is estimated for each OFDM symbol period instead of every third OFDM symbol. This reduces the channel estimation MSE and the prediction MSE.[7]

On the **uplink**, the active terminals taking part in the adaptive transmission are in competition for a part of the 20 MHz band, called a *contention band*. All active terminals assigned to a contention band send *overlapping pilot signals* during chunk time *i*. The predictors located at the access point then predict the uplink channels for all terminals at time *i+2*. The prediction is based on the latest and previously received uplink signals at the locations of the overlapping pilots.

The appropriate code and modulation rate that could be used by each terminal in each chunk is then determined. As for the downlink, this is done via tables of rate limits designed under a maximal bit error rate constraint, for a given average SINR and SINR prediction accuracy (variance). The adaptive resource scheduler then assigns the uplink transmission for time *i+2*. The control signals that report the scheduling decision are thereafter transmitted to the terminals by in-band signalling: This is the purpose of the later control symbols of the downlink chunk at time *i+1*, as described above. Each chunk within an antenna beam/stream is allocated exclusively to one flow.

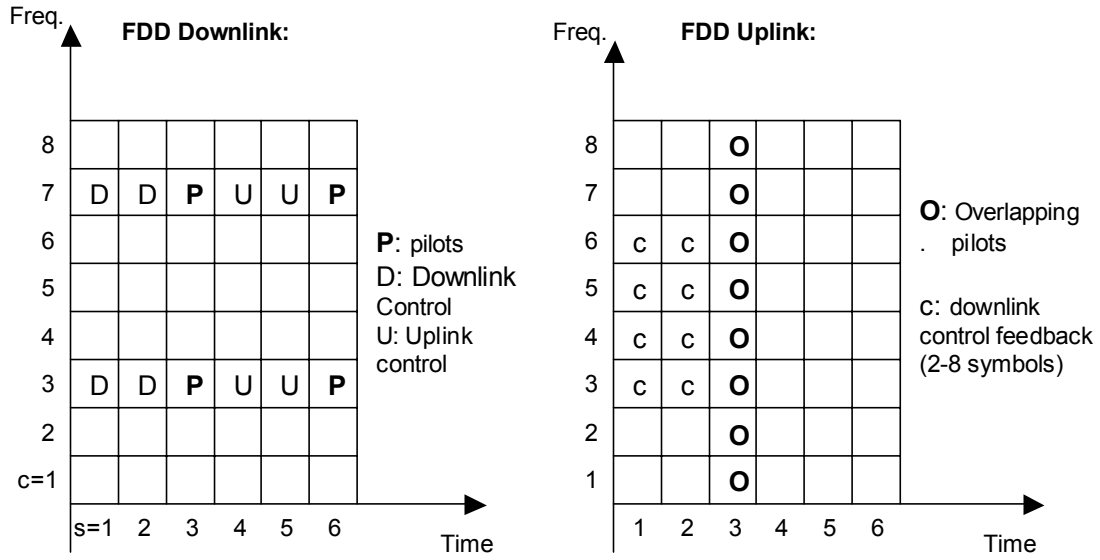


Figure 3.1: Pilot and control symbol patterns in FDD downlink and uplink

[7] Unfortunately, the payload symbols cannot be used in a decision-aided estimator to further improve the prediction. The reason is that the payload symbols within a chunk may be destined for a terminal that is much closer to the base station than the one who is performing the prediction. The modulation and coding of such payload symbols has been adjusted for that user, and may therefore be totally undetectable by all except the closest terminals.

The in-band control symbols on the uplink are a part of the control loop for the downlink. They carry the downlink channel prediction reports from all active terminals. Their number can be adapted to the requirements, varying from 1 to 8 appropriately coded 4-QAM symbols per chunk and it depends on the number of active terminals and on the terminal velocities. For timing reasons, such in-band control symbols are positioned early, in symbols 1 and 2 out of the six OFDM symbols, see Figure 3.1.

As mentioned above, all antennas from all terminals that might use this chunk at a future time transmit known pilot symbols simultaneously. All eight pilot symbols within OFDM symbol s = 3 are reserved for these overlapping pilots. The use of overlapping pilots in a Kalman filter that simultaneously estimates and predicts all channels is described in [SA05]. In multi-antenna receivers, the prediction is performed separately for all receiving antennas within a sector/cell.

The prediction horizon from symbol 3 in chunk *i* to the end of chunk *i+2* is 2.5 x 0.3373 ms = **0.843 ms**.

The resulting uplink pilot plus control overhead varies from 9 to 16 out of 48 symbols, giving an overhead factor of between 39/48= 0.812 to 32/48 = 0.667. Furthermore, a few more non-overlapping pilots may be required to be placed among the payload data to improve the channel interpolation in the time direction, when a coherent detection is performed, This topic has not yet been investigated.

3.1.2 Link adaptation and scheduling for TDD-based short-range cellular mode

The adaptation control loop for the short-range TDD cellular mode works in a very similar way as for the FDD case described above. TDMA/OFDMA is assumed to be used in both the uplink and downlink. The main difference to the FDD case described above is that the channel prediction for both uplink and downlink transmission is assumed to be performed by the terminals, based on downlink pilots and control symbols.[8] The uplink channel is predicted based on measurements of the downlink channel, assuming channel reciprocity in a TDD system. Thus, channel estimation based on the overlapping pilots from many terminals is avoided. On the other hand, the required feedback loop delay becomes longer in the TDD uplink than in the FDD uplink. In addition the noise plus interference level at the AP has to be predicted at the AP.

A TDD slot of duration 2 x 0.3372 ms = 0.6774 ms is shown in Figure 3.2 for one chunk bandwidth of 16 subcarriers, cf. Subsection 1.1.2. It contains an uplink chunk, a downlink chunk and two duplex guard intervals, or 30 OFDM symbols plus two guard intervals, numbered from 1 to 32. This design accommodates uplink/downlink asymmetry ratios between 1:2 to 2:1, related to the assumption R3:16 in [WIND71] that the long-term aggregate user traffic asymmetry is assumed to be within that interval.[9] The first 10 OFDM symbols always belong to the uplink, and the last 11 OFDM symbol belong to the downlink. Depending on the uplink-downlink limit, the uplink chunk size may vary between 16 x 10 = 160 and 16 x 19 = 304 constellation symbols. The pilots essential for prediction and control symbols used for scheduling allocation reports are placed in the later part of the downlink chunk, outside of the variable middle region.

The **downlink** is designed as follows. A TDD slot contains uplink chunk *i* and downlink chunk *i+1*, and the next slot contains uplink chunk *i+2* and downlink chunk *i+3*. Based on the pilot symbols in the downlink chunk *i+1*, all active terminals predict the channel quality in all chunks within a sub-band of interest at the downlink time *i+3*. These reports are source-coded and transmitted on the uplink control symbols within the uplink chunks at time *i+2*. The adaptive resource scheduler at the access point then allocates each chunk at time *i+3* exclusively to one of the flows. The allocation is reported to the users by the control symbols embedded within the so allocated downlink chunks at time *i+3*.

In the downlink chunks, six pilot symbols and eight appropriately coded 4-QAM in-band control symbols are assumed present, in the locations described by Figure 3.2. The purpose of the control symbols is exactly as for the FDD downlink described in the previous section: the earlier four symbols are used for controlling the downlink, while the later four symbols control the allocations in the next uplink slot. The

[8] Other alternatives are conceivable. For example, one could keep the uplink predictors at the access point, and also use them to predict the downlink channels, assuming channel reciprocity. Moving the predictors to the access points would result in lower complexity for the terminals. However, the downlink interference environment still has to be predicted at the terminals, and reported to the AP. Evaluation of this alternative scheme is a topic for further investigation.

[9] However, if the TDD asymmetry is varied, it would be strongly recommended to use the *same* asymmetry factor within all closely spaced cells of a cellular system, to avoid severe duplex interference.

control symbols located on the "pilot" subcarriers c = 6 and c = 14 are furthermore used to improve the channel estimation and prediction, doubling its update frequency by decision-aided channel estimation.

The prediction of the next *(i+3)* downlink chunk is performed based on the pilots and control symbols received until the last symbol of chunk *i+1*. The required downlink prediction horizon to the far end of chunk *i+3* is therefore one TDD slot length of **0.676 ms**. The states of the predictor cannot be updated during the uplink chunk, but have to be extrapolated using a statistical model of the channel time evolution. This is readily performed by Kalman estimators. Measurement-based updates then commence at the first pilot location in the next downlink chunk. With 6+8 = 14 pilot and control symbols, the downlink overhead factor due to pilots and signalling is 226/240 = 0.941 at asymmetry 1:1. When using adaptive modulation combined with convolutional coding, code and modulation rates between BPSK rate ½ and 64-QAM rate 5/6 are used as described in Section 2.2. With asymmetry 1:1, the corresponding number of payload bits per chunk varies between 114 and 1140.

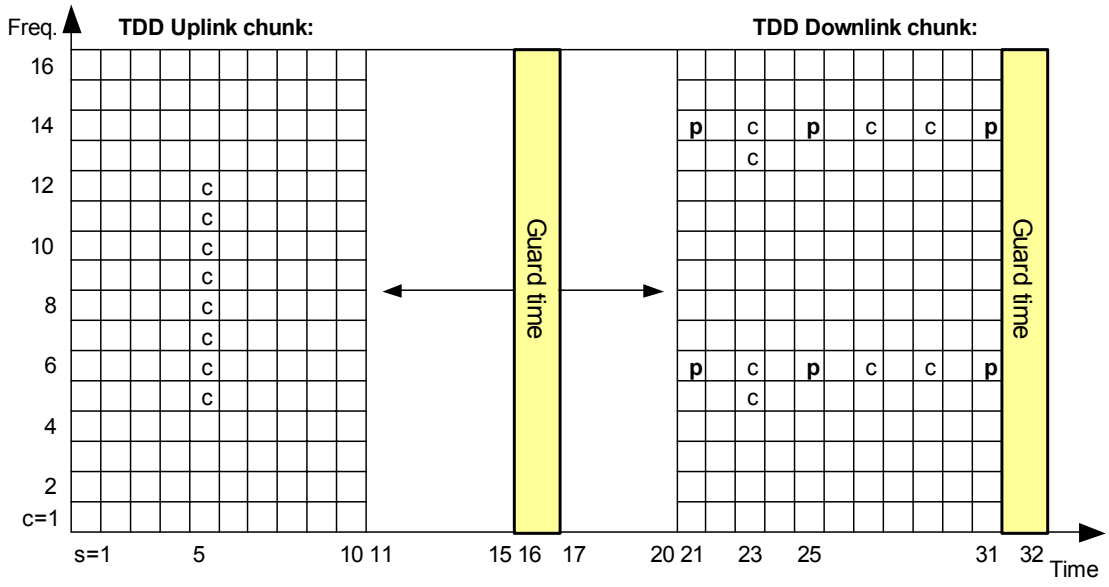


Figure 3.2: TDD Timeslot of 2 x 0.3372 ms, containing uplink and downlink chunk and guard times

The adaptation of the **uplink** works as follows: In addition to predicting the chunks at (downlink) slot *i+3*, the terminal also predicts their downlink channel at the subsequent (uplink) chunk *i+4*. The reports for chunk *i+3* and *i+4* are source-coded and transmitted on the uplink control symbols within the uplink chunks at time *i+2*. Assuming channel reciprocity, the downlink predictions for *i+4* are used as predictions of the uplink channel at *i+4*. At the time of the uplink chunk *i+2*, the access points also monitors the noise and interference situation and predicts it for the time of the uplink chunk *i+4*. The channel prediction and the noise-plus-interference prediction is combined into a SINR prediction for each user within each chunk at *i+4*. Based on the accuracy of the power predictions coming from each user, the code and modulation rates that each user could use in each chunk are determined, as described in Section 2.2. The adaptive resource scheduler at the access point then allocates each uplink chunk at time *i+4* exclusively to one of the flows. Bit interleaving is used within chunks. The resulting allocation is reported to the terminals by control symbols embedded within the downlink chunks at *i+3*.

Depending on the asymmetry ratio, the required prediction horizon from the end of chunk *i+1* to the far end of the uplink chunk *i+4* varies between 0.91 ms to 1.08 ms. For asymmetry 1:1, it is **1.00 ms**.

3.1.3 Channel prediction and predictor performance

The feedback loops for the FDD and the TDD physical layer modes that were outlined in the last two subsections were designed to be as fast as possible, within realistic constraints imposed by computation times and signalling delays. Still, the feedback loop requires 2.5 chunk times, corresponding to 15 OFDM symbols or 0.843 ms for the FDD mode. For the TDD mode, the required prediction horizon was 0.676 ms for the downlink and ranged from 0.91 ms to 1.08 ms for the uplink prediction.

What would be the appropriate methods for channel prediction? As was seen in Subsection 2.1.3.4, and will be seen further in Subsection 3.1.3.5 below, simple extrapolation of the present channel estimate will lead to large performance losses already when extrapolating in space over a small fraction of the carrier wavelength. To push the limits of performance of adaptive transmission techniques, and enable their use also for terminals with vehicular velocities, channel extrapolation must be replaced by prediction.

Extensive investigations of channel power predictors were performed in [EKS+99], [Ekm00], [Ekm02], [SEA01] and [ESA02]. Both theoretical considerations and evaluations on a large set of measured channels with 5 MHz bandwidth were taken into account. It was concluded that the class of channel power predictors that performed best on measured data was based on *linear prediction of the complex baseband channel*, followed by use of the *quadratic unbiased predictor* described in Section 2.2 to predict the channel power. Predictors of the channel gain that were based on nonlinear functions on the data performed significantly worse than linear functions. Predictors of the power that were based on previous samples of the measured power, instead of the channel gains or baseband signals, performed worse since phase information was lost. The noise level was found to be the crucial limiting factor for the attainable performance accuracy.[10] Use of more samples to average and suppress the noise will improve the prediction performance. It is therefore advantageous if a large fraction of symbols within the utilized subcarriers are either pilots, or can be used for decision-directed estimation. The feedback loops and pilot symbol positionings presented in the previous subsections have been designed with this in mind.

One way of performing channel prediction would be to perform it in the time domain. The most significant taps of the channel impulse would be predicted, for example with methods developed in [Ekm02]. The resulting predicted impulse response is then transformed to the frequency domain to obtain the predicted frequency-domain complex channel gains.

We will here take the alternative approach to perform the whole prediction in the frequency domain. We can then focus the prediction just on the subband of interest. A set of linear prediction filters, each responsible for its own subband of the total bandwidth, is utilized. Comparative evaluations of this approach with the time-domain approach are currently underway. In particular, the state space algorithm described in [SA03] has been utilized to predict the complex channel and the unbiased quadratic predictor described in Section 2.2 has then been used to predict the channel power. The algorithm in [SA03] starts by deriving a Kalman predictor. (For an overview of linear state space estimation using Kalman filters, see e.g. [AM79] or [KSH00].) The predictor utilizes the correlation of the channel in the frequency domain by predicting p pilot-containing subcarriers in parallel. It also utilizes the correlation in the time domain of the fading channel. The number p is a compromise between performance and computational complexity. In the evaluated design, we use $p = 8$, spanning 4 chunks in both the FDD and TDD designs. This means that 26 such Kalman estimators would be required to cover a complete band of 104 chunks in both cases.

It is then in [SA03] shown that the present OFDM channel prediction problem is ideally suited to the application of a novel low-complexity approximation of the Kalman algorithm, the Generalized Constant Gain algorithm [SLA02]. This class of algorithms avoids the need to update a quadratic state-space Riccati difference equation, which generates the dominant computational demand in a Kalman algorithm. In the present problem, the approximation leads to almost no performance loss.

Autoregressive models of order 4 are here used to model the temporal channel correlation properties. They are adjusted to the fading statistics. The state update equations of the Kalman/GCG estimators are based on these models. The predictor algorithms have been applied on channels with Urban Macro power delay profile [BGS+04] in the FDD wide-area downlink and to the IEEE 802.11n case ‘C’ non line of sight model [IEEE04] in the TDD short-range case. Their power delay profiles are shown in Table 4.1. White noise with known power was assumed. As described in the previous Subsections, the predictors are assumed to be applied in the following ways.

- In FDD downlinks, the channel state estimation is performed by the terminal. It is based on the pilot symbols, and on decision-directed estimation also at the downlink control symbols.
- In FDD uplinks, the channel prediction is performed at the access point, using overlapping pilots from the terminals. A generalization of the Kalman algorithm of [SA03], described in [SA05], is utilized. Its performance will be investigated below for the cases of 2 and 8 simultaneous users

[10] The Doppler spectrum of a fading channel describes a band-limited process, and band-limited processes can in principle be extrapolated (predicted) arbitrarily far into the future with perfect accuracy. However, the true situation is very different, since we are forced to work with noisy measurements of this band-limited process.

in each contention band, all with the same received power. (These results were obtained in [WIND26], using the ITU Vehicular-A channel described in Table 3.3 below.)

- For both TDD downlinks and uplinks, the channel power is predicted by the terminals. Reports are transmitted to the access point via the uplink control symbols. The prediction is performed based on the downlink pilots and control symbols that are used for decision-directed estimation.

Figure 3.3-Figure 3.6 show the resulting normalized prediction MSE $\tilde{\sigma}^2 = \sigma_{\varepsilon_h}^2 / \sigma_h^2$ that was defined in Section 2.2. Here, the complete Kalman algorithm, without the GCG approximation, has been used.

3.1.3.1 Performance of frequency domain Kalman/GCG channel prediction in FDD downlinks

Figure 3.3 shows the normalized prediction MSE $\tilde{\sigma}^2$ as function of the prediction horizon scaled in wavelengths, for different values of the SNR from 0 dB to 25 dB. The results are for full duplex terminals that use all timeslots for updating the predictor with measurements. The FDD downlink with Urban Macro channel is used. Note the large dependence of the performance on the SNR. The results for horizon zero represent the filter NMSE $E\,|h(t)-\hat{h}(t\,|\,t)|^2 \,/\, E\,|h(t)|^2$. When the prediction horizon L in $\hat{h}(t\,|\,t-L)$ increases, the prediction error grows.

Use of additional parallel pilot subcarriers in each estimator (increasing p) may improve the performance by better use of the channel correlation in the frequency direction. This effect is however limited by the frequency correlation of the particular channel. The value p = 8 used here spans 4 chunks of the FDD mode, or 625 kHz.

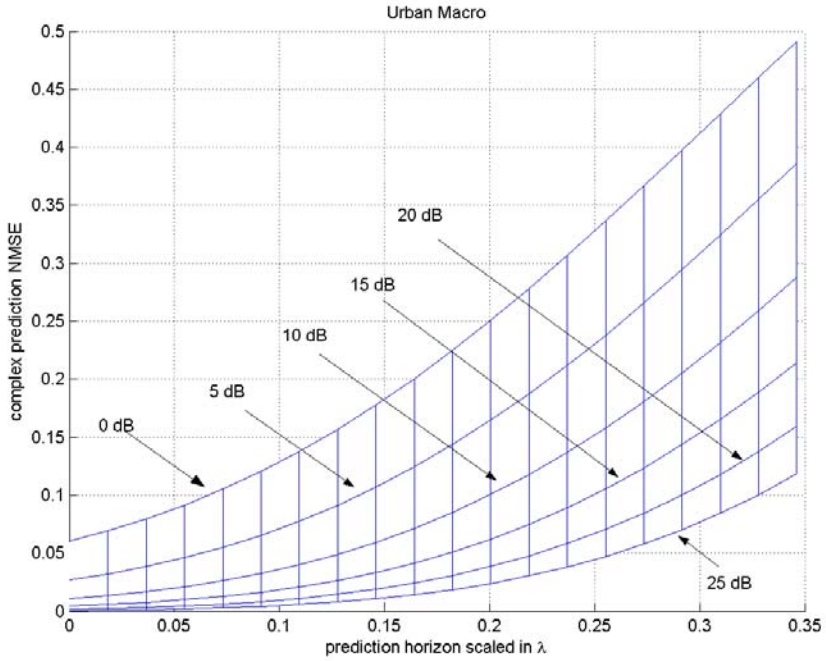


Figure 3.3: Prediction accuracy in terms of the normalized channel prediction MSE, as a function of the prediction horizon scaled in carrier wavelengths, and as function of the SNR. Results for FDD downlink over Urban Macro channels, using a Kalman algorithm that utilizes 8 subcarriers.

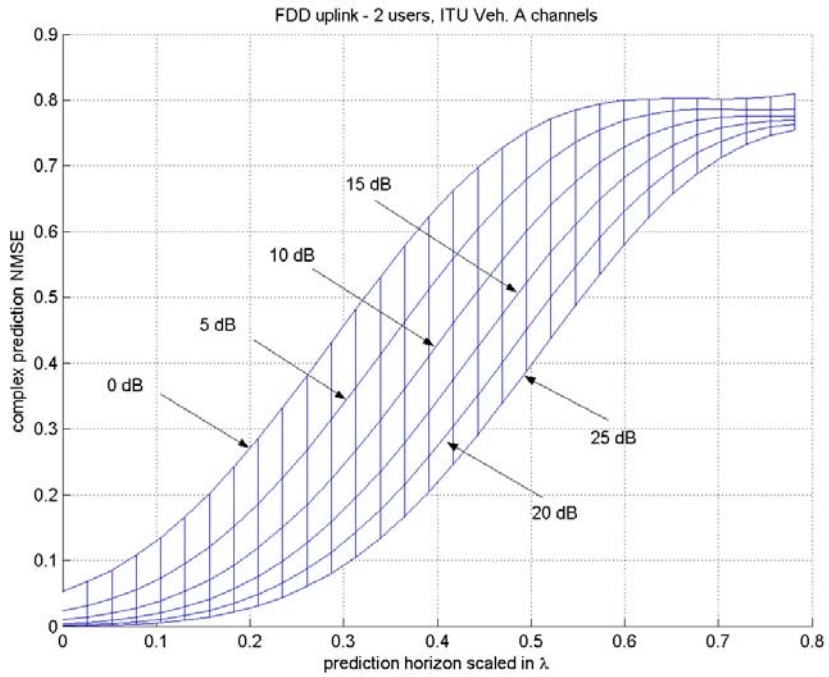


Figure 3.4: Prediction accuracy in terms of the normalized channel prediction MSE, as a function of the prediction horizon scaled in carrier wavelengths, and as function of the SNR. Results for FDD uplink over Vehicular-A channels, using a Kalman algorithm for overlapping uplink pilots, that utilizes 8 subcarriers. Average result for two simultaneous uplink users per contention band

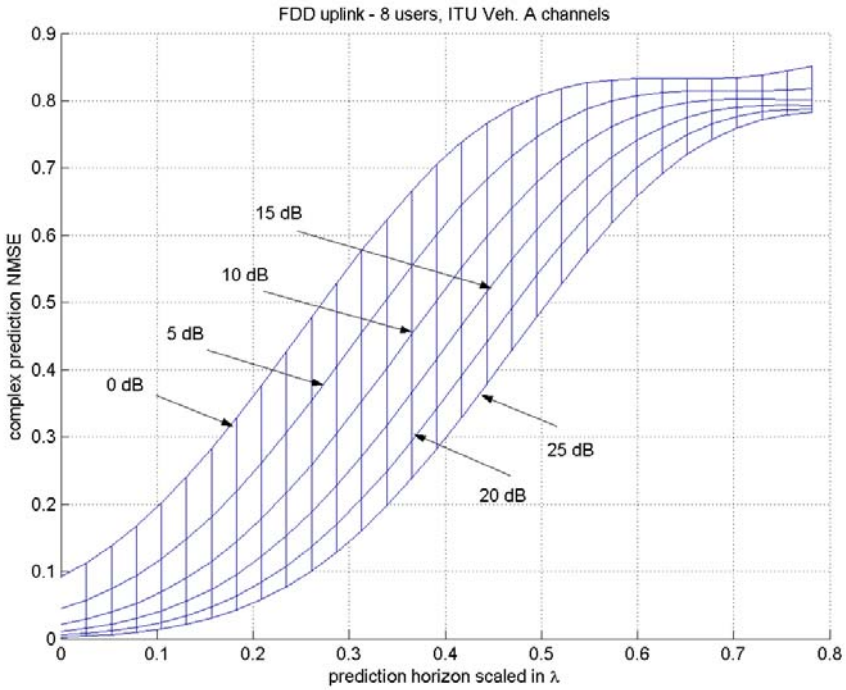


Figure 3.5: Prediction accuracy in terms of the normalized channel prediction MSE, as a function of the prediction horizon scaled in carrier wavelengths, and as function of the SNR. Results for FDD uplink over a Vehicular-A channel, using a Kalman algorithm for overlapping uplink pilots, that utilizes 8 subcarriers. Average results for eight simultaneous uplink users per contention band

3.1.3.2 Prediction of multiple FDD uplink channels based on overlapping pilots

Figure 3.4 and Figure 3.5 show results for FDD uplinks, where 2 users and 8 users simultaneously transmit overlapping pilots, all with the same average received power.

In a Kalman estimator based on overlapping pilots, separate sets of states are used for describing the channel of each user. The autoregressive models that describe the fading statistics of each user are adjusted individually to the velocity of the users [SA05]. The uplink control information could be used for improving the estimate by decision-directed methods. However, it is not utilized here, since the exact form of coding and signaling for the FDD uplink control data has not yet been specified. In the presented results, all users have the same velocity and travel through the same type of Vehicular-A propagation environment, but their channels were generated as independent realizations from this channel statistics. Furthermore, the average received power is equal for all users (slow power control is assumed to be used). Results in Figure 3.4 and Figure 3.5 are for the FDD uplink described in [WIND26], with 200 kHz chunk width and 50 kHz subcarrier spacing. Thus, the OFDM symbols that contain overlapping pilots contain 4 symbols per chunk, all of which are pilots. The Kalman predictor uses $p = 8$ subcarriers.

The results indicate that prediction based on overlapping pilots will decrease in accuracy with an increasing number of terminals, but this decrease is rather modest. Thus, the channel predictions in FDD uplinks in which not too many users occupy each contention band of the total bandwidth seems feasible. For more details on this state-space prediction algorithms, please see [SA05].

3.1.3.3 Performance of Kalman/GCG channel prediction in TDD downlinks and uplinks

Figure 3.6 shows the corresponding results for TDD downlinks and uplinks. In the proposed TDD system, both uplink and downlink channel prediction is performed at the terminal, based on the common pilot symbols that are transmitted from the access point. Thus, performance degradation due to use of overlapping pilots is avoided. Compared to the FDD downlink, the accuracy is somewhat worse, for a given prediction horizon. The reason is the TDD duty cycle. It reduces the availability of channel measurements: The channel estimator has to extrapolate its states over the uplink chunks, where the terminal switches to transmission, and no downlink pilots are available anyway. Note also that the required prediction horizon is higher for the TDD uplink than for any of the other cases. As a result, the TDD uplink has the worst performance for a given vehicle velocity among all the considered cases.

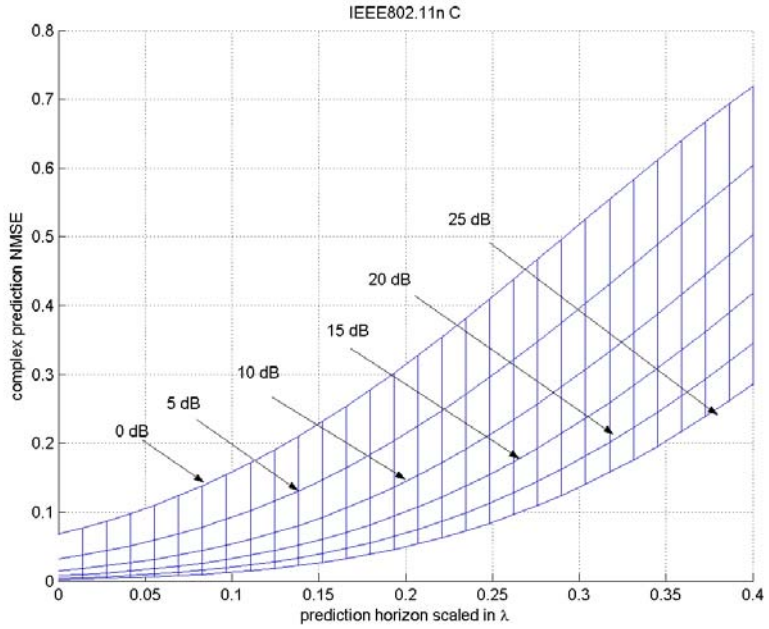


Figure 3.6: Prediction accuracy in terms of the normalized channel prediction MSE, as a function of the prediction horizon scaled in carrier wavelengths, as a function of the SNR. Results for TDD uplink and downlink over a Vehicular-A channel, using a Kalman algorithm based on using 8 subcarriers in parallel. The IEEE 802.11n case C non-line-of sight model is used.

3.1.3.4 Attainable prediction horizons and limiting vehicle velocities

The prediction accuracy depends on the horizon l scaled in wavelength, which in turn depends on the velocity v [m/s], the prediction horizon in time L [s] and the carrier wavelength λ_c via the relation

$$l = vL / \lambda_c . \tag{3.1}$$

The prediction accuracy will also depend on the SNR. In Sections 3.1.1 and 3.1.2, the required prediction horizons L scaled in chunk times or milliseconds were stated, for the assumed feedback loops. Table 3.1 below summarizes the corresponding prediction horizons scaled in wavelengths, when the considered TDD and FDD systems operate at 5 GHz. The relation (3.1) then gives the horizon as a fraction of the wavelength as

$$l\,[\lambda_c] = v\,[\text{km/h}]\; x\; L\,[\text{ms}]\; / \; 216 \quad @\ 5\ \text{GHz}. \tag{3.2}$$

Table 3.1: Maximal prediction horizons in the TDD and FDD adaptive multiple access schemes

	Maximal horizon *L*	**Fraction of λ_c at 30 km/h**	**Fraction of λ_c at 50 km/h**	**Fraction of λ_c at 70 km/h**
TDD downlink	0.676 ms	0.094	0.156	0.219
TDD uplink (1:1)	1.00 ms	0.150	0.250	0.350
FDD downlink	0.843 ms	0.117	0.195	0.273
FDD uplink	0.843 ms	0.117	0.195	0.273

Adaptive transmission to/from a terminal will be feasible up to a maximal velocity for a given SINR, or equivalently, down to a limiting SINR at a given velocity. For combinations of velocities and SINRs beyond such a boundary, non-adaptive transmission must be used. A preliminary estimate of the limiting SINR will be given here, based on the results in Figure 3.3-Figure 3.6. This calculation will then be evaluated by simulation in Subsection 3.1.5.2 below. These estimates are based on the assumptions and scaling of the presented uplinks and downlinks. They are conservative in that the prediction is performed to the *far end* of the chunk to be allocated[11]. The prediction accuracy to less distant symbols is higher.

It has been found from earlier investigations of the sensitivity of uncoded and Trellis-coded adaptive modulation to prediction errors, that if the rate limits are adjusted to take the prediction uncertainty into account, then a prediction NMSE of 0.1 leads to only a minor degradation in the attained spectral efficiency [FSE+04], [FSS+03b]. The coded M-QAM scheme is expected to be somewhat more sensitive to prediction errors, see the discussion in Section 2.2. We here provisionally use an upper limit of 0.15 for the allowed normalized MSE $\tilde{\sigma}^2 = \sigma^2_{\varepsilon_h} / \sigma^2_h$ of the prediction error. For higher values, a non-adaptive fallback mode is utilized instead of adaptive transmission. Table 3.2 shows the resulting limits for the SINR. Corresponding prediction horizons from Table 3.1 are indicated within parentheses.

Table 3.2: SINR limits for cases where the accuracy limit $\tilde{\sigma}^2 = 0.15$ allows the use of adaptive transmission, exemplified for three terminal velocities for a 5 GHz carrier frequency

	30 km/h	50 km/h	70 km/h
TDD downlink (Fig. 3.6)	< 0 dB (0.094 λ_c)	5 dB (0.156 λ_c)	10 dB (0.219 λ_c)
TDD uplink (Fig. 3.6)	5 dB (0.150 λ_c)	15 dB (0.25 λ_c)	> 25 dB (0.35 λ_c)
FDD downlink (Fig. 3.3)	< 0 dB (0.117 λ_c)	6 dB (0.195 λ_c)	12.5 dB (0.273 λ_c)
FDD uplink, 2 users (Fig. 3.4)	0 dB (0.117 λ_c)	7 dB (0.195 λ_c)	15 dB (0.273 λ_c)
FDD uplink, 8 users (Fig. 3.5)	3.5 dB (0.117 λ_c)	11 dB (0.195 λ_c)	20 dB (0.273 λ_c)

[11] However, this conservatism cannot be reduced by much. At 70 km/h, the channel state becomes outdated after only a few OFDM symbol times, resulting in a rise in the BER. See Figure 2.10 in Section 2.1.

The results of Table 3.2 are of interest for qualitative comparisons between the RI modes and their uplinks/downlinks. It is evident that adaptive transmission can be expected to work in the widest variety of situations in the proposed wide-area FDD downlinks and short-range TDD downlinks, while it works in the narrowest range of circumstances in the proposed short-range TDD uplink. The case of wide-area FDD uplinks, using overlapping pilots, falls somewhere in-between.

3.1.3.5 Comparison with extrapolation of the instantaneous measurement

The considered prediction algorithms have a significant computational load, and their use can be motivated only if they provide a significant performance advantage. The performance of the simple channel extrapolation predictor that was discussed in Section 2.1 is here presented in a form that is comparable with the above results.

As prediction of the channel gain *h(t)*, we thus use an outdated noisy estimate

$$\hat{h}(t \mid t-L) = h(t-L) + v(t-L) \ . \tag{3.3}$$

The zero mean noise *v(t-L)* represents an estimation error. Its variance would of course depend on how the channel estimate *h(t-L)* were obtained. For many estimators, the channel to estimation error power ratio is proportional to the SNR, as $E|h(t)|^2 / E|v(t)|^2$ [dB] = $SNR + G_N$ [dB]. For example, for the Kalman filter estimate (horizon zero) in Figure 3.3, $G_N \cong 10$ dB.

The normalized prediction MSE is in Figure 3.7 shown as a function of the horizon L scaled in carrier wavelengths, for the Urban Macro channel. The results are shown for different values of the channel to estimation error power ratio $E|h(t)|^2 / E|v(t)|^2$, from 0 dB (uppermost curves) in steps of 5 dB to a noise-free case (lowest curve). Comparing to the Kalman predictor performance for the corresponding case in Section 3.1.3.3, the performance is much worse. When extrapolating a noise-free sample of the channel, the prediction NMSE 0.15 would be surpassed already at 0.09 wavelengths. This corresponds to a horizon of 0.277 ms at 70 km/h according to (3.2). In the assumed FDD downlink, this corresponds to 5 OFDM symbols, which is less than one single chunk duration. In Figure 2.10, the BER increase for this case of 5 symbols delay was seen to indeed become unacceptably high.

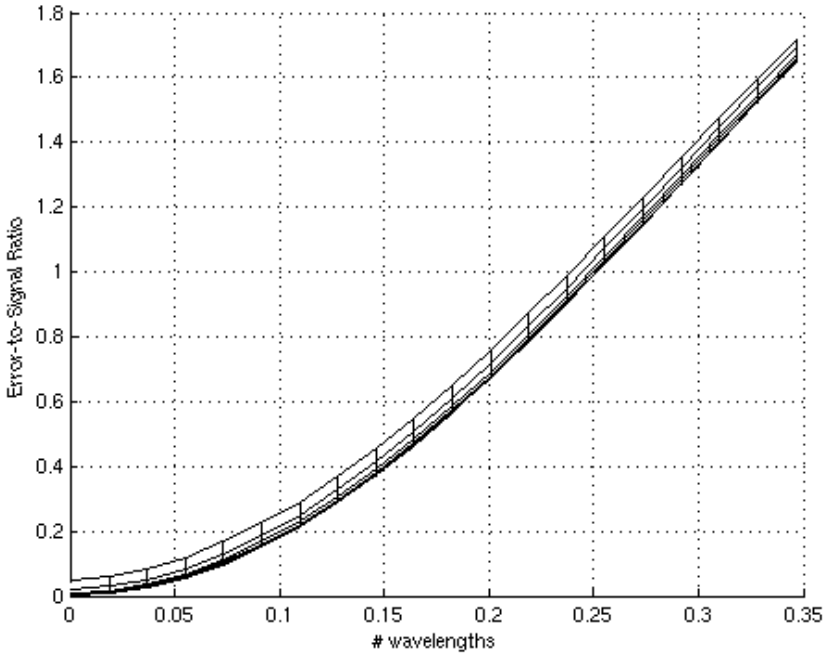


Figure 3.7: Prediction normalized MSE for the Urban Macro channel as function of the prediction horizon scaled in carrier wavelengths, when using a noisy outdated sample as predictor. Results for channel to estimation error ratios of 0 dB (uppermost curve) to 25 dB (lowermost curve)

3.1.4 Compression of feedback information

3.1.4.1 Required feedback rates without compression

Assume that one or several clients within a terminal are in competition for a sub-bandwidth of the total band comprising N chunks. K active terminals compete for this sub-bandwidth, or contention band. Assume that each terminal feeds back a proposed code and modulation rate for each chunk. For a scheme with r rates, each terminal then needs to feed back $N \log_2(r)$ bits per chunk duration T_{chunk}. For K users, the required total feedback data rate is then

$$R_f = KN \log_2(r) / T_{chunk} \text{ [bit/s]}. \quad (3.4)$$

For the full-band FDD downlink with N = 104 chunks and with r = 8 code and modulation rates, we would have a feedback overhead of 925 kbit/s per active user! Expressed in another way, for each chunk for which K downlink users compete, $K \log_2(r)$ feedback bits would have to be transmitted in the following uplink chunk. With only 40 non-pilot symbols per uplink chunk, K = 8 active terminals with r = 8 would consume 24/40 = 60% of the uplink bandwidth for control signalling, if one feedback bit/feedback symbol can be used on average. The situation described here is clearly unacceptable. Fortunately, there are several ways in which the required feedback rate can be reduced significantly.

First, all users need not be in competition for the whole band. A restriction of the contention bandwidth must of course be weighted against the corresponding loss of diversity and opportunity within the remaining chunks: With a smaller selection, there is less chance of finding a really good chunk. Expressed differently, the assignment of only a few users within each contention band will reduce the attainable multiuser scheduling gain. Here, significant research remains to be performed on the appropriate balancing of the resource allocation.

If the suggested code and modulation rates are fed back, one could furthermore consider feeding back not all rates, but only the most competitive ones. Indeed, M. Johansson in [Joh04a] showed that significant multiuser scheduling gains are attainable even when feeding back only one bit (r = 2). However, as explored in [Joh04b], one pays a heavy price for this feedback rate reduction: the scheme becomes very sensitive to the placement of the decision limit (rate limit) that determines the feedback variable. This scheme would become very sensitive to prediction errors.

Fortunately, there are additional tools at our disposal. The channel gains and SNRs at adjacent chunks will be highly correlated. (If they were not, the chunk widths would have been selected too wide, and we would have severe problems with channel variability within the chunks.) This correlation can be utilized to reduce the feedback rate. The channels are also correlated in time. Furthermore, it is likely that in most cells except those situated close to major roads, the large majority of terminals will not travel at vehicular speeds, but rather be stationary. For those terminals, very little feedback is required.

Methods that utilize the correlation in time and frequency are now presented. The results are explored for the ITU Vehicular A and ITU Pedestrian A channel statistics. The FDD downlink, with corresponding chunk sizes, is investigated. A modulation-coding scheme with r = 8 rates is assumed.

Table 3.3: ITU Vehicular A channel model

Delay [ns]	0	310	710	1090	1730	2510
Power [dB]	0	-1	-9	-10	-15	-20

Table 3.4: ITU Pedestrian A channel model

Delay [ns]	0	110	190	410
Power [dB]	0	-9.7	-19.2	-22.8

3.1.4.2 Discrete-amplitude compression: Lossless feedback of modulation format

In [Sha48] , Shannon showed that the ***entropy rate*** of a discrete-valued process gives us the minimum possible bit rate for lossless coding. The entropy rate is defined as

$$H(\chi) = \lim_{n\to\infty} H(X_n \mid X_{n-1},\ldots,X_1)\,. \tag{3.5}$$

In general, the entropy rate is difficult to compute, since estimation of a infinite number of conditional probabilities is needed to compute the conditional entropy. However, there are model assumptions that can be useful for this purpose. A good model for correlated discrete sources is the Markov model. In an *L*th order Markov model, the symbols are conditionally independent of the long-term history, so that (assuming stationarity)

$$H(X_n \mid X_{n-1},\ldots,X_1) = H(X_n \mid X_{n-1},\ldots,X_{n-L}) = H(X_{L+1} \mid X_L,\ldots,X_1) \tag{3.6}$$

and we can easily compute the entropy rate (see [CT91] for a detailed description). We have experimented with various orders of the Markov model, both in time and frequency, and the conclusion is that a 1st- or 2nd-order model in the frequency domain is enough, while a higher-order model in the time domain may be necessary to pick up all correlation. The table below shows estimates of the entropy rate of the two considered channels: the ITU Vehicular A channel and the ITU Pedestrian A channel.

Table 3.5: Entropy rate per chunk, or smallest lossless feedback rate in bits per chunk per user

Entropy rate (ITU Vehicular A)	$H(\chi) = 0.91$ [bits/chunk]
Entropy rate (ITU Pedestrian A)	$H(\chi) = 0.35$ [bits/chunk]

An adaptive ***arithmetic coder*** is a good choice of algorithm to exploit the correlation [Say00]. If the probabilities are correctly computed, then the arithmetic coder will achieve the entropy rate as estimated with the Markov model assumption. A disadvantage with arithmetic coding (or with any lossless coding algorithm of the modulation levels) is that the code will be variable rate, leading to a variable load in the feedback channel.

Thus, it seems realistic to attain a lossless feedback data rate of 1 bit/chunk for vehicular mobile users in the wide-area downlink. With a contention band consisting of 20 chunks, the required feedback rate would then be 59 kbit/s per active vehicular mobile user. A significantly larger reduction can be attained if the SNR values instead of the rates are fed back and lossy compression is used, as outlined in the next subsection.

3.1.4.3 Continuous-amplitude compression: Feedback of SNR values

To send the channel state information to the transmitter, we need to quantize the SNR values. We should seek to exploit the correlation in time and frequency to make the feedback transmission as efficient as possible. Three common methods for lossy compression are:

- *Vector quantization (VQ)*, where several symbols are quantized simultaneously to exploit correlation between them.
- *Linear prediction coding (LPC)*, where the process is linearly filtered in a decorrelating filter prior to quantization.
- *Transform coding (TC)*, where a de-correlating transform is applied to the source vectors prior to quantization.

Of these methods, vector quantization is the most powerful one, but the complexity is high, and in practice the other methods will in many cases give better results. Predictive coding and transform coding are in theory equally powerful, but they have slightly different properties. For example, transform coding is usually better when the coding rate is low (below 2 bits/sample). On the other hand, a transform coder works on segments of data, which may be a disadvantage. Predictive coding can be implemented without segmentation, but will not work well if the coding rate is too low.

In the problem at hand, we have separately studied how to exploit correlation in time and frequency.

Exploiting correlation in frequency

Transform coding has the potential to work very well in exploiting correlation in frequency between chunks. The blocking is no problem here, and while the bit rate may be too low for a linear predictor to work well, a transform coder can work well with arbitrarily low bit rates. For processes with high positive

correlation, the ***discrete cosine transform (DCT)*** is a good choice [Say00]. The DCT is very common in image and video compression, due to its low complexity (it can be implemented using the fast Fourier transform) and good properties in general.

After the transform, the DCT components should be quantized. Vector quantization would be of some use here, but the performance increase compared to scalar quantization would be fairly low since the components are already decorrelated by the DCT, and the complexity of VQ can be prohibitive. We have chosen to use scalar quantization in the performance evaluation. A bit allocation for the transform coefficients must be agreed upon by the transmitter and receiver. In practice, the bit allocation must be adapted to varying long-term channel conditions, but the adaptation can easily be implemented with a negligible additional bit rate, and will not be considered further here. The allocation of bits to each transform coefficient should be proportional to the component variance, and an algorithm for this purpose is described in e.g. [Say00].

A block diagram of the proposed scheme is given in Figure 3.8. Each SNR vector is DCT-transformed, and quantized in a scalar quantizer optimised for the pdf [Max60], [Llo82] of the component, and with a bit allocation according to the component variance. The indices of the quantizers are transmitted over the feedback channel, and the process is reversed in the base station, as shown in Figure 3.9.

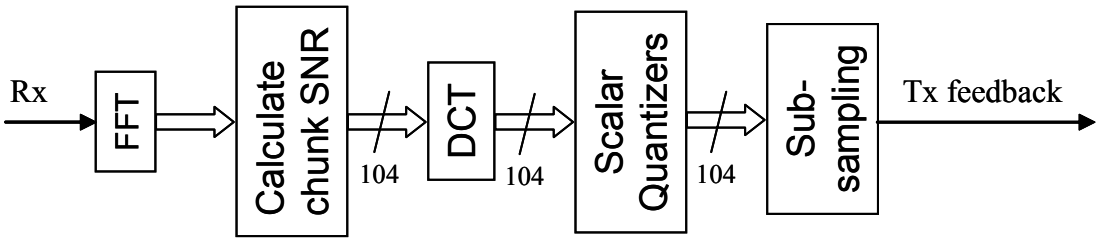


Figure 3.8: A block diagram of feedback handling in the terminal

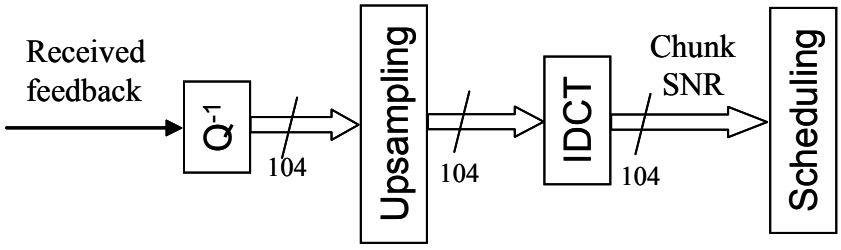


Figure 3.9: Block diagram of feedback handling at the base station

Exploiting correlation in time

The time-domain correlations are not as easy to exploit as the frequency domain correlations, due to that no blocking of data can be done; we cannot accept any time delay. Therefore, transform coding or vector quantization is not applicable. Furthermore, a linear predictor would not work well since the bit rate is very low (see the numerical results below). Instead, we suggest *sub-sampling the process*, by transmitting the quantizer indices every *N*:th block, and estimating the intermediate blocks with a MMSE estimator. The FIR MMSE estimator is defined as

$$\hat{X}_n = \sum_{i=1}^{M} a_i \tilde{X}_{n-n_i} \tag{3.7}$$

where the estimate is a linear combination of old values at arbitrary time-lags n_i. Note that the estimator relies on *quantized* values $\tilde{X}_{n-n_i}$ of the process, since the base station has no access to the original process. The optimal solution for the estimation coefficients $\{a_i\}$ is given by the Wiener solution. To obtain this solution, we must compute autocorrelations $R_{\tilde{x}\tilde{x}}(n_i - n_j)$ for the quantized process, and cross-correlations $R_{x\tilde{x}}(n_i)$ between the original and the quantized process, to find the optimal estimator coefficients. This can be done beforehand for a set of channel models, or it can be adapted in run-time.

We have applied the estimator separately for each quantized DCT components, since the components are approximately conditionally independent. For low-power DCT components where the bit allocation

algorithm has allocated zero bit, no estimator is computed; instead, we will just use the (pre-computed) time-average of the component.

3.1.4.4 Numerical results

We will now present results for the system described above.

Table 3.6: Rate limits for the adaptive modulation scheme in the SNR feedback coding example

Modulation format	Rate limit, SNR per symbol [dB]
BPSK	-∞
QPSK	7.75
8-PSK	12.81
16-QAM	16.16
32-Cross-QAM	19.81
64-QAM	22.95
128-Cross-QAM	26.24
256-QAM	29.33

Block diagrams of the coding and decoding process of feedback information is given in Figure 3.8 and Figure 3.9. An adaptive modulation scheme with eight uncoded rates is used. The rate limits, shown in Table 3.6, have been calculated assuming perfect SNR prediction. They are adjusted according to the criterion to maximize the number of bits in correctly received chunks (see Section 2.2). The scheme is basically the same as used in [WOS+03], with the rate limits adjusted for the chunk size. *K* terminals are competing for the whole band of 104 chunks. In the simulations, all terminals have the same average SNR 16 dB, and Maximum rate scheduling is used; Each chunk is given to the user with highest reported SNR.

The chunk SNRs are calculated as the average SNR in each chunk. At a given time instant, these chunk SNRs are collected in a vector of length 104 and coded using a DCT, compressed using optimised scalar quantizers and then subsampled before transmission. At the receiver, the inverse process is applied. The number of bits per chunk needed and the corresponding achieved throughput using a max SNR scheduler for the two scenarios (ITU Vehicular A and ITU Pedestrian A) at a mobile speed of 50 km/h are given the figures below, using a subsampling factor of 2.

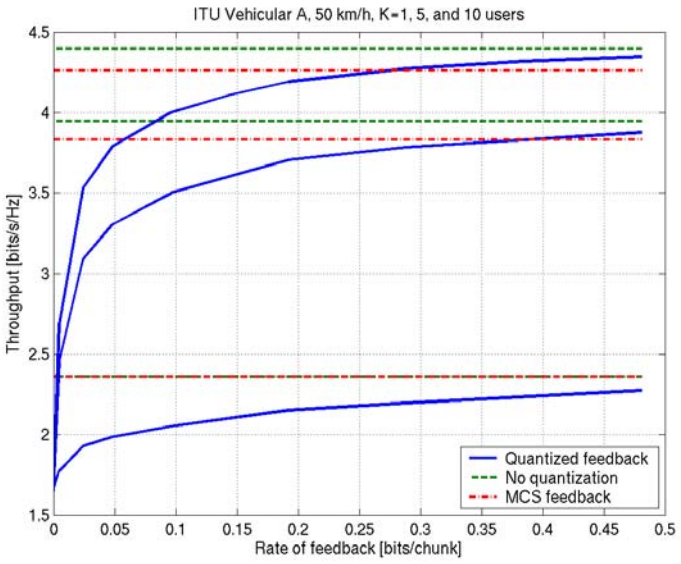


Figure 3.10: Throughput as function of the feedback rate, Vehicular A channel, 50 km/h

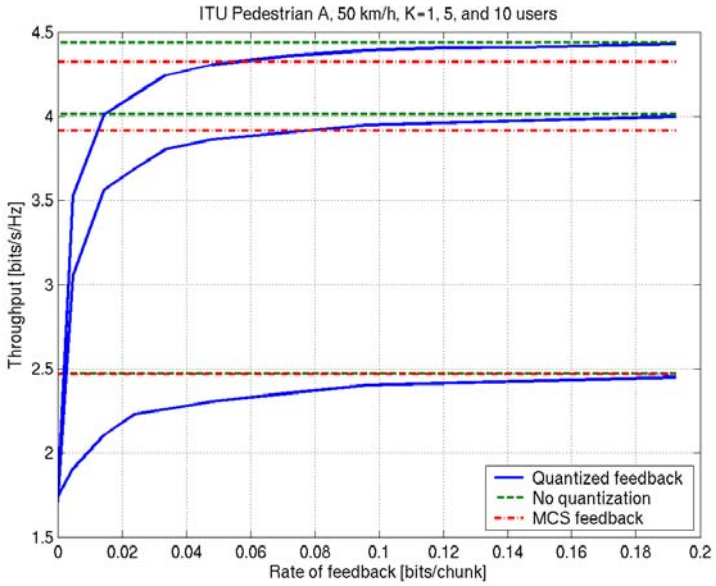


Figure 3.11: Throughput as function of the feedback rate, Pedestrian A channel, 50 km/h

As seen, it is possible to reduce the amount of feedback down to 0.1-0.4 bits/chunk from the 3 bits/chunk needed to signal the given modulation level and *still achieve a higher throughput* than the CM scheme that uses all 3 bits to signal the modulation level instead of the SNR. The reason is that with SNR feedback, the scheduler can distinguish between user terminals that can use the same code and modulation scheme. It can then select the one with the lowest error probability. At 50 km/h, the subsampling results in virtually no performance loss. If the mobile speed is reduced, it will be possible to further subsample the coded feedback information. E.g. at 5 km/h, it would be possible to reduce the feedback information by a factor of 10 compared to the values shown in the figures above.

With a feedback rate of 0.25 bits/chunk over the Vehicular A channel, we would only need 4 bits per chunk to accommodate 16 *vehicular* users per contention band, in addition to many more stationary users. Uplink feedback with only four 4-QAM symbols per chunk that use rate ½ coding is adequate for this. The uplink in Section 3.1.1, with 1-8 feedback symbols/chunk, was designed with these results in mind.

3.1.5 Assessment of multilink adaptation performance, limitations and feasibility

In the following multilink simulations, all terminals are assumed to have the same average SNR. All channels will have the same statistical properties, and all terminals will move with the same velocities. The interference is modelled by white Gaussian noise. The performance of the FDD system will be evaluated using the Urban Macro channel, with power delay profile described in the righthand column of Table 4.1. The adaptive resource scheduling and transmission follows the general scheme outlined at the beginning of Subsection 1.1.2.

The effects of channel estimation errors on the demodulation are not considered. The effects of channel prediction errors will be in focus. Unquantized feedback of the SNR is assumed and the feedback channel is assumed error-free.

The channels are rather flat, but not perfectly flat, within the chunks. In general, there is variability both with time and in the frequency direction. Within each chunk, the modulation and coding scheme potentially used by each user is determined by taking the average predicted SNR, (SNR_{av}) and the predicted SNR at the worst point within the chunk, SNR_w, for that user. The weighted average of these two values is then used as the *effective SNR,* which is reported

$$SNR_e = b\ SNR_{av} + (1\text{-}b)\ SNR_w\ . \tag{3.8}$$

The parameter b in (3.8) can be used to tune the performance of the scheme when there is significant channel variability within the chunks. With $b = 1$, a large variability would lead to a large increase in the BER, since the worst corner of the chunk generates most errors. With $b = 0$, we would have a

conservative scheme that tends to have less errors than the target BER. In all investigations reported here, we have used $b = 0.4$ in (3.8), which has performed well.[12]

The simulation is performed in the frequency domain. Simulated channels are generated and represent the true channel to/from each terminal. The predicted channels are generated by using the link-to-system interface model described in Subsection 6.1.6.1 of [WIND26]. It generates one complex random number with prescribed variance for each chunk, and adds this number to all baseband frequency-domain channel gains at all constellation symbol positions within the chunk. The predicted effective SNR for each of the 104 chunks is then calculated for each user terminal, using (3.8).

Based on the predicted effective SNR and the prediction uncertainty, a code and modulation rate is suggested for each user terminal and each chunk. We use either the coded or the uncoded scheme described in Section 2.2. Then, scheduling is performed. The scheduling strategy used is the Proportional Fair one, which in this case, where all users have the same average SNR, reduces to the Max Throughput strategy of giving the chunk to the stream that can use the highest code and modulation rate. If two streams can use the same rate, one of them is just randomly selected here. Thus, the potential performance gain obtainable by using SNR feedback instead of rate feedback that was discussed at the end of Subsection 3.1.4.4, is not utilized. (A lower error rate could be attained when SINR predictions are available at the scheduler by giving the channel to the user with lowest expected error rate.)

The resource scheduling buffer is assumed never to be emptied. Scheduling units with constant size of *512 bits* are used. The size is adjusted to that of small IP packets. Overhead due to the CRC code and sequence numbers for the link retransmission is not taken into account. Each scheduling unit is distributed among the allocated chunks, with superfluous payload symbols filled with zeros. A separate coding and modulation is used for each chunk. At the receiver, soft input Viterbi decoding is performed if the coded M-QAM scheme is used. If all bits belonging to a scheduling unit are received correctly, it is released to higher layers. Otherwise, a link retransmission would be generated. However, we have not implemented a link retransmission mechanism in the experiments below.

3.1.5.1 The multiuser scheduling gain with max rate scheduling and equal average SNR

Figure 3.12 below shows the ***throughput***, here defined and measured as the ***number of payload bits in correctly received 512 bit scheduling units, divided by the total number of transmitted payload symbols***. The throughput is shown as a function of the average SNR, in the wide-area FDD downlink. The CM scheme with convolutional coding and M-QAM from Section 2.2 is applied. Perfect channel predictions are assumed in this experiment, and the corresponding rate limits are used.

The (blue) curves with squares show the result when one user utilizes all the chunks, the good ones as well as the bad ones.

At velocity 5 km/h (solid line), the effect of channel variability within the chunks in the time direction is negligible. At 50 km/h we have some variability. Due to the use of (3.8), with $b = 0.4$, the effective SNR used to assign the appropriate code and modulation rate decreases somewhat, so the throughput decreases as well. The decrease at 50 km/h is rather small, equivalent to a 1 dB decrease of the SNR.

The (red) curves with rings show the results when eight users are scheduled. With all users having the same average SNR, we give each chunk to the user with the highest SNR. The resulting multiuser scheduling gain improves the throughput. Above SNR = 19 dB, the modulation-coding scheme saturates, and the highest rate (rate $R = 5/6$ 64-QAM) is always used by the appointed user of each chunk. The throughput is also improved significantly at low SNR values compared to the single user case, increasing to 0.85 bits/symbol at 0 dB and 5 km/h. In all experiments in the figure, the attained bit error rate is lower than the target maximum BER of 0.001.

[12] An alternative to be evaluated in future studies, which is likely to work better for very large channel variability within chunks, is to substitute the weighting of SNR values with the effective SNR weighting proposed in the WINNER Link-to-System Interface internal report.

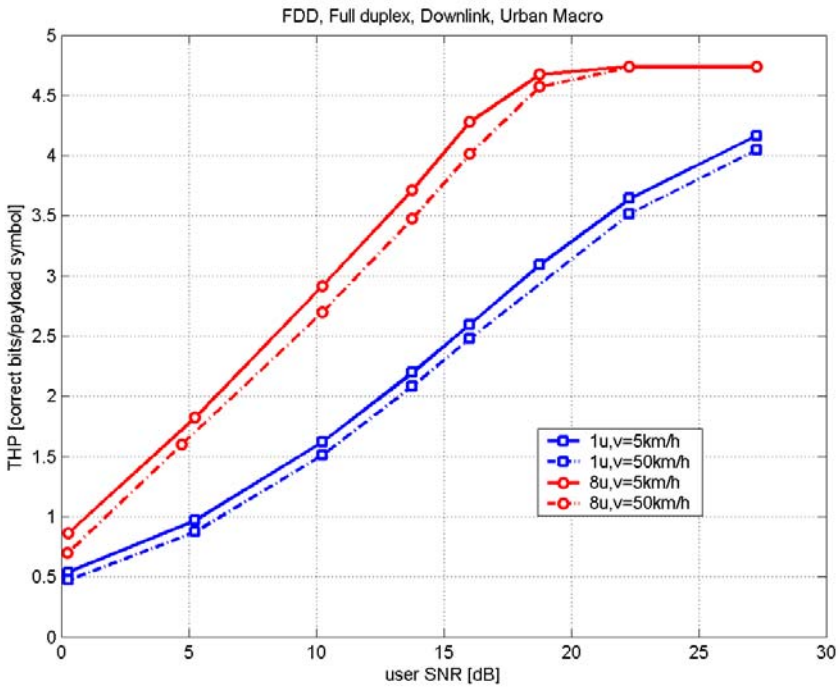


Figure 3.12: Throughput as a function of the SNR in the wide-area FDD downlink, when all users have the same average SNR and Urban Macro channel statistics

It is instructive to compare the throughput curve for one user at 5 km/h to the results presented in Figure 2.5, where a similar scheme has been optimised for target BER 0.001. At mid to high SNR levels, the throughput for one user in Figure 3.12 is lower than the number of bits per symbol attained in Figure 2.5. The difference at 15 dB is around 0.65 bits per symbol, or 21%. A part of the difference is due to the fact that Figure 3.12 counts the throughput within correctly received 512-bit scheduling units. However, the packet error rate is below 1%, so this difference is small. The remaining difference is due to three causes:

- The mapping of 512-bit scheduling units onto chunks results in some partially unfilled chunks. This effect causes around 5 percentage points of the 21% reduction at 15 dB. It also explains why the curve for eight users saturates at 4.75 bits/symbol at high SNR, rather than at the theoretical maximum of 5 bits/symbol.
- In the present experiment, no power control is used since this simplifies the resource allocation. In Figure 2.5, the power is adjusted in-between the rate limits. Comparing the performance of rate adaptation with and without power control in [FSE+04], it can be concluded that this causes around 10 percentage points of the 21% difference at 15 dB SNR.
- The remaining difference is due to the chunk size and the channel variability within chunks. In Figure 2.5, the rate and power allocation is performed on a per-subcarrier basis. In the present investigation, rate allocation is performed per chunk of 8 subcarriers. An Urban Macro channel has some variability within the 156 kHz chunk widths. When $b < 1$ in (3.8), we will use an effective SNR that is on average lower than the SNR that would be used if the allocation were performed on a per-subcarrier basis.

To convert the throughput values in Figure 3.12 to spectral efficiency, they should be multiplied by 36/48 due to the pilot and control overhead and by 832/1024 due to the assumed unused guardband subcarriers.

3.1.5.2 Effect of the prediction quality on throughput and error rates

In Figure 3.13, Figure 3.14 and Figure 3.15 below, we investigate the effect of different terminal velocities and the corresponding prediction uncertainties on the throughput, multiuser scheduling gain and bit error rate. The modulation-coding scheme with convolutional coding is used in Figure 3.13 and Figure 3.14. It is designed for a target maximal bit error rate of 0.001 at the prescribed level of prediction errors. (The attained *average* bit error rate should then be lower, since the maximal BER is targeted for the code and modulation rate limits. In-between the rate limits, the conditions are more advantageous.)

As in the previous Subsection, we test the FDD downlink, with different numbers of simultaneous users, all with the same average SNR and having the same channel statistics given by the Urban Macro model. Full duplex terminals are assumed. Figure 3.3 and Table 3.1 then describe the prediction uncertainties for different velocities, and corresponding horizons in wavelengths. Table 3.2 indicates that when performing experiments with all users having either 10 dB SNR or 19 dB SNR, the adaptation scheme should work rather well for all velocities up to 70 km/h at 19 dB, but that difficulties may be encountered at 70 km/h when the SNR is 10 dB. Figure 3.13, for 10 dB and Figure 3.14, for 19 dB show that this is indeed what happens. The dashed curves show the performance in the presence of prediction inaccuracy. For solid curves, predictions are assumed perfect but the effect of the channel variability within chunks due to the fading in time is taken into account, like in Figure 3.12. They are included for comparison.

The multiuser scheduling gain is considerable. Comparing one and eight users, it is 74% on average over all results for from 5 km/h to 50 km/h at 10 dB SNR. At 19 dB SNR, it is somewhat lower, 53%, since the code and modulation scheme begins to saturate at rate 5/6 64 QAM for eight simultaneous users.

The modulation-coding schemes are designed to fulfill the maximal BER constraints 0.001 in the presence of prediction errors. This is attained, with one exception: That of 70 km/h at 19 dB SNR. The corresponding packet error rates for 512-bit scheduling units are shown in Appendix 6.4. The packet error rates are low in all cases, below 1%, which indicates the throughput could be improved by tuning the scheme more aggressively. With an increasing velocity and corresponding increased prediction uncertainty, the rate limits are tuned more conservatively and the throughput is therefore decreased. Significant multiuser scheduling gains are however preserved also when rather long-range predictions are required, which is an interesting phenomenon. These result support the following conclusion:

With an advanced channel predictor, it is possible to adaptively utilize the short-term fading at vehicular velocities up to carrier frequencies of 5 GHz.

The glaring exception to this is the case of 70 km/h and 10 dB SNR presented in Figure 3.13, where the scheme fails. With a too high prediction uncertainty $\tilde{\sigma}^2 = \sigma_{\varepsilon_h}^2 / \sigma_h^2 = 0.178$, we can no longer "ride the peaks" of the fading channels for different terminals. This operating point (Velocity, SNR) is beyond the limit provisionally proposed in Table 3.2. The results strengthen the hypothesis that a simple velocity and SNR-dependent limit could be used to delineate if flows should be allocated via the adaptive or the non-adaptive algorithm within the resource scheduler outlined in Section 1.1.

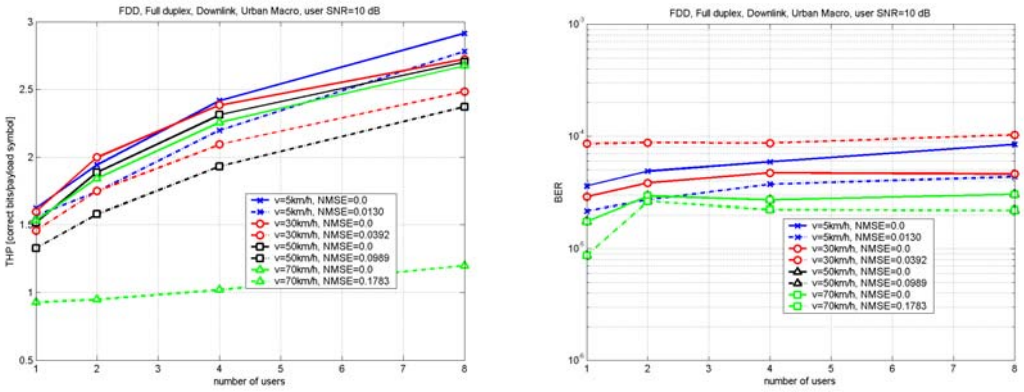


Figure 3.13: Throughput and bit error rates as functions of the number of active users, all with the same average SNR of 10 dB, when scheduling the chunks to the user with best instantaneous predicted rate. FDD wide-area downlink, 5GHz carrier frequency, 0.843 ms required prediction horizon. Convolutional coding and M-QAM designed under a maximum bit error constraint 0.001. Solid curves take channel variability within chunks into account but neglect the prediction uncertainty. Dashed curves take prediction uncertainty from Figure 3.3 into account

At the end of Section 2.2, the hypothesis was proposed that the relative impact of the prediction uncertainty should decrease with the number of active users: With many users, chunks are given to the users with relatively good channels, and correspondingly small prediction uncertainty. This assumption is in general consistent with the presented results, and with other similar results that have been obtained.

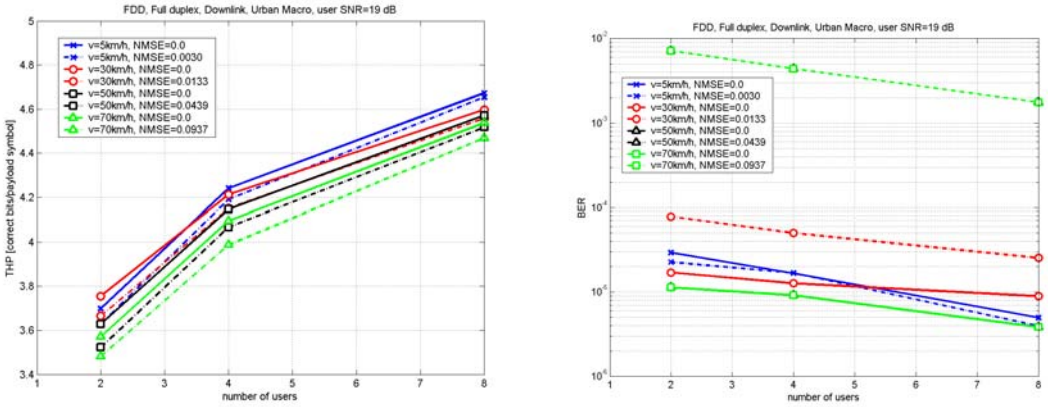


Figure 3.14: As in Figure 3.13, but for SNR = 19 dB for all users

We have so far utilized the convolutionally coded modulation scheme. The uncoded scheme using BPSK, QPSK, 16-QAM and 64-QAM should provide lower spectral efficiency, but should also be less sensitive to prediction errors. See Appendix 6.2 for its properties, when adjusted to attain maximal throughput under a maximal BER constraint 0.001 in the presence of prediction uncertainty. How these factors are played out in similar experiment setting as in Figure 3.13 (10 dB SNR) is shown in Figure 3.15. It is evident that the somewhat lower sensitivity of this scheme does not compensate for its significantly lower spectral efficiency. An uncoded scheme with additional rates, such as the one presented in Table 3.6, would perform significantly better at mid-to-high SNR, but would still be worse than the coded scheme at low SNR levels. In the following investigations, we will therefore use the convolutional coded M-QAM, instead of the uncoded scheme.

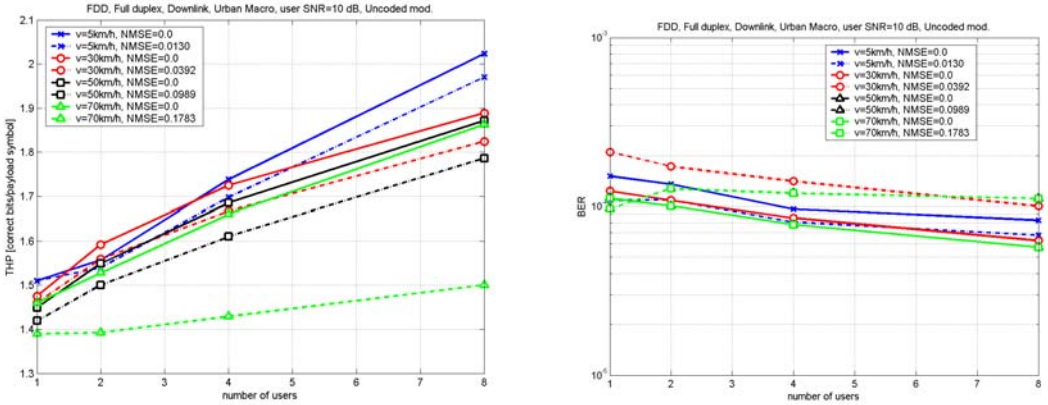


Figure 3.15: As in Figure 3.13, SNR = 10 dB, using uncoded BPSK,QPSK,16-QAM and 64-QAM

3.1.5.3 TDMA/OFDMA versus use of TDMA

We now investigate what happens when the following scheduling principle and constraint is applied: All users feed back their SNR predictions in each chunk as before. The scheduler selects the user who has the highest sum-of-rates capacity within all 104 chunks in the FDD downlink, and gives all the 104 chunks to this user. (With a full buffer assumption, the problem of this allocated resource unit being too large is ignored here.) The payload bits within the chunks are modulated and coded with chunk-specific rates, according to the feedback information. This experiment is designed to preserve the effect of the chunk-specific *link* adaptation, but to investigate the effect upon the multiuser scheduling gain that a TDMA scheduling constraint will have. The question is relevant in the investigation of different multiple access schemes.

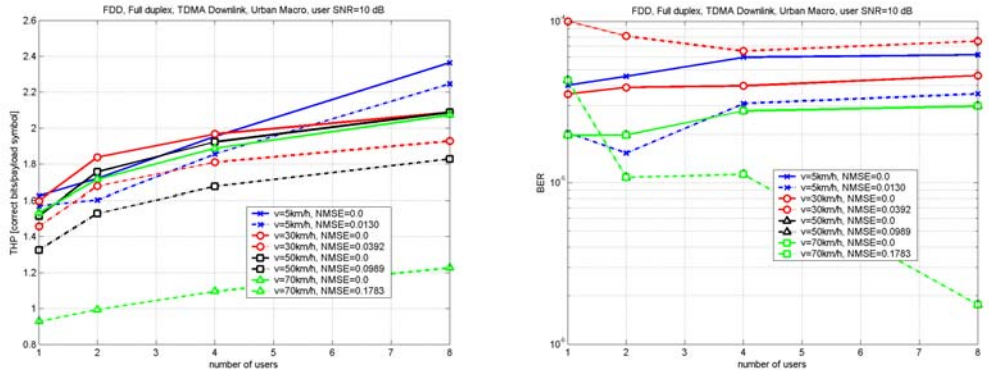


Figure 3.16: Use of TDMA scheduling constraint. Throughput and bit error rates as functions of the number of active users, all with the same average SNR of 10 dB, when scheduling the chunks to the user with best instantaneous predicted rate. FDD wide-area downlink, 5GHz carrier frequency, 0.843 ms required prediction horizon. Solid curves take channel variability within chunks into account but neglect the prediction uncertainty. Dashed curves take prediction uncertainty from Figure 3.3 into account. Compare to Figure 3.13 for the case of TDMA/OFDMA

The results are shown in Figure 3.16 and Figure 3.17 for the FDD downlink, under conditions that are otherwise identical to those used in Figure 3.13 and Figure 3.14, where all chunks could be allocated individually to different users. The results are of course unchanged for $K = 1$. However, comparing the results for $K = 8$ users who have the same average SNR, we see that the use of TDMA reduces the multiuser scheduling gain significantly. The throughput is reduced by around 22% at 10 dB and around 18 % at 19 dB. About *half of the total multiuser diversity effect is lost* when using TDMA instead of TDMA/OFDMA over the 16.25 MHz bandwidth. The reason is that there is less variability in the sum of the usable rates for the whole timeslot as compared to the rates in individual chunks. The varying fading at different subcarriers tends to average out the variability of the total TDMA channel. Less variability results in less multiuser diversity gain. However, the variability is not eliminated. A significant amount of multiuser diversity gain remains also in the TDMA case.

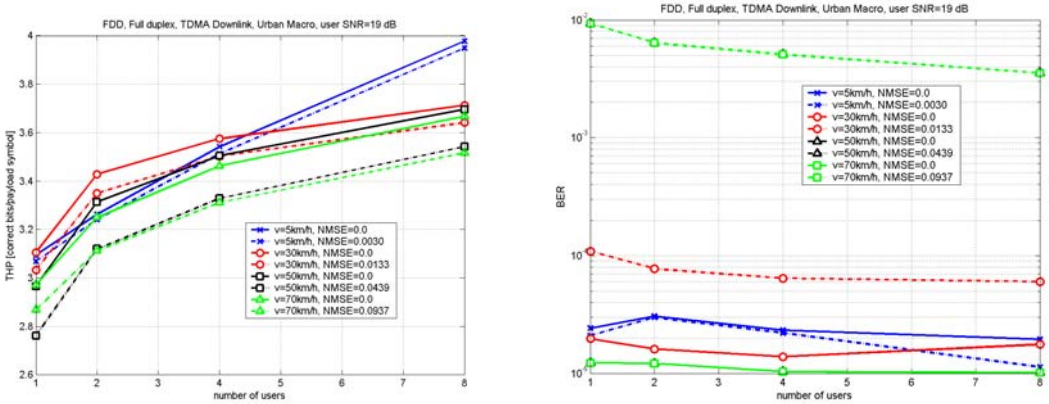


Figure 3.17: TDMA scheduling constraint. As for Figure 3.16, but for average SNR 19 dB for all users. Compare to Figure 3.14 for corresponding results for TDMA/OFDMA

The presented results represent a preliminary investigation. In subsequent investigations, a true OFDM-based TDMA scheme, which uses only one single code and modulation rate within the whole TDMA slot, needs to be compared to the chunk-adaptive TDMA scheme used above, and to the TDMA/OFDMA allocation. Furthermore, the effect on the delay statistics of having TDMA allocation instead of the more fine-grained TDMA/OFDMA allocation needs to be studied. However, more sophisticated traffic models than the full buffer model used here are required in such an investigation.

3.1.5.4 Effect of channel variability within chunks, with varying chunk size

In the design of the adaptation scheme for the wide area FDD mode, the chunks contain only 48 bits, and the overhead load is rather significant. Could the design be improved by increasing the chunk size? Investigating the interplay between the following three factors will provide the answer:

1. An increased chunk size will increase the channel variability within chunks. This will reduce the attainable performance of schemes that are basically designed based on an assumption of constant AWGN channels within chunks.
2. If the number of pilots and downlink control symbols within chunks is held constant, the downlink overhead can be reduced.
3. With a reduced total number of chunks, the uplink feedback rate is reduced. If no source coding of the information is used, the relative reduction would be proportional to the reduction of the number of chunks. If source coding is used, this relative reduction could be much smaller.

To investigate the problem, the experiment of Figure 3.13, for FDD downlinks with Urban Macro channels and users all having the same velocity and average SNR 10 dB, was repeated with the chunk widths doubled from 8 subcarriers (156.2 kHz) to 16 subcarriers (312.4 kHz). The chunk duration, and other tunings of the scheme, was held constant. The number of symbols per chunk is thus doubled from 48 to 96. With a constant number of 12 pilots and downlink control symbols, the downlink overhead factor would then be improved from 36/48 = 0.75 to 84/96 = 0.875.

The result is shown in Figure 3.18. As compared to the performance in Figure 3.13 with 48 symbols per chunk, we obtain a rather significant reduction of the spectral efficiency for the Urban Macro channel, around 9 % for one user and 13.5 % for eight users, on average over results for 5 km/h, 30 km/h and 50 km/h at 10 dB SNR. It is just barely compensated for by the improvement of the downlink overhead factor. Therefore, the extent and value of the reduction in uplink feedback data rate would determine if this rescaling is advantageous.

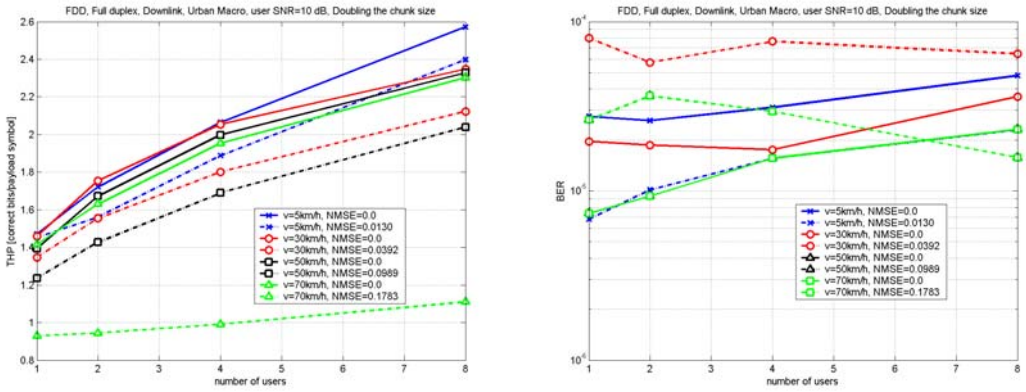


Figure 3.18: As for Figure 3.13, 10 dB SNR; but doubling the chunk widths from 8 subcarriers to 16 subcarriers in the wide-area FDD mode. Channels with Urban Macro statistics used

At the end of subsection 3.1.4, it was estimated that with efficient source coding, a feedback data rate of 4 bits per chunk would typically be adequate in the FDD uplink. If this information is carried by four coded 4-QAM control symbols in the uplink, it represents an overhead of 4/48. If the chunk sizes can be doubled, the overhead fraction is reduced to 4/96. This *absolute* reduction will thus be rather small. Furthermore, doubling of the chunk size will reduce the correlation between the SNRs, and thus reduce the effectiveness of the source coding. Therefore, more than 4 bits of feedback would be needed in the case with doubled chunks. The present results indicate that the benefits of doubling the chunk size in the FDD wide-area model would be insignificant at best, and are likely to be nonexistent.

3.1.6 Summary, conclusions and open issues

This investigation of adaptive TDMA/OFDMA has been aimed at overcoming the two key problems of adaptive transmission: The inaccurate and outdated channel state feedback and the potentially high feedback data rate. The study has been performed for a challenging test case: 5 GHz, with focus on wide area coverage over frequency selective channels, for terminals with vehicular velocities. We have furthermore focused on the FDD physical layer mode that has the smallest time-frequency chunks and the highest control overhead load.

The study leads to the following main conclusions:

- Adaptive TDMA/OFDMA transmission is feasible. No crucial show-stoppers have been identified in the present investigation.
- Predictive adaptation to the short-term fading and the frequency-domain channel variability leads to significant multiuser scheduling gains. With TDMA instead of TDMA/OFDMA, only half of these gains are realized for channels with Urban Macro statistics.
- For realistic SINR values, *transmission at 50 km/h will be feasible at 5 GHz carrier frequency* in the FDD downlink. This requires the use of channel prediction and a tight adaptation feedback loop. The limiting velocity at lower carrier frequencies is correspondingly larger.
- Predictive adaptation can use code and modulation rate boundaries adjusted so that bit error rate constraints are fulfilled in the presence of SINR prediction uncertainty. This method had been used and tested and has been shown to work in a variety of test scenarios. It enables efficient and robust co-optimisation of link adaptation and link retransmission strategies.
- The feasibility of adaptive transmission is limited by the channel prediction accuracy. For each link, this accuracy is determined by the SINR and by the velocity of the terminal. These two variables are thus the main variables that determine if flows may be adaptively allocated, or must instead use more traditional coding, interleaving and spreading to average the channel variations.
- The required channel state feedback data rate poses a potentially serious problem, in particular for the FDD wide-area downlink. An effective solution to this problem is to *feed back the predicted SNR, and source code it by a combination of transform coding in the frequency direction and subsampling in the time direction.*

While the results from this study are promising, several open issues require further investigation.

- The sensitivity of uplink TDMA/OFDMA transmission to **frequency synchronization errors** needs to be investigated. Frequency offsets result in inter-carrier interference both within chunks and between chunks allocated to different users. Interestingly, recent results presented within WINNER indicate that frequency synchronization accuracies well below 1 % of the subcarrier spacing are attainable. See Subsection 5.4.2.6 of [WIND23].
- The **size of chunks** in the FDD physical layer mode, with 36 payload constellation symbols in the downlink and a similar number in the uplink, is too small for forming a good code block length, when convolutional coding is performed block-wise. It was in Subsection 2.1.3.1 recommended to use at least 64 constellation symbols per coding block. Doubling of the chunk size was tested above, but was seen to be of doubtful value. A solution could be to perform coding over several chunks. A scheduling constraint is then imposed so that if the channel is to be given to a flow at all, at least two chunks, which accommodate equal modulation and code rate, must be allocated. The coding is then performed over the payload symbols of both chunks.
- The idea of allocating several chunks per user can be extended further, to **multi-chunk coding blocks**. We could adapt just the modulation format chunk-wise, and use coding with a few fixed rates over several chunks. Adaptive modulation would then make the channel within the multi-chunk block behave approximately like an AWGN channel with known noise power per bit. With larger code blocks, turbo coding or low density parity check (LDPC) codes become efficient. One problem when applying stronger coding schemes with low code rate is that the sensitivity to prediction errors will increase. Another problem with coding over multiple chunks is that a link-to-system interface is at present unavailable which would represent the performance of such a scheme in a system level simulator.
- The code block is naturally used as the scheduling unit and as the link retransmission unit. However, it is far from clear what **retransmission block size** is most appropriate, and also unclear on what considerations this choice should be based. Should one chose a block length

related to the IP packet size, a fixed block length or a variable block length determined by the numbers of payload bits per chunk? These issues are at present actively investigated within WINNER.

- Related to the issue of link retransmission is the following question: What is gained by using **link adaptation combined with retransmission**, rather than relying on retransmission only? The likely answer is that the use of link adaptation, with retransmission as a safety net, combines the best features of both schemes. Due to link adaptation, we gain a high throughput and low delay since the initial transmission is adjusted so that it will likely be received error-free. The presence of fast and efficient link retransmission enables the link adaptation scheme to be designed for high spectral efficiency, avoiding the need to set the bit error rate constraints at very low target levels. The presence of link retransmission provides robustness, in cases when link adaptation and channel prediction fails. The quantification of these qualitative observations, in particular in regard to the packet delay statistics, will require more sophisticated traffic models than the full buffer models utilized here.
- In the above simulation study, all users had the same average SNR and were moving with the same velocities. These idealized assumptions need to be exchanged for a more realistic system simulation setting. The same applies to the modelling of interference, which was represented by additive white Gaussian noise in the present study.
- The interplay of the multilink adaptation strategies with **more advanced resource scheduling** than the one considered here will also require more realistic system simulation studies.
- If each downlink terminal feeds back suggested code and modulation rates to be used in the next downlink transmission if that user is appointed, then the downlink control signalling is simplified. One can then broadcast just the identity number of the user appointed to each chunk. The CM rate to be used is implicitly that which was suggested by the appointed user. However, it was in Subsection 3.1.4 recommended to feed back the predicted SINR values, not the rates, since this enabled much more efficient coding of the feedback information. When the SINRs and their accuracies are available at the scheduler, this has additional advantages: Two users who can use the same CM scheme can be differentiated by their expected error rates. SINR values can furthermore be used to fine-tune the transmit power, if so desired. However, one might suspect that this scheme prevents implicit **downlink control signalling**: Would the scheduler have to broadcast not only the identity number, but also the utilized CM rate for each allocated chunk? This need not be the case. All users *could run algorithms that are identical to that run by the access point to determine the CM rates*, so explicit signalling of the rates may be unnecessary.
- The adaptation control loops were in the above studies assumed to be up and running. There is an **initial transient phase** that also needs to be considered and optimised. The terminal may have to be awakened from “sleep mode”, its oscillator frequency and time synchronization adjusted and the set of predictors needs to be initialised. Due to this rather heavy startup effort, it is likely to be inefficient to use adaptive scheduling for flows that transmit packets only rarely. Adaptive transmission is more appropriate for file transfer types of flows.
- The feasibility study above focused on the downlink. The uplink, in particular of the FDD mode has special features that need further study. One is the required precision of **slow power control**, used to equalize the received powers from different user terminals. How is the uplink channel estimation, based on overlapping pilots, affected by different received powers? The predictor has been tested for 10 dB differences in average gains, but higher values may cause problems.
- Finally, the best way to **combine MIMO transmission with adaptive TDMA/OFDMA** represents an interesting and important research topic for WINNER Phase I and Phase II. It is practical, although not globally optimal, to preserve orthogonality also when spatial dimensions are added to the time and frequency dimensions. Within a chunk, all AP antenna resources in a sector would then be allocated to the same terminal. For such systems, results so far indicated that *multiplexing*, i.e. transmitting different messages over the transmitters and separating them through channel inversion at the receiver, should be used when the channel is non-singular. Multiplexing requires no extra feedback and it preserves the strong multiuser scheduling gain that was obtained for SISO channels. This is *not* the case for e.g. Alamouti space-time coding. However, MIMO channel matrices will often be rank deficient, for example due to line-of-sight propagation environments. Multiplexing should then not be used. A switching strategy, which uses multiplexing where possible, and switches to another multi-antenna transmit scheme where appropriate, seems like the most promising combination.

3.2 Fundamental limits and adaptive OFDMA schemes for the uplink

The uplink in a single cell can be described in information-theoretic terms as a multiple access channel[13]. Since the transfer functions from the users to the base station are generally frequency-selective and the noise can be assumed to be Gaussian, the model of a Gaussian multiaccess channel with ISI as described by Cheng and Verdú applies [CV93]. The received signal at the base station of a cell with U users is

$$y(t) = \sum_{u=1}^{U} h_u(t) * x_u(t) + w(t)\,, \tag{3.9}$$

where $x_u(t)$ is the transmit signal of user u, $h_u(t)$ is the impulse response of user u, and $w(t)$ is Gaussian noise with PSD $S_w(\omega)$. The received signal can be expressed equivalently in the frequency domain as

$$Y(\omega) = \sum_{u=1}^{U} H_u(\omega) X_u(\omega) + W(\omega) \tag{3.10}$$

The channel gain to noise ratio (CNR) of the channel is thus defined as

$$T_u(\omega) = \frac{|H_u(\omega)|^2}{S_w(\omega)} \tag{3.11}$$

While the formulations (3.9) and (3.10) hold for all multiple access schemes, the channel model considerably simplifies by applying OFDM, which decomposes the frequency-selective channels into a set of parallel subchannels for each user:

$$y_n(k) = \sum_{u=1}^{U} H_{n,u} \cdot x_{n,u}(k) + w_n(k), \quad n = 1,\ldots,N \tag{3.12}$$

This is valid as long as the guard interval is longer than the longest delay spread and all users are synchronized in time and frequency. In pure OFDMA the subchannels are exclusively allocated to one user, i.e. if user u transmits over subchannel n all other users have to set this subchannel to zero.

3.2.1 Cell capacity

The capacity region of a Gaussian multiaccess channel with ISI and a method for achieving the maximum sum rate is described in [CV93]. In the following we define as *cell capacity* the sum of the channel capacities of all users. The transmit power of user u is denoted by

$$p_u = \mathrm{E}[|x_u(t)|^2] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_u(\omega) d\omega \tag{3.13}$$

where $S_u(\omega)$ is the PSD of user u's transmit signal. These transmit powers are limited to $p_{\max}$. For a single link with a frequency-selective channel, the capacity can be achieved by shaping the transmit PSD according to the water-filling theorem [Gal68], [CT91]. The extension of this result to the multi-user case is based on the introduction of the *equivalent channel* by $\hat{H}_u(\omega) = H_u(\omega) \big/ \sqrt{\lambda_u}$, which leads to the equivalent PSD $\hat{S}_u(\omega) = \lambda_u S_u(\omega)$ [CV93]. This scaling of the channel transfer function allows to adjust the water-filling diagrams of each user and combine them to a single multi-user water-filling diagram as depicted in Figure 3.19. Analytically, this is expressed as

$$\begin{aligned} \hat{S}_u(\omega) &= \begin{cases} \left[1 - \frac{\lambda_u}{T_u(\omega)}\right]^+ & \text{if } \frac{\lambda_u}{T_u(\omega)} \le \frac{\lambda_j}{T_j(\omega)} \ \forall j \neq u \\ 0 & \text{otherwise} \end{cases} \\ \lambda_u p_{\max} &= \tfrac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{S}_u(\omega) d\omega \end{aligned} \tag{3.14}$$

[13] Note that the downlink corresponds to a *broadcast channel*. This means that for the calculation of the fundamental limits (channel capacity), uplink and downlink are treated as rather distinct cases while for practical implementation the difference is less pronounced.

The multipliers λ_u are uniquely defined by this equation and can in principle be determined by numerical algorithms. A heuristic solution, presented in [MPS02], [WIND26], achieves an approximate solution. However, although this solution was found to perform well in numerous simulations, it lacks a formal proof of convergence.

The most important implication of equation (3.14) is that it gives an indication of the optimum multiple access scheme for achieving the maximum sum rate: The condition in (3.14) indicates that the available bandwidth is partitioned in frequency and each frequency band is allocated to the user with the smallest equivalent inverse CNR $\lambda_u / T_u(\omega)$. The multiple access scheme which achieves the cell capacity is thus FDMA, the bandwidth is partitioned among the users and the transmit PSDs are calculated according to (3.14). This is visualized by Figure 3.19, which is easy to draw; however it can also be observed from this figure that the calculation of the multipliers λ_u is not straightforward and highly nonlinear.

Once the transmit PSDs are known, the cell capacity is given by

$$C = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \log_2 \left(1 + \sum_{u=1}^{U} S_u(\omega) T_u(\omega) \right) d\omega \tag{3.15}$$

A good approximation to a general FDMA scheme is OFDMA, provided the size of the frequency bin is smaller than the smallest frequency band according to (3.14). Since an OFDM system is normally designed in such a way that the frequency bin is small in comparison to the channel variation, the approximation due to the discretisation of the frequency axis causes negligible loss.

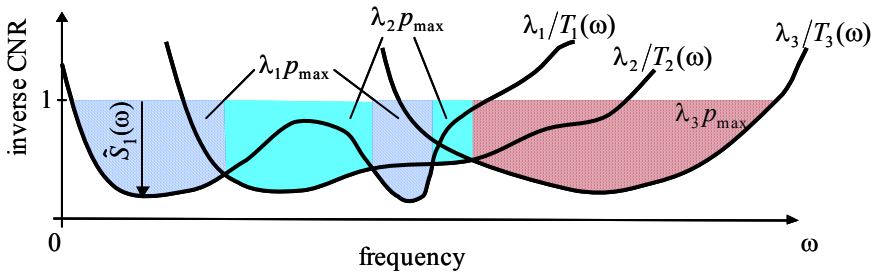


Figure 3.19 Multi-user water-filling diagram. The multipliers λ_u have to be scaled such that the water-levels for all users are unity. This allows to combine the diagrams of all users into a single one. The frequency bands are assigned to the user with lowest equivalent inverse CNR $\lambda_u / T_u(\omega)$.

An even more general formulation of the multiple-access channel has been provided recently by Yu et al. [YRB+04] and holds for a Gaussian multiple-access channel with vector inputs and vector output:

$$\mathbf{y} = \sum_{u=1}^{U} \mathbf{H}_u \mathbf{x}_u + \mathbf{w} \tag{3.16}$$

The *iterative water-filling* (ITW) algorithm described by Yu et al. maximises the sum capacity under a power constraint for each user. This leads to a solution, which is analogue to the multi-user water-filling theorem, but holds for the channel model (3.16) instead of (3.9). This formulation has the advantage that it can be applied directly to discrete frequencies and that a formal convergence proof exists.

The signal and noise covariance matrices are defined as $\mathbf{S}_u = \mathrm{E}[\mathbf{x}_u \mathbf{x}_u^{\mathrm{H}}]$ and $\mathbf{S}_w = \mathrm{E}[\mathbf{w}\mathbf{w}^{\mathrm{H}}]$, respectively, and the maximum sum rate, or *cell capacity*, is given as

$$C = \log_2 \left| \sum_{u=1}^{U} \mathbf{H}_u \mathbf{S}_u \mathbf{H}_u^H + \mathbf{S}_w \right| - \log_2 \left| \mathbf{S}_w \right| \tag{3.17}$$

The ITW algorithm finds the signal covariance matrix $\mathbf{S}_u$, which solves the optimisation problem:

$$\begin{aligned} &\text{maximize} \quad C \\ &\text{subject to} \quad \mathrm{trace}(\mathbf{S}_u) \le P_{\max}; \quad \mathbf{S}_u \text{ is positive semidefinite} \end{aligned} \tag{3.18}$$

This rather general formulation simplifies significantly for OFDMA and results in a surprisingly simple algorithm. If we define the transmit and receive signals as $\mathbf{x}_u = (x_{1,u}, \ldots, x_{N,u})^{\mathrm{T}}$, $\mathbf{y} = (y_1, \ldots, y_N)^{\mathrm{T}}$, the channel matrix $\mathbf{H}_u = \mathrm{diag}(H_{1,u}, \ldots, H_{N,u})$ and the noise as $\mathbf{w} = (w_1, \ldots, w_N)^{\mathrm{T}}$, the general formula (3.16)

is equivalent to (3.12). This makes the ITW algorithm applicable for OFDMA: The covariance matrices are diagonal matrices of the signal and noise power, respectively:

$$\begin{aligned} \mathbf{S}_u &= \mathrm{diag}\left(p_{1,u},\ldots,p_{N,u}\right), \quad \text{with } p_{n,u} = \mathrm{E}[|\, x_{n,u}\,|^2] \\ \mathbf{S}_w &= \mathrm{diag}\left(\sigma_1^2,\ldots,\sigma_N^2\right), \quad \text{with } \sigma_n^2 = \mathrm{E}[|\, w_n\,|^2] \end{aligned} \tag{3.19}$$

Hence, the power constraint simplifies to

$$\mathrm{trace}(\mathbf{S}_u) = \sum_{n=1}^{N} p_{n,u} \le p_{\max} \tag{3.20}$$

and the cell capacity (rate sum) to

$$C = \sum_{n=1}^{N} \log_2\left(1 + \frac{1}{\sigma_n^2}\sum_{u=1}^{U} |\, H_{n,u}\,|^2\; p_{n,u}\right), \tag{3.21}$$

which is the sum of the channel capacities per subchannel. This leads to the following formulation of the ITW algorithm:

Algorithm Iterative water-filling for OFDMA

1 $p_{n,u} = 0 \quad \forall\, n = 1,\ldots,N; u = 1,\ldots,U$

2 repeat

3 for $u = 1$ to U

4 $z_n = \sum_{\substack{j=1 \\ j\neq u}}^{U} |\, H_{n,j}\,|^2\; p_{n,j} + \sigma_n^2$

5 $\mathbf{p}_u = \underset{\mathbf{q}=(q_1,\ldots,q_N)}{\arg\max} \left\{ \sum_{n=1}^{N} \log_2\left(|\, H_{n,u}\,|^2\; q_n + z_n\right) \right\}$, subject to $\sum_{n=1}^{N} q_n \le p_{\max}, q_n \ge 0$

6 end

7 until desired precision is reached

This algorithm calculates the power allocation for all users and achieves asymptotically the maximum sum rate. Yu et al. show that the convergence of the algorithm is fast and that after very few iterations the solution is close to optimum [YRB+04]. In each step, it considers the transmit signals of all other users as noise (line 4) and calculates the single-user water-filling solution for user u (line 5). The objective function in line 5 can be replaced equivalently by

$$\sum_{n=1}^{N} \log_2\left(1 + \frac{|\, H_{n,u}\,|^2\; q_n}{z_n}\right), \tag{3.22}$$

which is the channel capacity for a single link with N parallel subchannels with transmit powers q_n and noise powers z_n, i.e. the CNR is given by $T_n = |\, H_{n,u}\,|^2 / z_n$. The optimisation problem in line 5 is hence the single-link rate maximization problem:

$$\begin{aligned} \text{maximize} \quad & \sum_{n=1}^{N} \log_2\left(1 + q_n T_n\right) \\ \text{subject to} \quad & \sum_{n=1}^{N} q_n \le P, q_n \ge 0 \end{aligned} \tag{3.23}$$

which has the well-known solution (e.g. [CT91], [Gal68])

$$q_n = \left[q_0 - \frac{1}{T_n} \right]^+, \quad q_0 \text{ such that } \sum_{n=1}^{N} q_n = p_{\max} \tag{3.24}$$

This solution is hence assigned to the power vector $\mathbf{p}_u = (p_{1,u} \cdots p_{N,u})$ of user u.

The ITW algorithm provides the power allocation for all users, which allows to calculate the cell capacity according to (3.21). Note that this cell capacity depends strongly on the channel transfer functions and thus is dominated by the path loss of the users, which in turn is determined by the distance of the users to the base station. In the following, we apply a deterministic distance distribution, which was suggested in [Eri04] and also applied in [WIND26]. The users are placed in the cell from 10% of the cell radius to the cell border and the user density is assumed to be constant. This gives the distances

$$d_u = \left(1 + 9\sqrt{\frac{u-1}{U-1}}\right)\frac{r_{\text{cell}}}{10}, \quad u = 1,\ldots,U \tag{3.25}$$

Figure 3.20 shows exemplary channel transfer functions and the corresponding power allocations for a short-range scenario with 4 users. Most subcarriers are allocated to the user which is closest to the base station, whereas the users which are further away concentrate their transmit power in few subchannels. Note that if the power constraint were not per user but globally, all power would be allocated to the closest user and the allocation problem would become a trivial one.

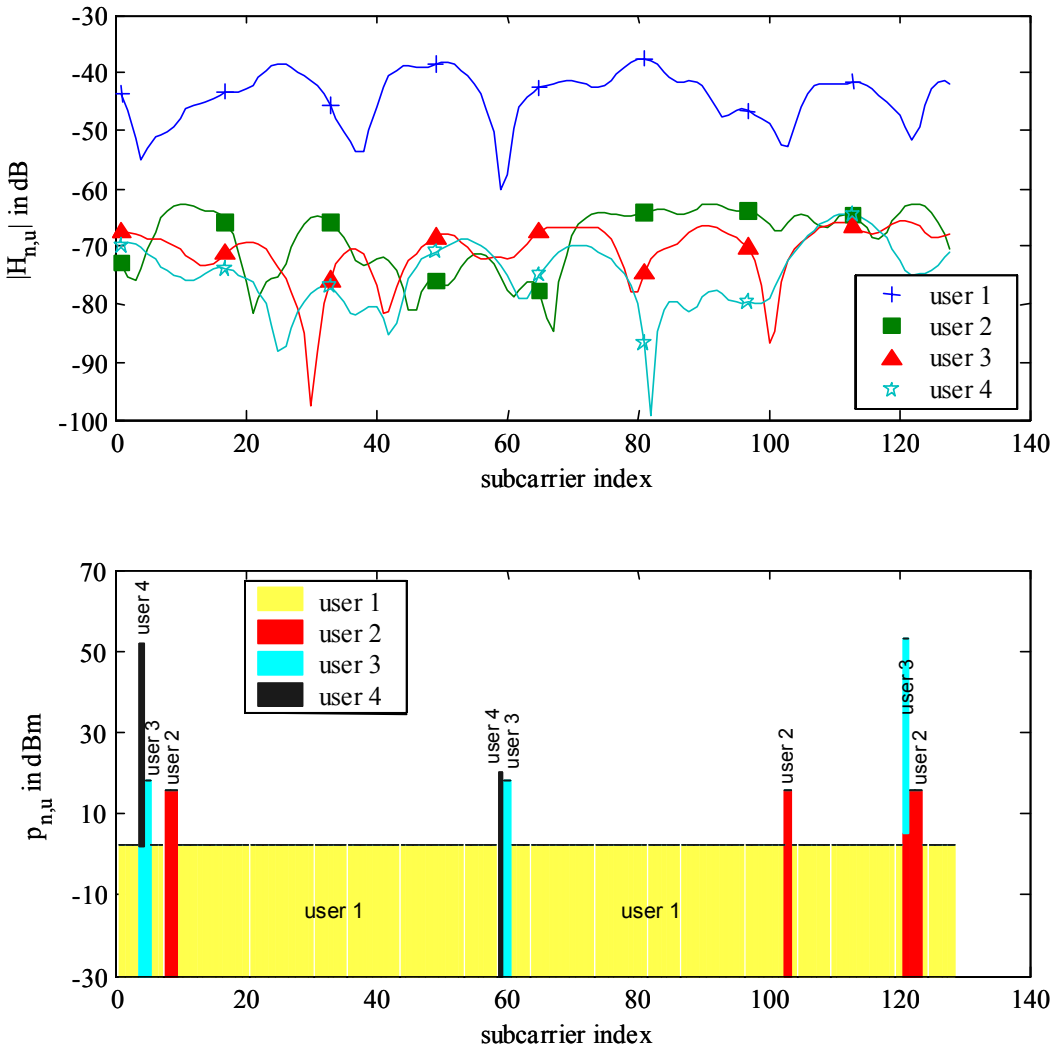


Figure 3.20: Channel transfer functions and power allocations for 4 users

We can also observe in Figure 3.20 that the subchannels are not exclusively allocated to only one user. Although the multi-user water-filling solution (3.14) requires pure FDMA, the discretization of the frequency axis into a finite number of bins implicates the sharing of some subchannels (see also [Dig96]). The resulting channel capacities for this example are listed in Table 3.7.

Table 3.7: Channel capacities of each user and cell capacity

user	1	2	3	4	**sum**
capacity [bps/Hz]	24.4	1.1	0.69	0.36	**26.6**

Figure 3.21 shows the average cell capacities as a function of the number of users. The users are distributed in the cell with constant density according to (3.25) and the cell capacity is averaged over 100 channel realisations.

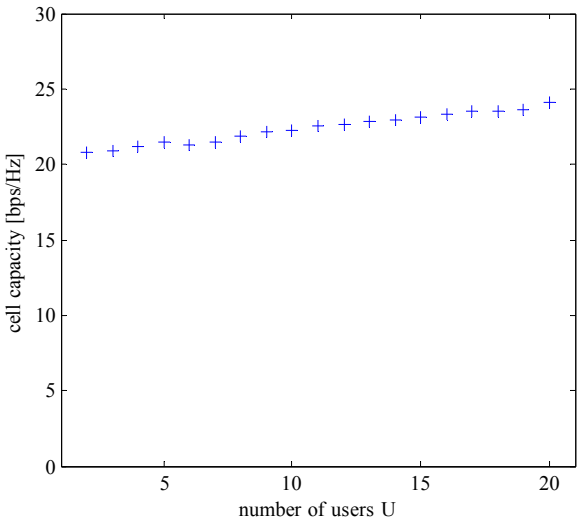


Figure 3.21: Average cell capacity as a function of the number of users. The users are evenly distributed in the cell, the transmit power per user is constant.

3.2.2 Subcarrier and bit and power allocation

In this section, we describe some algorithms for subcarrier, bit and power allocation in the OFDMA uplink. The following optimisation criteria will be considered:

- Maximum sum bitrate
- Minimum sum transmit power
- Minimum number of used subcarriers (minimum spectrum usage)

These optimisation problems can be simplified dramatically by a separation into two steps:

1. Subcarrier allocation: allocate to each user a set of dedicated subchannels
2. Bit and power allocation: for the set of dedicated subchannels, a single-user algorithm can be applied to derive the bit and power allocation for each subchannel

For the second step, apart from the optimum Hughes-Hartogs algorithm described in section 2.1.1, any other single-user bitloading algorithm can be applied.

3.2.2.1 Adaptation for maximum bitrate

The ITW algorithm calculates the optimum transmit covariance matrices for each user in order to maximize the sum capacity. For OFDMA, where each subcarrier is QAM modulated with a restricted number of bits per symbol, additional constraints apply. The optimisation problem with the basic implementation constraints is given by

$$\begin{aligned} \text{maximize} \quad & \sum_{u=1}^{U}\sum_{n=1}^{N} b_{n,u} \\ \text{subject to} \quad & \sum_{n} p_{n,u} \le p_{\max},\ p_{n,u} \ge 0 \\ & b_{n,u} \in \mathrm{B} \\ & \sum_{u=1}^{U} \operatorname{sgn}(p_{n,u}) \in \{0,1\} \\ & \Pi_{b,u} \le \Pi\beta_{b,\max} \end{aligned} \tag{3.26}$$

In addition to the power constraint, the number of bits per QAM symbol must be taken out of a finite, predefined set B, subchannels cannot be shared by several users and the bit error probability may not exceed a certain threshold. Rather than trying to solve this highly complex optimisation problem, we adapt the ITW algorithm to consider the additional constraints.

The main link between channel capacity and the achievable bitrate with QAM modulation is given by the gap approximation: the SNR gap Γ is the additional SNR that is required with QAM in comparison to the Shannon limit in order to achieve a symbol error probability Π_S:

$$\Gamma = \frac{1}{3}\left[Q^{-1}\left(\frac{\Pi_S}{4}\right)\right]^2 \tag{3.27}$$

The first step to adapt the IWT algorithm is to introduce the SNR gap in the calculation of the power allocation vector in line 5 of the ITW algorithm description. This can be done by simply defining the CNR for the optimisation problem (3.23) and its solution (3.24) as

$$T_n = \frac{|H_{n,u}|^2}{\Gamma \cdot z_n} \tag{3.28}$$

The algorithm provides power allocation vectors $\mathbf{p}_u$ for all users, which can be used to derive the subcarrier allocation by assigning a subchannel to the user with the highest allocated power:

$$a_n = \arg\max_u \{p_{n,u}\}, \tag{3.29}$$

where $\mathbf{a} = \{a_1, \ldots, a_N\}$ denotes the subcarrier allocation vector. This step eliminates the allocation of multiple users to one subchannel. Now, each user has a set of dedicated subchannels $\mathrm{N}_u = \{n = 1, \ldots, N : a_{n,u} = u\}$ and a single-user bitloading algorithm can be used to calculate the bit and power allocation for all users and subcarriers. In this step, the finite set of possible code rates will be considered.

The sum bitrate, which will be achieved by this method, will in all cases be upper-bounded by the cell capacity, described in section 3.2.1.

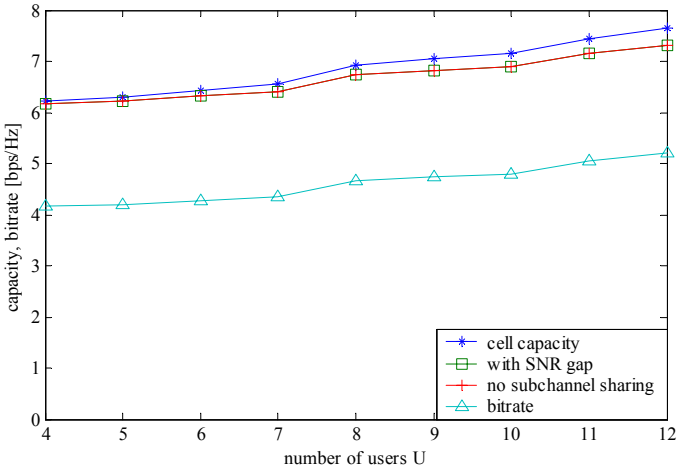


Figure 3.22: Maximum achievable bitrate in comparison with cell capacity and implementation constraints

In Figure 3.22, the difference between cell capacity and achievable bitrate, as well as the capacity loss due to the additional constraints are shown. The cell capacity is calculated with the ITW algorithm without any other constraint than a maximum transmit power of $p_{\max} = 200$ mW per user. This capacity is reduced, although not significantly, by introducing the SNR gap into the algorithm. In Figure 3.22, $\Gamma = 4$ was assumed, which corresponds to an uncoded symbol error probability of $\Pi_S = 10^{-3}$. The restriction to allocate a subchannel to not more than one user, involves no appreciable capacity loss. Finally, the restriction to a set of possible QAM alphabets involves a significant gap towards the cell capacity. In this example, the QAM constellations as defined within WINNER WP2 Task 2 for the short range cellular case have been applied: BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM, i.e. B = {1,2,4,6,8}.

It has to be noted that this still corresponds to an idealized setting, since many practical implementation restrictions and losses are not considered:

- imperfect, delayed or partial CSI
- user movement and thus time-variability of the channel
- non-linearities of the amplifier due to high PAPR
- sub-optimum bitloading algorithms
- limited packet sizes

It is expected that these effects will reduce the achievable bitrate significantly.

3.2.2.2 Adaptation for minimum transmit power

The previous optimisation criterion provides a valid upper bound for the achievable sum bitrate, but the distribution of the sum bitrate among the users follows no fairness criterion. A more realistic assumption is that each user has in addition to his limited transmit power a target bitrate $b_{\min}(u)$ and that the subcarrier, bit and power allocation is done to minimize the total transmit power:

$$\begin{aligned} \text{minimize} \quad & \sum_{u=1}^{U}\sum_{n=1}^{N} p_{n,u} \\ \text{subject to} \quad & \sum_{n} b_{n,u} \geq b_{\min}(u),\ b_{n,u} \in \mathrm{B} \\ & \sum_{n} p_{n,u} \leq p_{\max},\ p_{n,u} \geq 0 \\ & \sum_{u=1}^{U} \operatorname{sgn}(p_{n,u}) \in \{0,1\} \\ & \Pi_{b,u} \leq \Pi_{e,\max} \end{aligned} \tag{3.30}$$

The fundamental algorithm for this criterion has been formulated by Wong et al. [WCL+99] and is based on Lagrange optimisation. Wong's algorithm provides the best-known solution, but due to its complexity and slow convergence it is not practical for implementation. This problem has been adressed by various authors, beginning with the original authors themselves [WTC+99]. Another step towards a fast implementation was made by Yin and Liu [YL00] who partioned the task in two steps. This idea has been taken up by [PMS02], [KLL03] for further simplifications. Since an algorithm for this case is already described in detail in [WIND26], it is not reproduced here.

This optimization criterion is mentioned here for completeness and the fact that no simulation results have been produced is not meant to rest importance from it. However, the outcome of simulations with this optimisation criterion depends heavily on the chosen minimum bitrates $b_{\min}(u)$. Since it is not clear yet what values should be assumed for this parameter, the results would be not very meaningful or might even be misleading.

3.2.2.3 Adaptation for maximum spectral efficiency

The algorithm described in this section minimizes the number of used subcarriers while considering individual bitrate, power and PSD constraints [AP05]. There may be several reasons to allocate only as few subcarriers as necessary in a multi-user OFDM system. The unused subcarriers can be reserved for future users that might join the system without causing any reallocation to the existing users; or they might even be made available to another system, which might share this part of the spectrum without

causing interference to the primary system[14]. This approach is in line with recent advances like *spectrum pooling* [WJ04] and *spectrum agile radios* [MZC+04].

The `SpecEff` algorithm [AP05] minimizes the number of allocated subcarriers under the constraints of a minimum bitrate $b_{\min}(u)$ and a maximum transmit power $p_{\max}$ per user. Additionally, the transmit power per subchannel is limited to $S_{\max}$, which corresponds to a PSD mask. As in the other algorithms, each subchannel is assigned to only one user. Denoting by $a_{n,u}$ the allocation coefficient, which is $a_{n,u} = 1$ if subcarrier n is assigned to user u and $a_{n,u} = 0$ otherwise, we can formulate the optimisation problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{u=1}^{U}\sum_{n=1}^{N} a_{n,u} \\ \text{subject to} \quad & \sum_{n} b_{n,u} \geq b_{\min}(u),\ b_{n,u} \in \mathrm{B} \\ & \sum_{n} p_{n,u} \leq p_{\max},\ 0 \leq p_{n,u} \leq S_{\max} \\ & \sum_{u=1}^{U} a_{n,u} \in \{0,1\} \\ & \Pi_{b,u} \leq \Pi_{e,\max} \end{aligned} \tag{3.31}$$

The `SpecEff` algorithm, which calculates an (approximate) solution for this optimisation problem, can be described by the following steps:

1. `Sort all users according to their mean CNR in ascending order. The user with the worst average channel will be processed first.`
2. `for` $u = 1$ `to` U
3. `Denote by Φ the set of all not yet assigned subcarriers. This pool is sorted such that the subcarriers with higher CNR appear first.`
4. `Set` $m = 1$
5. `Apply a single-user bitloading algorithm to the first` m `subchannels of Φ.`
6. `If the bitrate, power and PSD constraints can be fulfilled with these` m `subcarriers, continue with step 2 for the next user. If not, increment` m `and continue with step 5.`

A more formal description along with examples can be found in [AP05].

3.2.3 Conclusions

The algorithms presented in this section optimise the subcarrier, bit and power allocation in a OFDM-based multi-user environment with respect to different basic criteria like bitrate, transmit power and required bandwidth. The transmission parameters are adapted to the channel, which is assumed to be static and perfect CSI at the transmitter is available. The adaptation will perform worse if the CSI is incaccurate or delayed and will cease to work for fast varying channels.

The described method to calculate the cell capacity can be used as an evaluation tool in order to determine the fundamental limits for the aggregated bitrate in the uplink. Basic implementation constraints like the set of available code and modulation rates can be included in this method in order to derive a tight upper bound on the sum bitrate. This upper bound can serve as a reference for other methods or simply show the limits of a given multiple-access channel.

The considered optimisation criteria lead to tractable problems and algorithms with reasonable complexity. However, they do not reflect the complex requirements that will appear in a real system. Especially in light of the expected complex mix of different traffic type, the above optimisation criteria seem overly simplified. On the other hand, it is not yet clear which traffic model would have to be applied to formulate adequate optimisation criteria. In any case, the traffic in a real system will be a mix of several traffic classes with very different properties. It will remain a challenge to jointly adapt the multi-access, coding and modulation to both the channel *and* the traffic type.

Another item to be investigated is the extension of this analysis and algorithm to multiple cells.

[14] There are some unresolved issue in the area of inter-system spectrum sharing, which are not addressed here, but are investigated in the WINNER project in WP6.

3.3 Adaptive MC-TF-CDMA scheme

A great number of link adaptation algorithms have been suggested and investigated in the past for single-user OFDM systems. Typically, these algorithms are based on the water-filling theorem [Pro02] and are referred to as bit and power loading algorithms. In most cases they assume that a fixed Bit Error Rate performance is required and either minimize the transmit power, while guaranteeing a fixed user rate, or maximize the user rate for a fixed transmit power.

For the case of multi-user systems the problem of link adaptation is much more difficult. Theoretical foundations of adaptive data transmission in downlink have been established in [GE01] and [LG01]. Basically, it is shown there that the *total* capacity of the channel can be achieved by frequency division multiple access (FDMA), but this does not ensure any QoS for individual users. In order to achieve some kind of fairness the users have to share the channel, i.e. more than one user may have to transmit the data over the same sub-carriers. This can be implemented by means of TDMA or CDMA.

An algorithm to perform adaptive user-to-sub-carrier allocation with bit and power loading has been proposed in [WIND26] for the downlink of multi-carrier CDMA systems with spreading both in time and frequency domain, referred to as MC-TF-CDMA systems. In particular, the algorithm yields an optimisation of the sub-carrier sharing that enables a considerable performance gain with respect to OFDMA systems, where each sub-carrier is exclusively assigned to one user [WIND26], [TCS04]. Therefore it is desired to further pursue this approach in the development of the WINNER air interface.

At this aim, we investigate here issues related to its implementation such as the optimisation complexity, the amount of auxiliary information to be transmitted, the sensitivity to the imperfectness of the channel state information and to channel time-variations under the assumption of TDD duplex mode.

Sections 3.3.1 present an overview of the considered system. The proposed link adaptation approach is reviewed in Section 3.3.2. Assessment of its performance and limitations is presented in Section 3.3.3. Finally, some conclusions are drawn in Section 3.3.4.

3.3.1 MC-TF-CDMA System Model

Let us consider a multi-carrier downlink transmission system based on OFDM modulation. In frequency domain the signal received by user *k* over sub-carrier *i* for the *j*th OFDM symbol can be expressed as

$$r_{ki}^{(j)} = \mu_{ki} s_i^{(j)} + \eta_{ki}^{(j)}, \qquad k = 1..K, \;\; i = 0..N-1, \tag{3.32}$$

where K is the number of active users and N is the total number of sub-carriers, $s_i^{(j)}$ is the signal transmitted by the base station over the ith sub-carrier in the j-th OFDM symbol period, μ_{ki} is the k-th users channel transfer function at sub-carrier i and $\eta_{ki}^{(j)}$ represents additive white Gaussian noise (AWGN) with variance σ^2. It is also helpful to introduce the sub-carrier channel-to-noise ratio (*CNR*) defined as

$$\xi_{ki} = \frac{|\mu_{ki}|^2}{\sigma^2}. \tag{3.33}$$

The sub-band channel-to-noise ratio may be defined as the (geometric) average of CNRs experienced by constituent sub-carriers.

Many different ways are known to map users' data to downlink signals s_i. We consider code division multiple access (CDMA) with spreading both in time and frequency, i.e. the transmitted signal is given by

$$s_{w_{pi}}^{(jL_t+s)} = \sum_{k=1}^{K} V_{kp} y_{kp}^{(j)} a_{k,iL_t+j}, \qquad j = 0\ldots L_t-1,\; i = 0\ldots L_f-1,\; p = 0\ldots P-1, \tag{3.34}$$

where $a_k = (a_{k0},...,a_{k,L_fL_t-1})$ is the spreading sequence used by user *k,* w_{pi} is the i-th sub-carrier in the p-th set of sub-carriers W_p (sub-band), L_t and L_f are the spreading factor in time and frequency, respectively, V_{kp} is the k-th user gain over the p-th sub-band, $y_{kp}^{(j)}$ is the complex modulated symbol transmitted by the base station to the k-th user over the p-th sub-band in the j-th OFDM symbol. We

assume that the set of sub-carriers has been partitioned into a number of disjoint sub-bands $W_1 \ldots W_P$ of size L_f, consisting of adjacent sub-carriers. If each user transmits data over all sub-bands, the total number of users is limited by $L=L_tL_f$ under the assumption of orthogonal spreading sequences. Notice that each sub-set W_p associated with L_tL_f orthogonal spreading sequences can be considered as a set of L_tL_f sub-channels. Such a system has been denoted in [WIND26] as multi-carrier time-frequency CDMA (*MC-TF-CDMA*).

3.3.2 Multi link adaptation algorithm

Let ρ_{kp} be the fraction of the p-th sub-band assigned to the k-th user, R_k be the number of bits to be transmitted by the k-th user within one OFDM symbol and $f(c) = \Gamma(2^c - 1)$ be the function specifying the SNR needed to achieve a particular error rate using some coding/modulation method with data rate c. Here Γ denotes the capacity gap which depends on the target error probability. Let us assume that sub-band sharing is implemented by means of CDMA, i.e. ρ_{kp} denotes the fraction of spreading sequences assigned to the k-th user. Let us consider the problem of finding sub-band sharing factors ρ_{kp}, coding/modulation formats c_{kp} and user gains V_{kp} such that the total transmitter power needed to achieve the target BER is minimized, while guaranteeing for each user its bit rate R_k. It can be shown that, after some derivations, the solution of this optimisation problem must satisfy the following system of non-linear equations and inequalities [TCS04]:

$$0 = (\beta_p^{(k)} - \beta_p)\rho_{kp} \tag{3.35}$$

$$R_k = \sum_{p=0}^{P-1} \rho_{kp} f'^{-1}(\lambda_k \xi_{kp}) \tag{3.36}$$

$$1 = \sum_{k=1}^{K} \rho_{kp} \tag{3.37}$$

$$\beta_p \le \beta_p^{(k)} = \frac{f(f'^{-1}(\lambda_k \xi_{kp})) - \lambda_k \xi_{kp} f'^{-1}(\lambda_k \xi_{kp})}{\xi_{kp}} \tag{3.38}$$

The following algorithm has been proposed in [WIND26] to solve this optimisation problem:

1. Produce some initial sub-carrier allocation $\{\rho_{kp}\}$. This can be done for example by assigning sub-carriers according to the single-user waterfilling theorem. Compute λ_k from (3.36) and substitute it into (3.38) to obtain $\beta_p^{(k)}$.
2. Find the worst allocated sub-carrier, i.e. the one given by $\arg\max_i \max_{k:\rho_{kp}>0} (\beta_p^{(k)} - \beta_p)$.
3. Reduce by $1/L$ the fraction of this sub-carrier $\rho_{k_{worst},p}$ occupied by the user k_{worst} yielding the greatest $\beta_p^{(k)}$, and increase by the same value the fraction $\rho_{k_{best},p}$ of the sub-carrier assigned to the user k_{best} with the smallest $\beta_p^{(k)}$.
4. Repeat steps 3 and 4 until the specified number of iterations J is exceeded or the value $\Delta = \max_{p,k:\rho_{kp}>0} (\beta_p^{(k)} - \beta_p)$ stops decreasing.

Having computed ρ_{kp} one can assign users to spreading sequences on each sub-band. This results in a number of equivalent sub-channels. Hence conventional single-user bit and power loading can be used to distribute power over these equivalent sub-channels and find appropriate coding/modulation formats for them. At this purpose, the algorithm suggested in [KRJ00] has been used here.

3.3.3 Assessment of complexity, performance and limitations

3.3.3.1 Simulation assumptions

WSSUS channel model [Hoh92] with pre-defined power delay profile according to [WIND23] has been used for all simulations. Uncoded M-QAM data transmission has been assumed and the target BER has been set equal to $2*10^{-3}$.

3.3.3.2 Auxiliary information

An important issue in any adaptive data transmission system is the amount of auxiliary information to be transmitted.

The implementation of the adaptive MC-TF-CDMA transmission approach described in Section 3.3.2 requires, in particular, the following auxiliary data:

1. Assignment of users to sub-bands, i.e. lists of users transmitting data over each sub-band.
2. Modulation/coding formats for each user and each sub-band.
3. User gains for each user and each sub-band.

Observe that user gains are selected so that the SNR at the receiver side is sufficient for achieving the target error rate and the function *f(c)* provides a one-to-one mapping from data rate *c* to this SNR. Hence user gains V_{kp} depend only on modulation/coding format c_{kp} and channel-to-noise ratio ξ_{kp} for user *k* and sub-band *p*, i.e. $V_{kp}^2 \xi_{kp} = f(c_{kp})$. Therefore the receiver can deduce $V_{kp,}$ given c_{kp} and estimated CTF. Under this assumption, the total amount of auxiliary information to be broadcasted by the base station can be estimated as

$$A = \frac{N}{L_f} L(\log_2(K) + \log_2 M_{\max}), \tag{3.39}$$

where M_{max} is the total number of different modulation formats used. Preliminary results indicate that the amount of auxiliary information can be reduced at least by a factor of 4 if appropriate source coding techniques are applied.

Another important issue in the design of an adaptive system is obtaining accurate channel state information (CSI) to be used by the optimisation algorithm. The CSI can be estimated by the user terminals and communicated back to the base station, or, in cases where channel reciprocity in uplink and downlink can be assumed, as in TDD systems, the CSI can be estimated at the base station on the basis uplink received symbols. The latter approach reduces the amount of auxiliary information by avoiding feedback signalling as well as the channel estimation delay. The amount of auxiliary information needed for uplink channel estimation in the considered system will be assessed in Section 3.3.3.4.

Furthermore, it has to be observed that the considered user gain estimation at the receiver may cause considerable performance degradation in case of imperfect channel state information as will be shown in Section 3.3.3.4.3.

3.3.3.3 Complexity

This link adaptation approach requires the following steps:

1. Evaluation of sub-carrier sharing factors performed by means of an iterative optimisation algorithm with *maximal* number of iterations denoted by *J*.

2. Evaluation of sub-carrier gains V_{kp} and modulation formats c_{kp} using conventional single-user bit and power loading algorithms for each *k*.

Hereafter, we will assume that the user-to-sub-carrier allocation algorithm proposed in [WIND26], [TCS04] as recalled in Section 3.3.2, is used for the first step, while the algorithm given in [KRJ00] is used for the second step.

Let us consider the computational overhead caused by the first step, hereafter referred to also as optimisation complexity. In order to avoid dependency on a particular simulation environment it is reasonable to consider the relative optimisation time defined as

$$\tau = \frac{T_1}{T_1 + KT_2} \tag{3.40}$$

where T_1 is the CPU time required by the first step and T_2 is the CPU time required for the second step.

The optimisation complexity depends on the number of users K, number of sub-bands P and the maximal number of iterations J. By adjusting the latter parameter one can select a tradeoff between the optimisation complexity and the quality of the achieved solution, that can be expressed as the relative transmit power required when the adaptive approach is used, defined as

$$P_r = \frac{E\left[\sum_{i=0}^{N-1} |s_i|^2\right]}{N\sigma^2} \tag{3.41}$$

The following figures present simulation results illustrating the impact of parameter J on the system performance in the different considered propagation scenarios and for different combinations of spreading in time and in frequency domains. The results are reported both in terms of relative optimisation time, in Figure 3.24, Figure 3.26 and Figure 3.28, and relative transmitter power, in Figure 3.23, Figure 3.25 and Figure 3.27**.** Here it is assumed that 32 users are active in the system, user data rate is equal to 128 and 64 bits per user per OFDM symbol for Short Range and Wide Area Cellular scenarios, respectively.

From Figure 3.23, Figure 3.25 and Figure 3.27 it can be seen that sub-carrier sharing provides 3-4 dB gain in required transmit power compared to the OFDMA case (L=1), so confirming the expectations and preliminary results of [WIND26]. Moreover, it can be in general observed that the required transmit power decreases by increasing the number of iterations J, since this affects the accuracy of sub-band sharing factors obtained at the first step of the optimisation algorithm.

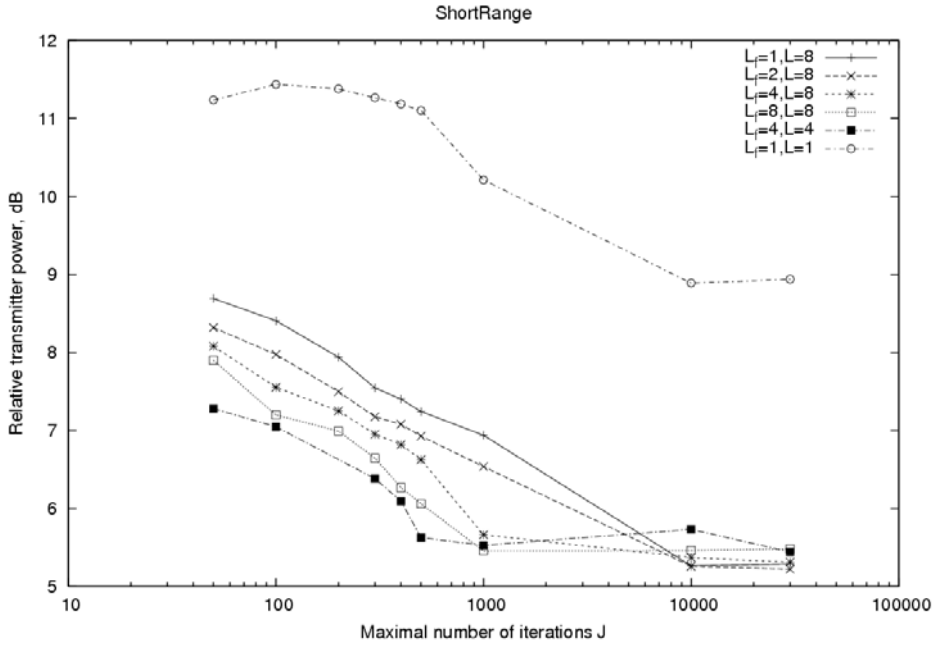


Figure 3.23: Relative transmitter power for Short Range scenario. Here *L=LtLf* where *Lt* and *Lf* are the spreading factors in time and frequency, respectively.

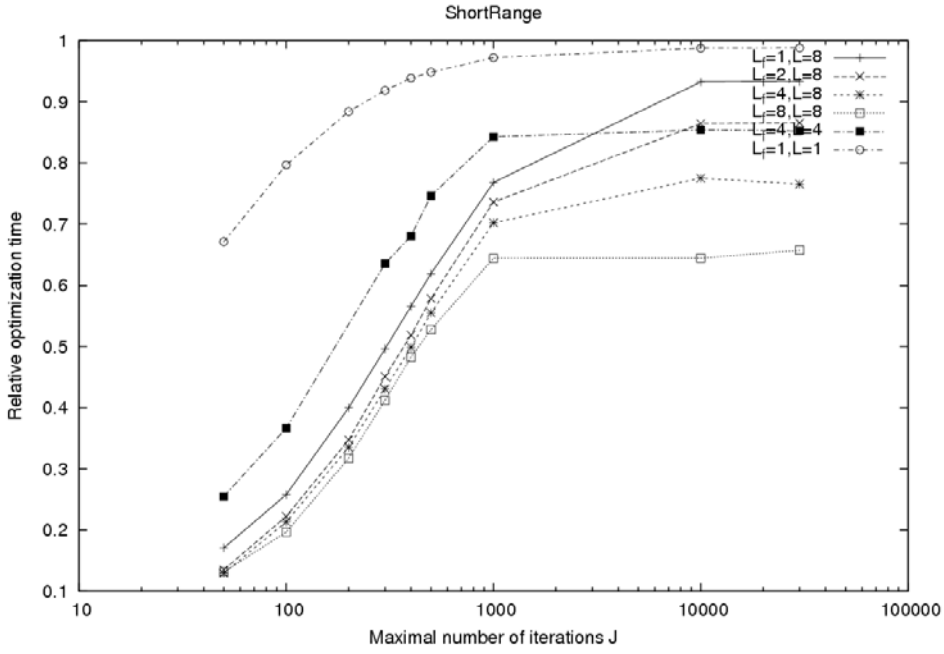


Figure 3.24: Relative optimisation time for Short Range scenario. Here $L=L_tL_f$ where L_t and L_f are the spreading factors in time and frequency, respectively.

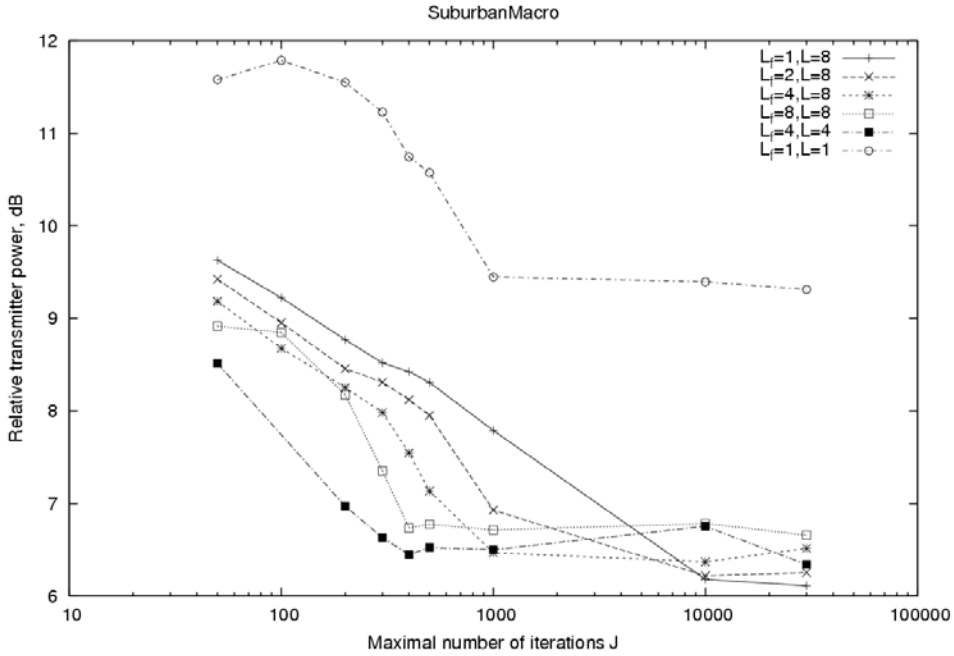


Figure 3.25: Relative transmitter power for Suburban Macro scenario. Here $L=L_tL_f$ where L_t and L_f are the spreading factors in time and frequency, respectively.

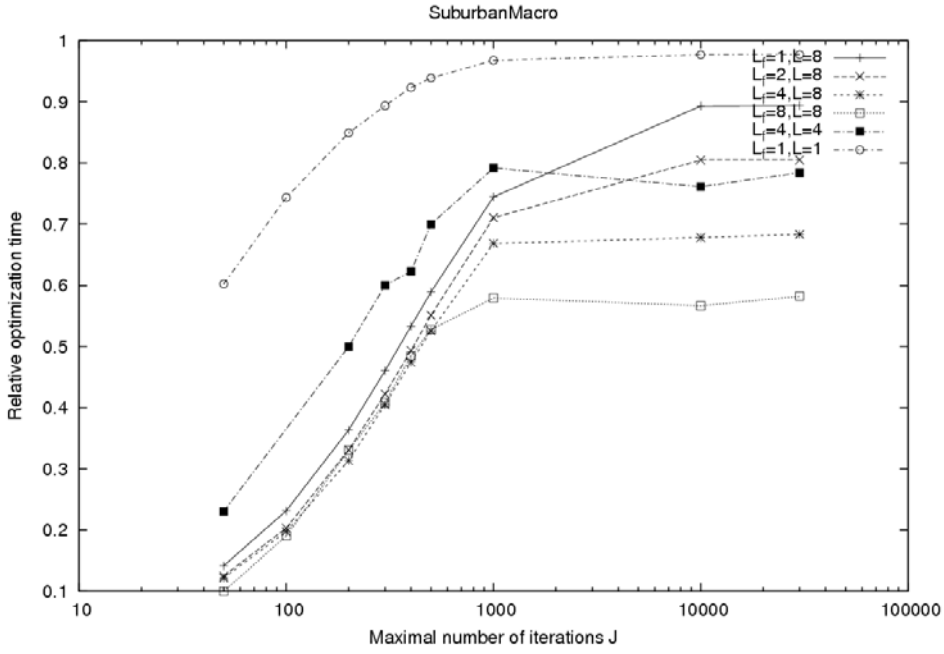


Figure 3.26: Relative optimisation time for Suburban Macro scenario. Here $L=L_tL_f$ where L_t and L_f are the spreading factors in time and frequency, respectively.

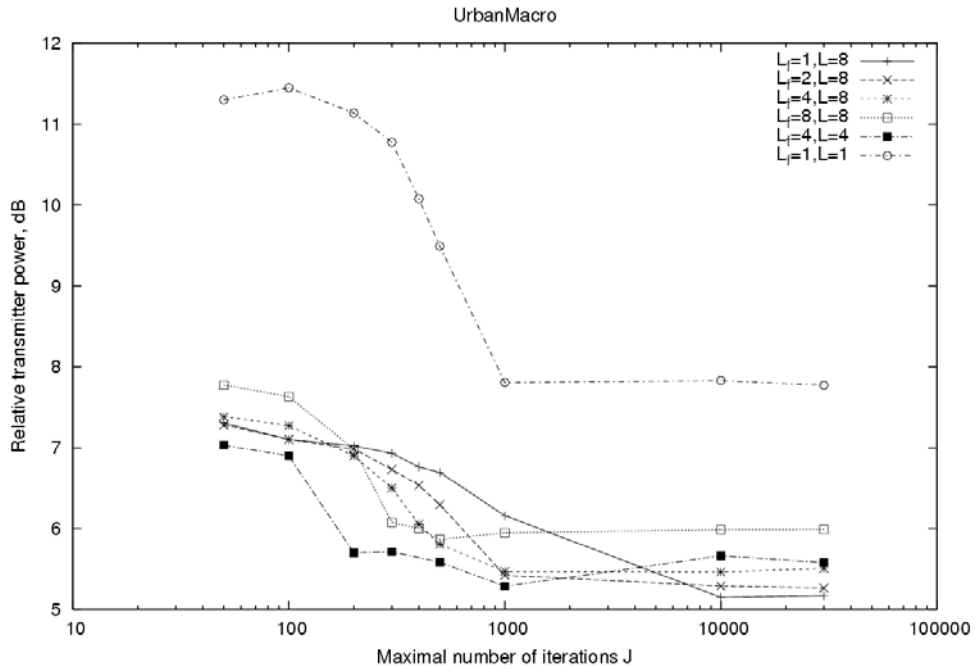


Figure 3.27: Relative transmitter power for Urban Macro scenario. Here $L=L_tL_f$ where L_t and L_f are the spreading factors in time and frequency, respectively.

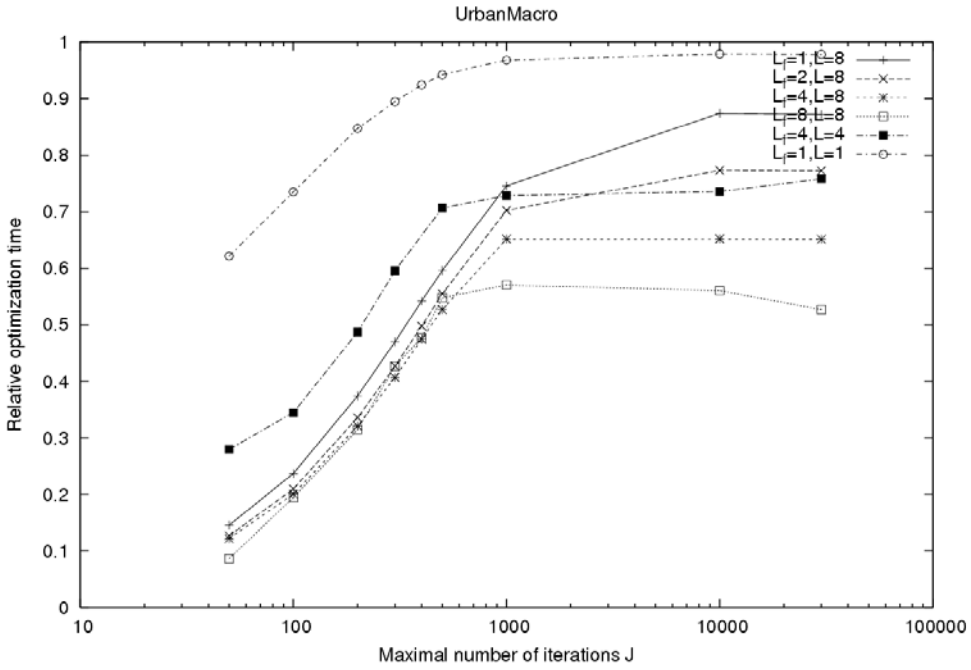


Figure 3.28: Relative optimisation time for Urban Macro scenario. Here $L=L_tL_f$ where L_t and L_f are the spreading factors in time and frequency, respectively.

We observe then that, for higher values of the spreading factor in frequency domain L_f, the curves saturate quicker. The reason for this is not related to the use of spreading in frequency rather than in time, which is ignored by the optimisation algorithm, but to the higher size of the sub-bands. For higher sub-band size, the number of sub-channels N/L_f seen by the optimisation algorithm decreases and hence the dimension of the optimisation problem reduces, so it becomes easier to solve. On the other hand, if one allows the algorithm to run until it fully converges, it appears that the higher L_f, the higher the required transmit power is needed. The reason is that performing optimisation over large sub-bands causes the average sub-band CNRs to be much closer to their global average values, i.e. the channel as it is seen by the optimisation algorithm becomes less frequency selective, thus reducing possible link adaptation gain.

Therefore, sub-band operation saves much computation time at the expense of some performance degradation that can be quantified in up to 0.8 dB, 0.7 dB and 0.2 for the urban macro, suburban macro and short range cases, respectively.

Same conclusions on the algorithm's convergence can be drawn from Figure 3.23, Figure 3.25 and Figure 3.27, where we can see, in particular, that if the maximal number of iterations J is chosen to be relatively small (i.e. below 1000), the complexity of the single user bit-loading dominates. However, for higher values of J the complexity due to the first step (solving Equations (3.35), (3.36) and (3.37)) becomes dominant. Since it depends on the total number of sub-channels, complexity curves for different values of L_f go to saturation at different values of τ. For $L_f \geq 4$, the computational overhead induced by the sub-band sharing optimisation step is comparable to the complexity of single-user bit loading step, i.e. $\tau \approx 0.5$.

From all figures it can be inferred that, besides the effect of the sub-band size, the total spreading factor L plays a role. By comparing the curves for L=4 and L=8, for smaller L the discrete approximation of the sharing factors with values from the set (0, 1/L, …., (L-1)/L,1) becomes less accurate and this affects the convergence of the algorithm.

It has to be observed that here L_f was always assumed to be equal to the sub-band size. However, some other spreading configurations should be investigated, while maintaining the per sub-band optimisation assumption, in order to study the tradeoff between spreading in frequency and in time domain. One example could be pure time domain spreading with spreading sequence length L over a number of adjacent sub-carriers constituting one sub-band.

3.3.3.4 Sensitivity to imperfect CSI

This section first briefly reviews the adopted joint channel estimation (JCE) method [SMW+01], [MWS+02] that is assumed to be applied at the base station in uplink as anticipated in Section 3.3.3.2. Some simulation results illustrating its performance are then reported. Based on these results, a suggestion for the number of OFDM symbols to be used as pilot symbols and on the number of estimated channel taps will be given for each considered propagation scenario.

Finally, simulation results related to the sensitivity of the suggested link adaptation algorithm to the imperfectness of the CSI are discussed.

3.3.3.4.1 Considered Channel Estimation Method

Let us first first consider a disjoint channel estimation method. Let $S_k = diag(s_{k0},...,s_{kn})$ be the diagonal matrix of pilot symbols transmitted by user k at some time instant and let $h_k = (\mu_{k0},...,\mu_{k,n-1})^T$ be the vector of CTFs experienced by this user. Assuming that other users do not transmit simultaneously on the same set of sub-carriers, the signal received by the base station can be represented as $r_k = S_k h_k + \eta_k$, where η_k is the vector of AWGN samples with variance σ^2. Then the least squares estimate of the vector of CTFs can be obtained as

$$\hat{h}_k = (S_k^H S_k)^{-1} S_k^H r_k \tag{3.42}$$

Since S_k matrix is a diagonal one, this equation implies that for each sub-carrier channel transfer function is estimated independently. Moreover, as well known, it is possible to transmit pilot symbols using only a few well-chosen $n < N$ sub-carriers and obtain estimates for the remaining ones using e.g. interpolation.

Observe, that all the CTFs are in one-to-one correspondence with the channel impulse response (IR) taps, i.e. $h_k = Fg_k$, where F is the DFT matrix and g_k is the vector of the k-th user's channel IR taps.

Since in general the number of leading component of the channel impulse response is much smaller compared to the number of sub-carriers, the estimation accuracy may be improved by resorting to time domain channel estimation, according to which only L first channel taps are assumed to be non-zero and their least squares estimate reads as $\hat{g}_k = (F^H S_k^H S_k F)^{-1} F^H S_k^H r_k$, where F denotes now a truncated $N \times L$ DFT matrix.

As shown in [SMW+01], this approach enables the application of a joint channel estimation (*JCE*) method, according to which many users can simultaneously transmit their pilots on the same set of sub-carriers In this case the received signal can be represented as

$$r = \sum_{k=0}^{K-1} S_k h_k + \eta = \widetilde{S} h + \eta\,, \tag{3.43}$$

where $\widetilde{S} = (S_0, S_1,..., S_{K-1})$ is the combined *N×(KN)* matrix of pilot signals and h is the combined vector of sub-carrier CTFs.

One can see that the number of unknowns (i.e. sub-carrier transfer factors of different users being estimated) in this system of equations is much more than the number of equations. However, if one considers this system of equations in time domain, it appears that the number of unknowns can be reduced to *N×(KL)* by exploiting the fact that only L first channel taps are non-zero.

Now the least-squares estimate of the sub-carrier transfer factors can be obtained as $h = \widetilde{F}(\widetilde{F}^H \widetilde{S}^H \widetilde{S} \widetilde{F})^{-1} \widetilde{F}^H \widetilde{S}^H r$, where $\widetilde{F}$ is the block-diagonal matrix with F matrices on its main diagonal. Estimation error is minimized if pilot signals satisfy the constraint

$$F^H S_i^H S_j F = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}.$$

This constraint is satisfied if Walsh codes are used to separate the users transmitting pilots on the same sub-carriers.

The maximal number of users which can be handled by this approach is equal to

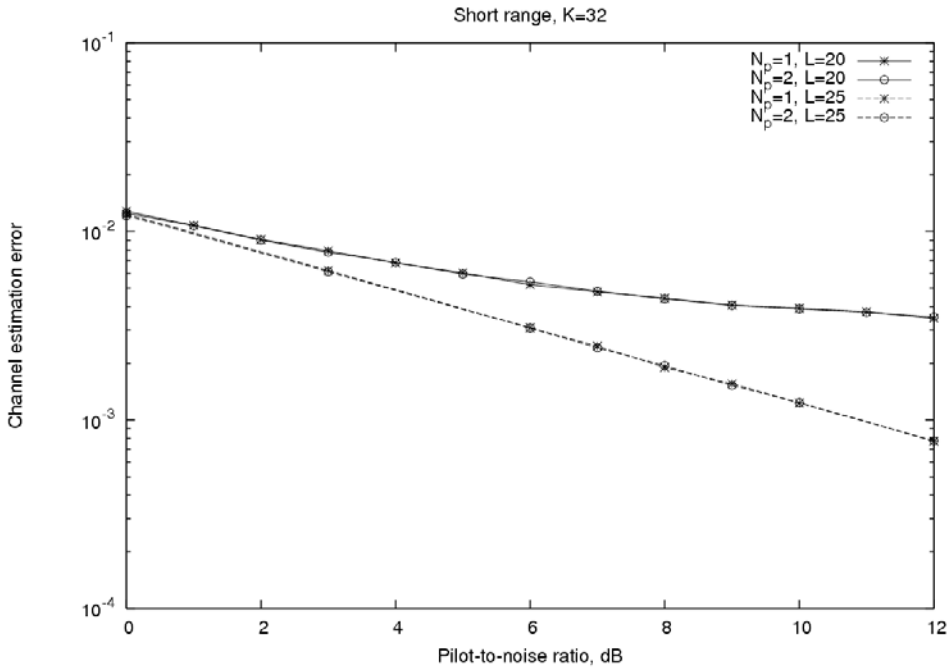
$$K_{\max} = \frac{N}{2^m}, 2^{m-1} < L \le 2^m, m \in Z \tag{3.44}$$

If the system has to support more than K_{max} users, it is necessary to partition them into N_p groups of size not more than K_{max} and perform channel estimation for them in N_p separate OFDM pilot symbols. However, if the channel is quickly varying, this may cause the CSI for different users to have substantially different quality as will be shown by simulation results.

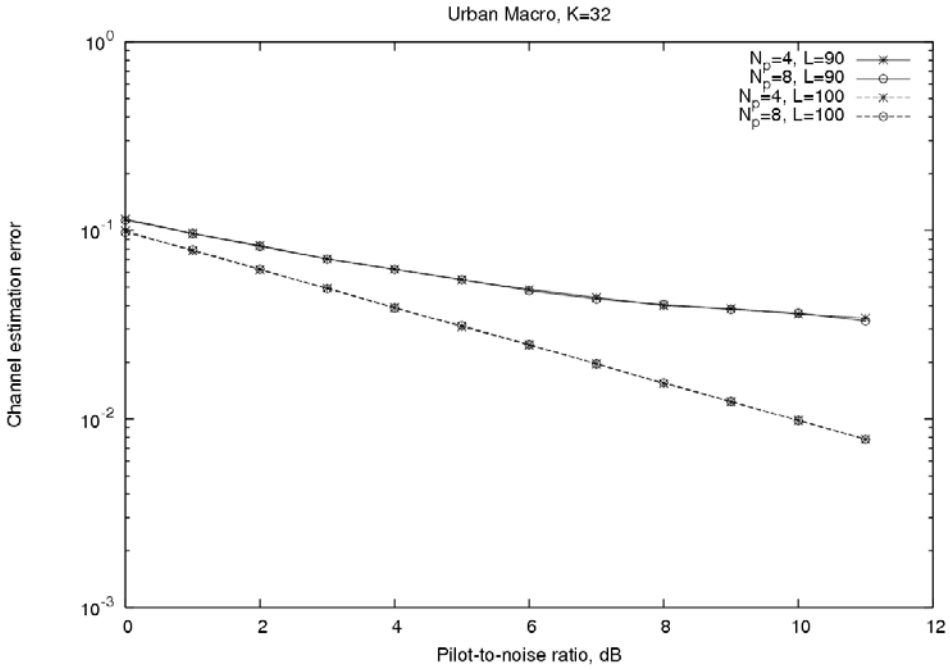
As discussed in [MWS+02], JCE achieves the same performance as disjoint channel estimation as far as proper pilot sequences, e.g. Walsh codes, are selected, but it exhibits the advantage of being more robust to inter-cell interference.

3.3.3.4.2 Performance of the channel estimation method

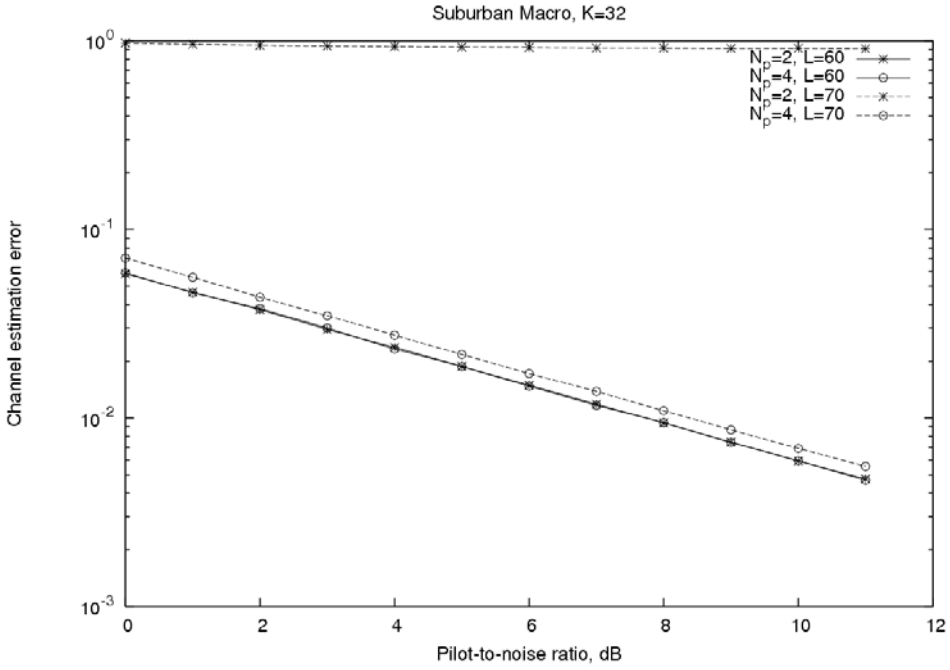
Figure 3.29 presents the simulation results illustrating the performance of JCE for different propagation scenarios. Results are reported in terms of mean squared error defined as $MSE = E[\frac{1}{KN}\sum_{k=0}^{K-1}\sum_{i=0}^{N-1}|\hat{\mu}_{ki} - \mu_{ki}|^2]$ and pilot-to-noise ratio defined as $E[\frac{|s_{ki}|^2}{N_0}]$, where N_0 is the noise PSD. The number of users was set to K = 32.



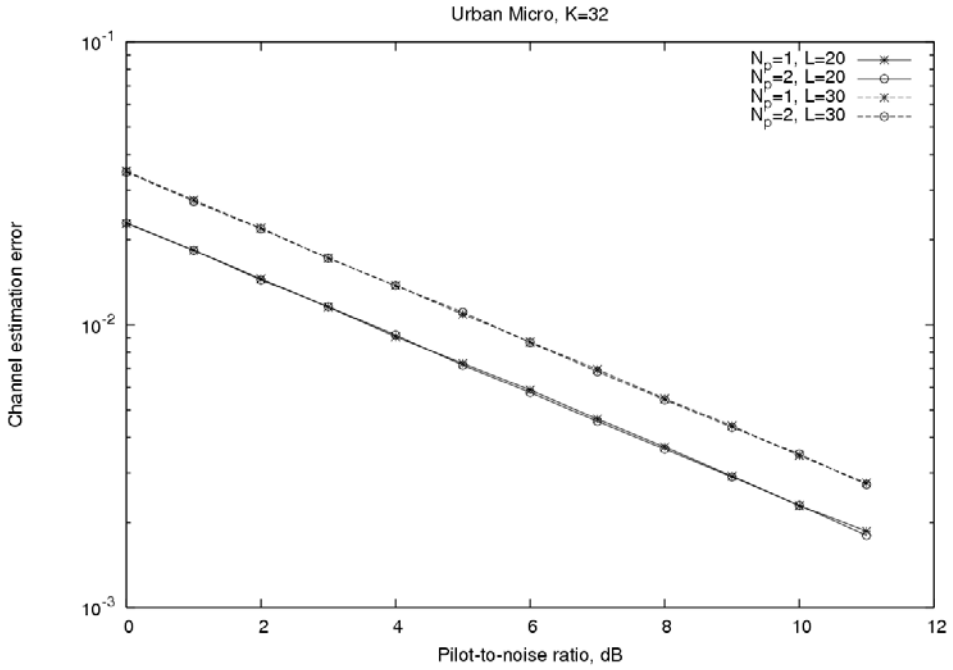
a) Short Range



b) Urban Macro



c) Suburban Macro



d) Urban Micro

Figure 3.29: JCE performance for different choices of the number of OFDM pilot symbols N_p and the number of estimated channel taps L.

These results clearly indicate that improper selection of parameters may result in considerable degradation of the system performance. The following table summarizes the channel estimation parameters chosen for various propagation scenarios.

Table 3.8: Required joint channel estimation parameters

Scenario	N_p	L
Short range	1	25
Suburban macro	2	60
Urban macro	4	100
Urban micro	1	20

As it was stated above, using excessive values of N_p in case of time-variant channel may cause the CSI of different users to have substantially different quality. Figure 3.30 illustrates the impact of channel variations on the JCE performance by averaging over all users in suburban macro scenario. It is assumed that users are partitioned into N_p groups. Each of these groups performs joint channel estimation using one of N_p OFDM symbols and the assignment of user group to OFDM symbol changes in time in order to ensure uniform performance of different user groups.

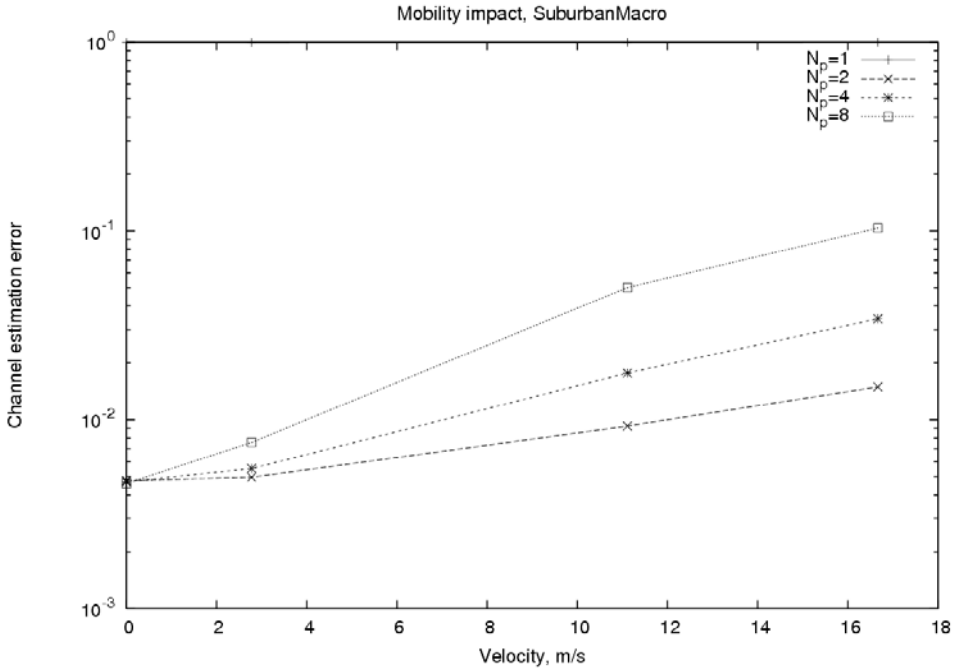


Figure 3.30: Impact of channel variations on the performance of the JCE method for different choices of the number of OFDM pilot symbols N_p.

It can be seen that while N_p=1 provides unacceptable performance (see Table 3.8), greater values of N_p provide quite good channel estimation quality in case of static channel. However channel estimation quality degrades with the increase of user terminal velocity and the degradation rate increases with N_p.

3.3.3.4.3 Sensitivity of the adaptive algorithm to imperfect CSI

This section presents the simulation results illustrating the sensitivity of the adaptive MC-TF-CDMA system to usage of the CSI obtained via the JCE method described in Section 3.3.3.4.1. The results are reported in terms of actual BER experienced by the user terminals, as opposed to the target BER $2 \cdot 10^{-3}$ assumed by the optimisation algorithm (see definition of capacity gap Γ), versus the pilot-to-noise ratio. Here it is assumed that the channel is static and pure frequency-domain spreading with L_f=L=8 is used. Moreover, we distinguish the case in which the user gains are transmitted by the AP, Figure 3.32, from that in which the user gains are derived by the UT receiver as explained in Section 3.3.3.2to reduce the amount of auxiliary information, Figure 3.31.

As it may be expected, if user gains are not transmitted, the system appears to be more sensitive to the channel estimation errors compared to the case of transmitting full auxiliary information. However, if the channel estimation errors are sufficiently small, the system performance in both cases becomes nearly indistinguishable.

From Figure 3.32 one can observe that the adaptive system exhibits considerably different sensitivity to the pilot-to-noise ratio in different propagation environments. This can be explained by observing (see Figure 3.30) that the mean squared channel estimation error decrease rate is considerably different for various propagation scenarios.

Comparing Figure 3.32, Figure 3.31 and Figure 3.29 one can also see that in order to achieve the target BER MSE around 0.01 is sufficient. This is quite close to the results presented in [PAE04].

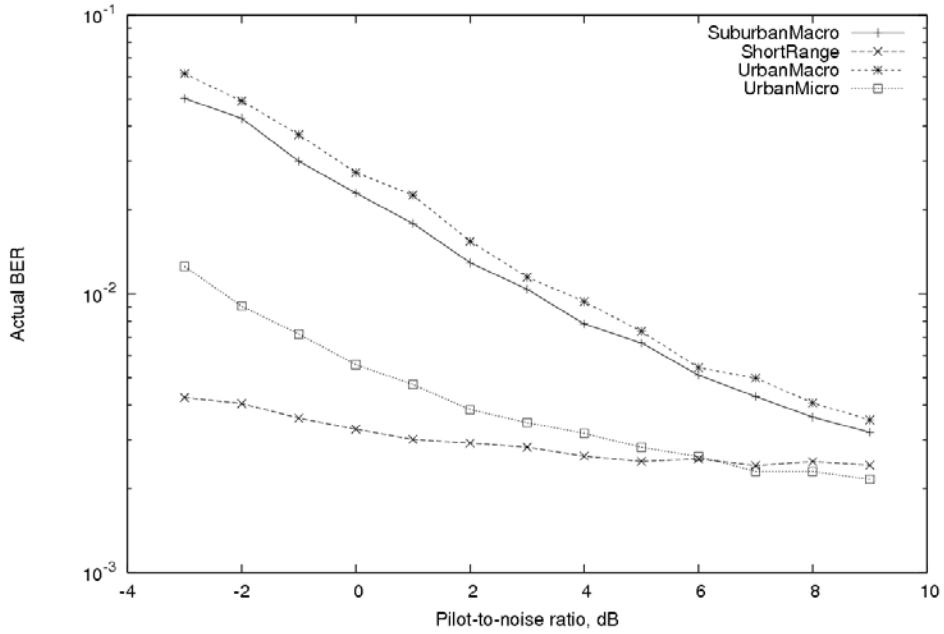


Figure 3.31: Sensitivity of the adaptive system to channel estimation errors (user gains deduced by the receiver)

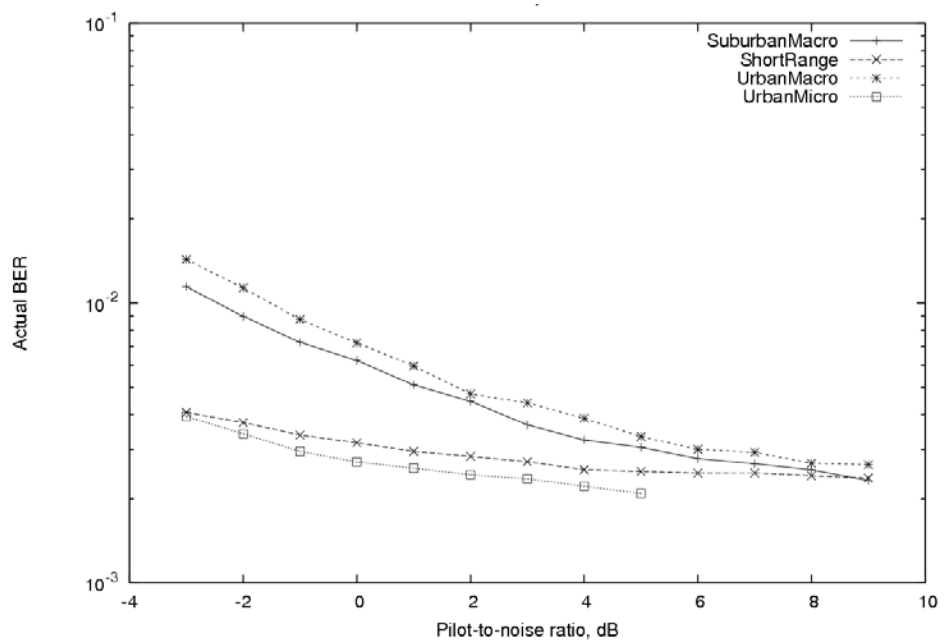


Figure 3.32: Sensitivity of the adaptive system to channel estimation errors (user gains transmitted by the access point)

3.3.3.5 Sensitivity to channel variations

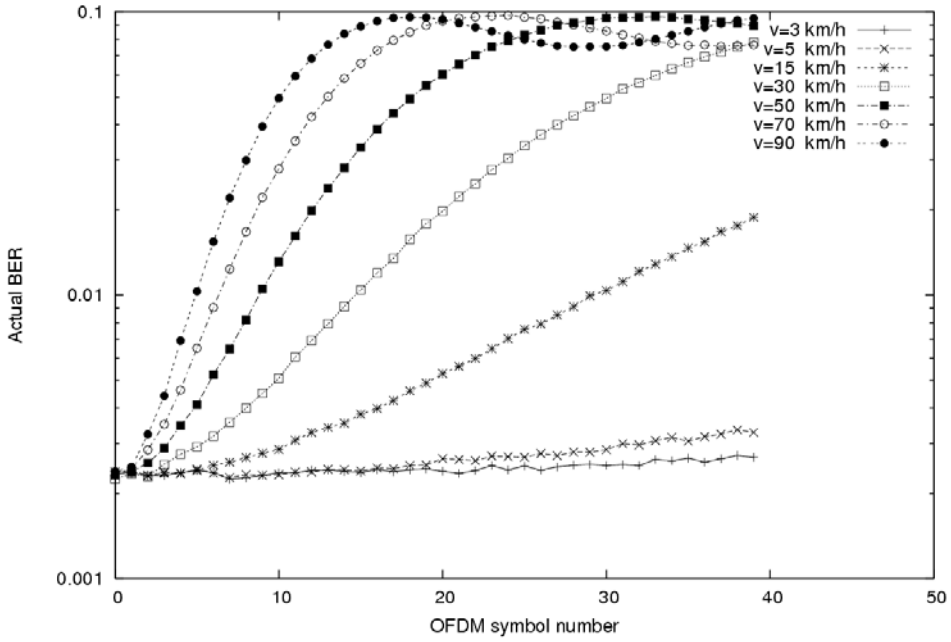
Any kind of link adaptation algorithm tries to adjust the data transmission scheme so that it fully exploits the current channel fading state. However, since wireless channels are time-varying, the data transmission scheme optimised for a particular channel state may quickly become highly non-optimal due to channel

variations. Hence one would need to periodically re-optimise the data transmission scheme. Therefore it is necessary to find the point at which the actual BER experienced by the receiver becomes too high. The following figures present simulation results illustrating the BER degradation in subsequent OFDM symbols in time-variant channels.

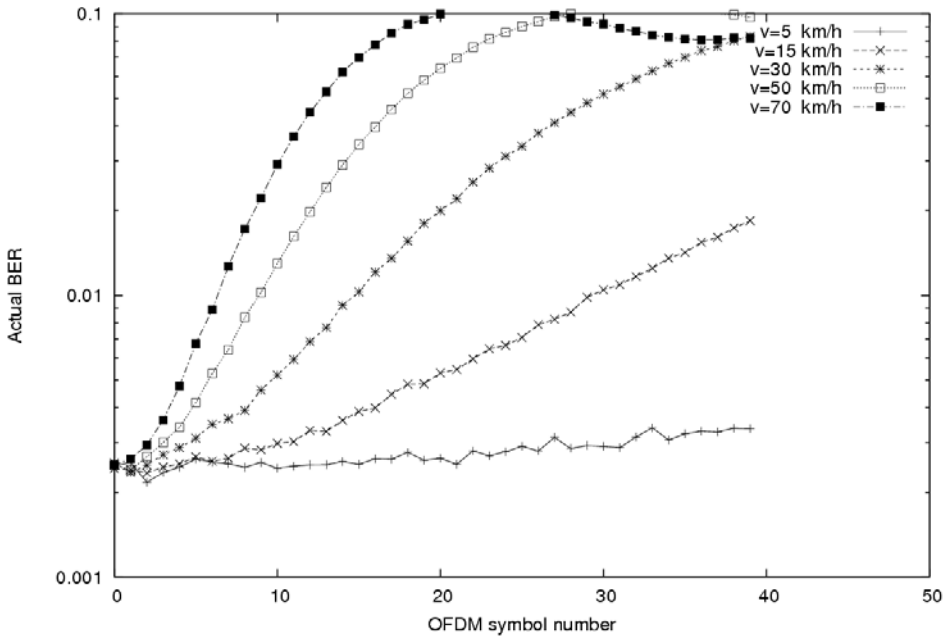
Optimisation of the data transmission scheme is performed immediately before transmission of the first OFDM symbol and the same scheme is used for all subsequent OFDM symbols. The actual BER increases with the OFDM symbol number. As it may be expected, BER degradation depends only on channel time autocorrelation properties and not on a particular channel power delay profile. Fixed a maximal acceptable BER $P_{0,}$ in order to properly select the re-optimisation interval for a given velocity, one has to find in Figure 3.33 the OFDM symbol number for which the actual BER curve correspondent to that velocity crosses with the straight line given by actual BER=P_0 , that is solve the equation BER(v,t)=P_0, where t (unknown) denotes time and v is the user terminal velocity. This gives a function t(v).

For example, Figure 3.34a presents curves illustrating the maximal length of the inter-optimisation interval t(v) in OFDM symbols for the case of P_0=5E-3. It can be seen that for the scenarios employing 20 MHz bandwidth with 1024 sub-carriers (Suburban macro, Urban macro, Urban micro) inter-optimisation interval cannot exceed 20 OFDM symbols for user terminal velocity 15 km/h. For the short range scenario employing 100MHz bandwidth with 2048 sub-carriers the maximal length of the inter-optimisation interval is approximately 2.5 times higher. Assuming that all required auxiliary information can be transmitted within one OFDM symbol (this is possible if appropriate source coding is used), this leads to approximately 5% and 2% overhead for these scenarios, respectively.

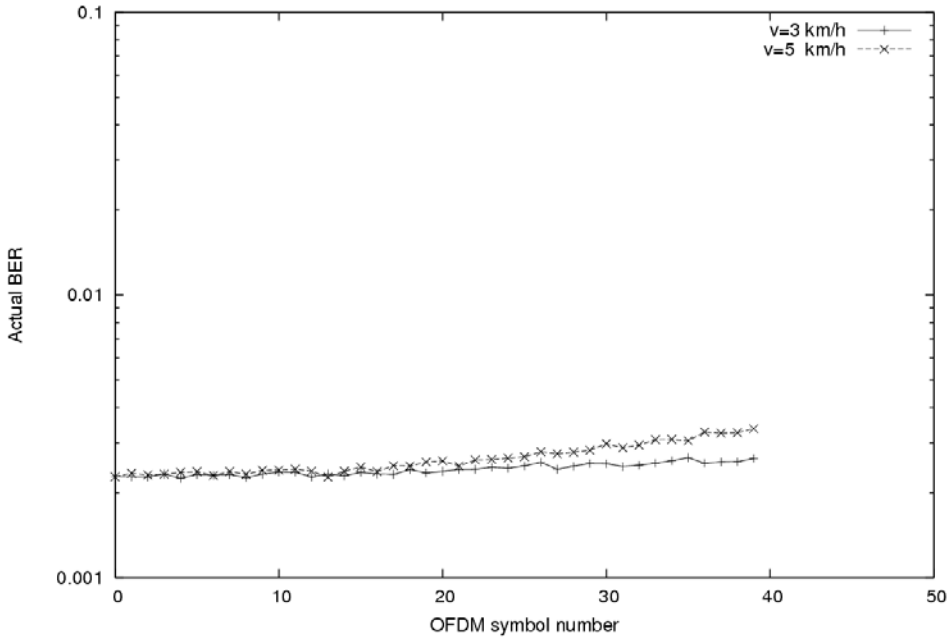
However, if the same results are considered in terms of the inter-optimisation *time* (Figure 3.34b), the system performance in both cases appears to be quite close.



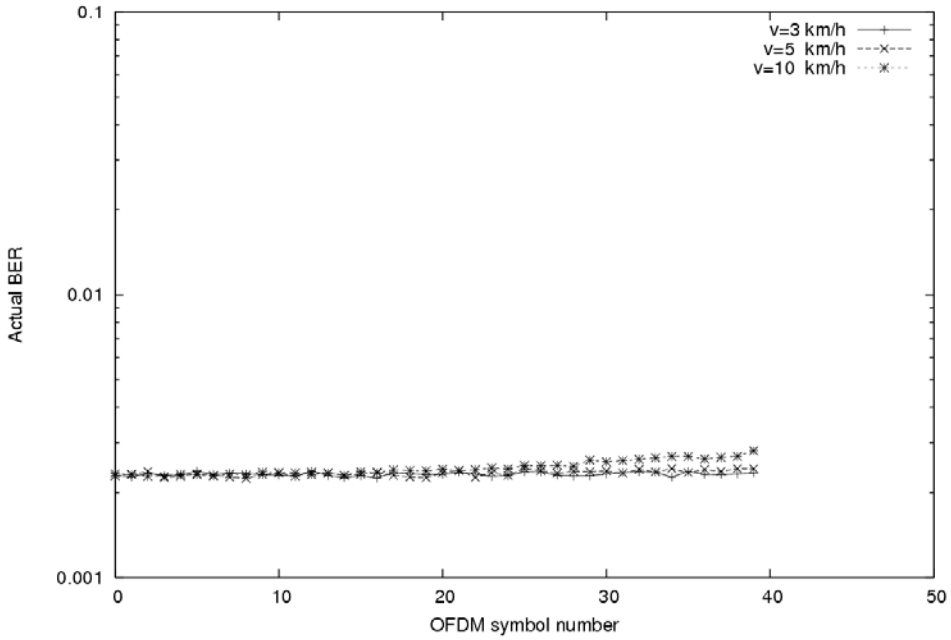
a) Suburban macro



b) Urban macro

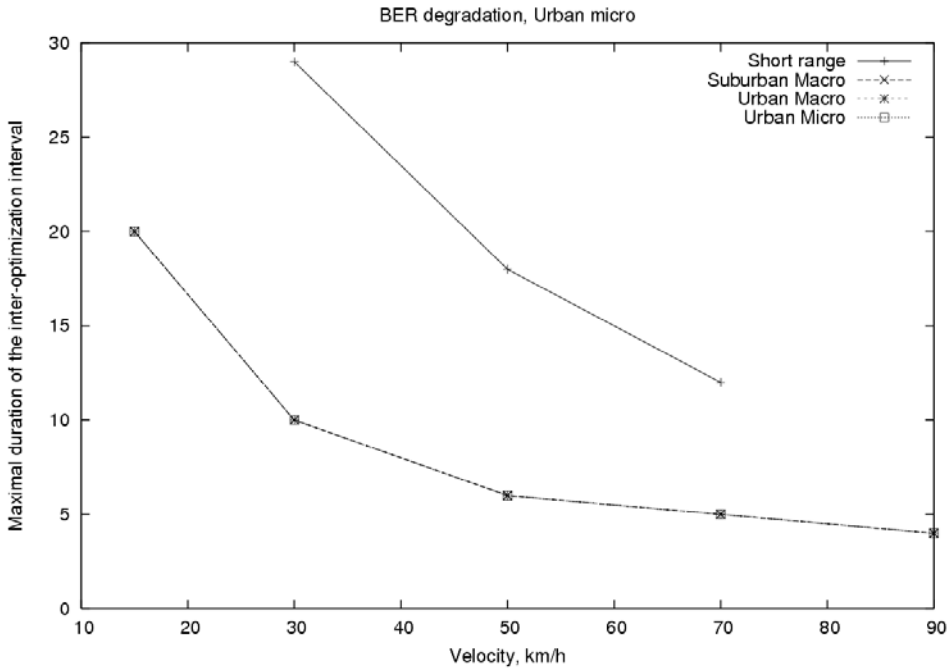


c) Urban micro

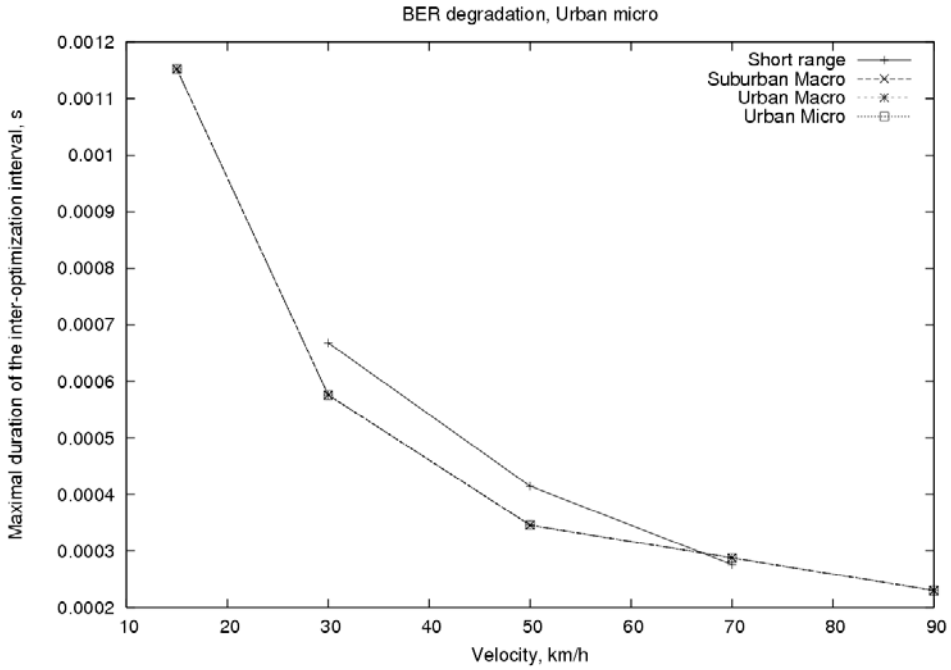


d) Short range

Figure 3.33: BER degradation in time-variant channels



a)



b)

Figure 3.34: Limitations of link adaptation in time-variant channels

3.3.4 Conclusions

In this section we have investigated by means of extensive simulation issues related to the implementation of an adaptive sub-carrier allocation, bit and power loading approach in a MC-TF-CDMA downlink system.

For the optimisation complexity, we have shown that in all considered scenarios evaluation of sub-carrier sharing factors takes approximately 55% of the total link adaptation computation time. From the point of view of tradeoff between achieved performance, optimisation complexity and amount of auxiliary signalling, adaptation over sub-bands of 8 sub-carriers appears to be the most suitable solution for all scenarios. Moreover, regarding the amount of auxiliary information, we have shown that it is possible to avoid transmission of user gains by exploiting the channel knowledge at the receiver.

For the sensitivity to channel estimation errors we observe that in all cases mean squared channel estimation error around 0.01 appears to be sufficient for link adaptation purposes. To achieve such accuracy, 1 OFDM pilot symbol is sufficient in Urban Micro and Short Range scenarios, while for Suburban Macro and Urban Macro 2 and 4 OFDM pilot symbols are necessary, respectively, under the assumption that JCE is employed and 32 users are simultaneously active in the system. In other words, for the channel estimation the urban macro scenario is the one requiring the highest pilot overhead, and this had to be expected given its high frequency selectivity.

For the sensitivity to channel variations, the same sensitivity was observed in all scenarios, as it is reasonably expected since the performance degradation is determined by the channel time correlation properties. However, since system parameters, including the OFDM symbol duration, are different in the short range and in the wide area cellular scenarios, it turns out that the overhead for update of the CSI measured in OFDM symbols is 2.5 times larger in wide area scenario.

4. Implications in WINNER Scenarios

4.1 Optimum multi-coded OFDM transmission

According to [WIND23], there is a limited set of different physical layer modes proposed currently. These are: wide area – cellular, wide area – cellular multi-hop, wide area – feeder links, short range – cellular and sort range – peer-to-peer. The optimum multi-coded OFDM transmission approach, presented in Section 2.1, seems to be most suitable for both short range modes due to lower maximum velocities and probably better channel estimation performance. Also, the wide area – cellular mode might be attractive for the presented adaptive approach, however it depends on the multiple access scheme and the mobility of the terminals.

It is important to comment here that the presented results have been based on the usage of the whole bandwidth (all subcarriers) for the adaptation for one link, so no multiuser case has been evaluated. However, it should be possible to modify the presented approach in order to apply it in multiuser cases, like those presented in Chapter 3, for instance. One could consider performing the code, modulation and power loading for a group of subcarriers, called chunks. Though, it requires further investigations to understand what are the pros and cons of such modification.

Another important thing, which has not been addressed in the presented evaluations is the signalling issue. It is crucial to design an optimum coding method of the feedback information and to consider if such technique provides enough performance improvement as compared to the additional overhead introduced.

4.2 Predictive adaptive resource scheduling using TDMA/OFDMA

The predictive adaptive TDMA/OFDMA scheme described in Section 3.1 is flexible. It can be applied for most types of flows in most considered deployment scenario. However, different parts of its features come to the forefront in different scenarios.

4.2.1 Scenario A: In and around buildings

In and around buildings, the TDD physical layer mode would be the primary system. Since terminals are stationary or move at most at pedestrian velocities, prediction performance and outdated channel state information are not of primary concern. Simpler channel estimation schemes could be applied in these situations.

In this scenario, there will be significant frequency selectivity in the channels over a wide band of up to 100 MHz. Therefore, adaptive allocation over frequency will provide significant multiuser scheduling gains. It will enable the flows to be allocated to the subbands most suitable for each terminal. Fading in time, and moving terminals, is not required to realize these multiuser scheduling gains.

Since MIMO transmission can be expected to be used extensively in this scenario, good combinations with MIMO schemes, primarily multiplexing-based, will be of interest.

4.2.2 Scenario B: Hot spot/area

The TDD physical layer mode would be of primary use in hot spot and area-wide but non-ubitiqous coverage. In most of these WINNER scenarios, terminals are stationary or slow moving. In the B1 scenario, with Typical Urban propagation conditions, vehicle speeds of up to 70 km/h are considered. The large majority of these terminals should be able to use the suggested adaptive scheme. If the cell radii are limited to 100-200 m, as presently assumed, then the channel frequency coherence bandwidth will be large. The large 780 kHz chunk bandwidths used in the present design of the TDD mode should then be suitable. If larger cell sizes are contemplated for the TDD mode, the chunk width should be reduced correspondingly.

An issue of particular interest in this scenario is how the proposed scheme interacts with relaying solutions. This will depend crucially on *where* the resource scheduler is located in a multihop transmission via relay nodes. With centralized scheduling, relaying would increase the feedback delays

considerably. It would probably make adaptive transmission at vehicular speeds impossible. If the interference environment is bursty and needs to be predicted, severe difficulties may also result for stationary or slow-moving users. If, on the other hand, the resource scheduling is performed at the last relay node, so that the adaptation feedback loop to the terminal can be terminated there, no extra delay is incurred.

4.2.3 Scenario C: Metropolitan

In the metropolitan scenario, we contemplate primarily the FDD mode. Populations of stationary users are there assumed mixed with vehicular users, with velocities of up to 70 km/h. The large majority of these terminals should be able to use the suggested adaptive scheme.

In the uplinks, the number of terminals per base station that can take part in adaptive transmission is limited by the capability of the uplink channel predictors at the access point. The predictor performance deteriorates with the number of users who simultaneously transmit overlapping pilots, see Subsection 3.1.3.2. When basing the algorithms on $p = 8$ parallel subcarriers, we could accommodate perhaps 8 simultaneous users per contention band, in each group of half-duplex FDD terminals. If the total 20 MHz uplink band is divided into four contention bands, and the two groups of terminals have the same number of terminals assigned to them, then $2 \times 4 \times 8 = 64$ active terminals that each have one or more adaptively allocated flows could be accommodated at each FDD access point.

There is one important type of flow that would mostly not be able to use adaptive transmission: Broadcast or multicast flows. They have to be detectable by all or a large subset of users, and therefore have to be transmitted in a safe non-adaptive mode. Broadcast transmission over multi-antenna access points also brings some special problems. In general, each such flow would therefore consume more radio spectrum resources than other flows.

4.2.4 Scenario D: Rural

In the rural scenario, there would be a mix of stationary and vehicular users. The FDD physical layer mode would be in use. Adaptive transmission should be successful except for three types of flows:

1. To/from vehicular users close to cell borders, where a low SINR would preclude sufficiently accurate channel prediction.
2. To/from high-speed vehicular users travelling at or over perhaps 70 km/h, where the limit depends heavily on the SINR.
3. Broadcast and most multicast flows, for reasons described above.

In all three cases, the non-adaptive fallback mode would be in use.

4.3 Adaptive approach to antenna selection and space-time block coding

The research on the adaptive approach to antenna selection and space-time block coding has been performed in the context of virtual antenna arrays constituted by mobile nodes in an ad-hoc manner. Out of the prioritised deployment scenarios and the corresponding channel models listed in [WIND52], the A1 and B1 seem to be the best suited ones for this adaptation strategy.

First of them is an indoor, NLOS channel model with low mobility (up to 5 km/h). In this case mobile users equipped with handheld devices and/or laptops, which may also act as relay stations, are taken into account. The proposed adaptation strategy would perfectly allow to decide which out of all the available users with relaying capabilities should be considered for the purposes of cooperative diversity exploitation.

The latter case is a typical urban LOS/NLOS channel with medium mobility (up to 70 km/h). The main constraint of this scenario is the relatively high maximum speed of mobile stations (e.g. in the form of buses equipped with antennas and relaying capabilities), which may result in the necessity for more sophisticated selection algorithm in comparison with the previous scenario.

4.4 Cell capacities and bitrate limits for the uplink in a single cell

In this section, the fundamental limits for the aggregate bitrate in the uplink have been calculated for several WINNER scenarios. The simulation parameters, based on several internal working documents,

- the comparison cases elaborated in WINNER T2.4 for [WIND26]
- the recommendations for the calibration of system simulator tools
- the preliminary WINNER air interface concept,

are summarized in the following table.

Table 4.1: Assumed simulation paramters

Parameter	Short range	Wide area
number of used subcarriers N_C	1664	864
number of guard samples N_G	256	100
sampling frequency f_A	102.4 MHz	20 MHz
guard interval T_G	2.5 μs	5 μs
subcarrier spacing Δf	50 kHz	19 531.25 Hz
noise figure F_N	6 dB	
noise power per subchannel $N_0 = k\,T\,F_N\,\Delta f$	$8.28 \cdot 10^{-16}$W = -120.8 dBm	$3.23 \cdot 10^{-16}$W-124.9 dBm
base station antenna gain G_{BS}	10 dBi	
cell radii	30m, 100m	400m, 500m, 1km, 2km
QAM constellations (bits per symbol)	B = {1, 2, 4, 6, 8}	B = {1, 2, 4}
path loss function (in dB, distance d in meters)	$20\lg(d) + 46.42, \quad d < 5$ $35\lg(d) + 21.96, \quad d > 5$	$28.3\lg(d) + 53.5$
power delay profile	τ_n = {0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 110, 140, 170, 200} ns $p_n = -${0, 2.1, 4.3, 6.5, 8.6, 10.8, 4.4, 6.6, 8.7, 10.9, 13.7, 15.8, 18.0, 20.2} dB	τ_n = {0, 10, 30, 250, 260, 280, 360, 370, 385, 1040, 1045, 1065, 2730, 2740, 2760, 4600, 4610, 4625} ns $p_n = -${3, 5.22, 6.98, 4.7184, 6.9384, 8.6984, 5.2204, 7.4404, 9.2004, 8.1896, 10.4096, 12.1696, 12.0516, 14.2716, 16.0316, 15.5013, 17.7213, 19.4813} dB

The following diagrams have been obtained for a fixed, deterministic user distribution according to (3.25) and have been averaged over 100 channel realizations. Note that the obtained results depend strongly on the user distribution and on the path loss function: if the users are located closer to the base station or the path loss is smaller, the cell capacity increase drastically and vice versa.

The results for the short range scenario show encouraging results: the cell capacity is above 20 bps/Hz and the maximum bitrate is limited by the maximum constellation size but not by the SNR.

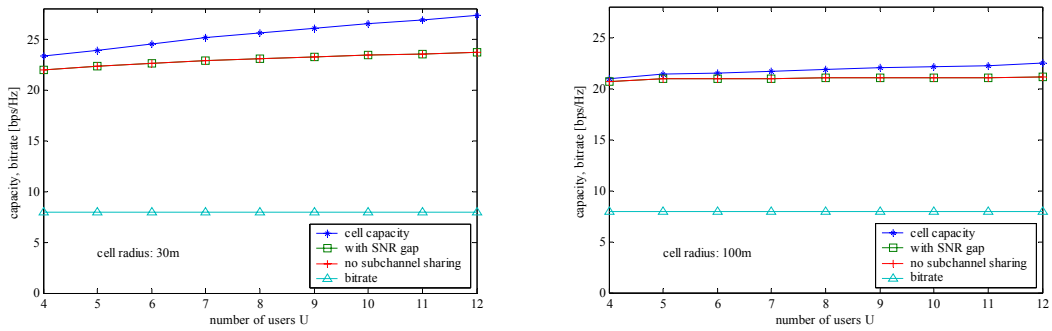


Figure 4.1: Cell capacities for short range scenario and cell radii of 30m and 100m

In the wide area case, the strong influence of the cell size is clearly visible. While for the smaller cells (400m and 500m), the bitrate is limited by 16-QAM as highest order modulation, for cells larger than 1km, the bitrate is clearly SNR-limited.

It has to be emphasized, though, that these results have been obtained with idealized assumptions and no fairness criterion has been considered. They should thus be understood as upper bounds rather than realizable values.

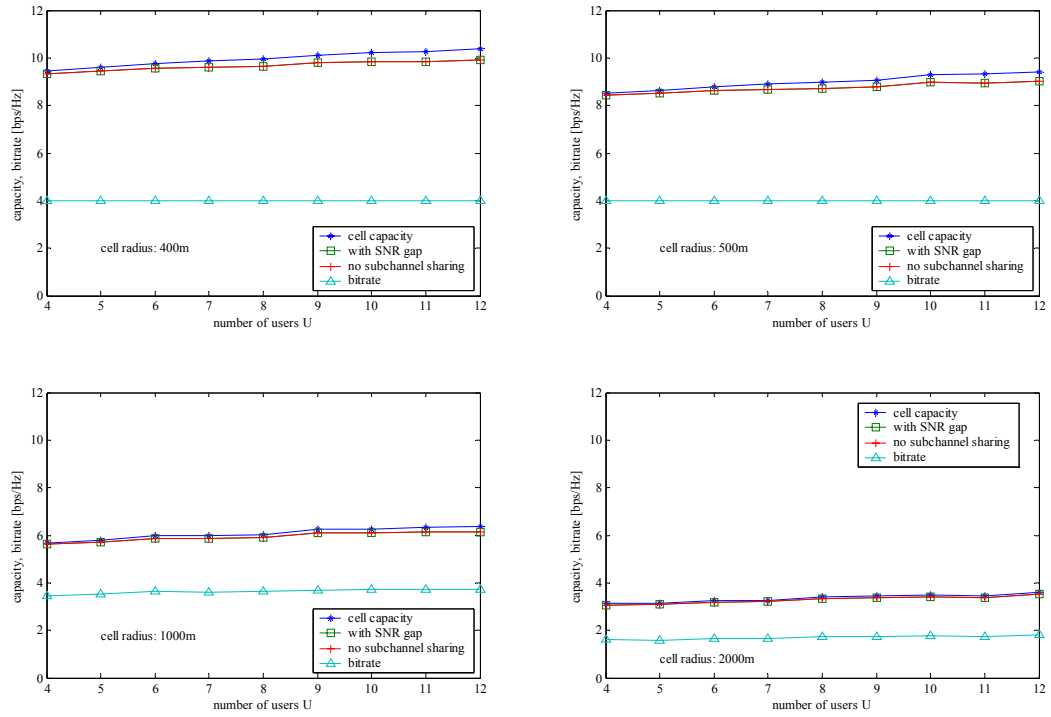


Figure 4.2: Cell capacities for wide area scenario

5. Summary and Conclusions

The area of adaptive transmission is vast, and the work reported in the present document has focused on a selected but important subset of problems. In particular, it has provided novel results in the following six areas:

1. Link adaptation based on as convolutionally coded M-QAM has been evaluated in the WINNER system setting.
2. Antenna selection algorithms have been proposed. It was outlined how this technique could be applied in a cooperative diversity context, where the antennas are distributed in space.
3. Adaptation requires feedback systems which give rise to delays, so outdated information will be used. A survey has been given of methods that enable the adaptation strategy to take this into account: Design constraints can be fulfilled also when the channel state is uncertain.
4. Detailed designs of adaptation schemes based on predictive TDMA/OFDMA were performed, both for the wide area FDD mode and for the short-range cellular mode based on TDD. Solutions were presented to the *feedback data rate problem* and to the problem of basing the adaptation on *outdated channel state information*.
5. A method for estimating an upper bound on the spectral efficiency of adaptively adjusted uplinks was proposed, and applied to WINNER short-range and wide area cellular uplinks.
6. A method for adaptive allocations of data streams in a multicarrier CDMA system with time-frequency spreading was evaluated by an extensive simulation study.

Furthermore, the place of adaptive transmission within the overall MAC layer architecture has been outlined in Chapter 1. Subsection 1.1.2 also presented the combinations of multiple access schemes selected as default assumptions by Task 2.4, to be used by other tasks during WINNER Phase I.

Link adaptation: Concerning the adaptive multi-coded OFDM transmission approach presented in Section 2.1, it was shown how to perform adaptation on both the code and modulation jointly with power loading in OFDM-based WINNER modes. Figure 2.11 gives a good description of the gain obtained in the single user case by adapting the subcarrier bit and power loading to the short-term Rayleigh fading statistics, as compared to using fixed modulation schemes. It was also outlined how to extend such an adaptive technique with space-time processing in the form of Alamouti space-time block codes at the transmitter and combining the signals from multiple antennas at the receiver.

A conclusion is that if the transmit power is adjusted, then the performance of the loading algorithm is not degraded by reducing the number of code and modulation schemes used for the adaptation. One might employ only the maximum and minimum code rates combined with all considered constellation schemes and observe no deterioration in overall code rate and bit error rate.

Another result is that the impact on the performance of variable lengths of encoded blocks should not be a crucial problem in case of the use of convolutional codes; The performance reduction is diminished by constraining the minimum block size to 64 QAM symbols.

Furthermore, the adaptive multi-coded OFDM system complexity is not increased substantially, as compared to traditional OFDM systems, because only one mother code with a few puncturing schemes needs to be implemented. Therefore, the increase in the complexity will basically be caused by the implementation of the loading algorithm. Channel estimation errors at the receiver do provide problems at the lowest SNR values. These problems are alleviated, but not fully solved, by taking the errors into account in the adaptation scheme. The code and modulation rate will be based on channel state information that must be propagated to the transmitter. At the time of reception, this information will be outdated. Channel state prediction errors were seen to pose potentially serious difficulties at vehicular velocities for realistic feedback delays. The systematic design of rate adaptation scheme that take the prediction uncertainty into account was proposed in Section 2.2 as a partial solution to this problem.

Adaptive TDMA/OFDMA transmission: Detailed designs were performed both for a wide-area cellular FDD-based mode and for a short-range TDD-based mode. A main aim was to overcome the two key problems of adaptive transmission in a multiuser cellular environment: The *inaccurate and outdated channel state feedback* for nonstationary users and the potentially *high required feedback data rate*. The design and evaluation had focused on a challenging test case: 5 GHz, with wide area coverage over frequency selective channels, for terminals with vehicular velocities. We have furthermore focused on the

FDD physical layer mode that has the smallest time-frequency chunks and the highest control overhead load.

The study results in the following main conclusions:

- Adaptive TDMA/OFDMA transmission is feasible. No crucial show-stoppers were identified in the present investigation.
- Predictive adaptation to the short-term fading and the frequency-domain channel variability leads to significant multiuser diversity (scheduling) gains: The cell throughput increases with the number of user terminals with independent channel realizations.
- For realistic SINR values, *transmission at 50 km/h will be feasible at 5 GHz carrier frequency* in FDD downlinks. This requires use of channel prediction and a tight adaptation feedback control loop. The limiting velocity at lower carrier frequencies is correspondingly larger.
- The multiuser scheduling gain is halved when TDMA is used instead of TDMA/OFDMA. The variability of the channel with frequency can then not be utilized, and the time-variability is averaged out.
- Predictive adaptation can use code and modulation rate boundaries adjusted so that bit error rate constraints are fulfilled in the presence of SINR prediction uncertainty. This method had been used and tested and has been shown to work in a variety of test scenarios. It enables efficient and robust co-optimisation of link adaptation and link retransmission strategies.
- The feasibility of adaptive transmission is limited by the channel prediction accuracy. For each link, this accuracy is determined by the SINR and by the velocity of the terminal. These two variables are thus the main variables that determine if flows may be adaptively allocated, or must instead use coding, interleaving and spreading to average the channel variations.
- The required channel state feedback data rate poses a potentially serious problem, in particular for the FDD wide-area downlink. An effective solution to this problem is to *feed back the predicted SNR, and source code it by a combination of transform coding in the frequency direction and subsampling in the time direction.*
- The time-frequency chunk size assumed in the TDMA/OFDMA FDD system was seen to be rather well adjusted to Urban Macro channel statistics, and to the vehicular velocities of interest for adaptive transmission. Doubling of the number of subcarriers included in a chunk was investigated, but turned out to be of questionable value.

It should be noted that the results were obtained under some idealized assumptions: All user terminals having the same average SNR, and a white Gaussian noise model for the interference. While the results from this study are promising, several open issues require further investigation. They have been outlined in Subsection 3.1.6. Of particular interest is to investigate the best ways to combine MIMO transmission with adaptive TDMA/OFDMA. When assigning all antenna resources within each chunk to a single terminal, results so far indicated that *multiplexing* is attractive. Multiplexing transmits different messages over the transmitters and separating them through channel inversion at the receiver. It should be used when the channel is non-singular. Multiplexing requires no extra feedback and it preserves the strong multiuser scheduling gain that was obtained for SISO channels. This is *not* the case for e.g. Alamouti space-time coding. However, MIMO channel matrices will sometimes be rank deficient, for example due to line-of-sight propagation environments. Multiplexing should then not be used. A switching strategy, which uses multiplexing where possible, and switches to another multi-antenna transmit scheme where appropriate, is a promising combination.

Single-cell performance bounds: In Section 3.2, a theoretical study was presented that was aimed at estimating the attainable cell performance when using multi-user waterfilling to adapt the transmission over uplinks. The described method to calculate the cell capacity can be used as an evaluation tool in order to determine the fundamental limits for the aggregated bitrate in the uplink. Basic implementation constraints like the set of available code and modulation rates can be included in this method in order to derive a tight upper bound on the sum bitrate. This upper bound can serve as a reference for other methods or simply show the limits of a given multiple-access channel.

The method has then been used to evaluate fundamental limits for the aggregate uplink bitrate capacity in WINNER cellular short-range and wide area modes, assuming that maximum rate scheduling is used and no fairness constraints are applied.

Multicarrier CDMA transmit power minimization. Algorithms were described in Subsection 3.2.2 that optimise the subcarrier, bit and power allocation in a OFDM-based multi-user environment with respect to different basic criteria like bitrate, transmit power and required bandwidth. This joint rate and power adaptation problem is more computationally demanding than the adaptive scheduling problem with fixed transmit power outlined in Subsection 1.1.2 and used in Section 3.1. The transmission parameters are adapted to the channel, which is assumed to be static and perfect CSI at the transmitter is available. As in Section 3.1, the adaptation will perform worse if the CSI is inaccurate or delayed and it will cease to work for fast varying channels.

Finally, issues related to the implementation of an adaptive sub-carrier allocation, bit and power loading approach in a MC-TF-CDMA downlink system were studied. Power minimization under a throughput constraint in OFDM-based downlinks was considered in Section 3.3. For the optimisation complexity, it was shown that in all considered scenarios, the evaluation of sub-carrier sharing factors takes approximately 55% of the total link adaptation computation time. From the point of view of the trade-off between achieved performance, optimisation complexity and the amount of auxiliary signalling, adaptation over sub-bands of 8 sub-carriers appears to be the most suitable solution for all scenarios. Moreover, regarding the amount of auxiliary information, we have shown that it is possible to avoid transmission of user gains by exploiting the channel knowledge at the receiver.

For the sensitivity to channel estimation errors, we observe that in all cases, mean squared channel estimation errors around 0.01 appear to be sufficient for link adaptation purposes. To achieve such accuracy, the use of one OFDM pilot symbol is sufficient in Urban Micro and Short Range scenarios, while for Suburban Macro and Urban Macro 2 and 4, OFDM pilot symbols are necessary, respectively, under the assumption that joint channel estimation is employed and 32 users are simultaneously active in the system. In other words, for the channel estimation the urban macro scenario is the one requiring the highest pilot overhead, and this had to be expected given its high frequency selectivity.

For the sensitivity to channel variations, the same sensitivity was observed in all scenarios. This result is expected, since the performance degradation is determined by the channel *time* correlation properties. However, since the system parameters, including the OFDM symbol duration, are different in the short range and in the wide area cellular scenarios, it turns out that the overhead in OFDM symbols for update of the CSI is 2.5 times larger in the wide area scenario.

This investigation, reported in Section 3.3 poses an intriguing question: When *throughput maximization* is the adaptation criterion, information-theoretic results demonstrate that exclusive (orthogonal) allocation (TDMA/OFDMA) of time-frequency resources is superior to sharing these resources, when perfect channel state information is available at the transmitter. This is one reason for TDMA/OFDMA being selected as the default multiple access scheme for adaptive transmission. However, the results of Section 3.3 seem to indicate that resource sharing by MC-CDMA leads to lower transmit power than orthogonal allocation by OFDMA. Does this imply that the *power minimization criterion* leads to a radically different conclusion than use of the throughput maximization criterion? Further studies are required here.

Outlook: The methods outlined here represent promising tools to be used in the evolution of the WINNER system concept. The continued design and evaluation work will be performed within assumptions represented by the evolving concept. It will also provide additional feedback to the concept work on the feasibility of different schemes in various scenarios. Work within other WP2 tasks need to be taken into account when evaluating new combinations of link adaptation, coding and scheduling:

- The link adaptation studied here has primarily utilized convolutional coding, to limit the effort. The use and performance when using other coding schemes such as Turbo coding and LDPC coding needs to be evaluated.
- The multi-link adaptation has been evaluated in combination with simple scheduling algorithms, using a full buffer assumption, no priorities among flows and mostly allocating based on either a maximal throughput or minimum transmit energy criterion. The considered optimisation criteria lead to tractable problems and algorithms with reasonable complexity. However, they do not reflect the complex requirements that will appear in a real system. Especially in light of the expected complex mix of different traffic type, the utilized optimisation criteria seem overly simplified. On the other hand, it is not yet clear which traffic model would have to be applied to formulate adequate optimisation criteria.

The traffic in a real system will be a mix of several traffic classes with very different properties. It will remain a challenge to jointly adapt the multi-access, coding and modulation to both the channel *and* the traffic type.

6. Appendix

6.1 Optimum multi-coded OFDM transmission

6.1.1 The description of a modified HH algorithm for COFDM

The steps of the modified HH algorithm can be described in a less formal notation as follows:

1. Let $b_n = 0$ and $p_{tot} = 0$, where b_n and p_{tot} denote the total number of information bits and the total power allocated for the transmission so far. Reset all subcarriers' power and CM scheme indexes c to zero.
2. Calculate the initial incremental power per bit $\Delta p_n = f(c,T_n)$ required to transmit an information bit using the strongest CM scheme (with index c equal to 0) for each subcarrier n. This initial incremental power depends on the actual channel state information for each subcarrier.
3. Increase a number of transmitted information bits until the total allocated power p_{tot} is equal to the power available in our system:
   - Find the smallest $\Delta p_{n,}$
   - $p_{tot} = p_{tot} + \Delta p_{,n,}$
   - assign to b_n the next bigger code rate and increment the index c (i.e. assign the next CM scheme to the n-th subcarrier)
   - Update the incremental power required to transmit one bit at the n-th subcarrier using the c+1 CM scheme in comparison with transmission using the c-th CM scheme $\Delta p_n = f(c_n,T_n)$

The subcarriers' power assignment guarantee that the BER for a given group of subcarriers assigned to a given CM scheme will not exceed an assumed value. This algorithm is optimum in the same sense as Hughes-Hartogs algorithm is – it uses the available power in the most effective way in order to transmit the highest number of information bits.
Each subcarrier can be a subject of a different attenuation and data can be transmitted on it with a different power level resulting in a different effective SNR. Thus, the subcarriers can be assigned to different modulators and encoders at the transmitter. If the effective SNR for a given subcarrier causes that the BER constraint is not met even for the strongest coding scheme this subcarrier is not used for data transmission (see the shaded area in Figure 6.1).

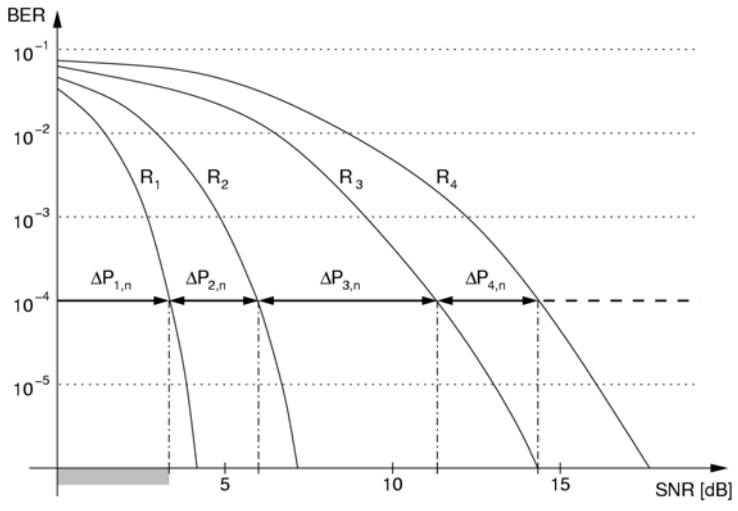


Figure 6.1: An example of reference curves for different code and modulation schemes.

The applicability of per-subcarrier modulation and coding schemes depends on the required coding gain. In order to achieve sufficient performance with respect to coding gain, an appropriately large code-block length might be necessary, and hence, a large number of OFDM symbols might be needed for signal detection.

6.1.2 Simulation results

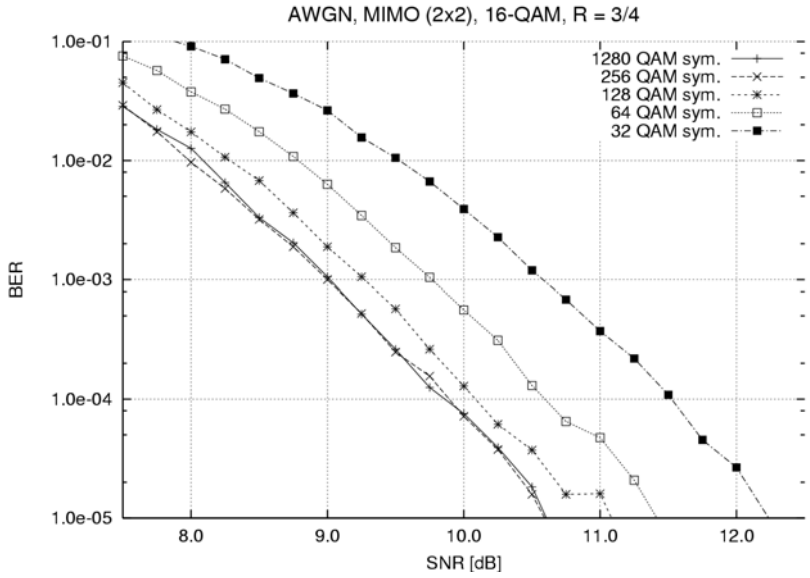


Figure 6.2: Impact of the codeword length on the bit error rate curves for MIMO transmission (2x2) in the AWGN channel (*R* = 3/4, 16-QAM)

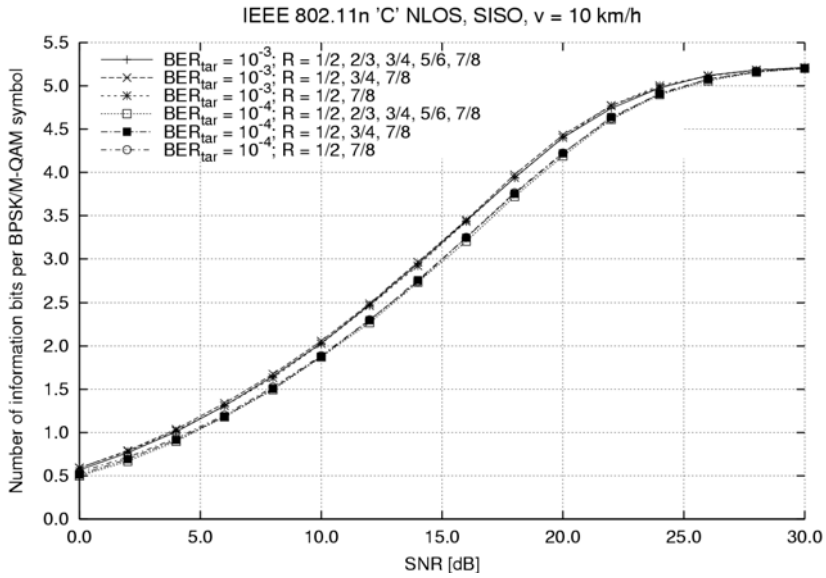


Figure 6.3: Impact of the number of coding schemes used in adaptation algorithm on the overall code rate for the IEEE 802.11n 'C' NLOS SISO channel model (*v* = 10 km/h, perfect CSI, no feedback channel delay)

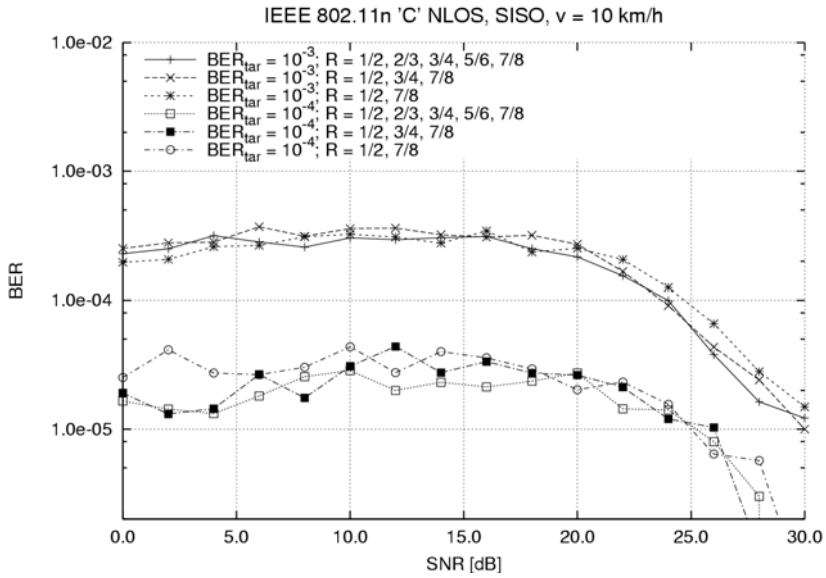


Figure 6.4: Impact of the number of coding schemes used in adaptation algorithm on the bit error rate for the IEEE 802.11n 'C' NLOS SISO channel model (v = 10 km/h, perfect CSI, no feedback channel delay)

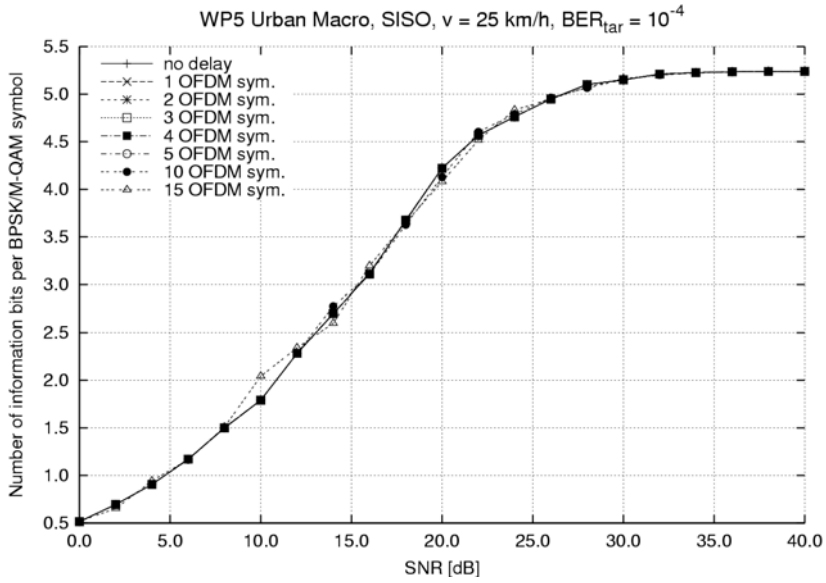


Figure 6.5: Overall code rate results for various delay of the feedback channel for the WP5 Urban Macro SISO channel model (v = 25 km/h, $BER_{tar} = 10^{-4}$, perfect CSI)

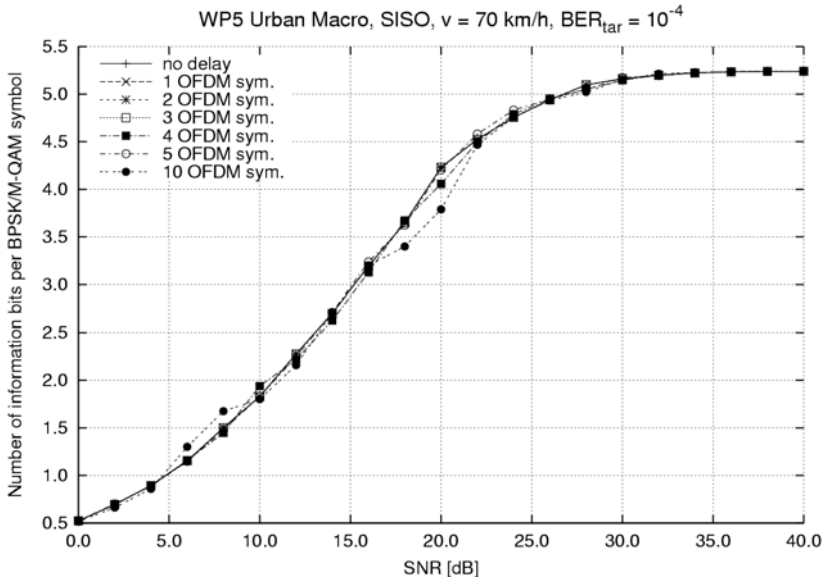


Figure 6.6: Overall code rate results for various delay of the feedback channel for the WP5 Urban Macro SISO channel model (v = 70 km/h, $BER_{tar} = 10^{-4}$, perfect CSI at Rx)

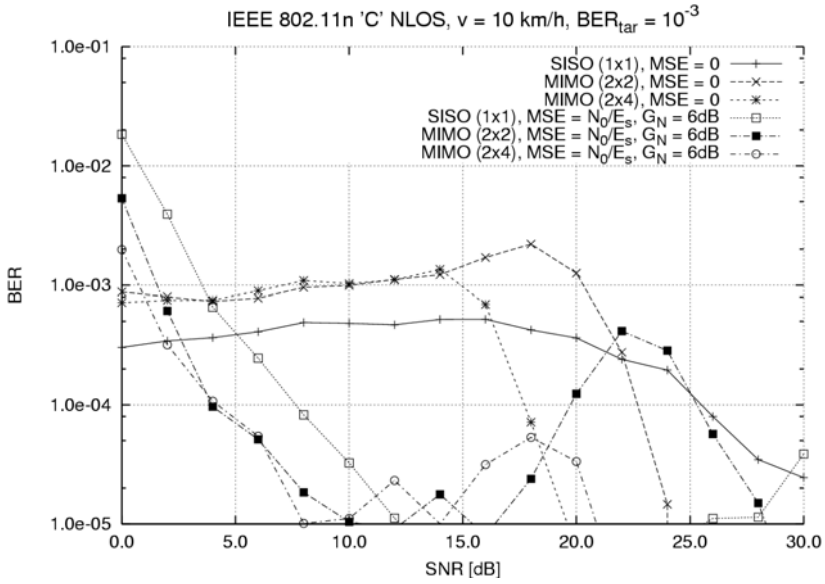


Figure 6.7: Estimation error impact on the bit error rate for the IEEE 802.11n 'C' NLOS SISO and MIMO channel models (v = 10 km/h, $BER_{tar} = 10^{-3}$, no feedback channel delay)

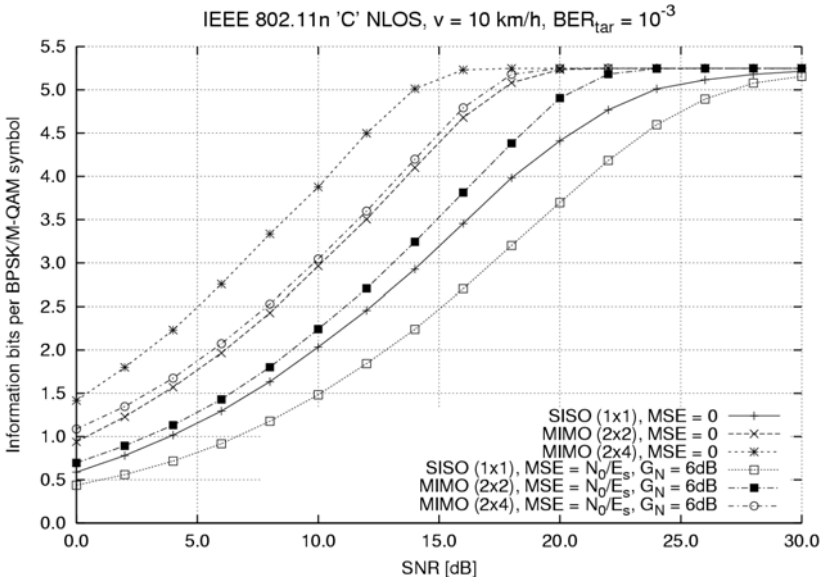


Figure 6.8: Overall code rate results degradation due to an estimation error as compared with a perfect CSI for the IEEE 802.11n 'C' NLOS SISO and MIMO channel models (v = 10 km/h, $BER_{tar} = 10^{-3}$, no feedback channel delay)

6.2 Adaptive modulation and coding based on channel prediction

The Figures below show two examples of the behaviour of rate limits for the uncoded modulation scheme discussed in Section 2.2, using BPSK, QPSK, 16-QAM and 64-QAM. The optimisation conditions are the same as for the figures in Section 2.2 for convolutional coding and modulation. Note that in the Figure below, the curves for BPSK with an imperfect prediction and QPSK with a perfect prediction overlap.

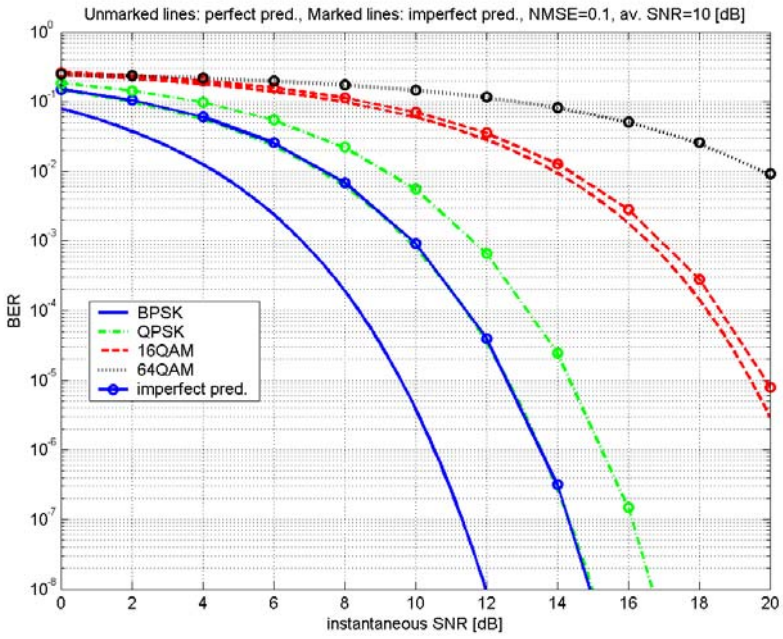


Figure 6.9: Bit error rates as functions of the instantaneous predicted SNR, for the four BPSK and M-QAM schemes. Solid lines: Results for perfect prediction. Dash-dotted lines: Attained BER, averaged over the true SNR, when using a specific rate at constant transmit power. Results for average SNR 10 dB and prediction error NMSE $\tilde{\sigma}^2$ = 0.10

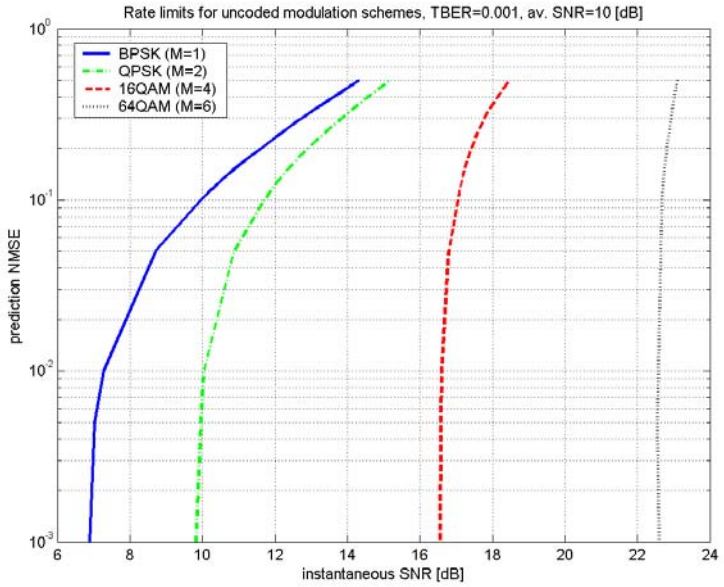


Figure 6.10: Rate limits in predicted SNR per symbol for adaptive uncoded BPSK/M-QAM, optimised for a target BER 0.001 at constant transmit power, shown as functions of the complex prediction error normalized MSE $\tilde{\sigma}^2$. Average SNR 10 dB

6.3 Adaptive approach to antenna selection and space-time block coding

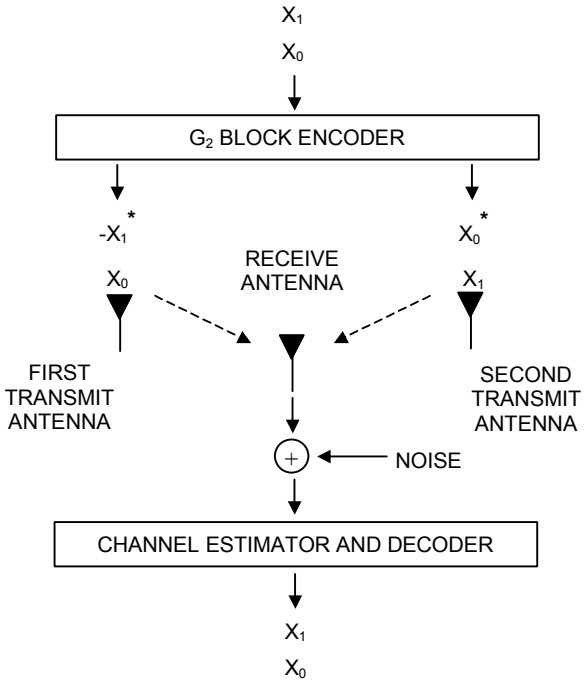


Figure 6.11: The diagram of the G_2 coded system

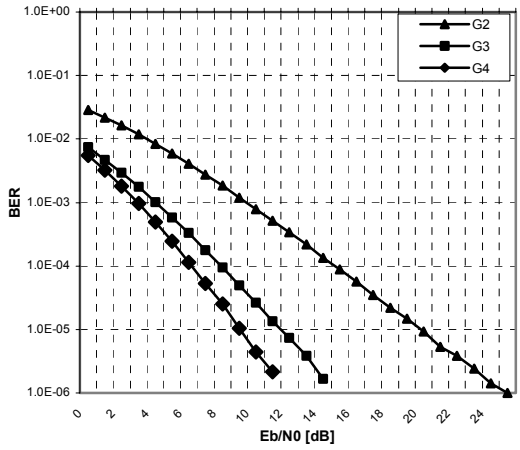


Figure 6.12: Reference results or the G_2 , G_3 ,and G_4 encoders

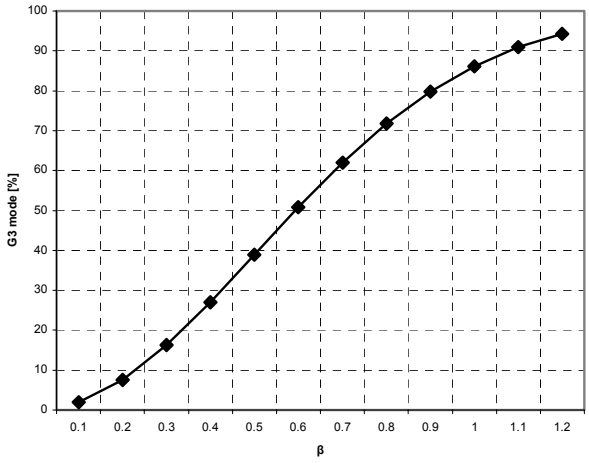


Figure 6.13: G_3 encoder usage percentage

6.4 Adaptive predictive TDMA/OFDM

6.4.1 Timing and computational complexity issues

The adaptive TDMA/OFDMA systems considered in Section 3.1 are based on three key algorithms: *prediction* of the frequency-domain baseband channel, *source coding* of the feedback information and *resource scheduling* that takes the channel quality into account. If any of these algorithms require a computational complexity that makes it infeasible to operate within the short chunk durations of 0.337 ms considered, the adaptation scheme as a whole would be infeasible. The computational complexity of the source coding scheme outlined in Subsection 3.1.4.3 is low, and need not be considered further. We will here discuss the complexity of the two other algorithms.

6.4.1.1 Computational complexity of GCG channel gain predictors

In the downlink of the considered FDD system and in the TDD system, prediction is performed at the terminals with the Generalized Constant Gain algorithm [SA03], [SLA02]. It can be formulated either as a Wiener estimation scheme or as a state estimator, and is in [SA03] formulated as an approximation of the Kalman state-space estimator, that uses a constant pre-computed gain in the prediction error feedback part instead of a time-varying gain obtained via a matrix Riccati difference equation. In the state-space

model that describes the channel evolution, fading in time is described by an autoregressive model of order n_D. Correlation in frequency is taken into account by using p pilot-bearing subcarriers in parallel. Only one transmit antenna per sector is assumed. Under these assumptions, eq. (21) in [SA03] gives the number of complex-valued multiplications + additions required per time step of the GCG algorithm as

$$n_p = 5pn_D + p^2 n_D .$$

With p = 8 pilot subcarriers per estimator and AR models of order n_D = 4, n_p = 416. With both the FDD downlink and the TDD chunk designs, use of p = 8 means that each predictor spans four chunk widths. A total of N = 26 predictors would be required to estimate the whole bandwidth that comprises 104 chunks in both the FDD and the TDD mode. Since each multiplication of two complex numbers is equivalent to four real-valued multiplications, and since six updates are performed per chunk (using either pilots or control symbols), the total required computational effort per chunk time is $24Nn_p = 259584$ real-valued multiplications+additions per update, when N = 26 and n_p = 416. With a chunk duration of 0.3372 ms, this corresponds to a computational load of 770 M real mult+add operations/s. This kind of computational load for a component of the baseband signal processing would be considered unrealistic in todays mobile phones. It should be well within realistic limits 10 years into the future for handheld devices as well as laptop-like terminals.

Part of the algorithm needs to be executed within 0.10 ms after the last measurement has been obtained, due to timing constraints in the feedback loop. This poses only relatively small additional requirements on the required computation performance.

The computational complexity is proportional to the number N of active predictors. When the terminal is in competition for a contention band that comprises only part of the total bandwidth, *the computation requirements decrease proportionally.*

The *FDD uplink* prediction, based on overlapping pilots, at the moment poses more of a problem. Due to the higher state-space dimensionality, the algorithm for computing the constant adaptation gain used by the GCG algorithm here encounters numerical difficulties. For that reason, we have utilized the full Kalman predictor solution so far. This algorithm is in its basic form too complex, even considering that much higher computational complexity can be accepted at FDD wide-area access points than in user terminals. However, there are good possibilities to reduce its complexity by using the special structure of the state updated matrices.

6.4.1.2 Computational complexity of algorithms for resource scheduling

Resource allocation and scheduling problems often lead to NP-hard integer programming problems of hopeless computational complexity. With a MAC slot that equals one single chunk time of 0.337 ms, a scheduling round was required to be completed in 0.060 ms. A guess of the available computational resources for scheduling at a WINNER access point could be $20*10^9$ operations per second, or $1.2*10^6$ operations for each scheduling round of 0.060 ms. The number of chunks within each MAC slot is 104 in both the wide-area and short range modes. A subset of these resources is available for use by the adaptive resource scheduler. The number of flows to be scheduled varies widely, but may be up to 100. However, each flow may be in competition for only a subset of the chunks, due to terminal complexity constraints and transmit power constraints. This gives some indications on the expected size of the scheduling problem.

Fortunately, there exist several powerful and efficient scheduling algorithms that have very low computational complexity. Most of them have the attractive property that the performance criterion can be evaluated locally for each chunk [Eri04]. A scheduling round would thus require up to 100 criterion evaluations (for different flows) that are performed locally for at most 104 chunks per spatial virtual beam. The required 10000 evaluations per spatial virtual beam should be easy to accomplish with $1.2*10^6$ operations available. For an example of an estimation and comparison of computational complexities, see Figure 6.4 in [Eri04].

6.4.2 Additional results: Packet error rates

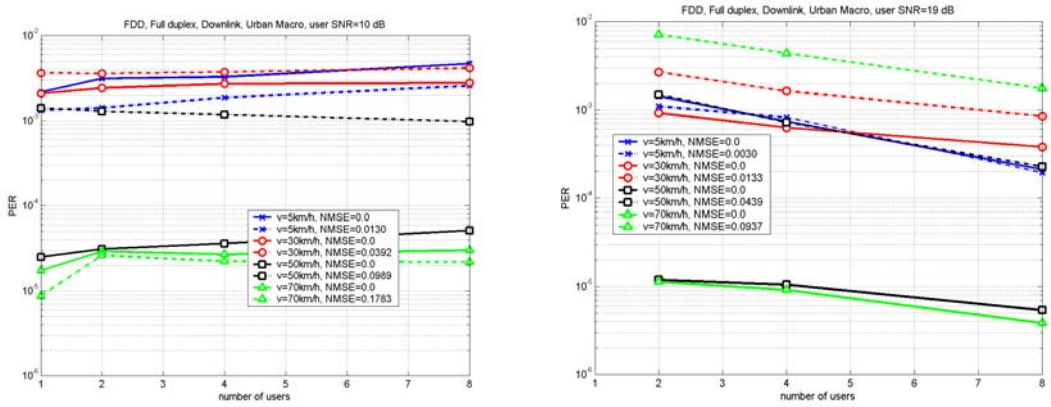


Figure 6.14: Packet error rates for TDMA/OFDMA corresponding to the experiments in Figures 3.13 and 3.14, for 10 dB SNR (left) and 19 dB SNR (right) for all users. Packets of size 512 bits are used

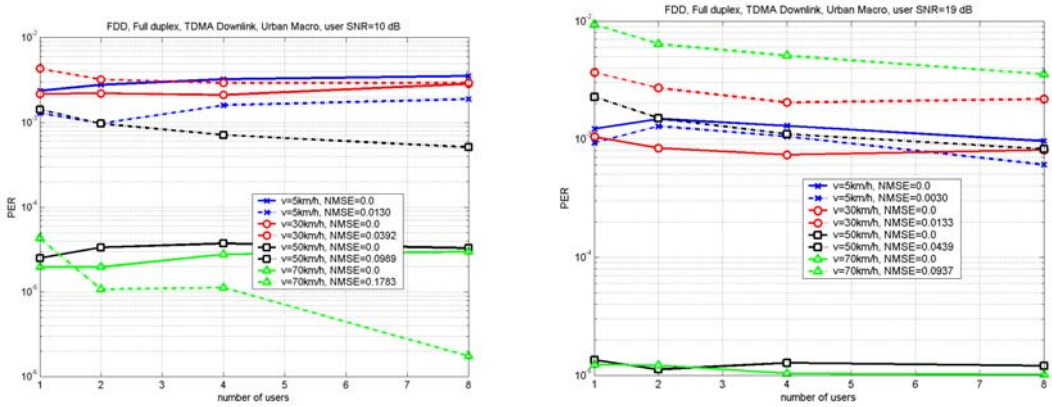


Figure 6.15: Packet error rates when using a TDMA scheduling constraint, corresponding to the experiments in Figures 3.16 and 3.17, for 10 dB SNR (left) and 19 dB SNR (right) for all users. Packets of size 512 bits are used

References

[ACN02] N. Al-Dhahir, A. R. Calderbank, and A. F. Naguib, “Space-time codes for wireless communications,” in The Wiley Encyclopedia of Telecommunications, New York: John Wiley and Sons Inc., 2002

[Ala98] S. M. Alamouti, “A simple transmitter diversity scheme for wireless communications,” *IEEE J. Select. Areas Commun*., vol. 16, pp. 1451–1458, October 1998.

[AM79] B.D.O. Anderson and J.B. Moore, Optimal Filtering. Prentice Hall, 1979.

[AP05] M. Aceña and S. Pfletschinger, “A spectrally efficient method for subcarrier and bit allocation in OFDMA”, *IEEE VTC Spring 05*, Stockholm, 3 May – 1 June 2005.

[BGS+04] D. Baum, G. Del Galdo, J. Salo, P. Kyösti, M. Milojevic “Tapped Delay Line (TDL) Models a.k.a. Link Level (LL) Models”, IST-2003-507581 WINNER, October 2004

[Bin90] J. A. C. Bingham, “Multicarrier Modulation for Data Transmission: an Idea whose Time Has Come”, *IEEE Comm. Mag*., Vol. 28, pp. 5-14, May 1990

[Cam99] J. Campello, “Practical bit loading for DMT”, *IEEE ICC*, June 1999

[CCB95] P. S. Chow, J. M. Cioffi, John A. C. Bingham, “A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels”, *IEEE Trans. Commun*., vol. 43, no. 2/3/4, pp. 773-775, Feb./Mar./Apr. 1995

[CG01] S. T. Chung and A. J. Goldsmith, “Degrees of freedom in adaptive modulation: an unified view”, *IEEE Trans. on Communication*, vol. 49, pp. 1561-1571.

[CT91] T. M. Cover, J. A. Thomas: *Elements of Information Theory*. John Wiley & Sons, 1991.

[CV93] R. Cheng and S. Verdú, “Gaussian multiaccess channels with ISI: capacity region and multiuser water-filling”, *IEEE Trans. on Information Theory*, vol. 39 , no. 3 , pp. 773 – 785, May 1993.

[Czy96] A. Czylwik, “Adaptive OFDM for wideband radio channels”, *IEEE Globecom’96*, pp. 713-718, Nov. 1996.

[Dig96] S. H. Diggavi, “Multiuser DMT: A multiple access modulation scheme”, *IEEE Globecom*, Nov. 1996.

[DPW02] Z. Dlugaszewski, A. Piatyszek, K. Wesolowski, “Adaptive Multi-Coded OFDM Transmission for Selective Fading Channels”, In Proceedings of the *IST Mobile & Wireless Communications Summit,* Aveiro, Portugal, June 2003, pp. 103-107

[DPW04] Z. Dlugaszewski, A. Piatyszek, K. Wesolowski, “Optimum Multi-Coded OFDM Transmission”, In Proceedings of the *9th International OFDM-Workshop*, Dresden, Germany, September 2004, pp. 187-191

[Ekm00] T. Ekman, Prediction of Mobile Radio Channels. Licenciate Thesis, Signals and Systems, Uppsala Univ., Sweden. Online: http://www.signal.uu.se/Publications/abstracts/l002.html

[Ekm02] T. Ekman, *Prediction of Mobile Radio Channels. Modelling and Design*. Ph.D. Th., Signals and Syst., Uppsala Univ. http://www.signal.uu.se/Publications/abstracts/a023.html

[EKS+99] T. Ekman, G. Kubin, M. Sternad and A. Ahlén, “Quadratic and linear filters for mobile radio prediction”, *IEEE Vehicular Technology Conference - VTC'99-Fall* Amsterdam, The Netherlands, September 19-22 1999, pp. 146-150.

[Eri04] N. C. Ericsson, „*Revenue Maximization in Resource Allocation. Applications in Wireless Communication Networks*”, Ph.D. Thesis, Uppsala University, Sweden, October 2004. Online: http://www.signal.uu.se/Publications/abstracts/a043.html

[ESA02] T. Ekman, M. Sternad and A. Ahlen, "Unbiased power prediction on broadband channels" *IEEE VTC 2002-Fall*, Vancouver, Canada, Sept. 2002.

[FH96] R. Fischer, J. Huber, “A new loading algorithm for discrete multitone transmission”. *IEEE Globecom ’96*, pp. 724-728, Nov. 1996.

[FSE+04] S. Falahati, A. Svensson, T. Ekman and M. Sternad, "Adaptive modulation systems for predicted wireless channels," *IEEE Trans. on Communications*, vol. 52, Feb. 2004, pp. 307-316.

[FSS+03a] S. Falahati, A. Svensson, M. Sternad, T. Ekaman, "Adaptive modulation systems for predicted wireless channels," *IEEE Globecom '03*, Dec. 2003.

[FSS+03b] S. Falahati, A. Svensson, M. Sternad and H. Mei, "Adaptive Trellis-coded modulation over predicted flat fading channels," *IEEE VTC 2003-Fall*, Orlando, Fla, Oct. 2003.

[Gal68] R. G. Gallager: *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.

[GC97] A. J. Goldsmith, S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Comm.*, vol. 45, no. 10, Oct. 1997.

[GE01] M. Effros and A. J. Goldsmith, "The Capacity Region of Broadcast Channels with Intersymbol Interference and Colored Gaussian Noise", *IEEE Transactions on Information Theory*, 47(1), January, 2001

[Hoh92] P. Hoher, "A statistical discrete-time model for the WSSUS multipath channel", *IEEE Transactionson Vehicular Technology*, 41(4), November 1992

[Hug87] D. Hughes-Hartogs: *Ensemble modem structure for imperfect transmission media*. US patent 4 679 227, filed May 20, 1985, issued July 7, 1987.

[HZF04] P. Herhold, E. Zimmerman, G. Fettweis, "On the Performance of Cooperative Amplify-and-Forward Relay Networks", *Proc. ITG Conference on Source and Channel Coding (SCC)*, Erlangen, Germany, 14-16 Jan 2004

[IEEE04] IEEE 802.11-03/940r2 "IEEE P802.11 Wireless LANs, TGn Channel Models", January 2004

[Joh04a] M. Johansson, "Diversity-enhanced equal access- Considerable throughput gains with 1-bit feedback," *IEEE SPAWC*, Workshop on Signal Processing in Wireless Communications, Lisbon, Portugal, July 11-14, 2004.

[Joh04b] M. Johansson, "Resource Allocation Under Uncertainty: Applications in Mobile Communications", Ph.D. Thesis, Signals and Systems, Uppsala Univ., Sweden, Oct. 2004. Online: http://www.signal.uu.se/Publications/abstracts/a041.html

[KLL03] D. Kivanc, G. Li, H. Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE Trans. Wireless Commun.*, vol. 2, no. 6, Nov. 2003.

[KRJ00] B. S. Krongold, K. Ramchandran, D. L. Jones: "Computationally efficient optimal power allocation algorithms for multicarrier communication systems". *IEEE Trans. Communications*, vol. 48, no. 1, pp. 23-27, Jan. 2000.

[KSH00] T. Kailath, A.H. Sayed, B. Hassibi, *Linear Estimation*. Prentice-Hall 2000.

[Lan03] J. N. Laneman, G. W. Wornell, "Distributed Space-Time Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks", *IEEE Transactions On Information Theory*, Vol. 49, No. 10, Oct. 2003, pp. 2415-2425

[LCL+99] S. K. Lai, R. S. Cheng, K. B. Letaief, R. D. Murch: "Adaptive trellis coded MQAM and power optimisation for OFDM transmission". *IEEE VTC Spring '99*, May 1999.

[LG01] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels–part I: Ergodic capacity", *IEEE Transactions on Information Theory*, 47(3), March, 2001

[Llo82] S P. Lloyd, "Least squares quantization in PCM." *IEEE Transactions on Information Theory*: 129-136

[Max60] J. Max, "Quantization for minimum distortion", *IRE Trans. on Information Theory*, 7-12.

[MPS02] G. Münz, S. Pfletschinger and J. Speidel, "An efficient waterfilling algorithm for multiple access OFDM", *IEEE Globecom '02*, Nov. 2002.

[MWS+02] I. Maniatis, T. Weber, A. Sklavos, Y. Liu, E. Costa, H. Haas and E. Schulz, "Pilots for joint channel estimation in multi-user OFDM mobile radio systems", In *Proceedings of the 7th International OFDM-Workshop*, 2002

[MZC+04] S. Mangold, Z. Zhong, K. Challapali, C.-T. Chou, "Spectrum agile radio: radio resource measurements for opportunistic spectrum usage", *IEEE Globecom '04*, Dec. 2004.

[PAE04] J. F. Paris, M. C. Aguayo-Torres, J. T. Entrambasaguas, "Impact of channel estimation error on adaptive modulation performance in flat fading", *IEEE Transactions On Communications*, 52(5), May 2004

[PMS02] S. Pfletschinger, G. Münz, J. Speidel: "Efficient subcarrier allocation for multiple access in OFDM systems". *7th International OFDM Workshop*, Hamburg, Sept. 2002.

[Pro02] J. Proakis, *Digital Communications*, McGraw Hill, 2002

[PWW05] A. Piatyszek, M. Wodczak, K. Wesolowski, "Multi-Coded OFDM Transmission Using Space-Time Block Codes", *XI National Symposium of Radio Science (URSI)*, Poznan, Poland, April 2005

[Rap96] T. S. Rappaport, *Wireless Communications Principles & Practice*. Prentice Hall, Upper Saddle River, New Jersey, 1996

[SA03] M. Sternad and D. Aronsson, "Channel estimation and prediction for adaptive OFDM downlinks," *IEEE VTC 2003-Fall*, Orlando, Fla, Oct. 2003.

[SA05] M. Sternad and D. Aronsson, "Channel estimation and prediction for adaptive OFDMA/TDMA uplinks, based on overlapping pilots", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2005)*. Philadelphia, PA, USA, March 19-23 2005. Online: http://www.signal.uu.se/Publications/abstracts/c0501.html

[Say00] K. Sayood, *Introduction to Data Compression.* Morgan Kaufman Publishers.

[SEA01] M. Sternad, T. Ekman and A. Ahlén, "Power prediction on broadband channels", *IEEE Vehicular Technology Conference VTC01-Spring*, Rhodes, Greece, May 6-9 2001.

[SF04] M. Sternad, S. Falahati, "Maximizing throughput with adaptive M-QAM based on imperfect channel predictions," *IEEE PIMRC 2004*, Barcelona, Sept. 2004.

[Sha48] C.E. Shannon, A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379-423, 623-656.

[SLA02] M. Sternad, L. Lindbom and A. Ahlén, "Wiener design of adaptation algorithms with time-invariant gains," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1895-1907, August 2002.

[SMW+01] A. Sklavos, I. Maniatis, T. Weber, P. W. Baier, E. Costa, H. Haas and E. Schulz, "Joint channel estimation in multi-user OFDM systems", In *Proceedings of the 6th International OFDM-Workshop*, 2001

[SOA+03] M. Sternad, T. Ottosson, A. Ahlen and A. Svensson, "Attaining both coverage and high spectral efficiency with adaptive OFDM downlinks". *IEEE VTC 2003-Fall*, Orlando, Fla, Oct. 2003.

[Son02] R. V. Sonalkar, "Bit- and power-allocation algorithm for symmetric operation of DMT-based DSL modems". *IEEE Trans. Communications*, vol. 50, no. 6, pp. 902-906, June 2002.

[SS00] R. V. Sonalkar, R. R. Shively, "An efficient bit-loading algorithm for DMT applications". *IEEE Communcations Letters*, vol. 4, no. 3, pp. 80-82, Mar. 2000.

[TCS04] P.Trifonov, E.Costa, and E. Schulz, "Adaptive user allocation, bit and power loading in multi-carrier systems", *In Proceedings of the 9th International OFDM-Workshop*, Dresden, 2004

[TJC99] V. Tarokh, H. Jafarhani, A. R. Calderbank, "Space-Time Block Coding for Wireless Communications: Performance Results", *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 3, pp. 451-460, March 1999

[WCL+99] C. Y. Wong, R. S. Cheng, K. B. Letaief, R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation". *IEEE JSAC*, vol. 17, no. 10, pp. 1747-1757, Oct. 1999.

[WIND21] IST-2003-507581 WINNER, "D2.1 Identification of radio link technologies", July 2004

[WIND23] IST-2003-507581 WINNER, "D2.3 Assessment of radio-link technologies", February 2005

[WIND26] IST-2003-507581 WINNER, "D2.6 Assessment of multiple-access technologies", October 2004

[WIND28] IST-2003-507581 WINNER, "D2.8 Assessment of key enhanced radio protocols", Ed. K. Eklund, Feb. 2005

[WIND32] IST-2003-507581 WINNER, "D3.2 Description of deployment concepts for future radio scenarios integrating different relaying technologies in a cellular infrastructure including definition, assessment and performance comparison of RAN protocols for relay base stations", February 2005.

[WIND52] IST-2003-507581 WINNER, "D5.2 Determination of Propagation Scenarios", July 2004

[WIND71] IST-2003-507581 WINNER, "D7.1 System requirements".

[WIP] Wireless IP project, Supported by the Swedish Foundation of Strategic Research (SSF). Online: http://www.signal.uu.se/Research/PCCwirelessIP.html

[WJ04] T. A. Weiss and J. K. Jondral, "Spectrum pooling: An innovative strategy for the enhancement of spectrum efficiency", *IEEE Commun. Mag.*. pp. S8-14, March 2004.

[Wod05] M. Wodczak, "On the Adaptive Approach to Antenna Selection and Space-Time Coding in Context of the Relay Based Mobile Ad-hoc Networks", accepted for presentation at the *XI National Symposium of Radio Science (URSI)* organised in Poznan, Poland, 7-8th of April, 2005

[WOS+03] W. Wang, T. Ottosson, M. Sternad, A. Ahlen and A. Svensson, "Impact of multiuser diversity and channel variability on adaptive OFDM*," IEEE VTC 2003-Fall,* Orlando, Fla, Oct. 2003.

[WTC+99] Ch. Y. Wong, C. Y. Tsui, R. S. Cheng, K. B. Letaief, "A real-time sub-carrier allocation scheme for multiple access downlink OFDM transmission". *IEEE VTC Fall '99*, Amsterdam, The Netherlands, pp. 1124-1128, Sept. 1999.

[YC01] W. Yu, J. M. Cioffi, "On constant power water-filling". *ICC '01*, June 2001.

[YL00] H. Yin, H. Liu: "An efficient multiuser loading algorithm for OFDM-based broadband wireless systems". *Globecom '00*, San Francisco, USA, pp. 103-107, Nov. 2000.

[YRB+04] W. Yu, W. Rhee, S. Boyd and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels", *IEEE Trans. on Information Theory*, vol. 50, pp. 145-152, Jan 2004.

[Yu02] W. Yu, *Competition and Cooperation in Multi-User Communication Environments*, Ph.D. Thesis, Stanford University 2002.

[Zha04] Y. Zhang, "Signalling overhead for ASBA in MC-CDMA system", Master's thesis, Munich University of Technology, 2004

[ZX03] Y. R. Zheng, C. Xiao, "Simulation Models with Correct Statistical Properties for Rayleigh Fading Channels", *IEEE Transactions on Communications*, Vol. 51, No. 6, pp. 920-928, June 2003